1 Wzór Taylora dla funkcji wielu zmiennych

1.1 Wzór Taylora

Niech $\mathcal{O} \subset \mathbb{R}^N$ – zbiór otwarty. Niech $f : \mathcal{O} \to \mathbb{R}$ – funkcja klasy C^r (tzn. różniczkowalna r razy i $r-$te pochodne są ciągłe). Niech $x, x_0 \in \mathcal{O}$, $h = x - x_0$, przy czym niech x, x_0 będą takie, aby $x_0 + \theta h \in \mathcal{O}$ dla $0 \leqslant \theta \leqslant 1$.

Utwórzmy funkcję pomocniczą

$$\varphi(t) = f(x_0 + th), \qquad t \in [0, 1]. \tag{1}$$

Z własności f wynika, że φ jest ciągła na $[0, 1]$ oraz różniczkowalna r razy w sposób ciągły na $]0, 1[$. Policzmy kolejne pochodne tej funkcji.

$$\varphi(t) = f(x_0^1 + th^1, x_0^2 + th^2, \dots, x_0^N + th^N),$$

$$\varphi'(t) = \sum_{i=1}^{N} \frac{\partial f}{\partial x^i}(x_0 + th)h^i,$$

$$\varphi''(t) = \sum_{i_1=1}^{N} \sum_{i_2=1}^{N} \frac{\partial^2 f}{\partial x^{i_1} \partial x^{i_2}}(x_0 + th)h^{i_1}h^{i_2},$$

$$\vdots$$

$$\varphi^{r-1}(t) = \sum_{i_1, i_2, \dots, i_{r-1}} \frac{\partial^{r-1} f}{\partial x^{i_1} \partial x^{i_2} \dots \partial x^{i_{r-1}}}(x_0 + th)h^{i_1}h^{i_2} \dots h^{i_{r-1}},$$

$$\varphi^r(t) = \sum_{i_1, i_2, \dots, i_r} \frac{\partial^r f}{\partial x^{i_1} \partial x^{i_2} \dots \partial x^{i_r}}(x_0 + th)h^{i_1}h^{i_2} \dots h^{i_r}.$$

Napiszmy dla φ wzór Taylora dla przyrostu argumentu równego 1 i z resztą w postaci Lagrange'a:

$$\varphi(1) = \varphi(0) + \frac{1}{1!}\varphi'(0) + \frac{1}{2!}\varphi''(0) + \dots + \frac{1}{(r-1)!}\varphi^{r-1}(0) + \frac{1}{r!}\varphi^r(\theta)$$

(tu $\theta \in [0, 1]$).

Wstawiając otrzymane wyżej wzory na pochodne φ, otrzymujemy:

$$f(x) = f(x_0) + \frac{1}{1!}\sum_{i=1}^{N} \frac{\partial f}{\partial x^i}(x_0)h^i + \frac{1}{2!}\sum_{i_1=1}^{N} \sum_{i_2=1}^{N} \frac{\partial^2 f}{\partial x^{i_1} \partial x^{i_2}}(x_0)h^{i_1}h^{i_2} + \dots$$

$$+ \frac{1}{(r-1)!} \sum_{i_1, i_2, \dots, i_{r-1}} \frac{\partial^{r-1} f}{\partial x^{i_1} \partial x^{i_2} \dots \partial x^{i_{r-1}}}(x_0)h^{i_1}h^{i_2} \dots h^{i_{r-1}} + R_r, \tag{2}$$

gdzie R_r jest *resztą $r-$tego rzędu*:

$$R_r = \frac{1}{r!} \sum_{i_1, i_2, \dots, i_r} \frac{\partial^r f}{\partial x^{i_1} \partial x^{i_2} \dots \partial x^{i_r}}(x_0 + \theta h)h^{i_1}h^{i_2} \dots h^{i_r}. \tag{3}$$

I to jest już kompletny wzór Taylora.

Przykł. Rozwinięcie w szereg Taylora funkcji $f(x, y) = \ln(1 + x + 2y)$. Dwie postaci zapisu wyższych pochodnych.

Czasem może nam przyjść ochota na oszacowanie reszty. Podamy tu takie proste oszacowanie.

1.2 Proste oszacowanie reszty

Stw. Weźmy kulę domkniętą $\mathcal{K} = \overline{K(x_0, \rho)}$, gdzie ρ jest takie, że $\mathcal{K} \subset \mathcal{O}$. Wtedy istnieje taka stała M, że

$$|R_r| \leqslant M||h||^r \tag{4}$$

dla wszystkich $x \in \mathcal{K}$.

Dow. Wiemy, że funkcje ciągłe na zbiorze zwartym są ograniczone; tak więc:

$$\sum_{i_1, i_2, \ldots, i_r} \frac{\partial^r f}{\partial x^{i_1} \partial x^{i_2} \ldots \partial x^{i_r}}(x) \leqslant M'$$

dla wszystkich $x \in \mathcal{K}$ i pewnej dodatniej stałej M'. Tak więc resztę R_r we wzorze (4) szacujemy przez

$$R_r \leqslant \frac{1}{r!} M' \sum_{i_1, i_2, \ldots, i_r}^{N} |h^{i_1}| \cdot |h^{i_2}| \cdot \cdots \cdot |h^{i_r}| = \frac{1}{r!} M' \sum_{i_1}^{N} |h^{i_1}| \cdot \sum_{i_2}^{N} |h^{i_2}| \cdot \cdots \cdot \sum_{i_r}^{N} |h^{i_r}|$$

$$= \frac{1}{r!} M' \left(\sum_{i=1}^{N} |h^i| \right)^r \leqslant N^{\frac{r}{2}} ||h||^r$$

Uzasadnienie ostatniej nierówności: Przypomnijmy sobie nierówność Schwarza: Zapodaje ona, że

$$\left| \sum_{i=1}^{N} a^i b^i \right| \leqslant \sqrt{\sum_{i=1}^{N} (a^i)^2} \cdot \sqrt{\sum_{i=1}^{N} (b^i)^2} = ||a|| \cdot ||b||;$$

więc mamy:

$$\sum_{i=1}^{N} |h^i| = \sum_{i=1}^{N} |h^i \cdot 1| \leqslant \sqrt{\sum_{i=1}^{N} (h^i)^2} \cdot \sqrt{\sum_{i=1}^{N} (1)^2} = ||h|| \cdot \sqrt{N}$$

Mamy więc

$$|R_r| \leqslant \frac{M' N^{\frac{r}{2}}}{r!} ||h||^r,$$

i oznaczając: $M = \frac{M' N^{\frac{r}{2}}}{r!}$, otrzymujemy wzór (4) czyli tezę.

CBDO

1.3 Morał

Podsumujmy: Wzór Taylora możemy zapisać w postaci:

$$f(x_0 + h) = [\text{wielomian stopnia } (r-1) \text{ od zmiennych } h^1, h^2, \ldots, h^N] + R_r,$$

gdzie R_r – mała stopnia wyższego niż $||h||^{r-1}$, tzn. spełniająca

$$\frac{R_r}{||h||^{r-1}} \xrightarrow{h \to 0} 0.$$

Wzór Taylora pozwala na przybliżenie skomplikowanych funkcji przez wielomiany, z którymi jest mieć do czynienia na ogół o wiele prościej.

W zastosowaniach najczęściej spotyka się zastępowanie funkcji przez wielomian pierwszego lub drugiego stopnia. Konieczność analizy wyższych potęg też zdarza się, ale rzadko.

Przykł. Energia drgań cząsteczki o dwu lub więcej atomach w pobliżu położenia równowagi. Kiedyś pisaliśmy wyrażenie na energię potencjalną układu N sprzężonych oscylatorów harmonicznych – była to *forma kwadratowa* w wychyleniach x^i z położenia równowagi:

$$Q(x^1, x^2, \dots, x^N = \sum_{i=1}^{N} \sum_{j=1}^{N} q_{ij} x^i x^j.$$

Nazywa się to *przybliżenie harmoniczne*. Czasem istnieje konieczność wyjścia poza to przybliżenie.

2 Ekstrema i punkty stacjonarne

Niech $\mathcal{O} \subset \mathbb{R}^N$ – zb. otwarty, niech $x_0 \in \mathcal{O}$, niech $f : \mathcal{O} \to \mathbb{R}$ **RYS.**

Def. Mówimy, że f ma w x_0 maksimum, jeżeli

$$\underset{x \in \mathcal{O}}{\forall} f(x) \leqslant f(x_0). \tag{5}$$

Def. Mówimy, że f ma w x_0 ścisłe maksimum, jeżeli

$$\underset{x \in \mathcal{O}}{\forall} f(x) < f(x_0). \tag{6}$$

Def. Mówimy, że f ma w x_0 maksimum lokalne, jeżeli

$$\underset{\rho > 0}{\exists} \quad \underset{x \in K(x_0, \rho)}{\forall} f(x) \leqslant f(x_0). \tag{7}$$

Uwaga. Analogicznie mówimy o minimum, minimum ścisłym, minimum lokalnym, jeśli zmienimy znaki nierówności w definicjach powyżej.

Stw. Niech $\mathcal{O} \subset \mathbb{R}^N$ – zb. otwarty, niech $f \in C^1(\mathcal{O})$, niech $x_0 \in \mathcal{O}$. Niech f ma w punkcie x_0 minimum lokalne. Wtedy

$$\frac{\partial f}{\partial x^i}(x_0) = 0, \qquad i = 1, \dots, N \tag{8}$$

(tzn. wszystkie pochodne cząstkowe są równe zeru w x_0).

Dow. Dla większej jasności zapiszmy tu jawnie współrzędne punktu x_0:

$$x_0 = (x_0^1, x_0^2, \dots, x_0^N).$$

Warunek (7) na minimum lokalne można przeformułować mówiąc, że $f(x_0+h)-f(x_0) \geqslant 0$ dla *dowolnego* wektora przyrostu h. Skoro tak, to weźmy wektor przyrostu posiadający tylko *pierwszą* składową różną od zera, a wszystkie pozostałe równe zeru. W ten sposób, $f(x_0+h) - f(x_0)$ jest funkcją tylko *jednej* zmiennej x^1. Przypomnijmy sobie teraz tw. dla funkcji jednej zmiennej $F(x)$ mówiące, że jeżeli $F(x)$ posiada w x^* maksimum lokalne, to $F'(x^*) = 0$. W naszej wersji oznacza to, że $\frac{\partial f}{\partial x^1}(x_0) = 0$.

Weźmy teraz wektor przyrostu h o niezerowej *drugiej* składowej, a wszystkich pozostałych równych zeru. Analogiczne rozumowanie prowadzi do wniosku, że $\frac{\partial f}{\partial x^2}(x_0) = 0$. Itd. W ten sposób otrzymujemy (8).

CBDO

Z teorii funkcji jednej zmiennej przypominamy sobie, że warunek $F'(x^*)$ był warunkiem *koniecznym*, ale *nie dostatecznym* na to, aby F posiadała w punkcie x^* maksimum. Analogicznie jest w przypadku funkcji wielu zmiennych: Warunek (8) jest warunkiem koniecznym, aby w x_0 istniało maksimum (mówi o tym powyższe Stwierdzenie), ale implikacja: (8) $\Longrightarrow$ (f posiada maksimum w x_0) na ogół *nie jest* prawdziwa.

Def. Niech $f \in C^1(\mathcal{O})$, $\mathcal{O}$ – zbiór otwarty w $\mathbb{R}^N$. Mówimy, że f ma w $x_0 \in \mathcal{O}$ *punkt krytyczny* (zwany też *stacjonarnym*), jeśli

$$\frac{\partial f}{\partial x^i}(x_0) = 0, \qquad i = 1, \ldots, N \tag{9}$$

W przypadku funkcji jednej zmiennej można było podać kryterium na to, aby punkt krytyczny był maksimum (minimum); był to warunek, aby *druga pochodna* funkcji w punkcie krytycznym była mniejsza (większa) od zera[1].

W przypadku funkcji wielu zmiennych również można podać warunek dostateczny na to, aby punkt krytyczny był maksimum (minimum). Jest to jednak bardziej skomplikowane niż w przypadku funkcji jednej zmiennej, i aby ten warunek podać, przypomnimy najsampierw kilka faktów z zakresu teorii *form kwadratowych.*

Def. Niech $k \in \mathbb{R}^N$. *Formą kwadratową* na $\mathbb{R}^N$ nazywamy funkcję

$$\omega(k) = \sum_{i,j=1}^{N} \omega_{ij} k^i k^j \tag{10}$$

Współczynniki występujące w powyższym wyrażeniu tworzą *macierz* formy kwadratowej:

$$\begin{pmatrix} \omega_{11} & \omega_{12} & \ldots & \omega_{1N} \\ \omega_{21} & \omega_{22} & \ldots & \omega_{2N} \\ \vdots & \vdots & \ldots & \vdots \\ \omega_{N1} & \omega_{N2} & \ldots & \omega_{NN} \end{pmatrix} \quad - \quad \text{macierz symetryczna:} \quad \omega_{ij} = \omega_{ji} \tag{11}$$

Dla macierzy formy kwadratowej (11) zdefiniujmy następujące liczby $D_1, D_2, \ldots, D_N$.

$$D_1 = \omega_{11} \tag{12}$$

$$D_2 = \det \begin{pmatrix} \omega_{11} & \omega_{12} \\ \omega_{21} & \omega_{22} \end{pmatrix} \tag{13}$$

$$D_3 = \det \begin{pmatrix} \omega_{11} & \omega_{12} & \omega_{13} \\ \omega_{21} & \omega_{22} & \omega_{23} \\ \omega_{31} & \omega_{32} & \omega_{33} \end{pmatrix} \tag{14}$$

itd.,

$$D_N = \det \begin{pmatrix} \omega_{11} & \omega_{12} & \ldots & \omega_{1N} \\ \omega_{21} & \omega_{22} & \ldots & \omega_{2N} \\ \vdots & \vdots & \ldots & \vdots \\ \omega_{N1} & \omega_{N2} & \ldots & \omega_{NN} \end{pmatrix} \tag{15}$$

Tw. (Kryterium dodatniej/ujemnej określoności form kwadratowych).

[1]Nie był to warunek najogólniejszy, ale tego ogólniejszego warunku nie będziemy tu przypominać, gdyż rozszerzenie go na przypadek funkcji wielu zmiennych wymaga znacznie bardziej zaawansowanej teorii

1. Jeśli wszystkie D_i są większe od zera: $D_i > 0$ dla $i = 1, 2, \ldots, N$, to dla dowolnego niezerowego wektora $k \in \mathbb{R}^N$ zachodzi:
$$\omega(k) > 0$$
(taką formę nazywamy *ściśle dodatnią*).
2. Jeśli zachodzi: $(-1)^i D_i > 0$ dla $i = 1, 2, \ldots, N$, to dla dowolnego niezerowego wektora $k \in \mathbb{R}^N$ zachodzi:
$$\omega(k) < 0$$
(taką formę nazywamy *ściśle ujemną*).

Dow. był on już niedawno w części 'algebraicznej' wykładu.

Przykł. Niech $N = 2$. Forma $\omega_+(k) = k_1^2 + k_2^2$ jest ściśle dodatnia, forma $\omega_-(k) = -k_1^2 - k_2^2$ jest ściśle ujemna, zaś forma $\omega_{+-} = k_1 k_2$ nie jest ani dodatnia, ani ujemna.

Def. (Przypomnienie – ostać kanoniczna formy kwadratowej): Jest to taka forma, że macierz formy jest macierzą diagonalną z liczbami: $1, 0, -1$ na przekątnej. Innymi słowy,
$$\omega(k) = \sum_{i=1}^{n} (k^i)^2 - \sum_{i=n+1}^{m} (k^i)^2 \qquad (m \leqslant N).$$

Tw. Każdą formę kwadratową można przez liniową zamianę zmiennych doprowadzić do postaci kanonicznej, która jest jedyna z dokładnością do przenumerowania zmiennych (tzn. ilość plusów i minusów w postaci kanonicznej formy jest jednoznaczna).

Tw. (warunek dostateczny istnienia ekstremum). Niech $f \in C^2(\mathcal{O})$, $\mathcal{O}$ – otwarty w $\mathbb{R}^N$. Niech $x_0 \in \mathcal{O}$ – punkt krytyczny funkcji f, tzn. $\frac{\partial f}{\partial x^i}(x_0) = 0, \quad i = 1, \ldots, N$. Niech
$$D_s(x_0) = \det\left(\frac{\partial^2 f}{\partial x^i \partial x^j}(x_0) > 0 \qquad (1 \leqslant i, j \leqslant s)\right)$$
dla $s = 1, 2, \ldots, N$. Wtedy f ma w x_0 ścisłe minimum lokalne.

Dow. Wypiszmy wzór Taylora dla f do 2. rzędu, uwzględniając, że x_0 jest punktem krytycznym:
$$f(x) = f(x_0) + \frac{1}{2} \sum_{i,j=1}^{N} \frac{\partial^2 f}{\partial x^i \partial x^j}(x_0 + \theta h) h^i h^j$$
Drugie pochodne f z założenia są ciągłe, a co za tym idzie – funkcje D_s też są ciągłe, więc istnieje $\rho > 0$ takie, że $D_s(x) > 0$ dla $x =\in K(x_0, \rho)$. Dla *wszystkich* x – a więc w szczególności dla $x = x_0 + \theta h$.

CBDO

3 Ekstrema związane (warunkowe)

3.1 Rozwiązanie przez funkcje uwikłane

Często w matematyce/fizyce mamy do czynienia z sytuacją, gdy musimy znaleźć ekstrema jakiejś funkcji przy nałożonym określonym warunku. Na przykład, chcemy znaleźć prostopadłościan o możliwie największej objętości przy warunku, że pole powierzchni tego prostopadłościanu jest ustalone.

Niech $f : \mathcal{O} \subset \mathbb{R}^N \to \mathbb{R}$, $\mathcal{O}$ – otwarty. Niech $P \subset \mathcal{O}$. (dalej zazwyczaj będziemy rozważać przypadki, gdzie P jest zadany jako poziomica pewnej różniczkowalnej funkcji g).

RYS.

Def. Niech $p_0 \in P$. Mówimy, że funkcja $f(x)$ ma w punkcie p_0 minimum lokalne przy warunku, że $x \in P$, jeśli istnieje otoczenie V punktu p_0 w $\mathcal{O}$ takie, że

$$\underset{p \in V \cap P}{\forall} f(p) \geqslant f(p_0).$$

Rozpatrzmy konkretniej przypadek $N = 2$. Niech $g : \mathcal{O} \to \mathbb{R}$, g jest klasy C^1. Niech $P = \{(x, y) \in \mathbb{R}^2 : g(x, y) = 0\}$ (tzn. P jest zadany jako zerowa poziomica g). Niech będzie dana funkcja $f \in C^1(\mathcal{O})$. Szukamy ekstremów funkcji $f(x, y)$ przy warunku, że $g(x, y) = 0$.

Sytuacja, gdy szukamy ekstremum funkcji bez żadnych warunków, na ogół różni się zasadniczo od tej, gdy szukamy ekstremów przy nałożeniu jakiegoś warunku.

Przykł. $f(x, y) = 2xy, g(x, y) = x^2 + y^2 - 1$. Gdy szukamy ekstremów f bez żadnego warunku, to ekstremów nie ma: jest jeden punkt krytyczny $(0, 0)$, który jest siodłem. Rozważmy teraz sytuację, gdy szukamy ekstremów f przy warunku $g(x, y) = 0$. Sparametryzujmy okrąg przez kąt ϕ we wsp. biegunowych: $x = \cos\phi$, $y = \sin\phi$; wtedy f obcięta do okręgu jest dana równaniem: $F(\phi) = f(\cos\phi, \sin\phi) = 2\sin\phi\cos\phi = \sin 2\phi$, i funkcja F ma cztery ekstrema: dwa maksima w $\phi \in \{\frac{\pi}{4}, \frac{5\pi}{4}\}$, co odpowiada punktom na płaszczyźnie: $(\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}})$; w tych punktach wartość f jest równa 1; i dwa minima w $\phi \in \{\frac{3\pi}{4}, \frac{7\pi}{4}\}$ tzn. $(\pm\frac{1}{\sqrt{2}}, \mp\frac{1}{\sqrt{2}})$ – w tych punktach wartość f jest -1. **RYS.**

Wróćmy do ogólnego przypadku funkcji zależnej od dwóch zmiennych $f(x, y)$ przy warunku $g(x, y) = 0$. Rozwiązanie problemu znajdowania ekstremum warunkowego możemy znaleźć, posługując się niedawno udowodnionym twierdzeniem o funkcjach uwikłanych. Będziemy zakładać, że równanie: $g(x, y) = 0$ da się (przynajmniej lokalnie) rozwikłać do postaci $y = y(x)$; da się tak zrobić, gdy $\frac{\partial g}{\partial y} \neq 0$. Wstawmy uzyskaną funkcję $y(x)$ do funkcji f. Zdefiniujmy: $F(x) = f(x, y(x))$. W ten sposób, badanie ekstremów funkcji $f(x, y)$ przy warunku $g(x, y) = 0$ sprowadza się do badania ekstremów funkcji $F(x)$.

Funkcja $F(x)$ posiada punkt x_0 podejrzany o ekstremum, gdy

$$\frac{\mathsf{d}F}{\mathsf{d}x}(x_0) = 0,$$

Policzmy pochodną funkcji F. Mamy

$$\frac{\mathsf{d}F}{\mathsf{d}x} = \frac{\partial f}{\partial x}(x, y(x)) + \frac{\partial f}{\partial y}(x, y(x)) \cdot \frac{\mathsf{d}y}{\mathsf{d}x}$$

oraz

$$\frac{\mathsf{d}y}{\mathsf{d}x} = -\frac{\frac{\partial g}{\partial x}(x, y(x))}{\frac{\partial g}{\partial y}(x, y(x))}$$

Punkt (x_0, y_0) (gdzie $y_0 = y(x_0)$) będzie podejrzany o ekstremum, gdy spełniona będzie równość

$$\frac{\partial f}{\partial x}(x_0, y_0) - \frac{\partial f}{\partial y}(x_0, y_0) \cdot \frac{\frac{\partial g}{\partial x}(x_0, y_0)}{\frac{\partial g}{\partial y}(x_0, y_0)} = 0. \tag{16}$$

Jest to *jedno* równanie na *dwie* liczby x_0 oraz y_0. Pamiętajmy jednak, że mamy też drugie równanie

$$g(x_0, y_0) = 0. \tag{17}$$

3.2 Przeformułowanie – metoda mnożników Lagrange'a

Wzór (16) nie wygląda miło. Lagrange podał schemat, który w znacznie bardziej przejrzysty sposób pokazuje sposób liczenia ekstremów warunkowych.

Uzyskuje się to wprowadzając dodatkową zmienną λ (λ jest nazywane *mnożnikiem Lagrange'a*), określoną jako:

$$\lambda = \frac{\frac{\partial f}{\partial y}(x_0, y_0)}{\frac{\partial g}{\partial y}(x_0, y_0)} \tag{18}$$

Równanie (16) można wtedy zapisać jako:

$$f_x - \lambda g_x = 0, \tag{19}$$

zaś definicję mnożnika Lagrange'a λ jako

$$f_y - \lambda g_y = 0. \tag{20}$$

(pamiętając, że cały czas mamy też trzeci warunek (17)).

Wprowadźmy teraz funkcję:

$$\Phi(x, y; \lambda) = f(x, y) - \lambda g(x, y). \tag{21}$$

Warunek konieczny istnienia ekstremum związanego możemy teraz zapisać jako

$$\begin{aligned} \Phi_x(x_0, y_0) = f_x(x_0, y_0) - \lambda g_x(x_0, y_0) &= 0 & (22) \\ \Phi_y(x_0, y_0) = f_y(x_0, y_0) - \lambda g_y(x_0, y_0) &= 0 & (23) \\ g(x_0, y_0) &= 0. & (24) \end{aligned}$$

Jest to układ 3 równań na 3 niewiadome; z tego rzadko potrzebujemy znać λ, rozwiązujemy więc go tak aby wyznaczyć tylko (x_0, y_0).

Przykł. $f(x, y) = 2xy, g(x, y) = x^2 + y^2 - 1$ raz jeszcze. Mamy: $\Phi(x, y) = 2xy - \lambda(x^2 + y^2 - 1)$. Warunek konieczny istnienia ekstremum jest:

$$\begin{aligned} \Phi_x(x_0, y_0) = 0 \Longrightarrow y_0 - \lambda x_0 &= 0, & (25) \\ \Phi_y(x_0, y_0) = 0 \Longrightarrow x_0 - \lambda y_0 &= 0, & (26) \\ g(x_0, y_0) &= 0. & (27) \end{aligned}$$

Mnożąc pierwsze z powyższych równań przez y_0, drugie przez x_0 i odejmując stronami, otrzymamy: $x_0^2 - y_0^2 = 0$, co daje $x_0 = \pm y_0$. Uwzględniając teraz trzecie równanie dostajemy: $(x_0, y_0) = (\pm\frac{1}{\sqrt{2}}, \pm\frac{1}{\sqrt{2}})$ oraz $(x_0, y_0) = (\pm\frac{1}{\sqrt{2}}, \mp\frac{1}{\sqrt{2}})$ – zgodnie z tym co dostaliśmy uprzednio.

3.3 Badanie warunku dostatecznego

W przypadku znajdowania ekstremów funkcji bez nałożonych żadnych warunków, po znalezieniu punktów krytycznych, jako podejrzanych o ekstrema, trzeba było je dodatkowo zbadać, aby zobaczyć, czy są ekstremami, czy nie. Gdy mamy kandydatów na ekstrema warunkowe, również powinno się przeprowadzić analogiczne badanie. Jest to zazwyczaj bardziej skomplikowane niż w przypadku kandydatów na ekstrema 'bezwarunkowe'. Są tu trzy zasadnicze sposoby postępowania.

1. Gdy zbiór, na którym szukamy ekstremów warunkowych, jest *zwarty* (tzn. domknięty i ograniczony), to możemy skorzystać z tw. Weierstrassa mówiącego, iż funkcja na zbiorze zwartym osiąga swoje kresy. W ten sposób, gdy z metody mnożników Lagrange'a znajdziemy punkty podejrzane o ekstrema warunkowe, to liczymy wartość funkcji w tych punktach; w ten sposób znajdujemy wartość największa i najmniejszą.

   Przykł. $f(x,y) = xy, g(x,y) = x^2+y^2-2$ raz jeszcze: Poziomicą zerową funkcji g jest okrąg, więc zbiór zwarty. W znalezionych już punktach $(1,1)$ i $(-1,-1)$ wartość f wy nosi $+1$, a w punktach $(-1,1)$ i $(1,-1)$ wynosi -1. Dwa pierwsze są więc maksimami, a dwa pozostałe – minimami.

2. Badamy kandydatów na ekstrema warunkowe, korzystając z teorii funkcji uwikłanych.

3. Wyznaczamy mnożniki Lagrange'a i liczymy drugą pochodną funkcji Φ z uzyskanymi mnożnikami w znalezionych punktach krytycznych, a następnie badamy określoność tej formy kwadratowej ograniczonej do płaszczyzny stycznej do poziomicy $g(x) = 0$.

Dwie ostatnie recepty brzmią być może dość abstrakcyjnie; podam konkretniejsze przykłady w wolnej chwili.

3.4 Przypadek gdy mamy M warunków

Może się wreszcie zdarzyć, że musimy znaleźć ekstremum funkcji f przy nałożonym nie jednym, a M warunkach. Sprecyzujmy to tak:

Niech $\mathcal{O} \subset \mathbb{R}^{N+M}$, niech $f \in C^1(\mathcal{O})$. Niech $g_1, g_2, \ldots, g_M \in C^1(\mathcal{O})$ i niech

$$\mathcal{P} = \{x \in \mathbb{R}^{N+M} : g_1(x) = 0 \wedge g_2(x) = 0 \wedge \cdots \wedge g_M(x) = 0\}.$$

RYS.

Podamy teraz (ale uzasadnienie sobie darujemy[2]) sposób, w jaki znajdujemy kandydatów na ekstrema w tym przypadku. (Tu również uzasadnienie – jako materiał nadobowiązkowy – planuję w wolnej chwili napisać.)

Utwórzmy mianowicie funkcję

$$\Phi(x, \lambda) = f(x) + \lambda_1 g_1(x) + \lambda_2 g_2(x) + \cdots + \lambda_M g_M(x); \tag{28}$$

występujące tu liczby $\lambda_1, \ldots, \lambda_M$ są parametrami, zwanymi *mnożnikami Lagrange'a*. Przyrównajmy następnie do zera pochodne:

$$\frac{\partial \Phi}{\partial x^1} = \frac{\partial f}{\partial x^1} + \lambda_1 \frac{\partial g_1}{\partial x^1} + \cdots + \lambda_M \frac{\partial g_M}{\partial x^1} = 0,$$

$$\frac{\partial \Phi}{\partial x^2} = \frac{\partial f}{\partial x^2} + \lambda_1 \frac{\partial g_1}{\partial x^2} + \cdots + \lambda_M \frac{\partial g_M}{\partial x^2} = 0,$$

...

$$\frac{\partial \Phi}{\partial x^{N+M}} = \frac{\partial f}{\partial x^{N+M}} + \lambda_1 \frac{\partial g_1}{\partial x^{N+M}} + \cdots + \lambda_M \frac{\partial g_M}{\partial x^{N+M}} = 0$$

(razem $N+M$ równań) oraz

$$g_1 = 0, \quad g_2 = 0, \ldots, g_M = 0.$$

[2] Podobnie jak w filmie 'Toy Story 2' bohater negatywny Al darował sobie prysznic przed wylotem do Tokio

Razem mamy $N + M + M$ równań na $N + M + M$ niewiadomych. Rozwiązując te równania dostaniemy zestaw $x^1, \dots, x^{N+M}$ oraz $\lambda_1, \dots, \lambda_M$ (tych ostatnich zazwyczaj nie potrzebujemy). W ten sposób mamy kandydatów na ekstrema.

Przykł. Na elipsie $\frac{x^2}{4} + \frac{y^2}{9} = 1$ znaleźć punkty najmniej odległy od prostej $3x - y + 9 = 0$.

Rozw. Niech $p_1 = (x_1, y_1)$ należy do elipsy, zaś $p_2 = (x_2, y_2)$ – do prostej. Musimy znaleźć najmniejszą wartość odległości pomiędzy punktami p_1 i p_2: $d(p_1, p_2) = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}$ *przy warunkach*, że p_1 należy do elipsy, zaś p_2 do prostej.

Łatwiej będzie rozwiązywąc równoważny problem badania *kwadratu* odległości d pomiędzy punktami p_1 oraz p_2: . Musimy zatem znaleźć ekstrema funkcji

$$f(x_1, x_2, y_1, y_2) = d^2(p_1, p_2) = (x_1 - x_2)^2 + (y_1 - y_2)^2$$

przy *dwóch* warunkach:

$$\frac{x_1^2}{4} + \frac{y_1^2}{9} - 1 = 0 \quad \text{przynależność do elipsy} \tag{29}$$

oraz

$$3x_2 - y_2 + 9 = 0 \quad \text{przynależność do prostej} \tag{30}$$

Postępując we wskazany wyżej sposób, tworzymy funkcję:

$$\Phi = (x_1 - x_2)^2 + (y_1 - y_2)^2 + \lambda \left(\frac{x_1^2}{4} + \frac{y_1^2}{9} - 1 \right) + \mu(3x_2 - y_2 + 9),$$

(λ, μ – mnożniki Lagrange'a), liczymy jej pochodne, przyrównujemy do zera i rozwiązujemy powstały układ równań.

No więc do dzieła. Liczymy pochodne Φ po kolejnych zmiennych i przyrównujemy do zera:

$$\begin{array}{rcccl} \frac{\partial \Phi}{\partial x_1} & = & 2x_1 - 2x_2 + \frac{1}{2}\lambda x_1 & = & 0 \\ \frac{\partial \Phi}{\partial x_2} & = & -2x_1 + 2x_2 + 3\mu & = & 0 \\ \frac{\partial \Phi}{\partial y_1} & = & 2y_1 - 2y_2 + \frac{2}{9}\lambda y_1 & = & 0 \\ \frac{\partial \Phi}{\partial y_2} & = & -2y_1 + 2y_2 - \mu & = & 0 \end{array} \tag{31}$$

Licząc 'na sztuki' niewiadome, mamy 4 powyższe równania plus jeszcze 2 równania (29) i (30): razem 6 równań, na 6 zmiennych: $x_1, y_1, x_2, y_2, \lambda, \mu$. Wygląda to, delikatnie mówiąc, niemiło. Ważne jest, aby przy analizie takich układów nie przestraszyć się, oraz zachować świeżość spojrzenia. Na ogół wtedy okazuje się, że 'nie taki diabeł straszny'... Dodajmy bowiem do siebie dwa pierwsze z równań (31), oraz równanie trzecie i czwarte. Otrzymamy wówczas:

$$\begin{array}{rcl} \frac{1}{2}\lambda x_1 + 3\mu & = & 0, \\ \frac{2}{9}\lambda y_1 - \mu & = & 0. \end{array}$$

Teraz, mnożąc drugie z powyższych równań przez 3 i dodając stronami, wyeliminujemy μ i dostaniemy:

$$\frac{1}{2}\lambda x_1 + \frac{2}{3}\lambda y_1 = 0,$$

skąd wynika, że

$$\lambda = 0 \quad \text{lub} \quad x_1 + \frac{4}{3}y_1 = 0.$$

Okazuje się, że pierwsza możliwość (tzn. $\lambda = 0$) prowadzi do sprzeczności (niech Czytelnik spróbuje to wykazać). Druga zaś daje:

$$x_1 = -\frac{4}{3}y_1,$$

skąd po wstawieniu do pierwszego równania więzów mamy

$$\frac{4}{9}y_1^2 + \frac{1}{9}y_1^2 = 1 \quad \Longrightarrow \quad y_1 = \pm\frac{3}{\sqrt{5}},$$

skąd dla x_1 dostajemy

$$x_1 = \mp\frac{4}{\sqrt{5}}.$$

Mamy zatem dwa punkty na elipsie, podejrzane o minimalizowanie odległości:

$$p = \left(-\frac{4}{\sqrt{5}}, \frac{3}{\sqrt{5}}\right), \quad p' = \left(\frac{4}{\sqrt{5}}, -\frac{3}{\sqrt{5}}\right).$$

Który z nich jest prawdziwym minimum? Najprościej uzyskamy odpowiedź, korzystając z argumentu 'zwartościowego' (nr 1 z listy powyżej). Do tego musimy policzyć odległości od prostej do punktów p oraz p'.

Podsumujmy: **Odp.**: Punktem na elipsie najmniej odległym od prostej jest $p_{min} = (x_1, y_1) = \left(-\frac{4}{\sqrt{5}}, \frac{3}{\sqrt{5}}\right)$.