RÓŻNICZKOWANIE FUNKCJI WIELU ZMIENNYCH:
rachunek pochodnych dla funkcji wektorowych

Pochodne cząstkowe funkcji rzeczywistej wielu zmiennych

$$(x_1, x_2, \ldots, x_n) \to f(x_1, x_2, \ldots, x_n)$$

wyliczamy według wzoru

$$\frac{\partial f}{\partial x_k} = f'_{x_k} = f_{x_k} = \lim_{h \to 0} \frac{f(x_1. \ldots, x_{k-1}x_k + h, x_{k+1}, \ldots, x_n) - f(x_1.x_2, \ldots, x_n)}{h},$$

gdzie granicę obliczamy, traktując zmienne różne od x_k jak ustalone parametry. Gdy zechcemy wyliczyć pochodne cząstkowe rzędu drugiego lub rzędów wyższych, to przyjmujemy oznaczenia

$$\frac{\partial^2 f(a_1, a_2, \ldots, a_n)}{\partial x_i x_k} = \frac{\partial^2 f}{\partial x_i x_k} = f_{x_i x_k}(a_1, a_2, \ldots, a_n) = f_{x_i x_k}.$$

Macierz

$$(f_{x_1}(a_1, a_2, \ldots, a_n), f_{x_2}(a_1, a_2, \ldots, a_n), \ldots, f_{x_n}(a_1, a_2, \ldots, a_n)) = f'(a_1, a_2, \ldots, a_n)$$

nazywamy *pochodną* funkcji f w punkcie $(a_1, a_2, \ldots, a_n)$: czasami taką pochodną oznaczamy krótko f'. W literaturze pochodną funkcji f w punkcie $(a_1, a_2, \ldots, a_n)$ bywa określane odwzorowanie liniowe, które wektorowi $(x_1, x_2, \ldots, x_n)$ przyporządkowywuje liczbę (wektor: w przypadku funkcji wielowartościowej)

$$f_{x_1}(a_1, a_2, \ldots, a_n)x_1 + f_{x_2}(a_1, a_2, \ldots, a_n)x_2 + \ldots + f_{x_n}(a_1, a_2, \ldots, a_n)x_n.$$

Funkcja wielowartościowa wielu zmiennych to funkcja F, która wektorowi $(x_1, x_2, \ldots, x_n)$ przyporządkowywuje wektor

$$\left(f^1(x_1, x_2, \ldots, x_n), f^2(x_1, x_2, \ldots, x_n), \ldots, f^k(x_1, x_2, \ldots, x_n)\right).$$

Innymi słowy: funkcja wielowartościowa wielu zmiennych to funkcja wektorowa, czyli funkcja która wektorowi przyporządkowywuje wektor. Macierz

$$F' = F^{(1)} = \begin{pmatrix} f^1_{x_1} & f^1_{x_2} & \cdots & f^1_{x_n} \\ f^2_{x_1} & f^2_{x_2} & \cdots & f^2_{x_n} \\ \vdots & \vdots & & \vdots \\ f^k_{x_1} & f^k_{x_2} & \cdots & f^k_{x_n} \end{pmatrix}$$

przedstawia pochodną takiej funkcji. Kolejne wiersze tej macierzy to pochodne funkcji rzeczywistych, które wektorowi $(x_1, x_2, \ldots, x_n)$ przyporządkowywują kolejne współrzędne wektora $F(x_1, x_2, \ldots, x_n)$. Dla funkcji rzeczywistej wielu zmiennych $f(x_1, x_2, \ldots, x_n)$ pochodna jest funkcją wektorowa: wartościami są wektory o tylu współrzędnych ile jest zmiennych. Stąd macierz

$$f" = f^{(2)} = \begin{pmatrix} f_{x_1x_1} & f_{x_1x_2} & \cdots & f_{x_1x_n} \\ f_{x_2x_1} & f_{x_2x_2} & \cdots & f_{x_2x_n} \\ \vdots & \vdots & & \vdots \\ f_{x_kx_1} & f_{x_kx_2} & \cdots & f_{x_kx_n} \end{pmatrix}$$

przedstawia drugą pochodną takiej funkcji. W analogiczny sposób wyznaczamy pochodne wyższych rzędów. O funkcji która ma pochodną mówimy, że jest *różniczkowalna.* Gdy taka funkcja na pochodną rzędu n, to mówimy, iż jest n-krotnie różniczkowalna.

Gdy $f(x, y, z) = xyz$, to pierwszą pochodną jest macierz

$$f' = (yz, xz, xy);$$

zaś macierz

$$f" = \begin{pmatrix} 0 & z & y \\ z & 0 & x \\ y & x & 0 \end{pmatrix}$$

jest drugą pochodną.

Twierdzenie. *Jeśli pochodne rzędu drugiego $f_{x_ix_k}$ oraz $f_{x_kx_i}$ są ciągłe na kostce do której należy punkt $(a_1, a_2, \ldots, a_n)$, to są równe w tym punkcie:* kostka to zbiór punktów (wektorów)

$$\{(y_1, y_2, \ldots, y_n) : b_1 < y_1 < c_1, b_2 < y_2 < c_2, \ldots \text{ oraz } b_n < y_n < c_n\}.$$

Twierdzenie to wymusza umiejętność sprawdzania: **Kiedy funkcja wektorowa jest ciągła?** Dopiero po sprawdzeniu ciągłości pochodnych cząstkowych drugiego rzędu można stosować to twierdzenie. Przykłady braku potrzebnej ciągłości można znaleść wśród niezbyt skomplikowanych funkcji wymiernych.

Gdy

$$f(x, y) = \begin{cases} xy\dfrac{x^2 - y^2}{x^2 + y^2}, & \text{gdy } x^2 + y^2 > 0; \\ 0, & \text{gdy x=0=y,} \end{cases}$$

to stosując wzory na różniczkowanie: w pierszym przypadku, oraz bezpośrednio z definicji: w drugim przypadku, wyliczamy, że

$$f_x(x, y) = \begin{cases} y\left(\dfrac{x^2 - y^2}{x^2 + y^2} + \dfrac{4x^2y^2}{(x^2 + y^2)^2}\right), & \text{gdy } x^2 + y^2 > 0; \\ 0, & \text{gdy x=0=y.} \end{cases}$$

Wzór na pochodną cząstkową f_y jest symetryczny z tym, że różni się o minus, a więc $f_x(0, y) = -y$ oraz $f_y(x, 0) = x$. Gdyby pochodne drugiego rzędu były ciągłe na kostce zawierającej punkt $(0,0)$, to mielibyśmy $f_{xy}(0,0) = -1$ oraz $f_{yx}(0,0) = 1$: sprzeczność z twierdzeniem. Brak równości tych pochodnych cząstkowych wyklucza ciągłości choć jednej z nich w punkcie $(0,0)$.

Dla różniczkowalnych funkcji wektorowych obowiązuje wzór na różniczkowanie złożenia. Mianowicie:

$$\text{Gdy } f(x) = h\left(g(x)\right), \qquad \text{to} \qquad f^{'}(x) = h^{'}\left(g(x)\right) \circ g^{'}(x),$$

gdzie symbol „$\circ$" oznacza mnożenie macierzy według wzoru $c_{ik} = \sum_{j=1}^{m} a_{ij} b_{jk}$.

Funkcję $f(x) = \sin x + \cos x$ można przedstawić jako złożenie funkcji $g : x \to (\sin x, \cos x)$ oraz $h : (x,y) \to x + y$. Wtedy $f(x) = h\left(g(x)\right)$, $g'(x) = \begin{pmatrix} \cos x \\ -\sin x \end{pmatrix}$ oraz $h'(x,y) = (1,1)$. Skąd

$$h^{'} \circ g^{'}(x) = (1,1) \circ \begin{pmatrix} \cos x \\ -\sin x \end{pmatrix} = \cos x - \sin x = f^{'}(x).$$

Funkcję $f(x) = x^x$ można przedstawić jako złożenie funkcji $g : x \to (x,x)$ oraz $h : (x,y) \to x^y$. Wtedy $f(x) = h\left(g(x)\right)$, $g'(x) = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$ oraz $h'(x,y) = (yx^{y-1}, x^y \ln x)$. Skąd mamy

$$h^{'}(g(x)) \circ g^{'}(x) = (xx^{x-1}, x^x \ln x) \circ \begin{pmatrix} 1 \\ 1 \end{pmatrix} = x^x(1 + \ln x) = f^{'}(x).$$

Dla różniczkowalnych funkcji wektorowych obowiązuje wzór Taylora, tzn. dla funkcji $(n+1)-$krotnie różniczkowalnej f zachodzi

$$f(x) = f(x_0) + \sum_{k=1}^{n} \frac{1}{k!} f^{(k)}(x_0) \circ (x - x_0)^k + \frac{1}{(n+1)!} f^{(n+1)}(x_0 + t(x - x_0)) \circ (x - x_0)^n,$$

gdzie $0 < t < 1$. W tym wzorze symbolami $f^{(k)}(x_0) \circ (x - x_0)^k$ oznaczono **stosowne:** wykonywalne przez odpowiednie oznaczanie wierszy i kolumn, mnożenie macierzy. W przypadku funkcji rzeczywistej wielu zmiennych, np. funkcji trzech zmiennych $f(x,y,z)$, oraz $k = 2$ będziemy mieli zapis mnożenia, w którym kolejność mnożenia nie będzie istotna

$$(x - x_0, y - y_0, z - z_0) \circ \begin{pmatrix} F_{xx} & F_{xy} & F_{xz} \\ F_{yx} & F_{yy} & F_{yz} \\ F_{zx} & F_{zy} & F_{zz} \end{pmatrix} \circ \begin{pmatrix} x - x_0 \\ y - y_0 \\ z - z_0 \end{pmatrix}.$$

Dla $k > 2$ problemy stosownego mnożenia macierzowego objaśnia tzw. *algebra odwzorowań wieloliniowych.*

Rozważmy równanie $F(x,y)=0$. Pomyślmy, że funkcja $x \to y(x)$ jest rozwiązaniem tego równania, czyli zawsze zachodzi $F(x,y(x))=0$. Gdy F jest funkcją różniczkowalną, to składamy ją z funkcją $x \to (x,y(x))$, a następnie różniczkujemy takie złożenie¿ Dostajemy

$$(F_x, F_y) \circ \begin{pmatrix} 1 \\ y^{'} \end{pmatrix} = 0, \qquad \text{czyli} \qquad F_x + y^{'}(x) F_y = 0.$$

Ostatnia równość to równanie różniczkowe rzędu pierwszego, dla którego dodatkowo zachodzi równość $F_{xy} = F_{yx}$: jest to konsekwencja twierdzenia, o ile pochodne cząstkowe rzędu drugiego występujące w tej równości są ciągłe. Ta równość może być używana przy badaniu funkcji $x \to y(x)$, o której mówimy, że jest zadana w formie *uwikłanej*.

Gdy badamy równanie różniczkowe $2x-y+(4y-x)y^{'}=0$, to kładziemy $F_x = 2x-y$ oraz $F_y = 4y-x$: bo mamy $F_{xy} = -1 = F_{yx}$. Dostajemy dwa równania różniczkowe: w pierszym y jest parametrem, zaś w drugim parametrem jest x. Wyliczamy rozwiązania takich równań różniczkowych, traktując parametry jak stałe. Dostajemy

$$F(x,y) = x^2 + Ay + D \qquad \text{oraz} \qquad F(x,y) = 2y^2 - xy + B.$$

Porównujemy te ostanie równania, pamiętając, że y albo x były traktowane jak stałe. Dostajemy

$$F(x,y) = x^2 + 2y^2 - xy + C = 0$$

i sprawdzamy, że jest to całka ogólna równania różniczkowego wyjściowego.

Gdy badamy równanie różniczkowe $1-x^2+x^2(y-x)y^{'}=0$, to mnożymy je stronami przez x^{-2}. Dostajemy równanie różniczkowe

$$\left(\frac{1}{x^2} - y\right) + (y-x)y^{'} = 0:$$

w literaturze zabieg taki bywa nazywany *mnożeniem przez czynnik całkujący*. Następnie kładziemy $F_x = \left(\frac{1}{x^2} - y\right)$ oraz $F_y = y - x$: gdyż

$$F_{xy} = -1 = F_{yx}.$$

Skoro parametry x albo y są traktowane jak stałe, to wyliczamy rozwiązania takich równań różniczkowych

$$F(x,y) = -x^{-1} - xy + D \qquad \text{oraz} \qquad F(x,y) = \frac{y^2}{2} - xy + \frac{y^2}{2} + B.$$

Porównujemy te ostanie równania i dostajemy

$$F(x,y) = \frac{-1}{x} - xy + \frac{y^2}{2} + C = 0.$$

Jest to całka ogólna równania różniczkowego wyjściowego.

Gdy badamy równanie różniczkowe $y^2 = (1 - xy)y^{'}$, to mnożymy je stronami przez y^{-1}. Dostajemy

$$y + \left(x - \frac{1}{y}\right) y^{'} = 0.$$

Następnie kładziemy $F_x = y$ oraz $F_y = x - \frac{1}{y}$: gdyż $F_{xy} = 1 = F_{yx}$ i wyliczamy rozwiązania takich równań różniczkowych

$$F(x,y) = xy + D \qquad \text{oraz} \qquad F(x,y) = xy - \ \ln y + +B.$$

Porównujemy te ostanie równania: y albo x były traktowane jak stałe, i dostajemy całkę ogólną równania wyjściowego

$$F(x,y) = xy + \ \ln y + C = 0.$$

Ekstrema lokalna. Mówimy, że funkcja rzeczywista $F(x_1, x_2, \ldots, x_n)$ ma ekstremum lokalne w punkcie $(x_1, x_2, \ldots, x_n)$ gdy istnieje kostka

$$\{(y_1, y_2, \ldots, y_n) : b_1 < y_1 < c_1, b_2 < y_2 < c_2, \ldots \text{ oraz } b_n < y_n < c_n\},$$

do której punkt ten należy oraz: – wartość funkcji F na dowolnym wektorze z tej kostki jest niewiększa niż $F(x_1, x_2, \ldots, x_n)$, wtedy takie ekstremum nazywamy *nimimum lokalnym*; – wartość funkcji F na dowolnym wektorze z tej kostki jest niemniejsza niż $F(x_1, x_2, \ldots, x_n)$, wtedy takie ekstremum nazywamy *maksimum lokalnym.* Gdy badamy ekstrema lokalne funkcji rozwijalnej według wzoru Taylora, to ekstrema mogą istnieć jedynie w tych punktach, gdzie macierz pierwszej pochodnej to macierz złożona z samych zer. Wtedy o istnieniu ekstremum decyduje druga pochodna lub pochodne wyższych rzędów.

Twierdzenie. *Gdy dla* 2*-krotnie różniczkowalnej funkcji rzeczywistej*

$$(x_1, x_2, \ldots, x_n) \to F(x_1, x_2, \ldots, x_n),$$

w punkcie $(a_1, a_2, \ldots, a_n)$ *zachodzi*

$$(F_{x_1}(a_1, a_2, \ldots, a_n), F_{x_2}(a_1, a_2, \ldots, a_n), \ldots, F_{x_n}(a_1, a_2, \ldots, a_n)) = (0, 0, \ldots, 0)$$

oraz: – forma kwadratowa

$$(x_1, x_2, \ldots, x_n) \to (x_1, x_2, \ldots, x_n) \circ F^{"}(a_1, a_2, \ldots, a_n) \circ \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{pmatrix}$$

jest dodatnio określona: przyjmuje wartości dodatnie za wyjątkiem punktu $(0, 0, \dots, 0)$, *to funkcja F ma w punkcie $(a_1, a_2, \dots, a_n)$ minimum lokalne; – zaś gdy ta forma kwadratowa jest ujemnie określona*: przyjmuje wartości ujemne za wyjątkiem punktu $(0, 0, \dots, 0)$, *to funkcja F ma w punkcie $(a_1, a_2, \dots, a_n)$ maksimum lokalne; – zaś gdy ta forma kwadratowa nie jest określona*: przymuje wartości dodatnie lub ujemne za wyjątkiem $(0, 0, \dots, 0)$, *to funkcja F w punkcie $(a_1, a_2, \dots, a_n)$ nie ma ekstremum; – w pozostałych przypadkach funkcja F w punkcie $(a_1, a_2, \dots, a_n)$ może mieć ekstemum lokalne, ale nie musi mieć.*

Gdy szukamy ekstremów lokalnych funkcji

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

wpierw rozwiązujemy układ równań

$$\begin{cases} F_x = 4x^3 - 4x + 4y = & 0; \\ F_y = 4y^3 + 4x - 4y = & 0. \end{cases}$$

Dodając stronami te równania wyliczamy, że $x = -y$. Następnie rugujemy y i otrzymujemy równanie $x(x^2-2) = 0$. Skąd wyliczamy, że musimy zbadać zachowanie funkcji F w punktach $(0, 0)$, $(\sqrt{2}, -\sqrt{2})$ oraz $(-\sqrt{2}, \sqrt{2})$. Obliczamy pochodne cząstkowe rzędu drugiego

$$F_{xx} = 12x^2 - 4, \qquad F_{xy} = 4 = F_{yx} \text{ oraz } F_{yy} = 12y^2 - 4.$$

Dla punktów $(\sqrt{2}, -\sqrt{2})$ oraz $(-\sqrt{2}, \sqrt{2})$ badamy formę kwadratową

$$(x, y) \circ \begin{pmatrix} 20 & 4 \\ 4 & 20 \end{pmatrix} \circ \begin{pmatrix} x \\ y \end{pmatrix} = 20x^2 + 20y^2 + 8xy = 20 \left(x + \frac{y}{10} \right)^2 + \frac{99}{5} y^2.$$

Jest to forma kwadratowa dodatnio określona, a więc w punktach tych funkcja F przyjmuje minima lokalne:

$$F(-\sqrt{2}, \sqrt{2}) = -8 = F(\sqrt{2}, -\sqrt{2}).$$

Dla punktu $(0, 0)$ stosowna forma kwadratowa to

$$(x, y) \circ \begin{pmatrix} -4 & 4 \\ 4 & -4 \end{pmatrix} \circ \begin{pmatrix} x \\ y \end{pmatrix} = -4x^2 - 4y^2 + 8xy = -4(x - y)^2.$$

Podpada ona pod część twierdzenia „przypadki pozostałe” i w punkcie tym ekstremum nie ma. Gdy założymy

$$x = y \text{ oraz } x \neq 0, \text{ to mamy } F(x, x) = 2x^4 > 0.$$

Zaś gdy założymy

$$x = -y \text{ oraz } 0 < x^2 < 2, \text{ to mamy } F(x, -y) = 2x^2(x^2 - 2) < 0.$$

Gdy szukamy najmniejszej i największej wartości funkcji

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

w kole $x^2 + y^2 \leq 4$, to rozwiązujemy układ równań

$$F_x = 4x = 0 \qquad \text{oraz} \qquad F_y = -4y = 0.$$

Wyliczamy $x = 0 = y$. Skoro

$$F_{xx}(0,0) = 4, F_{xy}(0,0) = 0 = F_{yx}(0,0) \text{ oraz } F_{yy}(0,0) = -4,$$

to forma kwadratowa

$$(x,y) \circ \begin{pmatrix} -4 & 0 \\ 0 & -4 \end{pmatrix} \circ \begin{pmatrix} x \\ y \end{pmatrix} = 4x^2 - 4y^2$$

jest nieokreślona: funkcja F we wnętrzu rozważanego koła nie ma ekstremów. Funkcja ta przyjmuje wartości ekstremalne na okręgu $x^2 + y^2 = 4$: wynika to z tzw. *własności topologicznych funkcji ciągłych*, a więc pozostała do zbadanie funkcja

$$F(x, \sqrt{4 - x^2}) = 4x^2 - 8.$$

Gdy $x = 0$, to $F(0,2) = -8$; zaś gdy $x = 2$, to $F(2,0) = 8$: znaleźliśmy minimum oraz maksimum funkcji F w kole $x^2 + y^2 \leq 4$.

Gdy rozważymy funkcję uwikłaną $x \to y(x)$ zadaną równaniem

$$F(x,y) = xy + \ln x + \ln y = 0,$$

to wyliczamy pochodne cząstkowe

$$F_x = \frac{1 + xy}{x} \qquad \text{oraz} \qquad F_y = \frac{1 + xy}{y}.$$

Z równości $F_x + y' F_y = 0$ dostajemy równanie różniczkowe $xy' = -y$. Całką ogólną tego równania różniczkowego jest rodzina hiperbol $C = yx$. Zaś rozważana funkcja uwikłana to hiperbola $xy = C$, gdzie C jest stałą, która spełnia równanie $C + \ln C = 0$.

Metoda czynnika nieoznaczonego Lagrange'a. Gdy szukamy ekstremów lokalnych funkcji $(x,y) \to H(x,y)$ przy warunku $g(x,y) = 0$, to badamy funkcję

$$F(x,y) = H(x,y) + \lambda g(x,y).$$

Punkty w których mogą być ekstrema lokalne są zawarte wśród rozwiązań układu równań $F_x(x,y) = 0$, $F_y(x,y) = 0$ oraz $g(x,y) = 0$. Przykładowo, gdy

$$H(x,y) = x^2 + y^2 \qquad \text{oraz} \qquad g(x,y) = xy - 1 = 0,$$

to badamy funkcję

$$F(x,y) = x^2 + y^2 - \lambda(xy - 1).$$

Dostajemy do rozwiązania układ równań $2x - \lambda y = 0$, $2y - \lambda x = 0$ oraz $1 = xy$. Odejmując stronami pierwsze równanie od drugiego wyliczamy, że $x = y$. Po wyrugowaniu y z równania trzeciego dostajemy dwa punkty $(1,1)$ oraz $(-1,-1)$, w których mamy minima lokalne funkcji $H(x,y) = x^2+y^2$ przy warunku $g(x,y) = xy - 1 = 0$. Dlaczego ta metoda zawsze prowadzi do poprawnego wyniku?