21 Pochodne wyższych rzędów

Licząc pochodne cząstkowe pierwszego rzędu możemy do otrzymanych funkcji ponownie stosować operacje liczenia pochodnej- otrzymując pochodne rzędu 2, następnie rzędu 3 itp. Na przykład, dla funkcji dwu zmiennych $f(x, y)$ pochodna cząstkowa z $\frac{\partial}{\partial x}$ w punkcie (x_0, y_0) stosowana do $\frac{\partial f}{\partial x}$ daje pochodną drugiego rzędu: $\frac{\partial^2 f}{\partial x^2}(x_0, y_0)$ (inny zapis: $f''_{xx}(x_0, y_0)$, lub niepreferowany przeze mnie: $f_{xx}(x_0, y_0)$. Jest to druga pochodna w punkcie x_0 funkcji zmiennej x oznaczmy ją np. G, gdzie

$$G(x) = f(x, y_0).$$

Podobnie, dla

$$H(y) := f(x_0, y)$$

mamy

$$\frac{d^2}{dy^2}H(y_0) = \frac{\partial^2 f}{\partial y^2}(x_0, y_0).$$

Ale można też rozważać tzw. pochodne mieszane, np. gdy najpierw liczymy pochodną względem x, potem względem y. Kolejność operacji różniczkowania zapisuje się tak, jak w przypadku składania odwzorowań. We wspomnianym przypadku

$$\frac{\partial^2 f}{\partial y \partial x}(x_0, y_0) = \frac{\partial}{\partial y}\left\{\frac{\partial}{\partial x} f(x_0, y_0)\right\}.$$

Można mieć pewne zastrzeżenia do precyzji ostatniego zapisu[1] - bo w celu wyliczenia pochodnej "zewnętrznej" należy tę "wewnętrzną" wyliczyć w punktach (x_0, y), gdzie y przebiega pewne otoczenie y_0. Najlepiej zobaczmy to na paru prostych przykładach.

Przykład 1. Dla $f(x, y) = xy^4 + 3x^2y^2$ mamy $f'_x = y^4 + 6xy^2$, $f'_y = 4xy^3 + 6x^2y$, więc $f''_{xy} = \frac{d}{dx}(4xy^3 + 6x^2y) = 4y^3 + 12xy$. Natomiast $f''_{yx} = \frac{d}{dy}(y^3 + 6xy^2) = 4y^3 + 12xy$. Jak za chwilę zobaczymy, to nie jest przypadek, że pochodne mieszane są równe czyli operatory różniczkowe $\frac{\partial f}{\partial x}$ oraz $\frac{\partial f}{\partial y}$ są na odpowiednich klasach funkcji przemienne. Jednak następny przykład pokaże, że na ogół wcale tak nie musi być.

Przykład 2,

Pochodne mieszane zależą od kolejności różniczkowań następującej funkcji

$$f(x, y) = \begin{cases} xy\frac{x^2-y^2}{x^2+y^2}, & x^2 + y^2 > 0; \\ 0, & \texttt{dla} \quad \text{x=y=0}. \end{cases} \tag{1}$$

Proszę sprawdzić, że w punkcie P o współrzędnych $(0, 0)$ mamy $f^{''}_{xy}(P) = -1$, natomiast $f^{''}_{yx}(P) = 1$.

Pochodnych wyższych rzędów raczej nie będziemy używali (mają one znaczenie przy przybliżaniu wartości f przy użyciu wzoru Taylora i wtedy wprowadzimy dla nich tzw. zapis wielowskaźnikowy). Natomiast pochodne cząstkowe rzędu 2 liczy się dość często podczas badania ekstremów lokalnych, do czego przejdziemy w następnym wykładzie. Najpierw poznamy warunki gwarantujące równość pochodnych mieszanych. Podał je Hermann Schwarz (1843-1921), autor m. inn. nierówności dla iloczynu skalarnego i ciekawych twierdzeń z analizy zespolonej. Zacznijmy od definicji macierzy drugiej różniczki.

Definicja. Dla funkcji n zmiennych określonej w otoczeniu punktu P macierzą Hessego $H_f(P)$ nazywamy macierz kwadratową $n \times n$ postaci

$$H_f(P) = \begin{pmatrix} \frac{\partial^2 f}{\partial x_1^2}(P) & \dots & \dots & \frac{\partial^2 f}{\partial x_1 \partial x_n}(P) \\ \frac{\partial^2 f}{\partial x_2 \partial x_1}(P) & \dots & \dots & \dots \\ \dots & \dots & \dots & \dots \\ \frac{\partial^2 f}{\partial x_n \partial x_1}(P) & \dots & \dots & \frac{\partial^2 f}{\partial x_n^2}(P) \end{pmatrix}$$

[1]chodzi o prawą stronę tej ostatniej równości

Twierdzenie Schwarza o symetrii. Jeśli drugie pochodne mieszane są ciągłe w danym punkcie P, to są one równe. Macierz Hessego jest więc symetryczna w przypadku funkcji klasy C^2, czyli takich, dla których pochodne cząstkowe $\frac{\partial^2 f}{\partial x_j \partial x_k}$ są w tym punkcie ciągłe.

Dowód. Przy liczeniu pochodnych cząstkowych rzędu 2. pozostałe (z wyjątkiem dwu) zmienne są ustalone. Bez straty ogólności wystarczy więc wykazać tezę dla funkcji dwu zmiennych. Oznaczenie ich, jak zwykle przez x, y pozwala uniknąć kłopotliwych wskaźników. Mamy więc wykazać, że gdy są ciągłe w punkcie $P = (x, y)$, to pochodne mieszane $\frac{\partial^2 f}{\partial x \partial y}(P)$ oraz $\frac{\partial^2 f}{\partial y \partial x}(P)$ są równe.

Punktem wyjścia będzie następująca liczba: Dla $h, k \neq 0$ dostateczie małych rozważmy wartość:

$$\Lambda := f(x+h, y+k) - f(x, y+k) - f(x+h, y) + f(x, y) \tag{2}$$

oraz zdefiniujmy dwie funkcje jednej zmiennej:

$$\phi(x) := f(x, y+k) - f(x, y), \qquad \psi(y) := f(x+h, y) - f(x, y). \tag{3}$$

Wówczas okazuje się, że

$$\phi(x+h) - \phi(x) = \Lambda = \psi(y+k) - \psi(y).$$

Ponieważ ϕ oraz ψ są ciągłe i różniczkowalne w otoczeniu punktów x, y odpowiednio, dla h, k dostatecznie małych spełniają one założenia twierdzenia Lagrange'a o wartości średniej. Punkty z odcinka otwartego o końcach $x, x+h$ można opisać przy użyciu parametru $\alpha \in (0, 1)$ w postaci $x_\alpha := x + \alpha h$. (Może być zarówno $h < 0$, jak i może być $h > 0$, więc lepiej mówić o odcinku, niż o przedziale, gdzie narzucona jest nierówność między jego końcami.)

Iloraz różnicowy $\frac{\phi(x+h)-\phi(x)}{h}$ jest więc równy pochodnej: $\phi'(x_\alpha)$ w pewnym takim punkcie pośrednim x_α. Podobnie, $\exists_{\beta \in (0,1)}\ \psi'(y_\beta) = \frac{\psi(y+k)-\psi(y)}{k}$, przy czym liczniki obydwu tych ilorazów różnicowych są równe Λ. Stąd

$$\frac{\Lambda}{h} = \phi'(x_\alpha), \quad \frac{\Lambda}{k} = \psi'(y_\beta). \tag{4}$$

Z definicji funkcji ϕ, ψ, z definicji pochodnych cząstkowych i z liniowości operacji różniczkowania wynika, że pochodne "zwykłe" ϕ', ψ' są przyrostami odpowiednich pochodnych cząstkowych dla f. I do tych przyrostów ponownie stosujemy twierdzenie Lagrange'a o wartości średniej (tym razem dzielimy $\phi'(x_\alpha)$ przez k (odpowiednio, $\psi'(y_\beta)$ przez h) czyli -przez mianowniki ilorazów różnicowych). Dla pewnych $\delta, \gamma \in (0, 1)$ mamy

$$\frac{1}{k}\phi'(x_\alpha) = \frac{1}{k}\left(\frac{\partial f}{\partial x}(x_\alpha, y+k) - \frac{\partial f}{\partial x}(x_\alpha, y)\right) = \frac{\partial^2 f}{\partial y \partial x}(x_\alpha, y_\delta).$$

Lewa strona jest dzięki (4) równa $\frac{\Lambda}{hk}$. Analogiczne rozumowanie dla funkcji ψ prowadzi do równości

$$\frac{\Lambda}{hk} = \frac{\partial^2 f}{\partial x \partial y}(x_\gamma, y_\beta)$$

dla pewnej wartości $\gamma \in (0, 1)$. Ponieważ odległości punktów (x_α, y_δ) od P wynoszą $\sqrt{\alpha^2 h^2 + \beta^2 k^2} \leqslant \|(h, k)\|$, podobnie $\|(x_\gamma, y_\beta) - (x, y)\| \leqslant \|(h, k)\|$, to przy zmierzaniu (h, k) do wektora zero -te odległości też zmierzają do zera i przechodząc do granicy w otrzymanych przed chwilą równościach

$$\frac{\partial^2 f}{\partial y \partial x}(x_\alpha, y_\delta) = \frac{\partial^2 f}{\partial x \partial y}(x_\gamma, y_\beta),$$

otrzymamy dowodzoną tezę, czyli równość tych pochodnych mieszanych w punkcie (x, y). $\square$

Powstaje pytanie, czy możemy zdefiniować drugą różniczkę jako różniczkę pierwszej różniczki? Dla uproszczenia przyjrzyjmy się sytuacji funkcji o wartościach w $\mathbb{R}$. Jeśli : $D \subset \mathbb{R}^n \to \mathbb{R}$ jest różniczkowalna w każdym punkcie zbioru otwartego D, to różniczka $d_P f$, jako odwzorowanie, którego zmienną jest P, przyjmuje wartości w $\mathcal{L}(\mathbb{R}^n, \mathbb{R})$, wiec różniczka "powtórna" (iterowana): $d(df)$ przyjmie wartości w przestrzeni wektorowej $Z := \mathcal{L}(\mathbb{R}^n, \mathcal{L}(\mathbb{R}^n, \mathbb{R}))$. Wygodniejsza od przestrzeni Z jest (kanonicznie izomorficzna z Z przestrzeń Z_1 wszystkich odwzorowań dwuliniowych $T : \mathbb{R}^n \times \mathbb{R}^n \ni (\vec{x}, \vec{y}) \to T(\vec{x}, \vec{y}) \in \mathbb{R}$. Takie odwzorowania są jednoznacznie określone przez macierz o wyrazach $T(e_j, e_k), j, k \in \{1, \ldots, n\}$. Dzięki twierdzeniu Schwarza można wykazać, że odwzorowania otrzymane w ten sposób przez powtórne różniczkowanie dla f klasy C^2 są symtryczne (gdyż odpowiadają macierzy symetrycznej). Z kolei odwzorowania dwuliniowe symetryczne $T : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}$ tworzą przestrzeń wektorową izomorficzną z przestrzenią wielomianów jednorodnych stopnia 2, zwanych formami kwadratowymi. Wspomniany izomorfizm przypisuje odwzorowaniu dwuliniowemu symetrycznemu $T(\cdot, \cdot)$ formę kwadratową $q(\vec{x}) = T(\vec{x}, \vec{x})$. Izomorfizm odwrotny (nieco trudniejszy w konstrukcji) dany jest przez tzw. wzór polaryzacyjny:

Gdy $q : \mathbb{R}^n \to \mathbb{R}$ jest formą kwadratową, czyli wielomianem jednorodnym stopnia 2, to definiujemy

$$T(x, y) = \frac{1}{4}(Q(x + y) - Q(x - y)). \tag{5}$$

Tym razem, dla uproszczenia, pomijam strzałki nad wektorami. Okazuje się, że dopiero zastosowanie tych izomorfizmów do iterowanej różniczki -da nam właściwą (czyli wygodną w zastosowaniach) drugą różniczkę.

Definicja. Dla funkcji $f : D \subset \mathbb{R}^n \to \mathbb{R}$ klasy C^2 definiujemy jej drugą różniczkę w punkcie $P \in D$ jako formę kwadratową:

$$d_P^2 f(t_1, t_2, \ldots t_n) = \sum_{j=1}^{n} \sum_{k=1}^{n} \frac{\partial^2 f}{\partial x_j \partial x_k}(P) t_j t_k.$$

W zapisie bezargumentowym pierwsza różniczka miała postać

$$d_P f = \sum_{j=1}^{n} \frac{\partial f}{\partial x_j}(P)\, dx_j,$$

gdzie dx_j reprezentuje odwzorowanie liniowe "rzut na j-tą oś",

$$(dx_j)(t_1, t_2, \ldots, t_n) = t_j,\ dx_j : \mathbb{R}^n \to \mathbb{R}.$$

Iloczyn pary takich odworowań, zapisywany jako $dx_j dx_k$ nie jest już odwzorowaniem liniowym, ale wielomianem jednorodnym stopnia 2 (czyli formą kwadratową), takie odwzorowania $dx_j dx_k$ tworzą bazę wektorową przestrzeni form kwadratowych.

W pewnym sensie uzasadnieniem takiej, a nie innej definicji będzie (stanowiący istotę dowodu wzoru Taylora) następujący fakt.

Lemat. Dla funkcji f klasy C^2 w otoczeniu punktu $P_0 = (x_0, y_0)$ i dla niezerowego wektora $\vec{\Delta} = (h, k)$ niech $P_s := P_0 + s\vec{\Delta}$. Wówczas druga pochodna w punkcie s z funkcji $\Phi(s) := f(P_0 + s\vec{\Delta}) = f(x_0 + sh, y_0 + sk)$ dana jest wzorem

$$\Phi^{''}(s) = f^{''}_{xx}(P_s)h^2 + 2f^{''}_{xy}(P_s)hk + f^{''}_{yy}(P_s)k^2.$$

Jest to więc wartość drugiej różniczki w punkcie P_s na wektorze o współrzędnych (h, k).

Zanim dowód pojawi się na następnym wykładzie, proszę spróbowac to samodzielnie przeliczyć.