# Funkcje wielu zmiennych – różniczkowalność

Zajmiemy się teraz różniczkowaniem funkcji wielu zmiennych. Zaczniemy od pojęcia pochodnej cząstkowej, bo jest ono najważniejszym i zarazem najprostszym z tych, którymi przyjdzie nam się zająć. W tym wykładzie, jeśli nie piszemy wyraźnie, że jest inaczej funkcja $f:\mathbf{G}\to\mathbf{R}^l$ będzie określona na zbiorze otwartym $\mathbf{G}\subset\mathbf{R}^k$. Będziemy starali się przenieść twierdzenia użyteczne dla optymalizacji funkcji o wartościach rzeczywistych, czyli dla znajdowania ich wartości najmniejszych i największych. W niektórych przypadkach pojęcie pochodnej cząstkowej nam wystarczy, a w niektórych zmuszeni zostaniemy do użycia pojęcia różniczki funkcji, którego zdefiniowanie chwilowo odkładamy.

**Definicja pochodnej cząstkowej**

Pochodna cząstkową pierwszego rzędu odwzorowania $f:\mathbf{G}\to\mathbf{R}^l$ ze względu na zmienną $x_i$, $1\le i\le k$, w punkcie $\mathbf{p}\in\mathbf{G}$, nazywamy granicę $\lim\limits_{h\to 0}\dfrac{f(\mathbf{p}+h\mathbf{e}_i)-f(\mathbf{p})}{h}$, o ile istnieje; $\mathbf{e}_i\in\mathbf{R}^k$ to wektor, którego wszystkie współrzędne z wyjątkiem i-tej są równe 0 a i-ta równa jest 1: $\mathbf{e}_i=(0,,...0,1,0,...,0)$. Tę pochodną cząstkową oznaczamy symbolem $f_{x_i}(\mathbf{p})$, zamiast archaicznego oznaczenia stosowanego jeszcze dzisiaj (głównie przez fizyków) - $\dfrac{\delta f}{\delta x_i}(\mathbf{p})$.

**Przykłady**

**1.** Niech $f(\mathbf{x})=x_1+2x_2^3+x_3e^{x_4}$. Z definicji pochodnej cząstkowej wynika, że

$$f_{x_1}(\mathbf{x})=\lim_{h\to 0}\frac{f(\mathbf{x}+h\mathbf{e}_1)-f(\mathbf{x})}{h}=\lim_{h\to 0}\frac{f(x_1+h,x_2,x_3,x_4)-f(x_1,x_2,x_3,x_4)}{h}=$$

$$=\lim_{h\to 0}\frac{x_1+h+2x_2^3+x_3e^{x_4}-(x_1+2x_2^3+x_3e^{x_4})}{h}=\lim_{h\to 0}\frac{h}{h}=1.$$

Pochodną $f_{x_1}$ funkcji $f$ obliczamy traktując $x_1$ jako argument funkcji przy jednoczesnym traktowaniu zmiennych $x_2,x_3,x_4$ jako stałych (parametrów).
Licząc analogicznie, otrzymujemy jeszcze trzy równości (proszę sprawdzić!)

$$f_{x_2}(\mathbf{x})=6x^2,\ f_{x_3}(\mathbf{x})=e^{x_4},\ f_{x_4}(\mathbf{x})=x_3e^{x_4}.$$

**2.** Niech $f\begin{pmatrix} r \\ \varphi \end{pmatrix}=\begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}$ - tym razem współrzędne punktów piszemy pionowo, co – jak się później okaże - ma sens.
Obliczmy pochodną względem zmiennej $r$.

$$f_r\begin{pmatrix} r \\ \varphi \end{pmatrix} = \lim_{h\to 0}\frac{f\begin{pmatrix} r+h \\ \varphi \end{pmatrix} - f\begin{pmatrix} r \\ \varphi \end{pmatrix}}{h} = \lim_{h\to 0}\frac{\begin{pmatrix} (r+h)\cos\varphi \\ (r+h)\sin\varphi \end{pmatrix} - \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}}{h} = \lim_{h\to 0}\frac{\begin{pmatrix} h\cos\varphi \\ h\sin\varphi \end{pmatrix}}{h} = \begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}.$$

Teraz kolei na pochodną względem zmiennej $\varphi$.

$$f_\varphi\begin{pmatrix} r \\ \varphi \end{pmatrix} = \lim_{h\to 0}\frac{f\begin{pmatrix} r \\ \varphi+h \end{pmatrix} - f\begin{pmatrix} r \\ \varphi \end{pmatrix}}{h} = \lim_{h\to 0}\frac{\begin{pmatrix} r\cos(\varphi+h) \\ r\sin(\varphi+h) \end{pmatrix} - \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}}{h} =$$

$$= \begin{pmatrix} \lim\limits_{h\to 0}\dfrac{r\cos(\varphi+h) - r\cos\varphi}{h} \\ \lim\limits_{h\to 0}\dfrac{r\sin(\varphi+h) - r\sin\varphi}{h} \end{pmatrix} = \begin{pmatrix} \lim\limits_{h\to 0}\dfrac{-2r\sin(\varphi+\frac{h}{2})\sin\frac{h}{2}}{h} \\ \lim\limits_{h\to 0}\dfrac{2r\sin\frac{h}{2}\cos(\varphi+\frac{h}{2})}{h} \end{pmatrix} = \begin{pmatrix} -r\sin\varphi \\ r\cos\varphi \end{pmatrix}.$$

Widzimy, że w przypadku odwzorowania o wartościach w $\mathbf{R}^2$ otrzymaliśmy wektor a nie liczbę. Rezultat ten jest dokładnie taki, jaki należało się spodziewać. Jeżeli funkcja o wartościach w przestrzeni $\mathbf{R}^l$ ma w jakimś punkcie pochodną względem którejś ze swych $k$ zmiennych, to ta pochodna cząstkowa jest wektorem $l$ wymiarowym.

Właściwie na tym można by zakończyć, ale warto jeszcze otrzymany rezultat zinterpretować fizycznie Można myśleć, że wartością funkcji $f$ jest punkt płaszczyzny oddalony o $r$ od punktu $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ lub wektor zaczynający się w punkcie $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ i kończący się w punkcie $f\begin{pmatrix} r \\ \varphi \end{pmatrix} = \begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}$ - traktujemy więc liczby $r$ i $\varphi$ jako tzw. współrzędne biegunowe punktu płaszczyzny. Przy obliczaniu pochodnej względem $r$ traktujemy zmienną $\varphi$ jako stałą. Możemy interpretować zmienną $r$ jako czas. Po zmianie czasu o $h$ znajdujemy się w punkcie $f\begin{pmatrix} r+h \\ \varphi \end{pmatrix} = \begin{pmatrix} (r+h)\cos\varphi \\ (r+h)\sin\varphi \end{pmatrix}$. Znaleźliśmy się więc w punkcie leżącym na tej samej półprostej wychodzącej z punktu $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$, ale w innej odległości od początku układu współrzędnych. Zmiana odległości równa jest zmianie czasu. Wobec tego prędkość skalarna powinna być równa 1, a wektor prędkości powinien być równoległy do półprostej, po której porusza się punkt. Wektor $\begin{pmatrix} \cos\varphi \\ \sin\varphi \end{pmatrix}$ jest równoległy do półprostej wychodzącej z punktu $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ i przechodzącej przez punkt $\begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}$. Jego długość wynosi 1. Jest to tzw. wektor prędkości wektorowej poruszającego się punktu. Podobnie można zinterpretować pochodną względem $\varphi$. Tym razem $r$ się nie zmienia, natomiast zmienia się kąt jaki tworzy wektor o początku $\begin{pmatrix} 0 \\ 0 \end{pmatrix}$ i końcu $\begin{pmatrix} r\cos\varphi \\ r\sin\varphi \end{pmatrix}$ z osią odciętych (poziomą osią układu współrzędnych). W tej sytuacji $\varphi$ oznacza zarówno czas jak i ten kąt. Wobec tego ruch odbywa się po okręgu o środku $\begin{pmatrix}0\\0\end{pmatrix}$ i promieniu $r$. Chwilowa prędkość wektorowa jest więc wektorem stycznym do tego okręgu. Długość tego wektora wynosi $r$, bo prędkość kątowa jest równa 1. Wektorowi $f_\varphi\begin{pmatrix}r\\ \varphi\end{pmatrix} = \begin{pmatrix}-r\sin\varphi\\ r\cos\varphi\end{pmatrix}$ przysługują obie te własności. To właśnie jest wektor prędkości w tym ruchu w momencie $\varphi$.

**3.** Niech $f\begin{pmatrix}x\\y\end{pmatrix} = \begin{Bmatrix} 0, \text{ jesli x=0=y} \\ \dfrac{\mathrm{xy}}{\mathrm{x}^2+y^2}, \text{ jesli x} \neq 0 \text{ lub y} \neq 0 \end{Bmatrix}$

Funkcja ta nie jest ciągła w punkcie $\begin{pmatrix}0\\0\end{pmatrix}$, bowiem dla $x \neq 0$ mamy $f\begin{pmatrix}x\\x\end{pmatrix} = \frac{1}{2}$ i jednocześnie $f\begin{pmatrix}0\\0\end{pmatrix} = 0 \neq \frac{1}{2}$. Oznacza to, że jeśli zbliżamy się do punktu $\begin{pmatrix}0\\0\end{pmatrix}$ wędrując wzdłuż prostej o równaniu $y = x$, to wartości badanej funkcji nie dążą do $0 = f\begin{pmatrix}0\\0\end{pmatrix}$. Jest to jedyny punkt nieciągłości tej funkcji. Zbadamy teraz kwestię istnienia pochodnych cząstkowych funkcji $f$. We wszystkich punktach z wyjątkiem punktu $\begin{pmatrix}0\\0\end{pmatrix}$ pochodne cząstkowe istnieją, co wynika z twierdzeń pozwalających na obliczanie pochodnej funkcji jednej zmiennej rzeczywistej. Również w punkcie $\begin{pmatrix}0\\0\end{pmatrix}$ funkcja $f$ ma pochodne cząstkowe. Wykażemy to.

Mamy $\lim\limits_{h\to 0}\dfrac{f\begin{pmatrix}h\\0\end{pmatrix} - f\begin{pmatrix}0\\0\end{pmatrix}}{h} = \lim\limits_{h\to 0}\dfrac{0-0}{h} = 0.$ Wykazaliśmy, że $f_x\begin{pmatrix}0\\0\end{pmatrix} = 0.$ W taki sam sposób wykazujemy, że $f_y\begin{pmatrix}0\\0\end{pmatrix} = 0.$ Zauważmy jeszcze, że jeśli $x \neq 0$ lub $y \neq 0,$ to $f_x\begin{pmatrix}x\\y\end{pmatrix} = \dfrac{y^3 - x^2y}{(x^2+y^2)^2}$ - wynika to z twierdzenia o pochodnej ilorazu dwu funkcji jednej zmiennej. Analogicznie $f_y\begin{pmatrix}x\\y\end{pmatrix} = \dfrac{x^3 - xy^2}{(x^2+y^2)^2}.$ Zachęcamy studentów do samodzielnego sprawdzenia tych wzorów oraz do sprawdzenia, że pochodne cząstkowe, które właśnie znaleźliśmy są nieciągłe w punkcie $\begin{pmatrix}0\\0\end{pmatrix}$.

Przykład **3.** pokazuje, że stwierdzenie istnienia pochodnych w jakimś punkcie, a nawet w całej dziedzinie funkcji nie pozwala jeszcze zbyt wiele na temat tej funkcji wywnioskować-za istnienia pochodnych cząstkowych nie wynika nawet ciągłość funkcji. Jasne jest

, ze potrzebne nam są własności na stwierdzanie ciągłości funkcji i co więcej na stwierdzanie, że jej zachowanie w małym otoczeniu punktu różniczkowalności jest w przybliżeniu takie jak funkcji liniowej. To jest podstawowa idea w rachunku różniczkowym. Stosowaliśmy rozumowania oparte na tej właśnie idei wielokrotnie w przypadku funkcji jednej zmiennej. To one doprowadziły nas do sformułowania twierdzeń pozwalających na ustalanie w jakich przedziałach funkcja różniczkowalna jest monotoniczna, w jakich punktach może mieć lokalne ekstrema itd. Musimy podobne rozumowania przenieść na funkcje wielu zmiennych. Podamy teraz definicję różniczkowalności funkcji wielu zmiennych i warunek konieczny i wystarczający dla różniczkowalności.

**Definicja funkcji różniczkowalnej w punkcie**

Funkcja $f:\mathbf{G}\to\mathbf{R}^l$ jest różniczkowalna w punkcie $\mathbf{p}\in\mathbf{G}$ wtedy i tylko wtedy, gdy istnieje przekształcenie liniowe $L:\mathbf{R}^k\to\mathbf{R}^l$, takie, że $\lim_{\mathbf{h}\to 0}\frac{f(\mathbf{p}+\mathbf{h})-f(\mathbf{p})-L\mathbf{h}}{\|\mathbf{h}\|}=\mathbf{0}$. Wtedy przekształcenie liniowe $L$ nazywamy różniczką funkcji w punkcie $\mathbf{p}$ i oznaczamy symbolem $Df(\mathbf{p})$ lub $df(\mathbf{p})$ lub $f'(\mathbf{p})$.

Studenci ambitni sprawdzą, że z warunek nałożony na różniczkę może być spełniony przez co najwyżej jedno przekształcenie liniowe. **PREMIA** za dowód tego stwierdzenia.

**Warunek wystarczający dla różniczkowalności**

Jeśli funkcja $f:\mathbf{G}\to\mathbf{R}^l$ określona na otwartym podzbiorze przestrzeni $\mathbf{R}^k$ ma pochodne cząstkowe względem zmiennych $x_1, x_2, ..., x_k$ w każdym punkcie pewnej kuli otwartej $\mathbf{B}(\mathbf{p},\varepsilon)$ o środku w punkcie $\mathbf{p}$ i wszystkie one są ciągłe punkcie $\mathbf{p}$ to funkcja jest różniczkowalna w punkcie $\mathbf{p}$ i zachodzi następujący wzór: $Df(\mathbf{p})\mathbf{h}=f_{x_1}(\mathbf{p})h_1+f_{x_2}(\mathbf{p})h_2+...+f_{x_k}(\mathbf{p})h_k$.
Dowód tego twierdzenia pomijamy, można go znaleźć np. w znakomitej książce
Andrzej Birkcholc Analiza Matematyczna . Funkcje wielu zmiennych PWN. 1986.

Szczególnie istotnym przypadkiem są funkcje wielu zmiennych o wartościach rzeczywistych i takimi tylko się zajmujemy. W tym przypadku często mówimy o gradiencie funkcji zamiast o jej różniczce w punkcie.

**Definicja gradientu funkcji o wartościach rzeczywistych**

Jeśli $f:\mathbf{G}\to\mathbf{R}$ jest funkcja określoną na podzbiorze otwartym $\mathbf{G}$ przestrzeni $\mathbf{R}^k$ różniczkowalną w punkcie $\mathbf{p}\in\mathbf{G}$, to gradientem funkcji $f$ w punkcie $\mathbf{p}$ nazywamy taki wektor $grad\, f(\mathbf{p})$, że dla każdego wektora $\mathbf{h}\in\mathbf{R}^k$ zachodzi równość $Df(\mathbf{p})\mathbf{h}=grad\, f(\mathbf{p})\cdot\mathbf{h}$.

Różnica między gradientem i różniczką wydaje się różnicą minimalną: chodzi o to, że gradient jest wektorem k-wymiarowym, natomiast różniczka jest przekształceniem liniowym z przestrzeni $\mathbf{R}^k$ w jednowymiarową przestrzeń $\mathbf{R}$.

Ponieważ stosujemy standardowe bazy w przestrzeni $\mathbf{R}^k$, więc współrzędne wektora $grad\, f(\mathbf{p})$ są równe odpowiednim współrzędnym $Df(\mathbf{p})$. To nasz wybór, naturalny w przypadkach rozpatrywanych w tym wykładzie. Gdybyśmy jednak rozważali kwestie ogólniejsze – nie byłoby żadnego „naturalnego” wyboru bazy, pojęcie standardowej bazy straciłoby sens i utożsamianie gradientu z różniczką za pomocą współrzędnych nie byłoby możliwe.

Pochodna cząstkowa obliczana jest po to, by uzyskać informacje o tym jak zmienia się funkcja w kierunku jednej z osi układu współrzędnych. Różniczkę, o ile istnieje obliczamy po to, by dowiedzieć się jak zachowuje się funkcja w całym otoczeniu punktu. Pojęciem pośrednim jest pochodna kierunkowa.

**Definicja pochodnej kierunkowej**

Pochodną kierunkową funkcji $f:\mathbf{G}\to\mathbf{R}^l$ w punkcie $\mathbf{p}$ w kierunku wektora $\mathbf{v}$ nazywamy granicę $\lim\limits_{t\to 0}\dfrac{f(\mathbf{p}+t\mathbf{v})-f(\mathbf{p})}{t}$, jeśli ta granica istnieje. Tę pochodną oznaczamy symbolem $f_{\mathbf{v}}(\mathbf{p})$.

Jest jasne, że uogólniliśmy pojęcie pochodnej cząstkowej $f_{x_i}(\mathbf{p})=f_{e_i}(\mathbf{p})$. Pochodna kierunkowa w kierunku wektora $\mathbf{v}$ obliczana jest po to, by ocenić tempo zmian funkcji w otoczeniu punktu $\mathbf{p}$ na prostej przechodzącej przez punkt $\mathbf{p}$ i równoległej do wektora $\mathbf{v}$.
W punktach różniczkowalności funkcji, pochodną kierunkową można nieraz łatwiej znaleźć po obliczeniu różniczki funkcji niż korzystając bezpośrednio z jej definicji.

**Twierdzenie o istnieniu pochodnej kierunkowej a punktach różniczkowalności funkcji**

Jeśli funkcja $f:\mathbf{G}\to\mathbf{R}^l$ jest różniczkowalna w punkcie $\mathbf{p}\in\mathbf{G}$, $\mathbf{v}\in\mathbf{R}^k$, to funkcja na w punkcie $\mathbf{p}$ pochodną kierunkową w kierunku wektora $\mathbf{v}$ i zachodzi równość $f_{\mathbf{v}}(\mathbf{p})=Df(\mathbf{p})\mathbf{v}$.

**Dowód.** Mamy

$$\lim_{t\to 0}\frac{f(\mathbf{p}+t\mathbf{v}-f(\mathbf{p})}{t}=\lim_{t\to 0}\left(\frac{f(\mathbf{p}+t\mathbf{v})-f(\mathbf{p})-Df(\mathbf{p})(t\mathbf{v})}{\|t\mathbf{v}\|}\cdot\frac{\|t\mathbf{v}\|}{t}+Df(\mathbf{p})\mathbf{v}\right)=Df(\mathbf{p})\mathbf{v}$$

Skorzystaliśmy tu z tego, że wyrażenie jest ograniczone, więc po pomnożeniu przez wyrażenie dążące do $\mathbf{0}$ oraz z tego, że $Df(\mathbf{p})(t\mathbf{v})=tDf(\mathbf{p})\mathbf{v}$ i oczywiście z tego, że $f$ jest różniczkowalna w punkcie $\mathbf{p}$, z czego wynika, że $\lim\limits_{t\to 0}\dfrac{f(\mathbf{p}+t\mathbf{v})-f(\mathbf{p})-Df(\mathbf{p})(t\mathbf{v})}{\|t\mathbf{v}\|}=\mathbf{0}$.

W ten sposób zakończyliśmy dowód tego twierdzenia.

Z tego twierdzenia wynika w szczególności, że przy ustalonym punkcie $\mathbf{p}$ pochodna $f_{\mathbf{v}}(\mathbf{p})$ jest liniową funkcją wektora $\mathbf{v}$, pod warunkiem różniczkowalności funkcji $f$ w punkcie $\mathbf{p}$.

Na zakończenie wykładu powtórzmy: z różniczkowalności funkcji w punkcie wynika istnienie pochodnych kierunkowych w tym punkcie we wszystkich kierunkach, w szczególności istnienie pochodnych cząstkowych. Z istnienia pochodnych cząstkowych nie

wynika nawet ciągłość funkcji – widzieliśmy to na przykładzie **3.** Można podać przykład funkcji, która w pewnym punkcie ma pochodne we wszystkich kierunkach i to równe 0 i jednocześnie nie jest ciągła w tym punkcie. Oznacza to, że zbadanie zachowania się funkcji na prostych przechodzących przez dany punkt to jedynie wstęp do zbadania zachowania się tej funkcji w otoczeniu tego punktu.
Tych kwestii nie będziemy dokładnie analizować, bo to wykracza znacznie poza potrzeby inżyniera.