(1)

Zaczynamy teraz kolejny fragment wykładu, tym razem związany z rachunkiem różniczkowym funkcji wielu zmiennych. Chciałabym, żebyśmy przypomnieli sobie i przećwiczyli następujące zagadnienia (1) pochodna funkcji wielu zmiennych rzeczywistych, (2) znajdowanie i badanie ekstremów funkcji wielu zmiennych rzeczywistych (3) funkcje zadane w sposób uwikłany (materiał szczególnie ważny z punktu widzenia wykładu z fizyki statystycznej) (4) ekstrema funkcji na powierzchniach.

POCHODNA FUNKCJI WIELU ZMIENNYCH materiał ten nie jest dla Państwa całkowicie nowy. Z całą pewnością potrafią państwo przynajmniej obliczać pochodne cząstkowe funkcji wielu zmiennych. Wykład dzisiejszy ma służyć usystematyzowaniu i uporządkowaniu wiadomości. Polecam także przeczytanie notatek zamieszczonych w pliku
www.fuw.edu.pl/~konieczn/wyklad4.pdf
który jest na podobny temat.

Ciągłość: Standardowo omawiając różniczkowanie funkcji wielu zmiennych dyskutuje się odwzorowanie z $\mathbb{R}^n$ do $\mathbb{R}^m$ czyli odwzorowania postaci

$$F: \mathbb{R}^n \longrightarrow \mathbb{R}^m$$

$$F(x^1, \dots, x^n) = \left(F^1(x^1 \dots x^n),\ F^2(x^1 \dots x^n),\ \dots,\ F^m(x^1 \dots x^n) \right)$$

My zaczniemy skromnie od sytuacji kiedy $n = 2$, $m = 1$; mówić więc będziemy o rzeczywistych funkcjach dwóch zmiennych.

konieczność obliczania pochodnych wynika zazwyczaj z potrzeby badania jak mniej więcej zachowuje się funkcja blisko ustalonego punktu bez wnikania we wszystkie szczegóły kształtu wykresu tej funkcji. Będziemy chcieli opisać w przybliżeniu zachowanie funkcji $f: \mathbb{R}^2 \to \mathbb{R}$ w otoczeniu punktu (x_0, y_0) w taki sposób, aby różnica między prawdą a tym przybliżeniem nie była zbyt duża.

Najpierw zajmiemy się pojęciem ciągłości.

Co to znaczy, że funkcja jednej zmiennej jest ciągła? Formalnie:

$g: \mathbb{R} \to \mathbb{R}$ jest ciągła w x_0 $\iff$ def. Cauchy'ego

$$\forall \varepsilon > 0 \ \exists \delta > 0 \quad |x - x_0| < \delta \Rightarrow |g(x) - g(x_0)| < \varepsilon$$

albo

$g: \mathbb{R} \to \mathbb{R}$ jest ciągła w $x_0 \iff$ def. Heinego

$$\forall (x_n): \lim_{n\to\infty} x_n = x_0 \quad f(x_n) \xrightarrow{n\to\infty} f(x_0)$$

Obie te formalne definicje oznaczają to samo: W pobliżu punktu x_0 wartości funkcji f różnią się niewiele od wartości tej funkcji w punkcie x_0. Albo: podchodząc wystarczająco blisko do punktu x_0 widzę wartości funkcji nie odbiegające od $f(x_0)$ bardziej niż o mały ε.

Intuicyjnie mówimy, że funkcja jest ciągła, jeśli jej wykres można narysować nie odrywając długopisu od kartki. Zazwyczaj jest to dobre przybliżenie, tak długo jak zapominamy o okrutnych przykładach funkcji takich jak diabelskie schody.

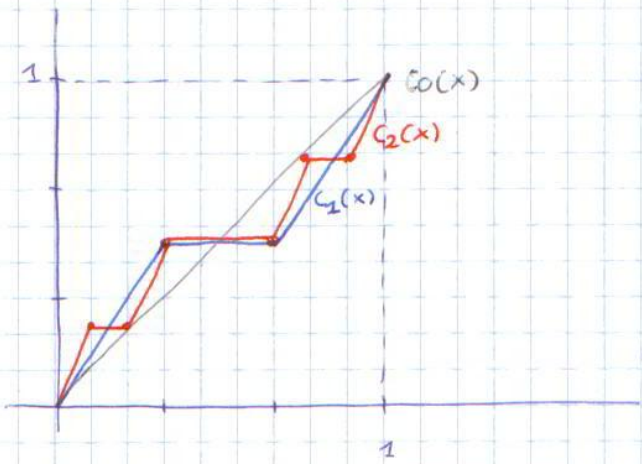
Oto różne przykłady funkcji ciągłych i nieciągłych ③

$x \mapsto \sin x$ ciągła wszędzie

$x \mapsto \sin \frac{1}{x}$, $0 \mapsto 0$ ciągła wszędzie poza 0.

$x \mapsto x \sin x$, ciągła wszędzie, choć dziwna

i na końcu diabelskie schody - też funkcja ciągła wszędzie ale baaaardzo dziwna:



$C_0(x) = x$

$$C_1(x) = \begin{cases} \frac{3}{2}x & 0 \leq x < \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} \leq x \leq \frac{2}{3} \\ \frac{3}{2}x - \frac{1}{2} & \frac{2}{3} \leq x \leq 1 \end{cases}$$

$C_2(x) = \dots$

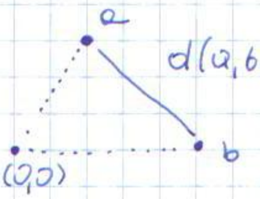
$$C_{n+1}(x) = \begin{cases} \frac{1}{2}C_n(3x) & 0 \leq x < \frac{1}{3} \\ \frac{1}{2} & \frac{1}{3} \leq x < \frac{2}{3} \\ \frac{1}{2}C_n(3x-2) + \frac{1}{2} & \frac{2}{3} \leq x \leq 1 \end{cases}$$

Punkty w których coś tu się dzieje to punkty tzw. zbioru Cantora.

$\lim_{n\to\infty} C_n(x) = C(x)$. $C(x)$ jest ciągła w każdym punkcie $[0,1]$, różniczkowalna wszędzie poza zbiorem Cantora i ma pochodną równą zero wszędzie tam gdzie jest różniczkowalna .... bardzo dziwna funkcja...

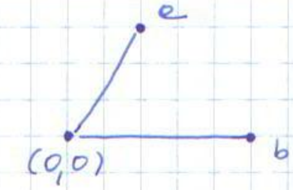
Żeby powiedzieć co to znaczy „funkcja ciągła" w przypadku funkcji wielu zmiennych musimy najpierw wiedzieć co znaczy „w otoczeniu punktu" czy „w pobliżu punktu" a to z kolei oznacza, że pojęcie ciągłości zależy od sposobu mierzenia odległości. Generalnie nie biorąc, oczywiście, nadal funkcja ciągła to jest taka, której wartości w pobliżu punktu (x_0, y_0) różnią się niedużo od wartości funkcji w tym punkcie. Następujący przykład pokazuje, że sposób mierzenia odległości ma znaczenie:

Rozważmy dwa sposoby mierzenia odległości punktów na płaszczyźnie: euklidesowy



$$d(a,b)=\sqrt{(x_a-x_b)^2+(y_a-y_b)^2}$$

i kolejowy



$$\rho(a,b)=\sqrt{x_a^2+y_a^2}+\sqrt{x_b^2+y_b^2}$$

(chyba, że a, b leżą na jednej prostej wraz z zerem)

Rozważmy funkcję, która przypisuje punktowi $p=(x,y)$ wartość

$$f(p) = r(p)\,\varphi(p)$$

gdzie $r(p)$ jest euklidesową odległością p od zera a φ kątem jak we współrzędnych biegunowych, $\varphi(p) \in [0, 2\pi[$

Rozważmy $p=(1,0)$. Funkcja f jest nieciągła w p względem metryki euklidesowej. Odległość euklidesowa od punktu $(1,-a)$ do punktu p jest równa a. Różnica wartości funkcji w $(1,-a)$ i p jest bliska 2π. W metryce kolejowej za to nie wolno podchodzić do punktu p w poprzek! Można tylko wzdłuż torów czyli po osi Ox. Funkcja f jest ciągła względem metryki kolejowej.

W dalszym ciągu będziemy zajmować się ciągłością względem metryki euklidesowej. Wolno więc będzie

Podchodzić do punktu z każdej strony o odległość będziemy mierzyć „zwyczajnie"
Zapiszmy więc definicję ciągłości dla funkcji na $\mathbb{R}^n$:

- $f: \mathbb{R}^n \to \mathbb{R}$ jest ciągła w $p=(x_0^1, \dots, x_0^n)$ jeśli $\forall \varepsilon>0 \ \exists \delta: \quad d(x,p) < \delta \Rightarrow |f(p)-f(x)| < \varepsilon$

Jest też wersja ciągowa:

- $f: \mathbb{R}^n \to \mathbb{R}$ jest ciągła w $p=(x_0^1, \dots, x_0^n)$ jeśli dla każdego ciągu (x_n) takiego, że $\lim_{n\to\infty} x_n = x_0$ $\lim f(x_n) = f(p_0)$.

Popatrzmy jakie komplikacje wiążą się z tym, że do punktu w $\mathbb{R}^n$ $(\mathbb{R}^2)$ można „podchodzić z różnych stron i na różne sposoby" a nie tylko z lewej czy z prawej.

PRZYKŁAD:
Rozważmy funkcję

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2} & (x,y) \neq (0,0) \\ 0 & (x,y) = (0,0) \end{cases}$$

Na osiach współrzędnych, czyli $\{(0,y)\}$ i $\{(x,0)\}$ f jest równa zero. Wędrując wzdłuż osi widzimy więc funkcję jako ciągłą. Wędrując wzdłuż innej prostej np $t \mapsto (t, at)$ widzimy

$$f(t,at) = \frac{at^2}{t^2+a^2t^2} = \frac{a}{1+a^2} \neq 0$$

Na innych prostych f też jest stała, ale nie równa 0, więc widzimy nieciągłość.

PRZYKŁAD: Teraz popatrzmy na funkcję

$$g(x,y)=\begin{cases}\dfrac{xy^2}{x^2+y^4} & (x,y)\neq(0,0)\\ 0 & (x,y)=(0,0)\end{cases}$$

Na osiach $\{(0,y)\}$ i $\{(x,0)\}$ g jest zero - nie widzimy nieciągłości; na dowolnej prostej

$$t\longmapsto(t,at) \quad \text{mamy}$$

$$g(t,at)=\frac{ta^2t^2}{t^2+a^4t^4}=\frac{a^2t}{1+a^4t^2}\xrightarrow{t\to 0}0$$

gdy jednak zmierzamy do $(0,0)$ wzdłuż krzywej

$$t\longmapsto(at^2,t) \quad \text{mamy}$$

$$g(at^2,t)=\frac{at^2t^2}{a^2t^4+t^4}=\frac{a}{a^2+1}\neq 0.$$

Nie prezentujemy przykładów okropnych w tym sensie jak diabelskie schody, choć one też istnieją.

Przejdźmy teraz do różniczkowania: Poszukiwanie pochodnej funkcji oznacza poszukiwanie pierwszego przybliżenia, które zawiera informację nie tylko o wartości funkcji (zerowe przybliżenie) w danym punkcie ale także trochę o zmienności tej funkcji. Pierwsze przybliżenie nie może też być zbyt skomplikowane. Przybliżać będziemy więc funkcję, której wykresem jest płaszczyzna. Zwyczajowo formułujemy to tak:

$$f(p+h) = a(p+h) + R(p,h) \quad (*)$$

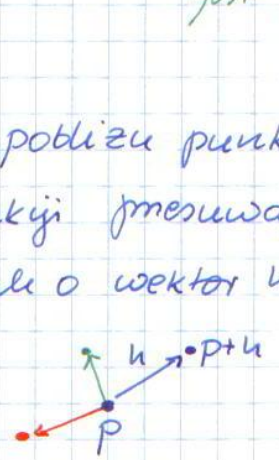
punkt w $\mathbb{R}^2$

przyrost (wektor)

reszta

odwzorowanie afiniczne – jego wykresem jest płaszczyzna

W pobliżu punktu $p \in \mathbb{R}^2$ badamy zachowanie funkcji przesuwając się trochę w różnych kierunkach o wektor h



$\mathbb{R}^2 \to \mathbb{R}$

Odwzorowanie afiniczne jest to takie odwzorowanie, którego wykresem jest płaszczyzna w $\mathbb{R}^3$.

Płaszczyzna przechodząca przez zero jest wykresem odwzorowania liniowego. Afiniczne odwzorowanie może mieć w $(0,0)$ inną wartość niż 0.

$$a(p+h) = a(p) + A(h)$$

Każde odwzorowanie afiniczne można tak zapisać. „A" nazywa się częścią liniową odwzorowania afinicznego.

z równości L i P oraz własności reszty wynika, że $a(p) = f(p)$

Reszta musi mieć następującą własność

$$\lim_{h \to 0} \frac{|R(p,h)|}{\|h\|} = 0,$$ czyli nie tylko

sama reszta jest coraz mniejsza gdy zbliżamy się do p (czyli z h do zera) ale także ona podzielona przez $\|h\|$!

Ostatecznie przybliżenie wygląda tak:

$$f(p+h) = f(p) + A(h) + R(p,h)$$

to jest pochodna!

Pochodna funkcji określonej na $\mathbb{R}^2$ jest odwzorowaniem liniowym z przestrzeni przyrostów ($\mathbb{R}^2$) do $\mathbb{R}$ czyli macierzą jednowierszową

$$A = [\ast \quad \ast]$$

Ogólniej: Pochodna funkcji określonej na $\mathbb{R}^n$ jest odwzorowaniem liniowym określonym na $\mathbb{R}^n$ (przyrosty) o wartościach w $\mathbb{R}$, czyli macierzą jednowierszową $[\underbrace{\ast \quad \ast \quad \cdots \quad \ast}_{n}]$

Jeszcze ogólniej: Pochodna ~~funkcji~~ odwzorowania określonego na $\mathbb{R}^n$ o wartościach w $\mathbb{R}^m$ jest odwzorowaniem liniowym z $\mathbb{R}^n$ do $\mathbb{R}^m$ czyli macierzą mającą n kolumn i m wierszy

$$\underbrace{\begin{bmatrix} \ast & \ast & \cdots & \ast \\ \ast & \ast & \cdots & \ast \\ \vdots & \vdots & \ddots & \\ \ast & \ast & \cdots & \ast \end{bmatrix}}_{n} \left.\vphantom{\begin{bmatrix} \ast \\ \ast \\ \vdots \\ \ast \end{bmatrix}}\right\} m$$

$$\boxed{\text{Wyrazy macierzy zależą od punktu } p \text{ w jakim wyznaczamy pochodną to znaczy mogą być funkcjami współrzędnych!}}$$

Jak to może wyglądać w praktyce?

Rozważmy funkcję $f: \mathbb{R}^2 \longrightarrow \mathbb{R}$

$f(x,y) = x^2 y$ i spróbujmy ją przybliżyć w otoczeniu punktu $p(1,2)$

$$f(p+h) = f(1+h_x, 2+h_y) = (1+h_x)^2(2+h_y) =$$
$$= (1+2h_x+h_x^2)(2+h_y) = 2+4h_x+2h_x^2+h_y+2h_xh_y+h_x^2h_y =$$
$$= 2 + \underbrace{(4h_x+h_y)}_{A(h)} + \underbrace{(2h_x^2+2h_xh_y+h_x^2h_y)}_{\text{Reszta}} =$$
$$= 2 + \begin{bmatrix} 4 & 1 \end{bmatrix}\begin{bmatrix} h_x \\ h_y \end{bmatrix} + (2h_x^2+2h_xh_y+h_x^2h_y)$$

Sprawdźmy, czy to jest dobra reszta

$$\frac{|R(p,h)|}{\|h\|} = \frac{|2h_x^2+2h_xh_y+h_x^2h_y|}{\sqrt{h_x^2+h_y^2}} \leq \frac{|2h_x^2+2h_xh_y+h_x^2h_y|}{|h_x|} =$$
$$= |2h_x+2h_y+h_xh_y| \xrightarrow{h\to 0} 0$$

Własność reszty jest zachowana!

Odwzorowanie $A = \begin{bmatrix} 4 & 1 \end{bmatrix}$ (zapis w bazie standardowej) jest pochodną f w punkcie $(1,2)$. Piszemy

$$f'(1,2) = \begin{bmatrix} 4 & 1 \end{bmatrix}$$

Powyższy przykład jest na tyle prosty, że można było znaleźć pochodną wprost analizując postać funkcji w otoczeniu punktu. Zazwyczaj tak jednak nie jest, a definicja pochodnej nie jest konstruktywna. Musimy więc sami wymyśleć jak znajdować wyrazy macierzowe

Odwzorowania liniowego będącego pochodną.
Przypominamy sobie że kolumny macierzy odwzorowania to wartości odwzorowania na wektorach bazowych (zapisane w bazie docelowej). Żeby znaleźć pochodną trzeba zatem patrzeć jak zmienia się funkcja w kierunku wektorów bazowych.

$f: \mathbb{R}^2 \to \mathbb{R}$

badamy $f(p+te_x)$ i $f(p+te_y)$

$$A = [A(e_x), A(e_y)]$$

$$f(p+te_x) = f(p) + A(te_x) + R(p, te_x) =$$
$$= f(p) + tA(e_x) + R(p, te_x)$$

$$f(p+te_x) - f(p) = tA(e_x) + R(p, te_x) \quad /:t$$

$$\frac{f(p+te_x)-f(p)}{t} = A(e_x) + \frac{R(p,te_x)}{t}$$

$$\frac{f(p+te_x)-f(p)}{t} \longrightarrow \frac{\partial f}{\partial x}(p) \qquad \frac{R(p,te_x)}{t} \xrightarrow{t \to 0} 0$$

z własności reszty

$\frac{\partial f}{\partial x}(p)$ ← pochodna cząstkowa

$p = (x, y)$

$$\frac{\partial f}{\partial x}(p) = \lim_{t \to 0} \frac{f(x+t, y) - f(x,y)}{t} = A(e_x)$$

Podobnie $A(e_y) = \frac{\partial f}{\partial y}(p) = \lim_{t \to 0} \frac{f(x, y+t) - f(x,y)}{t}$

Wyrazy macierzowe macierzy pochodnej to pochodne cząstkowe funkcji f obliczone w tym punkcie w którym obliczamy pochodną.

Pewna szczególna notacja: Pochodna funkcji zależnej od wielu zmiennych nazywana jest też RÓŻNICZKĄ. Stosuje się także specyficzną notację. Na przykład dla funkcji

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

normalnie napisalibyśmy

$$f'(x,y) = \left[\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y} \right] = [2xy, x^2]$$

Według tej szczególnej notacji zamiast

$\cancel{\phi}$ $f'(x,y)$ piszemy df

Po prawej stronie ~~de~~ też wprowadzamy szczególne oznaczenie: Zauważmy, że

$$[2xy, x^2] = 2xy \cdot [1,0] + x^2 [0,1]$$

Oznaczając $\cancel{2xy}$ $[1,0] = dx$ i $[0,1] = dy$

Dostajemy

$$df = 2xy \cdot dx + x^2 dy$$

Ogólnie $df = \frac{\partial f}{\partial x} dx + \frac{\partial f}{\partial y} dy$

Oznaczenia dx i dy są sensowne. Możemy bowiem rozważać współrzędne jako szczególne funkcje na $\mathbb{R}^2$ $(x,y) \longmapsto x$ $(x,y) \longmapsto y$; pochodne tych funkcji to

odpowiednio $[1, 0]$ i $[0\ 1]$ niezależnie od (12)
tego w którym punkcie je obliczamy. Istotnie
więc $dx = [1\ 0]$, $dy = [0, 1]$