NI-MPI přednáška 1

Analýza I: funkce více proměnných

23. 9. 2024

FIT ČVUT

Norma a vzdálenost – připomenutí

Definice 1.1 (Norma *norm*)

Norma na vektorovém prostoru V (nad $\mathbb{R}$ nebo $\mathbb{C}$) je zobrazení $\|\cdot\| : V \to \mathbb{R}_0^+$ splňující:

1. $\|\mathbf{x}\| = 0 \Rightarrow \mathbf{x} = \mathbf{0}$,
2. $\|\alpha\mathbf{x}\| = |\alpha| \cdot \|\mathbf{x}\|$,
3. $\|\mathbf{x}+\mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (trojúhelníková nerovnost),

pro všechna $\mathbf{x}, \mathbf{y} \in V$ a všechny skaláry α.

Máme-li normu $\|\cdot\|$ na V, definujeme vzdálenost vektorů $\mathbf{x}, \mathbf{y} \in V$ jako $d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$. Zřejmě platí

- $d(\mathbf{x}, \mathbf{y}) = 0 \Leftrightarrow \mathbf{x} = \mathbf{y}$;
- $d(\mathbf{x}, \mathbf{y}) = d(\mathbf{y}, \mathbf{x})$ (symetrie);
- $d(\mathbf{x}, \mathbf{z}) \leq d(\mathbf{x}, \mathbf{y}) + d(\mathbf{y}, \mathbf{z})$ (trojúhelníková nerovnost).

Příklady norem

Pro $\mathbf{x} = (x_1, x_2, \ldots, x_n) \in \mathbb{R}^n$(nebo $\mathbb{C}^n$)

- $\|\mathbf{x}\|_2 := \sqrt[2]{\sum_{i=1}^n |x_i|^2}$ (eukleidovská norma)
- $\|\mathbf{x}\|_1 := \sum_{i=1}^n |x_i|$ (součtová norma)
- Obecně, pro libovolné $p \geq 1$

$$\|\mathbf{x}\|_p := \sqrt[p]{\sum_{i=1}^n |x_i|^p}$$

- $\|\mathbf{x}\|_\infty := \max\left\{|x_i| \,\middle|\, i \in \{1, \ldots, n\}\right\}$ (maximová norma)

(Ověření trojúhelníkové nerovnosti není vždy triviální!)

Okolí bodu

Zvolme nějakou normu $\|\cdot\|$ na $\mathbb{R}^n$.

Buď $\mathbf{x} \in \mathbb{R}^n$ a $\delta \in \mathbb{R}^+$, **δ-okolí bodu $\mathbf{x}$** je množina

$$H_\delta(\mathbf{x}) = \{\mathbf{b} \in \mathbb{R}^n \colon \|\mathbf{x} - \mathbf{b}\| < \delta\}.$$

(Taková množina se také nazývá otevřená koule o středu $\mathbf{x}$ a poloměru δ.)

Obecně pro okolí: pokud není třeba specifikovat parametr δ budeme psát jednoduše $H(\mathbf{x})$.

Hromadný bod

Mějme množinu $M \subset \mathbb{R}^n$.

Řekneme, že $\mathbf{x} \in \mathbb{R}^n$ je **hromadným bodem** M (*limit/cluster/accumulation point of* M), pokud pro všechna $r > 0$ platí $(H_r(\mathbf{x}) \setminus \{\mathbf{x}\}) \cap M \neq \emptyset$.

Bod $\mathbf{x} \in M$, který není hromadný, se nazývá **izolovaný**.

Limita posloupnosti

Mějme posloupnost bodů z $\mathbb{R}^n$, tedy

$$(\mathbf{x}_i)_{i=0}^{+\infty},$$

kde $\mathbf{x}_i \in \mathbb{R}^n$.

Řekneme, že posloupnost $(\mathbf{x}_i)_{i=0}^{+\infty}$ **má limitu** $L \in \mathbb{R}^n$, pokud

$$\forall \epsilon > 0 \quad \exists N \quad \forall n > N \quad \mathbf{x}_n \in H_\epsilon(L).$$

Značení:

$$\lim_{n \to \infty} \mathbf{x}_n = L$$

Funkce více proměnných

Reálnou funkcí více reálných proměnných rozumíme zobrazení

$$f : D_f \to \mathbb{R},$$

kde $D_f \subseteq \mathbb{R}^n$ (pro n kladné celé).

Tedy funkce, která má n reálných parametrů a vrací taktéž reálnou hodnotu.

D_f je **definiční obor** (*domain*).

$f(D_f)$ je **obor hodnot** (*range*).

Graf funkce

Graf funkce f je množina

$$\Gamma_f = \{(b_1, b_2, \ldots, b_n, f(b_1, b_2, \ldots, b_n)) \colon (b_1, b_2, \ldots, b_n) \in D_f\} \subset \mathbb{R}^{n+1}.$$

Limita funkce více proměnných

Řekneme, že funkce $f : D_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^n$, **má limitu** $L \in \mathbb{R}$ **v hromadném bodě** b **množiny** D_f pokud

$$\forall H(L) \quad \exists H(\mathbf{b}) \quad \mathbf{x} \in (D_f \cap H(\mathbf{b})) \setminus \{\mathbf{b}\} \Longrightarrow f(\mathbf{x}) \in H(L)$$

Značení:

$$\lim_{\mathbf{x} \to \mathbf{b}} f(\mathbf{x}) = L.$$

Limity pomocí posloupnosti

Věta 1.2

Mějme funkci $f : D_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^n$. *Funkce* f *má v hromadném bodě* $\mathbf{b}$ *množiny* D_f *limitu* L, *tedy* $\lim_{\mathbf{x}\to\mathbf{b}} f(\mathbf{x}) = L$, *právě tehdy, když pro všechny posloupnosti* $(\mathbf{x}_i)_{i=0}^{+\infty} \subset D_f$, $\mathbf{x}_i \neq \mathbf{b}$, *platí*

$$\lim_{n\to+\infty} \mathbf{x}_n = \mathbf{b} \quad \Longrightarrow \quad \lim_{n\to+\infty} f(\mathbf{x}_n) = L$$

Spojitá funkce

Řekneme, že funkce $f : D_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^n$, je **spojitá v bodě** $\mathbf{x}_0 \in D_f$ pokud

$$\forall \epsilon > 0 \quad \exists \delta > 0 : \quad x \in (D_f \cap H_\delta(\mathbf{x}_0)) \implies f(x) \in H_\epsilon(f(\mathbf{x}_0)).$$

Funkce f je **spojitá** pokud je spojitá ve všech bodech z D_f.

Lze formulovat i pomocí limity: funkce f je spojitá, pokud pro všechny neizolované body $\mathbf{x}_0 \in D_f$ platí

$$\lim_{\mathbf{x} \to \mathbf{x}_0} f(\mathbf{x}) = f(\mathbf{x}_0).$$

Lokální extrémy

Definice 1.3

Řekneme, že reálná funkce f má v bodě $\mathbf{b} \in D_f$

1. **lokální minimum**, pokud $\exists \delta > 0, \forall \mathbf{x} \in (D_f \cap H_\delta(\mathbf{b})), f(\mathbf{x}) \geq f(\mathbf{b})$;
2. **ostré lokální minimum**, pokud $\exists \delta > 0, \forall \mathbf{x} \in (D_f \cap H_\delta(\mathbf{b})) \setminus \{\mathbf{b}\}, f(\mathbf{x}) > f(\mathbf{b})$;
3. **globální minimum**, pokud $\forall \mathbf{x} \in D_f, f(\mathbf{x}) \geq f(\mathbf{b})$.

Bod $\mathbf{b}$ se nazývá postupně bodem lokálního minima, bodem ostrého lokálního minima a bodem globálního minima funkce f.

Pro globální minimum (na množině $D \subset D_f$) se též používá **argument minima** s tímto (zneužitým) značením:

$$\mathbf{b} = \operatorname*{argmin}_{\mathbf{x} \in D} f(\mathbf{x}).$$

(Obecně by argmin měl být vzor minima, tj. množina bodů.)

(Ostré) lokální maximum a globální maximum se definuje analogicky.

Extrémy na omezených a uzavřených množinách

Definice 1.4

Množina $D \subset \mathbb{R}^n$ se nazývá

- **omezená**, pokud je podmnožinou nějaké otevřené koule;
- **otevřená**, pokud s každým svým bodem obsahuje i nějaké jeho okolí;
- **uzavřená**, pokud $\mathbb{R}^n \setminus D$ je otevřená (obsahuje i svou hranici, tedy body, v jejichž každém okolí leží bod z D a bod mimo D).

(Uzavřenou množinu lze také charakterizovat tím, že obsahuje všechny své hromadné body.)

Věta 1.5

*Je-li $D_f \subset \mathbb{R}^n$ **omezená** a **uzavřená**, pak má spojitá funkce $f : D_f \to \mathbb{R}$ v D_f globální minimum a globální maximum.*

Definice parciální derivace

Označme jednotlivé proměnné jako $x_1, x_2, x_3, \ldots, x_n$.

Definice 2.1

Parciální derivace funkce f ve směru x_i (nebo podle x_i) v bodě $\mathbf{b} = (b_1, b_2, \ldots, b_n) \in D_f$ takovém, že $\exists H(\mathbf{b}) \subset D_f$, je

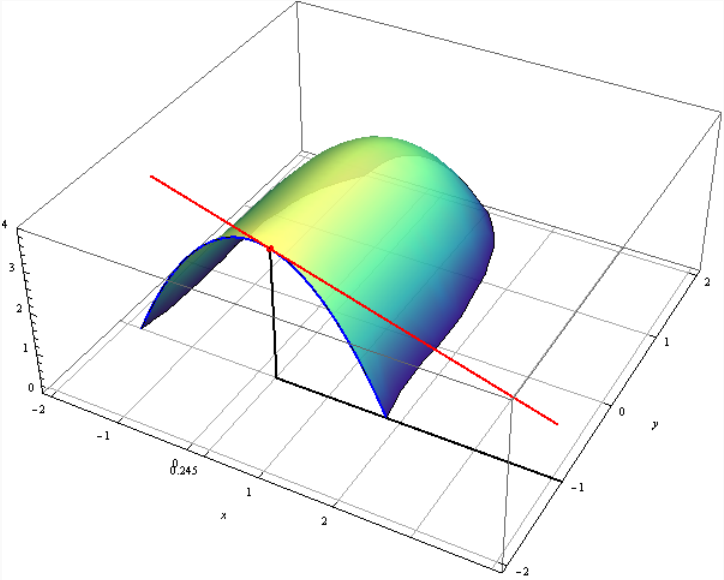
$$\lim_{h \to 0} \frac{f(b_1, b_2, \ldots, b_i + h, \ldots, b_n) - f(b_1, b_2, \ldots, b_i, \ldots, b_n)}{h} = L \in \mathbb{R},$$

pokud tato limita existuje.

Značení: $\frac{\partial f}{\partial x_i}(\mathbf{b}) = L$. (Jiná častá značení: $\partial_{x_i} f(\mathbf{b})$ a $f_{x_i}(\mathbf{b})$.)

Jedná se o směrnici tečny ke grafu funkce f ve směru osy x_i.

Parciální derivace - ilustrace



Gradient

Definice 2.2

Gradient funkce f v bodě $\mathbf{b} \in D_f$ je (řádkový) vektor

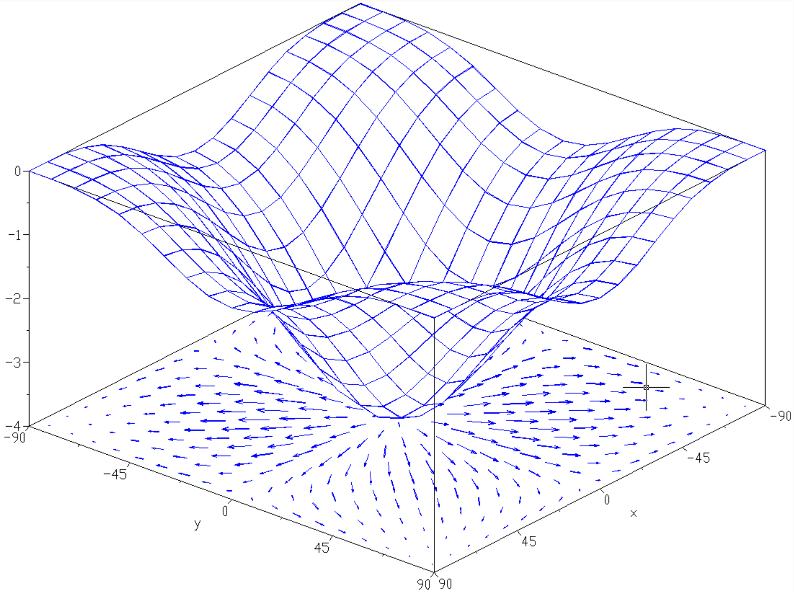
$$\nabla f(\mathbf{b}) = \left(\frac{\partial f}{\partial x_1}(\mathbf{b}), \frac{\partial f}{\partial x_2}(\mathbf{b}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{b}) \right).$$

Poznámka: jelikož se gradient opírá o pojem parc. derivace, bod $\mathbf{b}$ je nutně vnitřním bodem D_f (v D_f leží i nějaké jeho okolí).

Poznámka 2: uvedená definice je zjednodušená: lze ji použít pokud jsou všechny parciální derivace funkce f spojité na nějakém okolí bodu $\mathbf{b}$. (Nezjednodušená definice gradientu je v handoutu.)

Geometrický význam: gradient ukazuje směr (v D_f) nejvyššího růstu funkce f.

Gradient - ilustrace



Derivace ve směru

Parciální derivace je vždy ve směru některé ze souřadnicových os x_i. Co směrnice v jiných směrech (v D_f)?

Definice 2.3

Nechť $\mathbf{v} \in \mathbb{R}^{n,1} = \mathbb{R}^n$, $\|\mathbf{v}\| = 1$. **Derivace funkce f ve směru $\mathbf{v}$ v bodě $\mathbf{b}$** $\in D_f$ takovém, že $\exists H(\mathbf{b}) \subset D_f$, je

$$\nabla_{\mathbf{v}} f(\mathbf{b}) = \lim_{h \to 0} \frac{f(\mathbf{b} + h\mathbf{v}) - f(\mathbf{b})}{h}.$$

Věta 2.4

Jsou-li všechny parciální derivace funkce f na nějakém okolí bodu $\mathbf{b}$ spojité, pak platí

$$\nabla_{\mathbf{v}} f(\mathbf{b}) = \nabla f(\mathbf{b}) \cdot \mathbf{v}.$$

Poznámka: někdy se nevyžaduje, aby směr $\mathbf{v}$ byl jednotkový, tj. $\|\mathbf{v}\| = 1$. (Potom se o $\mathbf{v}$ nemluví jako o směru, ale o obecném vektoru.)

Tečná nadrovina

Obdobně jako tečna pro $n = 1$, tak tečnou nadrovinou rozumíme objekt sídlící ve stejném prostoru jako graf funkce (tedy v $\mathbb{R}^{n+1}$), který je lineární varietou kodimenze 1 a který, v nějakém smyslu, nejlépe postihuje chování přírustku funkce f v daném bodě $\mathbf{b} \in D_f$.

Toto chování umíme popsat pro jednotlivé směry $\mathbf{v} \in \mathbb{R}^n$: pro každý směr umíme v tomto směru najít tečnu funkce f zúžené na tento směr, tj. zúžené na $D_f \cap \{\mathbf{x} \colon \mathbf{x} = \mathbf{b} + t\mathbf{v}, t \in \mathbb{R}\}$. Toto zúžení lze chápat jako funkci jedné proměnné, tedy $\mathbb{R} \to \mathbb{R}$, a u té umíme najít tečnu.

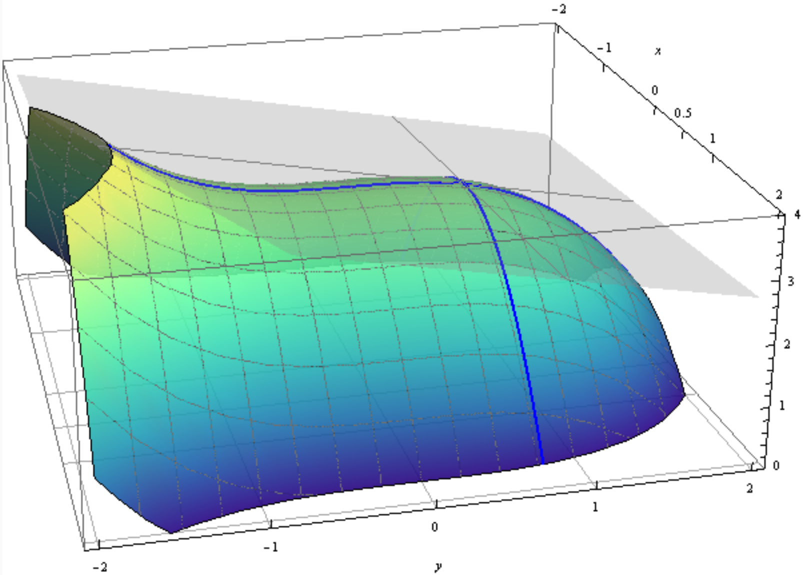
Sjednotíme-li tečny ve všech směrech (v bodě $\mathbf{b} \in D_f$), dostaneme **tečnou nadrovinu funkce f v bodě $\mathbf{b}$**. (Podmínkou pro existenci je existence gradientu f v bodě $\mathbf{b}$.)

Rovnice této nadroviny je

$$z = \frac{\partial f}{\partial x_1}(\mathbf{b})\,(x_1 - b_1) + \frac{\partial f}{\partial x_2}(\mathbf{b})\,(x_2 - b_2) + \cdots + \frac{\partial f}{\partial x_n}(\mathbf{b})\,(x_n - b_n) + f(\mathbf{b}).$$

Její normálový vektor je $\left(\frac{\partial f}{\partial x_1}(\mathbf{b}), \frac{\partial f}{\partial x_2}(\mathbf{b}), \ldots, \frac{\partial f}{\partial x_n}(\mathbf{b}), -1\right)$.

Tečná nadrovina - ukázka



Věta 3.1 (Nutná podmínka lokálního extrému)

Nechť má funkce $f : D_f \to \mathbb{R}$, $D_f \subset \mathbb{R}^n$, v bodě $\mathbf{b}$ parciální derivaci podle i-té proměnné.

Pokud f má v bodě $\mathbf{b}$ lokální extrém, potom

$$\frac{\partial f}{\partial x_i}(\mathbf{b}) = 0.$$

Důkaz Věty 3.1.

Zavedeme funkci jedné proměnné:

$$g(x_i) = f(b_1, b_2, \ldots, b_{i-1}, x_i, b_{i+1}, \ldots, b_n).$$

Funkce g je v bodě b_i diferencovatelná a pro její derivaci platí

$$g'(b_i) = \frac{\partial f}{\partial x_i}(\mathbf{b}).$$

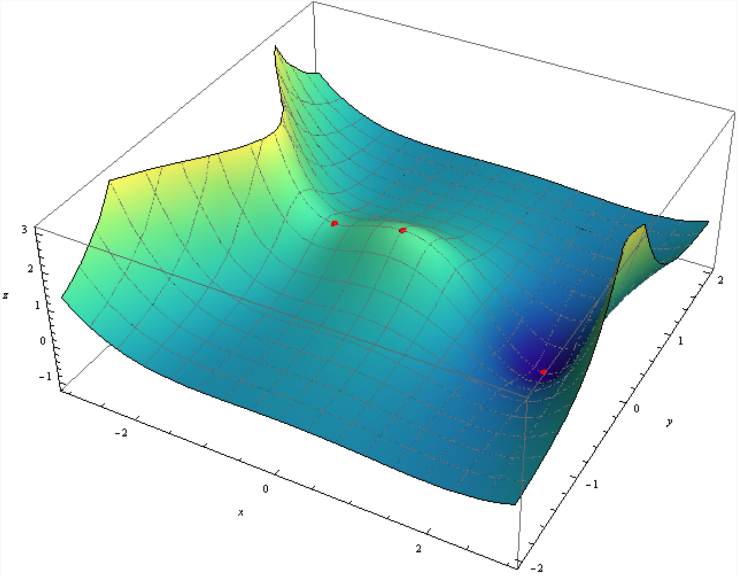
Protože f má v $\mathbf{b}$ lokální extrém, má g v b_i lokální extrém, a tedy $g'(b_i) = 0$. $\square$

Důsledek: pokud existuje gradient funkce f v bodě $\mathbf{b}$, pak existence lokálního extrému implikuje

$$\nabla f(\mathbf{b}) = \mathbf{0}.$$

Body $\mathbf{b} \in D_f$ splňující $\nabla f(\mathbf{b}) = 0$ se nazývají **stacionární**.

Stacionární body



Kritické body

Mějme funkci $f : D_f \to \mathbb{R}$, kde $D_f \subseteq \mathbb{R}^n$.

V úloze hledání (lokálních) extrémů funkce f jsou body podezřelé z extrému buď

- body stacionární

anebo

- body, ve kterých gradient f neexistuje.

Takové body se souhrně nazývají **kritické body**.