

PV027 Optimalizace

Radka Svobodová

Definice optimalizace

Již v minulosti řada matematiků a přírodovědců dospěla k přesvědčení, že přírodní děje lze popsat jako optimalizační procesy.

Euler: *„Na světě se nestane nic, v čem by nebylo vidět smysl nějakého maxima nebo minima“.*

Leibnitz: *„Náš svět je nejlepší ze všech možných světů, a proto lze jeho zákony vyjádřit extrémálními principy“.*

Definice optimalizace II

Optimalizaci lze definovat například následovně:

Obor zabývající se určením nejlepšího řešení jistého matematicky definovaného problému.

Definice optimalizace III

Postup při optimalizaci:

- Nastudování optimalizačních kritérií problému
- Nalezení metody řešení problému
- Matematický popis řešení (pomocí funkce)
- **Nalezení minima funkce**

Definice optimalizace IV

Předmět PV027 se zabývá matematickou optimalizací, tedy **minimalizací reálných funkcí**, tj. úlohami typu:

$$\min_{x \in M} f(x)$$

$$\text{kde: } f: R^n \rightarrow R$$

$$M \subseteq R^n$$

Definice optimalizace V

Poznámka:

Není nutné zabývat se samostatně také maximalizací, neboť ji lze převést na minimalizaci pomocí vztahu:

$$\max_{x \in M} f(x) = - \min_{x \in M} (-f(x))$$

Definice optimalizace VI

Speciální oblasti optimalizace:

- Diskrétní optimalizace:

Využívají se v případech, kdy mají význam pouze celočíselná řešení.

- Speciální úlohy:

Existuje určitý typ úloh, jejichž řešení nemá příliš smysl popisovat jako $\min_{x \in M} f(x)$.

Příklad: problém obchodního cestujícího

Sylabus

- Úvod
 - Definice optimalizace
 - Informace o předmětu
 - Sylabus předmětu
 - Vstupní požadavky předmětu
 - Požadavky ke zkoušce a k zápočtu
 - Materiály ke studiu
 - Aplikace optimalizačních metod
 - Motivační příklady
 - Matematické základy

Sylabus II

- Optimalizace bez omezení (unconstrained)
 - Nelder-Meadova metoda
 - spádové metody
 - metody sdruženého gradientu
 - newtonovské metody
 - metody s omezeným krokem
 - úloha nejmenších čtverců

Sylabus III

- Optimalizace s omezením (constrained)
 - Lineární programování:
 - grafické řešení úloh
 - přímá metoda
 - simplexová metoda
 - Nelineární programování:
 - Obecný přístup
 - kvadratické programování
 - Celočíselné programování
 - metoda větví a mezí

Sylabus IV

- Globální optimalizace:
 - simulované žíhání
 - genetické algoritmy
 - metoda difuzní rovnice
- Praktický projekt
 - Aplikace vybrané optimalizační metody pro řešení konkrétního problému

Vstupní požadavky

- **Znalosti z oblastí:**
 - Lineární algebra
 - Matematická analýza

Požadavky ke zkoušce a hodnocení

Požadavky: znalosti v rozsahu přednášek +
vypracovaný projekt

Hodnocení:

- A: 100 - 90 %
- B: 90 - 80 %
- C: 80 - 70 %
- D: 70 - 60 %
- E: 60 - 50 %
- F: 49 - 0 %

Požadavky k zápočtu

Požadavky:

Vypracovat **zápočtový projekt** (zadání projektu získá student po domluvě s učitelem).

Hodnocení:

z projekt splňuje požadavky, domluvené při zadávání

n jinak

Poznámka:

Předmět PV027 nelze ukončit kolokviem.

Aplikace optimalizačních metod

Optimalizační úlohy se vyskytují všude tam, kde je možné vybírat si z více možných rozhodnutí a přitom kvalitu jednotlivých rozhodnutí ohodnotit nějakým reálným číslem. Konkrétní oblasti využití optimalizace:

Matematika: teorie aproximace, optimalizace numerických procesů atd.

Geometrie: geodézie, minimální křivky a plochy, optimální oblasti a tvary, atd.

Ekonomické a politologické teorie: využívání zdrojů a zásob, optimální skladba výroby, tvorba cen, rozložení rizika, strategické hry, teorie eskalace konfliktu, atd.

Fyzika: mechanika, geometrická optika, teorie pružnosti, hydrodynamika, teorie relativity, atd.

Aplikace optimalizačních metod II

Přírodní vědy: modely fyzikálních, chemických a biologických procesů, atd.

Teorie řízení: optimální řízení, optimální systémy, hierarchické řízení, koordinační strategie, atd.

Teorie konstruování: optimalizace konstrukcí, optimalizace tvarů, optimální odhad neznámých parametrů, optimalizace dynamických vlastností mechanických systému, optimalizace spolehlivosti a rizika konstrukcí, atd.

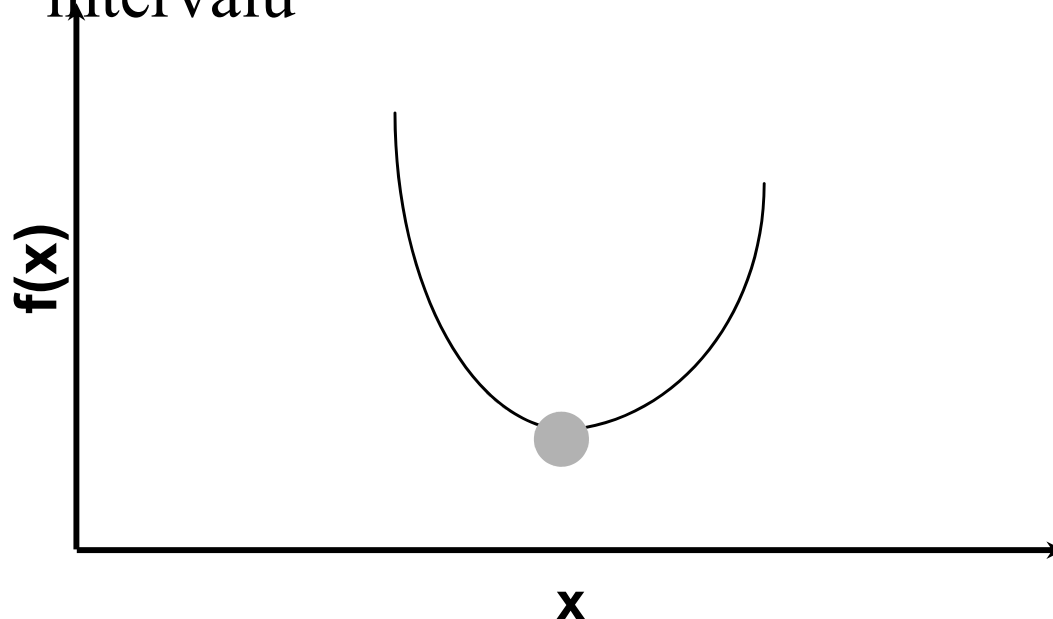
Další aplikace: při správě vodních toků (vypouštění a napouštění nádrží), v zemědělství (např. optimální krmná směs pro zvířata), v dopravě a logistice a kdekoliv jinde, kde se nám podaří zformulovat smysluplnou optimalizační úlohu

Typy optimalizací

Lokální X globální:

Lokální optimalizace:

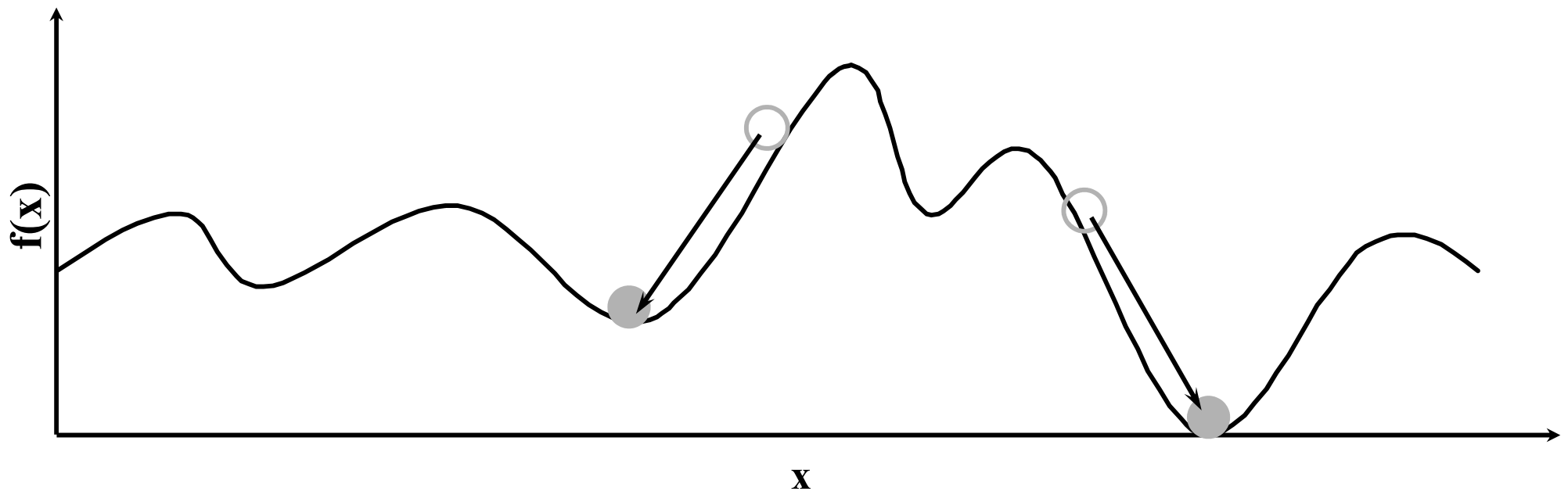
- Nalezení jediného minima, nacházejícího se v určitém intervalu



Typy optimalizací II

Lokální optimalizace – další možný přístup:

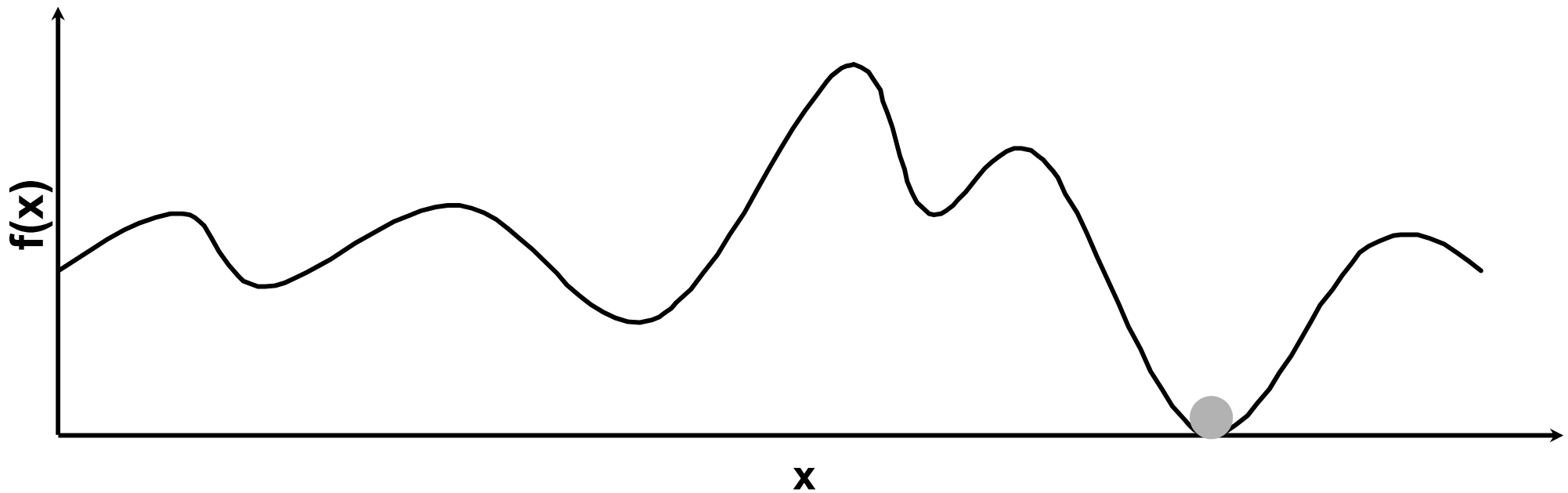
- Nalezení nejbližšího minima do kterého lze sestoupit ze vstupního bodu



Typy optimalizací III

Globální optimalizace:

- Hledání nejhlubšího minima v daném intervalu



Typy optimalizací IV

Bez omezení X s omezeními:

Bez omezení (unconstrained)

Kromě podmínky minimality funkce $f(\mathbf{x})$ neexistuje žádná jiná podmínka, kterou by měla hledaná optimální hodnota \mathbf{x} splňovat.

S omezeními (constrained)

Vedle podmínky minimality pracujeme ještě s dalšími podmínkami.

Lineární X nelineární:

Lineární

$f(\mathbf{x})$ je lineární funkcí \mathbf{x}

Nelineární

jinak

Motivační příklady

Aproximace dat: nelineární model bez omezení

Najděte nejlepší aproximaci pomocí součtu čtverců pro funkci:

$$\phi(t, \mathbf{x}) = \left(1 - \frac{\mathbf{x}_1 t}{\mathbf{x}_2} \right)^{\frac{1}{\mathbf{x}_1 c} - 1}$$

kde $c = 96,05$ a aproximované body dány tabulkou

Motivační příklady II

Plánování výroby: lineární model s omezeními

Výrobce používá materiál (m) a práci (l) k výrobě nejvýše čtyř výrobků (a až d). Požadavky na jednotlivé výrobky a zisk z jejich prodeje jsou dány následovně:

k výrobě 1 kusu výrobku a je třeba $4m + 2l$ zdrojů a zisk je 50 Kč/kus

k výrobě 1 kusu výrobku b je třeba $m + 5l$ zdrojů a zisk je 80 Kč/kus

k výrobě 1 kusu výrobku c je třeba $2m + l$ zdrojů a zisk je 30 Kč/kus

k výrobě 1 kusu výrobku d je třeba $2m + 3l$ zdrojů a zisk je 40 Kč/kus

Během jednoho dne je k dispozici nejvýše 30 jednotek materiálu a 50 jednotek (např. hodin) práce. V jakém počtu mají být produkovány jednotlivé výrobky tak, aby bylo dosaženo maximálního zisku?

Motivační příklady III

Konstrukční problém: nelineární model s omezeními

Navrhněte válcovitou plechovku o objemu 1 dm^3 , na kterou se spotřebuje co nejméně kovu.

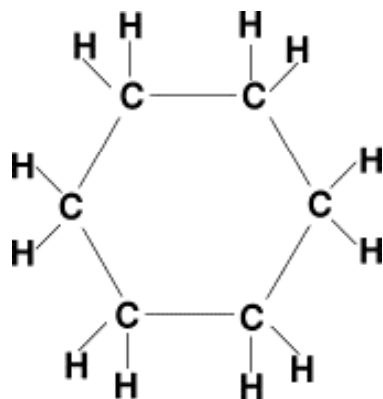
Motivační příklady IV

Aplikace v chemii: nelineární model bez omezení (s více lokálními minimy)

Je dána molekula, urči konformace této molekuly, které jsou v definovaném chemickém prostředí nejstabilnější.

Konformace = uspořádání atomů v prostoru.

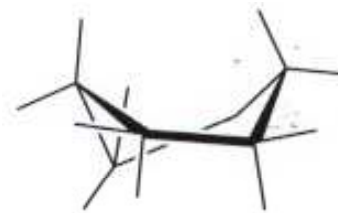
Strukturní vzorec:



Konformace:



Židličková



Zkřížená židličková



Vaničková



Poloviční židličková

Motivační příklady IV b)

Čím je konformace stabilnější, tím nižší má potenciální energii.

Potenciální energie = energie daného uspořádání atomů v prostoru.

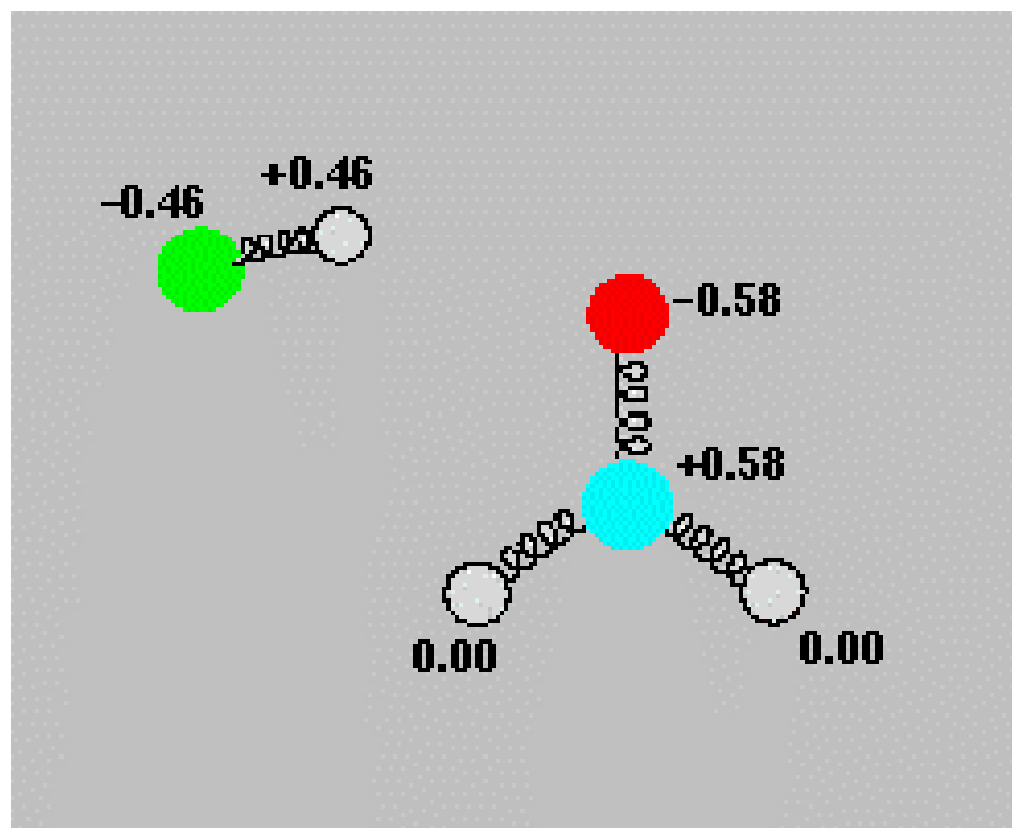
$E_{\text{pot}} = \phi(\text{souřadnic atomů})$

kde ϕ je potenciálová funkce

$\phi : \mathbf{R}^{3N} \rightarrow \mathbf{R}$, kde N je počet atomů v molekule.

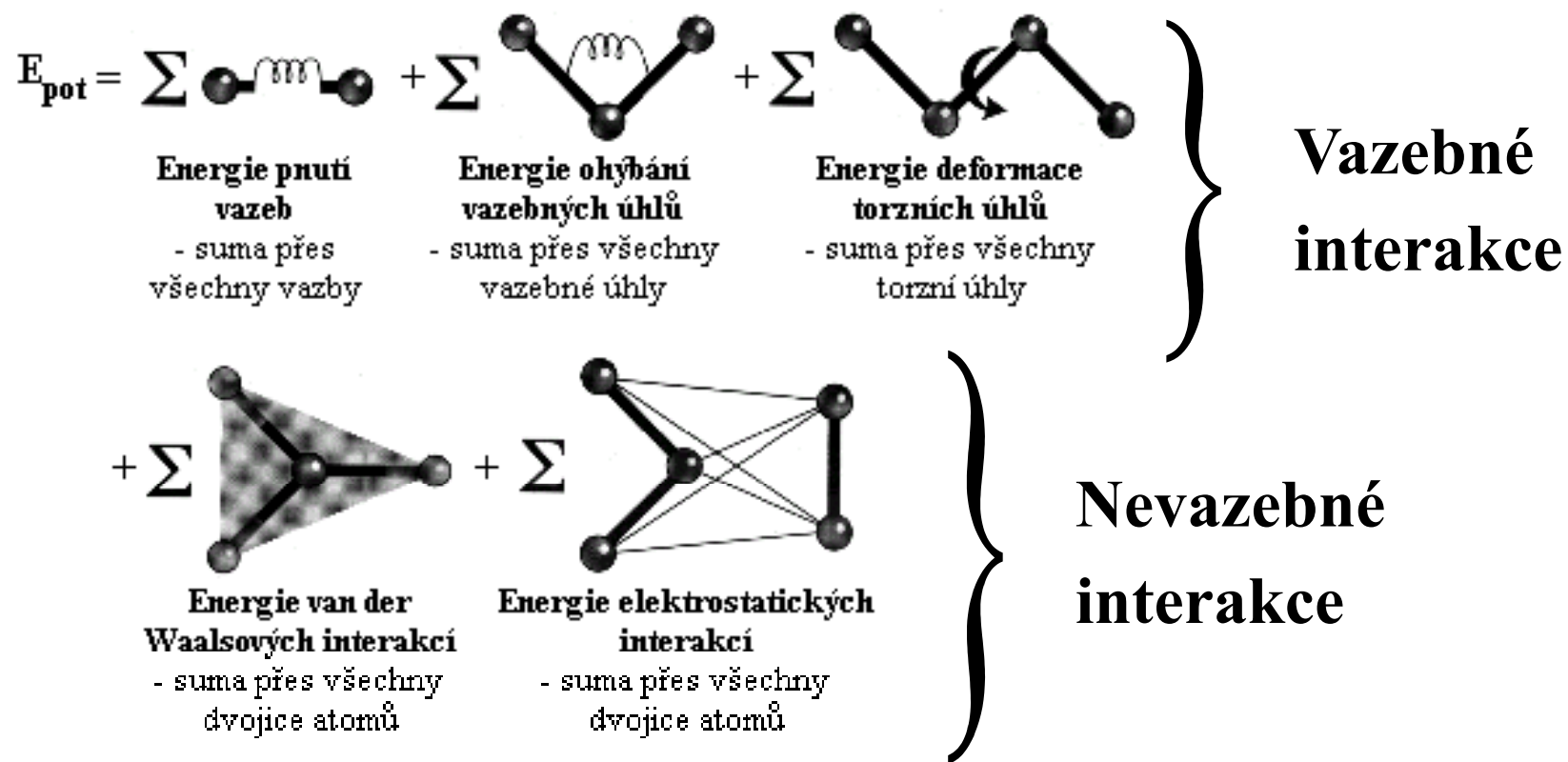
Motivační příklady IV c)

Vytvoříme model molekuly:
Nabité koule, spojené pružinami.



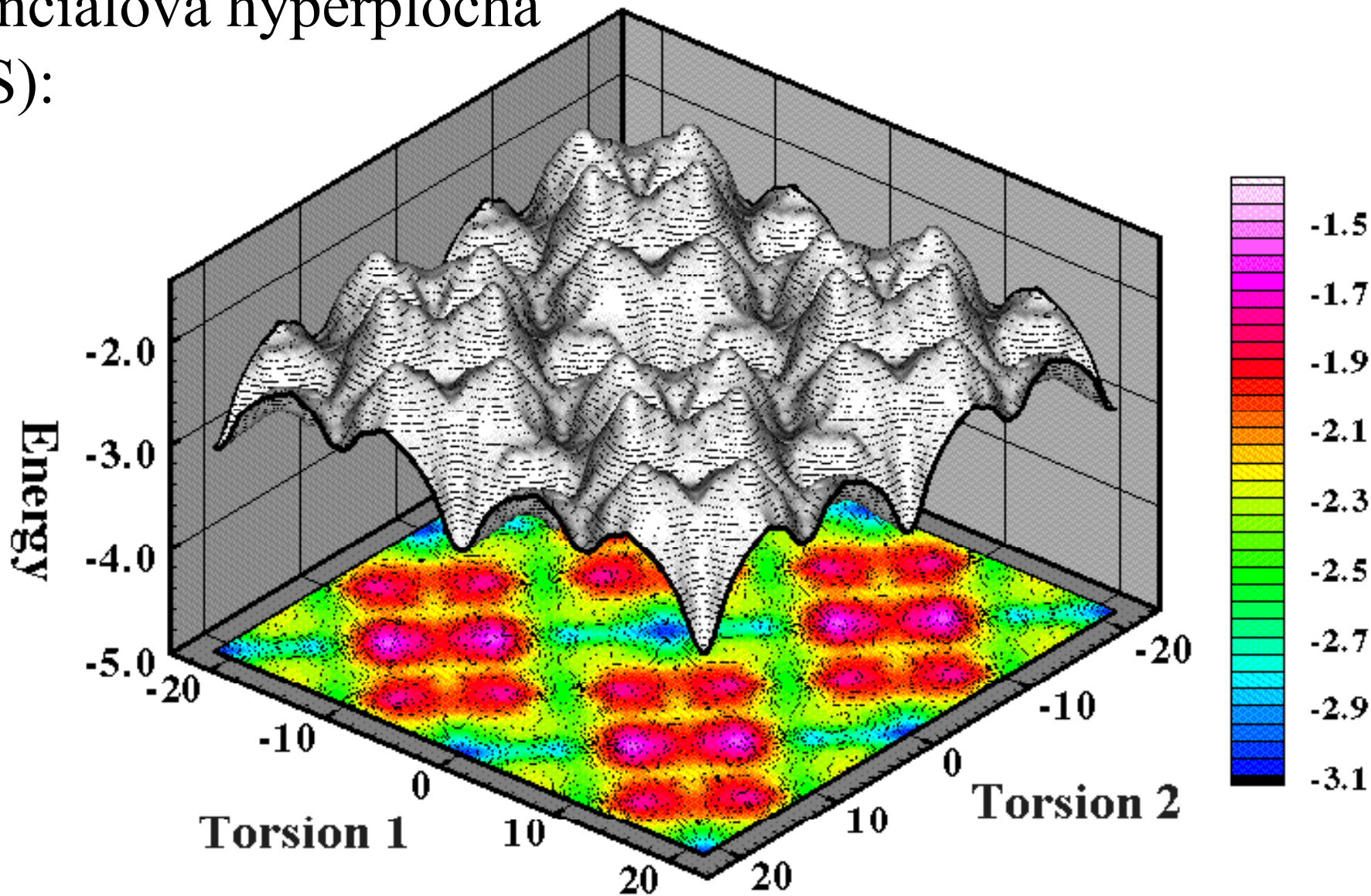
Motivační příklady IV d)

Popíšeme vztah mezi souřadnicemi a E_{pot} :



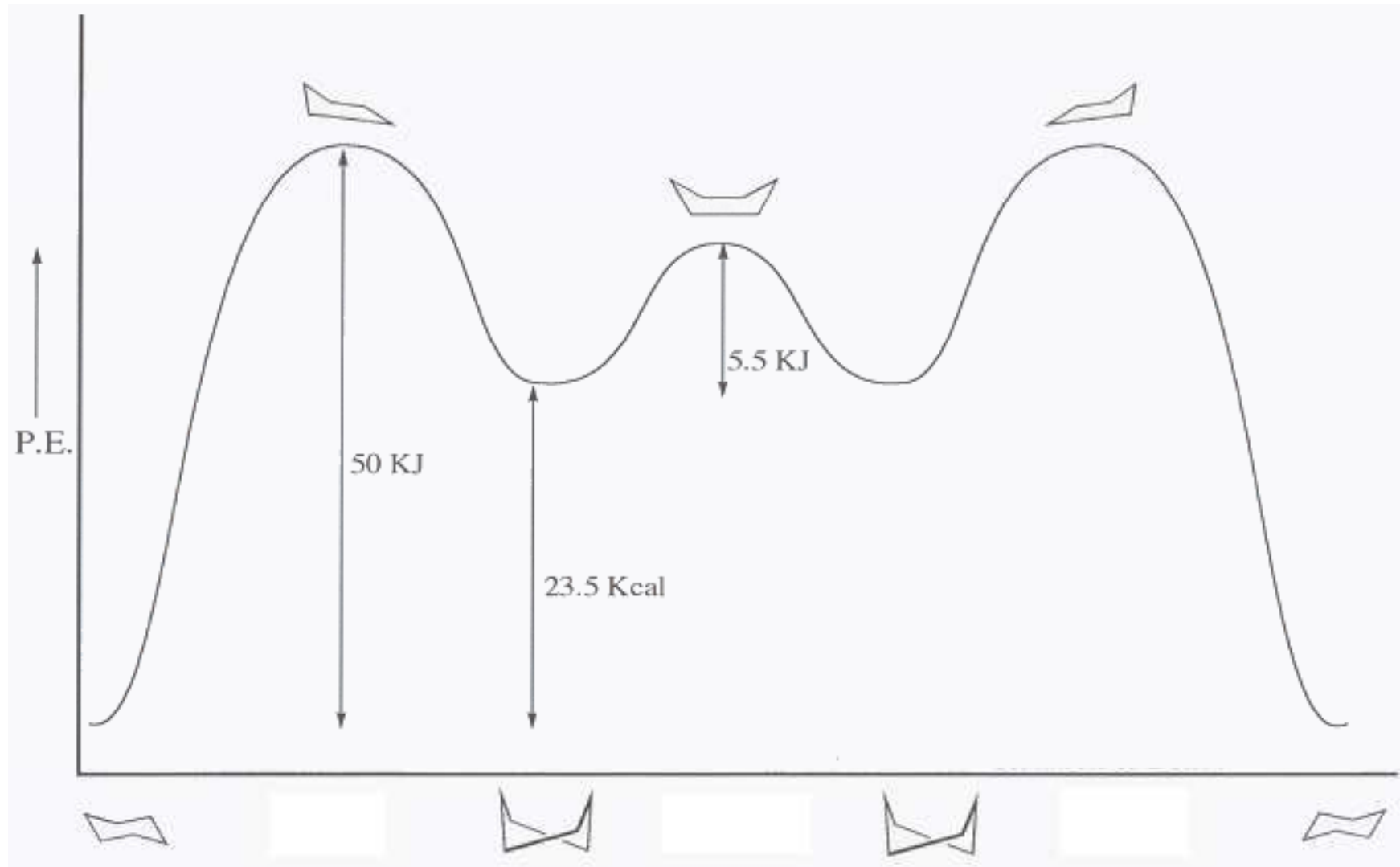
Motivační příklady IV e)

Grafem potenciálové funkce je
potenciálová hyperplocha
(PES):



Motivační příklady IV f)

Hledáme minima potenciálové funkce.



Matematické základy

Vektory:

- budeme pracovat hlavně se sloupcovými vektory
- vektor budeme označovat malými tučnými písmeny
- transpozici vektoru označíme pomocí písmene T v horním indexu

$$\mathbf{a} = \begin{pmatrix} a_1 \\ a_2 \\ \dots \\ a_n \end{pmatrix}$$

$$\mathbf{a}^T = (a_1, a_2, \dots, a_n)$$

Matematické základy II

Matice:

- budou se označovat velkými tučnými písmeny
- transpozici matice označíme pomocí písmene T v horním indexu

$$\mathbf{B} = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} & \dots & \mathbf{B}_{1n} \\ \mathbf{B}_{21} & \mathbf{B}_{22} & \dots & \mathbf{B}_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_{m1} & \mathbf{B}_{m2} & \dots & \mathbf{B}_{mn} \end{pmatrix}$$

$$\mathbf{B}^T = \begin{pmatrix} \mathbf{B}_{11} & \mathbf{B}_{21} & \dots & \mathbf{B}_{m1} \\ \mathbf{B}_{12} & \mathbf{B}_{22} & \dots & \mathbf{B}_{m2} \\ \vdots & \vdots & \ddots & \vdots \\ \mathbf{B}_{1n} & \mathbf{B}_{2n} & \dots & \mathbf{B}_{mn} \end{pmatrix}$$

Matematické základy III

Skalární součin:

- skalární součin vektorů \mathbf{a} a \mathbf{z} zapíšeme jako:

$$\mathbf{a}^T \mathbf{z} = \mathbf{z}^T \mathbf{a} = \sum_i a_i z_i$$

Posloupnost bodů z \mathbb{R}^n :

- jednotlivé body v posloupnosti budeme rozlišovat horním indexem v závorce. Budeme je tedy psát například jako $\mathbf{x}^{(1)}$, $\mathbf{x}^{(2)}$... nebo $\mathbf{x}^{(k)}$.
- speciální význam bude mít hvězdička (například v \mathbf{x}^*). Budeme tak označovat bod, který je řešením vyšetřovaného problému.

Matematické základy IV

Přímka v \mathbb{R}^n :

- je definována jako množina bodů:

$$\mathbf{x} = f(\alpha) = \mathbf{x}' + \alpha \cdot \mathbf{s}$$

pro všechna $\alpha \in \mathbb{R}$, kde \mathbf{x}' je jistý bod, ležící na přímce a $\mathbf{s} \in \mathbb{R}^n$ je směr přímky.

- kvůli jednoznačnosti je někdy vhodné směr normalizovat, takže například při euklidovské normě platí:

$$\mathbf{s}^T \mathbf{s} = \sum_i s_i^2 = 1$$

Matematické základy V

Funkce, kterými se budeme zabývat, budou mít spojitě derivace až do druhého řádu.

Gradient:

- Vektor prvních parciálních derivací funkce f v bodě $\mathbf{x} = (x_1, x_2, \dots, x_n)$
- Označujeme ho $\nabla f(\mathbf{x})$:

$$\nabla f(\mathbf{x}) = \left(\frac{\partial f}{\partial x_1}, \frac{\partial f}{\partial x_2}, \dots, \frac{\partial f}{\partial x_n} \right)$$

Matematické základy VI

Hessova matice (hessián):

- Matice druhých parciálních derivací funkce f v bodě $x = (x_1, x_2, \dots, x_n)$
- Označujeme ho $\nabla^2 f(x)$

$$\nabla^2 f(x) = \begin{pmatrix} \frac{\partial^2 f}{(\partial x_1)^2} & \frac{\partial^2 f}{\partial x_1 \partial x_2} & \cdots & \frac{\partial^2 f}{\partial x_1 \partial x_n} \\ \frac{\partial^2 f}{\partial x_2 \partial x_1} & \frac{\partial^2 f}{(\partial x_2)^2} & \cdots & \frac{\partial^2 f}{\partial x_2 \partial x_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial^2 f}{\partial x_n \partial x_1} & \frac{\partial^2 f}{\partial x_n \partial x_2} & \cdots & \frac{\partial^2 f}{(\partial x_n)^2} \end{pmatrix}$$

- Je zřejmé, že Hessova matice je symetrická

Matematické základy VII

Gradient a hessián jako zobrazení:

- Operátor ∇ je zobrazením, které každé diferencovatelné funkci f přiřazuje jinou funkci ∇f .
- Tedy výrazy $\nabla f(\mathbf{x})$ a $\nabla^2 f(\mathbf{x})$ lze psát také jako $(\nabla f)(\mathbf{x})$ a $(\nabla^2 f)(\mathbf{x})$

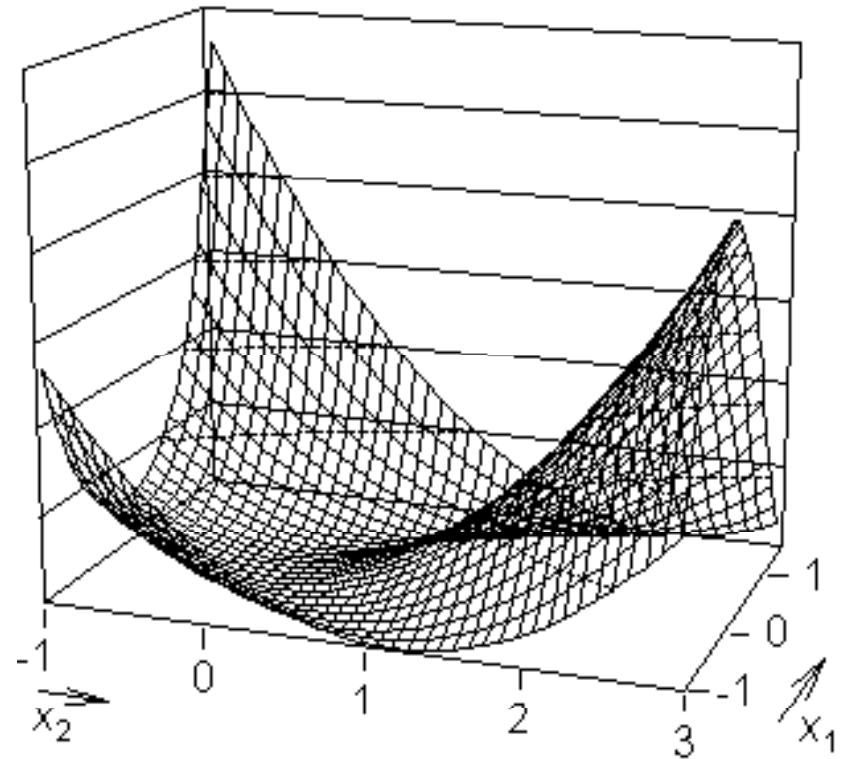
Matematické základy VIII

Gradient a hessián Rosenbrockovy funkce:

– Rosenbrockova funkce:

$$f(\mathbf{x}) = 100(x_2 - x_1^2)^2 + (1 - x_1)^2$$

– Graf má tvar banánovitěho údolí se strmými srázy v jednom směru a s plochým dnem ve druhém směru.

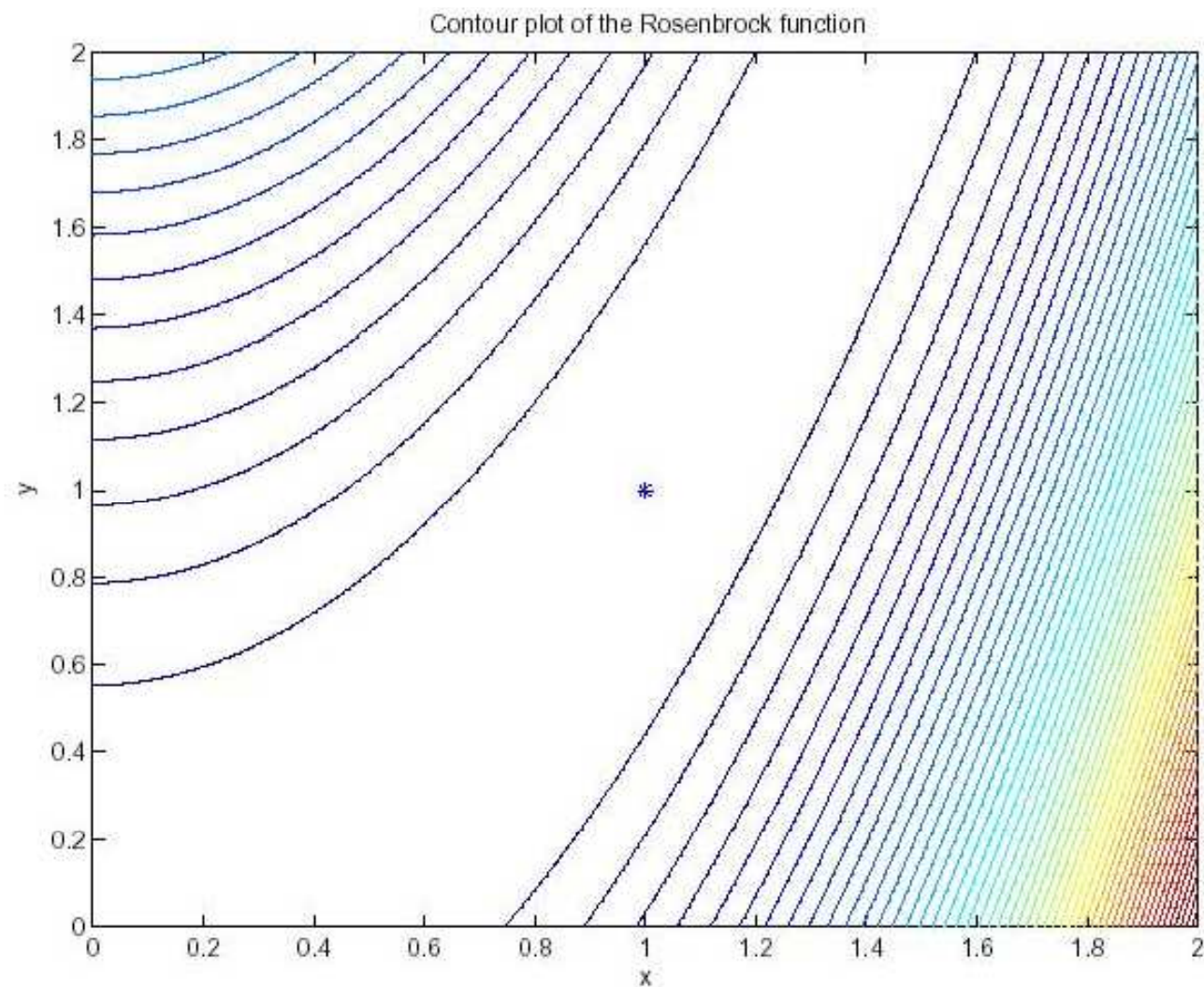


Matematické základy IX

Poznámka:

Globálním minimem Rosenbrockovy funkce je bod $(1, 1)$ a leží na dlouhé pomalu se svažující oblasti, obklopené prudkými srázy.

Je velmi obtížné ji optimalizovat => používá se k testování optimalizačních metod.



Matematické základy X

Gradient a hessián Rosenbrockovy funkce:

Gradient:

$$\nabla f(\mathbf{x}) = \left(-400x_1(x_2 - x_1^2) - 2(1 - x_1), \quad 200(x_2 - x_1^2) \right)^T$$

Hessián:

$$\nabla^2 f(\mathbf{x}) = \begin{pmatrix} -1200x_1^2 - 400x_2 + 2 & -400x_1 \\ -400x_1 & 200 \end{pmatrix}$$

Typy extrémů

Lokální minimum a lokální maximum (často se označuje pouze minimum a maximum):

Minimum:

$$\exists \Omega \forall \mathbf{x} \in \Omega(\mathbf{x}^*): f(\mathbf{x}) \geq f(\mathbf{x}^*)$$

Maximum:

$$\exists \Omega \forall \mathbf{x} \in \Omega(\mathbf{x}^*): f(\mathbf{x}) \leq f(\mathbf{x}^*)$$

Kde $\Omega(\mathbf{x}^*)$ je (vícerozměrné) okolí bodu \mathbf{x}^* .

Typy extrémů III

Globální minimum:

Pokud $f(\mathbf{x}) \geq f(\mathbf{x}^*)$ pro všechna \mathbf{x} z definičního oboru funkce f , pak je \mathbf{x}^* globálním minimem funkce f .

Obdobně je definováno globální maximum funkce f .

Typy extrémů IV

Při práci s minimy a maximy, definovanými na předchozích slidech se může stát, že pro některé funkce dostaneme „paradoxní“ výsledky.

Např. $x = 0$ je lokálním minimem funkce:

$$f(x) = \min(1 + x, 0, 1 - x)$$

a zároveň je také globálním maximem dané funkce.

