Optimalizace

1. Co a k čemu je optimalizace

Tom Werner, Tom Kroupa

FEL ČVUT

Co znamená ‘optimalizace’?

- Latinsko-anglický slovník:
  optimus (adj.) =
  - very good, best
  - excellent
  - most beneficial, most advantageous

- Merriam-Webster dictionary:
  optimization =
  - An act, process, or methodology of making something (such as a design, system, or decision) as fully perfect, functional, or effective as possible.
  - Specifically, the mathematical procedures (such as finding the maximum of a function) involved in this.

- **Mathematical optimization** (or **mathematical programming**) is the selection of a best element, with regard to some criterion, from some set of available alternatives.
  — George Dantzig

Obecná úloha matematické optimalizace

Najdi minimum dané funkce $f \colon X \to \mathbb{R}$ na dané množině X:

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

Funkce f **nabývá minima** na množině X v prvku $x^* \in X$, jestliže $f(x^*) \leq f(x)$ pro každé $x \in X$. Množina všech takových prvků x^* se značí

$$\operatorname*{argmin}_{x \in X} f(x)$$

Názvosloví:

- X je množina **přípustných řešení**.
- f je **účelová (cílová, kriteriální, ...) funkce**.
- x^* se nazývá **argument minima** funkce f na množině X, nebo **optimální řešení** úlohy.
- $f(x^*) = \min_{x \in X} f(x)$ se nazývá **minimální hodnota** f na X, nebo **optimální hodnota** úlohy.

Analogicky pro **maxima**. Maxima a minima se nazývají **extrémy** nebo **optima**.

Úplně jednoduchý příklad

Zadání:

- $X = \{A, 1, \beta, \spadesuit, \bigstar\}$
- Hodnoty f jsou tyto:

x	A	1	β	$\spadesuit$	$\bigstar$
$f(x)$	0	-1	5	2	-1

Řešení:

$$\min_{x \in X} f(x) = -1$$

$$\operatorname*{argmin}_{x \in X} f(x) = \{1, \bigstar\}$$

Má úloha optimální řešení?

- Je-li množina X konečná (tj. má konečný počet prvků), úloha má vždy optimální řešení.
- Funkce $f(x) = x$ na množině $X = \mathbb{R}$ nenabývá minima.
- Funkce $f(x) = 1/x$ na množině $X = [0, \infty)$ nenabývá minima (je na ní neomezená).

Dělení optimalizačních úloh

- **kombinatorická optimalizace**:
  X je konečná (ale obvykle obrovská)
  ($X \subseteq \{0,1\}^n$ nebo X obsahuje permutace, textové řetězce, grafy, konfigurace Rubikovy kostky, ...)

- **spojitá optimalizace**:
  $X \subseteq \mathbb{R}^n$ je nekonečná nespočetná (hlavní náplň kursu)

- **variační počet**:
  X je množina funkcí (na nějakém pevném definičním oboru $D \subseteq \mathbb{R}^n$)

Příklad kombinatorické optimalizace: Problém obchodního cestujícího

Máme n měst.

Mezi každou dvojicí měst i, j je silnice o známé nezáporné délce $d(i, j)$.

Najdi nejkratší trasu, která navštíví každé město právě jednou a vrátí se nazpět do výchozího města.

- X je množina všech permutací n prvků,
  tj. n-tic $(i_1, \ldots, i_n)$ kde $i_1, \ldots, i_n \in \{1, \ldots, n\}$ jsou různé.
- Pro permutaci $(i_1, \ldots, i_n)$ je

$$f(i_1, \ldots, i_n) = \sum_{k=1}^{n-1} d(i_k, i_{k+1}) + d(i_n, i_1)$$

  Příklad pro $n = 4$: $\quad f(2, 4, 3, 1) = d(2, 4) + d(4, 3) + d(3, 1) + d(1, 2)$

- NP-těžká úloha:
  - P≠NP $\implies$ neexistuje algoritmus řešící NP-těžké úlohy v polynomiálním čase (velikosti úlohy, zde n).
  - teorie výpočetní složitosti (computational complexity theory)
  - intuice: “prokletí dimenzionality”, “kombinatorická exploze”, ...

Příklad spojité optimalizace: Bod na křivce nejbližší danému bodu

- $X \subseteq \mathbb{R}^2$ je množina dvojic $(x, y) \in \mathbb{R}^2$ splňujících soustavu

$$
\begin{aligned}
xy &= 1 \\
x &\geq 0
\end{aligned}
$$

- $f(x, y) = \sqrt{(x - x_0)^2 + (y - x_0)^2}$

Význam: Najdi bod na kladné větvi hyperboly s rovnicí $xy = 1$, který je nejblíže bodu (x_0, y_0).

Příklad variačního počtu: Nejkratší křivka spojující dva body

Najdi nejkratší křivku spojující dva dané body $(x_1, y_1), (x_2, y_2) \in \mathbb{R}^2$.

- X je množina všech diferencovatelných funkcí (předpokládáme $x_1 \neq x_2$)

$$\varphi\colon [x_1, x_2] \to \mathbb{R}$$

  splňujících $\varphi(x_1) = y_1$ a $\varphi(x_2) = y_2$.

- Minimalizuj délku křivky

$$f(\varphi) = \int_{x_1}^{x_2} \sqrt{1 + \varphi'(x)^2}\, dx$$

Řešením je afinní funkce, jejímž grafem je úsečka procházející danými dvěma body.

Další klasické úlohy variačního počtu:

- brachistochrona
- řetězovka

Spojitá optimalizace

Standardní tvar úlohy spojité optimalizace

$X \subseteq \mathbb{R}^n$ je množina řešení soustavy rovnic a nerovnic.

$$\begin{aligned}
\min \quad & f(x_1, \ldots, x_n) \\
\text{za podmínek} \quad & g_i(x_1, \ldots, x_n) \le 0, \quad i = 1, \ldots, m \\
& h_i(x_1, \ldots, x_n) = 0, \quad i = 1, \ldots, \ell \\
& x_1, \ldots, x_n \in \mathbb{R}
\end{aligned}$$

Názvosloví:

- $\mathbf{x} = (x_1, \ldots, x_n)$ jsou **proměnné** úlohy
- $g_i(\mathbf{x}) \le 0$ a $h_i(\mathbf{x}) = 0$ jsou **omezující podmínky (omezení)**
- $g_i(\mathbf{x}) \le 0$ jsou **omezení typu nerovnosti** a $h_i(\mathbf{x}) = 0$ jsou **omezení typu rovnosti**

Jiný zápis úlohy:

$$\min\{\, f(\mathbf{x}) \mid \underbrace{\mathbf{x} \in \mathbb{R}^n,\ \mathbf{g}(\mathbf{x}) \le \mathbf{0},\ \mathbf{h}(\mathbf{x}) = \mathbf{0}}_{X} \,\}$$

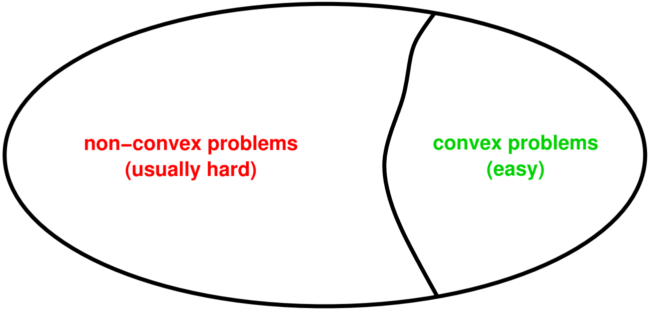
kde $\mathbf{g}\colon \mathbb{R}^n \to \mathbb{R}^m$ a $\mathbf{h}\colon \mathbb{R}^n \to \mathbb{R}^\ell$.

Klíčová role konvexity

Úloha je **konvexní**, jestliže funkce f a g_i jsou konvexní a funkce h_i jsou afinní.

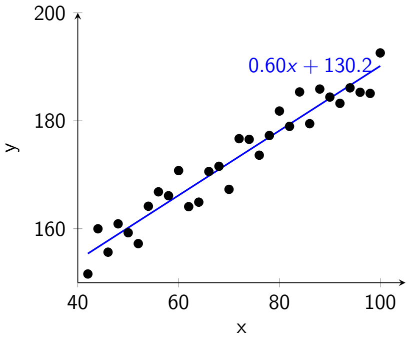
- Pak každé lokální minimum je zároveň globální.
- Tyto úlohy jsou většinou snadné (na rozdíl od nekonvexních úloh).

In fact, the great watershed in optimization isn't between linearity and nonlinearity, but convexity and nonconvexity.
— R. Tyrrell Rockafellar, 1993



Příklad: Prokládáme body přímkou

Odhadujeme funkční závislost výšky y [cm] na váze x [kg] člověka z m měření $(x_i, y_i) \in \mathbb{R}^2$.



- Modelujme vztah **afinní funkcí** $y = \theta_1 + \theta_2 x$.
- Minimalizujeme součet čtverců **svislých** vzdáleností bodů od přímky

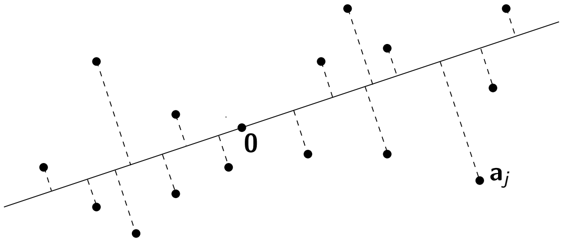
$$f(\boldsymbol{\theta}) = f(\theta_1, \theta_2) = \sum_{i=1}^{m} (\theta_1 + \theta_2 x_i - y_i)^2 = \|\mathbf{A}\boldsymbol{\theta} - \mathbf{y}\|^2$$

  To je **lineární úloha nejmenších čtverců**.
- Podmínka na optimum: **normální rovnice** $\mathbf{A}^T\mathbf{A}\boldsymbol{\theta} = \mathbf{A}^T\mathbf{y}$

Příklad: Prokládáme body přímkou/rovinou apod., ale jinak

Dáno n bodů v $\mathbb{R}^m$, tvořící sloupce matice $\mathbf{A} \in \mathbb{R}^{m \times n}$

Najdi podprostor dané dimenze tak, aby součet čtverců vzdáleností bodů k němu byl minimální.



- Řešení: pomocí spektrálního rozkladu matice $\mathbf{A}\mathbf{A}^T$, tj. pomocí SVD matice $\mathbf{A}$
- Nejde nijak převést na řešení soustavy lineárních rovnic!

Příklad: Minimalizujeme Himmelblauovu funkci

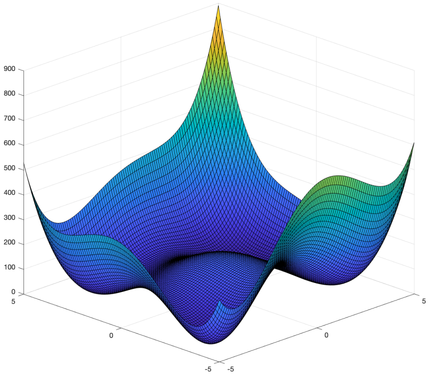
Funkce

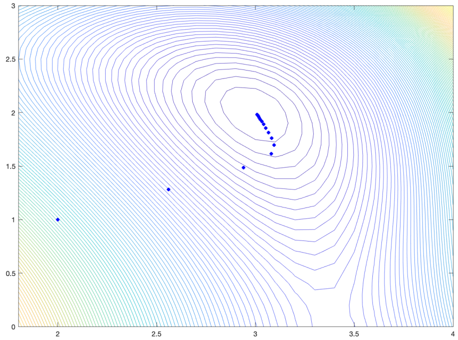
$$f(\mathbf{x}) = f(x_1, x_2) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$$

má 4 lokální minima (např. (3,2)) a 1 lokální maximum.

- **Gradientní metoda** z bodu $\mathbf{x}_0 = (2, 1)$ s krokem $\alpha = 0.01$:

$$\mathbf{x}_{k+1} = \mathbf{x}_k - \alpha \nabla f(\mathbf{x}_k)$$





Příklad: Optimální směs zeleniny

Pro 3 druhy zeleniny jsou dány měrné obsahy živin a minimální požadavky pro jednu přílohu oběda:

	mrkev	*zelí*	*okurka*	**požadavek**
vitamín A [mg/kg]	35	0.5	0.28	0.5 mg
vitamín C [mg/kg]	60	300	80	15 mg
vláknina [g/kg]	30	20	10	4 g
Cena [Kč/kg]	26	22	60	

Najdi hmotnosti zelenin, které splní výživové požadavky při minimální celkové ceně.

- Formulace problému (**lineární program**):

$$
\begin{array}{rrcrcrcr}
\text{min} & 26x_1 & + & 22x_2 & + & 60x_3 & & \\
\text{za podmínek} & 35x_1 & + & 0.5x_2 & + & 0.28x_3 & \geq & 0.5 \\
& 60x_1 & + & 300x_2 & + & 80x_3 & \geq & 15 \\
& 30x_1 & + & 20x_2 & + & 10x_3 & \geq & 4 \\
& & & & & x_1, x_2, x_3 & \geq & 0
\end{array}
$$

- Optimální řešení je $x_1 \doteq 0.12$, $x_2 \doteq 0.03$, $x_3 = 0$ za cenu 3.59 Kč.
- Při požadavku $x_3 \geq 0.1$ (okurka!) je řešení $x_1 \doteq 0.097$, $x_2 \doteq 0.004$, $x_3 = 0.1$ za 8.62 Kč.

Příklad: Těžiště

Jsou dána čísla $a_1, \dots, a_m \in \mathbb{R}$. Najdi minimum (na $\mathbb{R}$) funkce

$$f(x) = \sum_{i=1}^{m} (x - a_i)^2$$

- f je konvexní kvadratická funkce
- Optimální řešení je **aritmetický průměr** $x^* = \frac{1}{m} \sum_{i=1}^{m} a_i$

Jsou dány body $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$. Najdi minimum (na $\mathbb{R}^n$) funkce

$$f(\mathbf{x}) = f(x_1, \dots, x_n) = \sum_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|^2 = \sum_{i=1}^{m} \sum_{j=1}^{n} (x_j - a_{ij})^2 = \sum_{j=1}^{n} \underbrace{\sum_{i=1}^{m} (x_j - a_{ij})^2}_{f_j(x_j)}$$

- Rozpadá se na n nezávislých úloh.
- Optimální řešení je **těžiště** $\mathbf{x}^* = \frac{1}{m} \sum_{i=1}^{m} \mathbf{a}_i$

Příklad: Medián, geometrický medián

Jsou dána čísla $a_1, \dots, a_m \in \mathbb{R}$. Najdi minimum (na $\mathbb{R}$) funkce

$$f(x) = \sum_{i=1}^{m} |x - a_i|$$

- f je konvexní po částech afinní funkce, není diferencovatelná v bodech a_i
- optimální řešení je **medián** čísel $a_1, \dots, a_m$

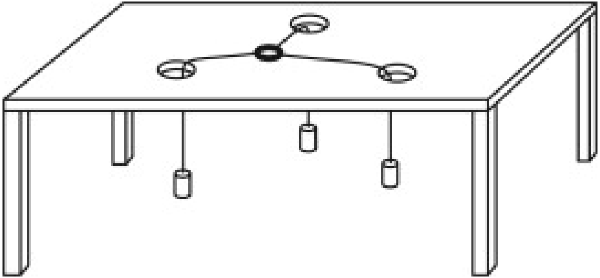
Fermatův-Weberův problém:

Jsou dány body $\mathbf{a}_1, \dots, \mathbf{a}_m \in \mathbb{R}^n$. Najdi minimum (na $\mathbb{R}^n$) funkce

$$f(\mathbf{x}) = f(x_1, \dots, x_n) = \sum_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|$$

- f je konvexní funkce n proměnných, není diferencovatelná v bodech $\mathbf{a}_i$
- optimální řešení se nazývá je **geometrický medián**

Varignon frame: mechanický počítač na hledání geometrického mediánu v rovině (tj. pro $n = 2$)



Závěrečné poznámky

Historie optimalizace

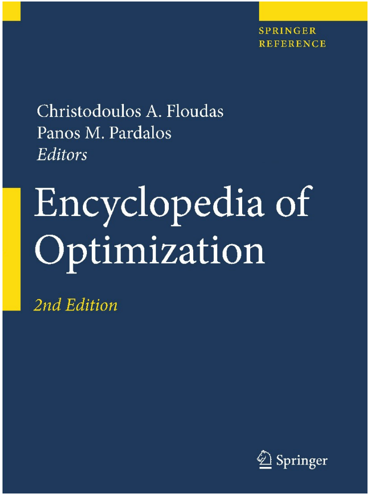
Klasická doba:

- infinitezimální počet, mat. analýza (Newton, Leibniz)
- variační počet (Newton, Leibniz, Euler)
- podmínky na volné lokální extrémy (Fermat)
- podmínky na lokální extrémy vázané rovnostmi (Lagrange)

Moderní optimalizace (po 2. světové válce s nástupem počítačů, důraz na algoritmy):

- lineární programování:
  - teorie, formulace (Kantorovič, Koopmans – Nobelova cena 1975!)
  - simplexová metoda (Dantzig)
  - dualita, teorie her (von Neumann)
- podmínky na lok. extrémy vázané nerovnostmi (Karush-Kuhn-Tucker)
- moderní kombinatorická optimalizace:
  - řezy a toky v grafu (Ford, Fulkerson)
  - celočíselné lin. programování (Gomory, Chvátal), polyhedrální metody
- polynomiální algoritmus na LP (Chačian), algoritmy vnitřního bodu (Karmarkar)
- semidefinitní programování (SDP)

Viz např. `http://www.mitrikitti.fi/opthist.html`



Více než 4000 stránek!

Mnoho aplikací

- **ekonomie a finance:** minimální riziko, maximální zisk, nastavení cen, ...
- **logistika:** doprava, průmysl, zásobování, ...
- **řízení (control engineering):** výtahu, robota, vlaku, letadla, aktivní budovy, ...
- **rozvrhování a plánování (scheduling):** školní rozvrh, výrobní kroky, cesta mobilního robota, sled úkonů robotického manipulátoru, aircrew scheduling
- **floor planning:** návrh integrovaných obvodů (VLSI design) a plošných spojů
- **routing:** IDOS, navigace v autě, návrh počítačové sítě, ...
- **pravděpodobnost a statistika:** princip maximální věrohodnosti, princip maxima entropie, regrese (modelování funkční závislosti náhodných proměnných), rozhodování za neurčitosti
- **počítačové vidění:** rekonstrukce scény z obrazů (multiview geometry), segmentace obrazu pomocí řezů v grafu (graph cuts), hledání tváří v obraze (AdaBoost), ...
- **porozumění/zpracování signálu** (audio, EKG, EEG, ...): separace zdrojů, auditory scene analysis, ...
- **rozpoznávání a strojové učení:** minimální trénovací chyba, nejjednodušší model
- **návrh mechanických struktur:** most, jeřáb, hák, křídlo letadla
- **molekulární modelování:** např. protein folding
- **teorie her**
- přiřazování radiových frekvencí v mobilní síti
- ...

Příroda optimalizuje

Nothing takes place in the world whose meaning is not that of some maximum or minimum.
— Leonhard Euler

Fyzikální zákony je (skoro?) vždy možno formulovat ve variačním tvaru:

- systém v rovnováze je v lokálním minimu potenciální energie
- Fermatův princip nejkratšího (přesněji: lokálně extrémního) času v optice
- princip nejmenšího účinku v klasické mechanice
- ...

O čem bude B0B33OPT

1. Použití lineární algebry v optimalizaci
   - Lineární úloha nejmenších čtverců
   - PCA
   - Maticové rozklady (QR, spektrální, Cholesky, SVD)
2. Analýza a numerické iterační metody
   - Derivace vektorových a maticových výrazů
   - Podmínky optimality na volné lokální extrémy
   - Iterační numerické metody na hledání lok. minim (gradientní, Newtonova, Gaussova-Newtonova)
   - Lok. extrémy vázané rovnostmi, Lagrangeovy multiplikátory
3. Lineární programování
   - Formulace úloh LP
   - Něco o algoritmech na LP
   - Dualita v LP
4. Konvexní optimalizace
   - Konvexní množiny a funkce
   - Třídy konvexních opt. úloh