Optimalizace

1. Optimalizační úlohy

Tomáš Kroupa   Tomáš Werner

2022 LS

Fakulta elektrotechnická
ČVUT v Praze

O čem je optimalizace?

Optimization

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

Optimalizační úloha je určena:

- množinou prvků, ze kterých vybíráme
- kritériem, které prvky ohodnocuje reálným číslem
- požadavkem na minimalizaci/maximalizaci toho kritéria

Aplikace

Nepřeberné množství úloh ve strojovém učení, umělé inteligenci, rozpoznávání, informatice, statistice, fyzice a ekonomii:

- Hledáme optimální trasu robota
- Učíme neuronovou síť
- Konstruujeme nejlevnější transportní síť
- Minimalizujeme náklady na výrobu produktu
- Predikujeme budoucí vývoj náhodné veličiny

Z pohledu studenta

Cíle

1. Naučit se matematicky formulovat úlohy vyžadující optimalizaci jistého kritéria při zadaných omezeních
2. Porovnat varianty zadaného problému a jejich obtížnost
3. Navrhnout vhodnou metodu řešení

Které oblasti matematiky využijeme?

- lineární algebra
- matematická analýza

Optimalizační úloha minimalizace

$$\text{Minimalizuj } f(x) \text{ za podmínky } x \in X$$

- Množina přípustných řešení $X \subseteq X'$
- Účelová funkce $f\colon X' \to \mathbb{R}$
- Minimum funkce f na množině X je prvek $x^* \in X$, pro který platí $f(x^*) \leq f(x)\ \forall x \in X$
- Množinu všech minim funkce f na X značíme jako

$$\arg\min_{x \in X} f(x) = \{x^* \in X \mid f(x^*) \leq f(x)\ \forall x \in X\}$$

Optimalizační úloha maximalizace

$$\text{Maximalizuj } f(x) \text{ za podmínky } x \in X$$

- Maximum funkce f na množině X je prvek $x^* \in X$, pro který platí $f(x^*) \geq f(x)\ \forall x \in X$
- Množinu všech maxim funkce f na X značíme jako

$$\arg\max_{x \in X} f(x) = \{x^* \in X \mid f(x^*) \geq f(x)\ \forall x \in X\}$$

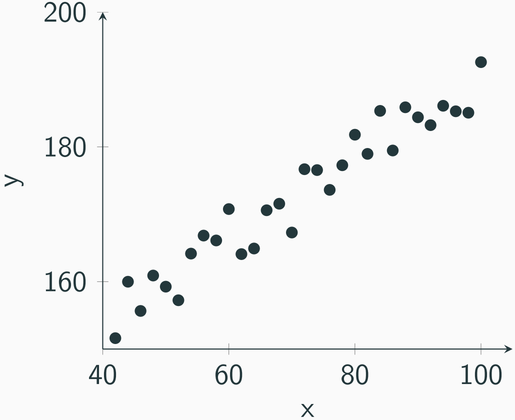
Převod Min na Max

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

Příklady optimalizačních úloh

Prokládáme body přímkou

Modelujeme vztah váhy x [kg] a výšky y [cm] na základě dat.



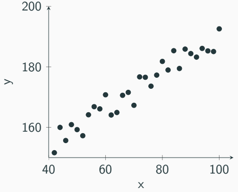
Cíl

Hledáme přímku, která co nejtěsněji proloží černé body.

Prokládáme body přímkou – formulace modelu

Máme m měření $(x_i, y_i) \in \mathbb{R}^2$ váhy a výšky.

Vztah vyjádříme lineární funkcí



$$f(x, \theta_1, \theta_2) = \theta_1 + \theta_2 x,$$

kde θ_1 a θ_2 jsou neznámé parametry.

Soustava lineárních rovnic $y_i = \theta_1 + \theta_2 x_i$, $i = 1, \ldots, m$, s neznámými θ_1 a θ_2 je vlivem náhody přeurčená.

Úloha nejmenších čtverců

$$\text{Minimalizuj} \quad \sum_{i=1}^{m} (y_i - f(x_i, \theta_1, \theta_2))^2 \text{ za podmínky } \theta_1, \theta_2 \in \mathbb{R}$$

Prokládáme body přímkou – formulace modelu maticově

$$\mathbf{y} = \begin{bmatrix} y_1 \\ \vdots \\ y_m \end{bmatrix} \in \mathbb{R}^{m\times 1} \quad \mathbf{A} = \begin{bmatrix} 1 & x_1 \\ \vdots & \vdots \\ 1 & x_m \end{bmatrix} \in \mathbb{R}^{m\times 2} \quad \boldsymbol{\theta} = \begin{bmatrix} \theta_1 \\ \theta_2 \end{bmatrix} \in \mathbb{R}^{2\times 1}$$

Úloha nejmenších čtverců

$$\text{Minimalizuj} \quad \|\mathbf{y} - \mathbf{A}\boldsymbol{\theta}\|^2 \ \text{ za podmínky } \boldsymbol{\theta} \in \mathbb{R}^2$$

Prokládáme body přímkou – řešení

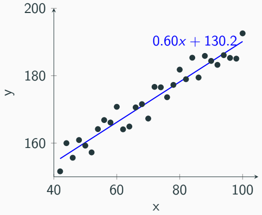
Pomocí lineární algebry lze úlohu reformulovat jako hledání řešení soustavy lineárních rovnic

$$\mathbf{A}^\top \mathbf{A}\boldsymbol{\theta} = \mathbf{A}^\top \mathbf{y}.$$

Ta má v našem případě jediné řešení $\boldsymbol{\theta}^* = (\mathbf{A}^\top \mathbf{A})^{-1}\mathbf{A}^\top \mathbf{y}$.

Optimální řešení úlohy nejmenších čtverců:

$$\theta_1^* = 130.2, \quad \theta_2^* = 0.6.$$



Hledáme optimální směs zeleniny

Pro 3 druhy syrové zeleniny udává tabulka výživové hodnoty, ceny a nejmenší předepsaný obsah živin v jedné příloze jídla.

	Mrkev	*Bílé 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	

Cíl

Nalézt množství každého druhu zeleniny [kg], které minimalizuje cenu přílohy jídla při splnění předepsaných výživových limitů.

Hledáme optimální směs zeleniny – formulace modelu

	Mrkev	*Bílé 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	

Úloha lineárního programování

$$
\begin{aligned}
\min \quad & 26x_1 + 22x_2 + 60x_3 & & \\
\text{za podmínek} \quad & x_i \geq 0, & & i = 1, 2, 3 \\
& 35x_1 + 0.5x_2 + 0.28x_3 \geq & & 0.5 \\
& 60x_1 + 300x_2 + 80x_3 \geq & & 15 \\
& 30x_1 + 20x_2 + 10x_3 \geq & & 4
\end{aligned}
$$

Hledáme optimální směs zeleniny – řešení

Úloha lineárního programování

$$\begin{aligned} \min \quad & 26x_1 + 22x_2 + 60x_3 \\ \text{za podmínek} \quad & x_i \geq 0, & i = 1, 2, 3 \\ & 35x_1 + 0.5x_2 + 0.28x_3 \geq & 0.5 \\ & 60x_1 + 300x_2 + 80x_3 \geq & 15 \\ & 30x_1 + 20x_2 + 10x_3 \geq & 4 \end{aligned}$$

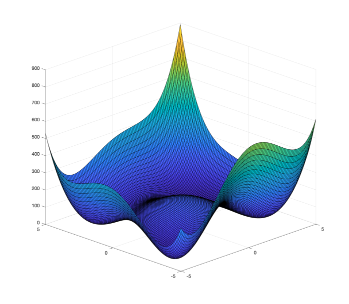
- Optimální řešení je $(0.115, 0.027, 0)$ za cenu 3.592
- Při požadavku na okurku $x_3 \geq 0.1$ dostaneme řešení $(0.097, 0.004, 0.1)$ za cenu 8.618

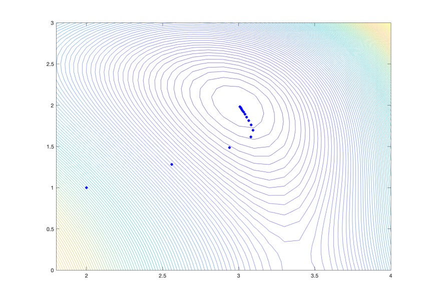
Minimalizujeme Himmelblauovu funkci

Funkce $f(\mathbf{x}) = (x_1^2 + x_2 - 11)^2 + (x_1 + x_2^2 - 7)^2$ má 4 lokální minima (např. (3,2)) a 1 lokální maximum.

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

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





Nejkratší křivka

Cíl

Nalezněte nejkratší křivku spojující 2 body v rovině.

Intuice napovídá, že řešením je úsečka.

Nejkratší křivka – formulace a řešení úlohy

- Uvažujme $\mathbf{x}, \mathbf{y} \in \mathbb{R}^2$ a předpokládejme, že $x_1 \neq y_1$.
- Křivka spojující $\mathbf{x}$ a $\mathbf{y}$ je grafem spojitě diferencovatelné funkce $f\colon [x_1, y_1] \to \mathbb{R}$, kde $f(x_1) = x_2$ a $f(y_1) = y_2$.
- Délka křivky je $D(f) = \int_{x_1}^{y_1} \sqrt{1 + f'(t)^2}\, dt$

Úloha variačního počtu

$$\min \quad D(f)$$

kde f je spojitě diferencovatelná funkce vyhovující omezením výše

Úlohu řeší afinní funkce procházející body $\mathbf{x}$ a $\mathbf{y}$.

Kategorie optimalizačních úloh

Typ množiny přípustných řešení X definuje tyto kategorie:

- spojitá optimalizace – X je nespočetná množina vektorů v $\mathbb{R}^n$ vyjádřená jako množina řešení rovnic a nerovnic
- diskrétní optimalizace – X je konečná/spočetná
- variační počet – X obsahuje reálné funkce

V tomto kurzu se budeme zabývat spojitou optimalizací.

Obecně o úloze spojité optimalizace

Spojitá optimalizace

Úloha v obecném tvaru

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

Úsporněji píšeme:

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

Základní otázky

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

Je úloha přípustná?

Je množina X neprázdná?

Existuje globální minimum?

- Nabývá funkce f na X minima, neboli $\operatorname*{arg\,min}_{\mathbf{x} \in X} f(\mathbf{x}) \neq \emptyset$?
- Jak velká je množina $\operatorname*{arg\,min}_{\mathbf{x} \in X} f(\mathbf{x})$?

V aplikacích se často spokojíme s lokálním minimem.

Formy řešení

Analytický tvar

Globální minimum úlohy nejmenších čtverců pro lineární regresi je vektor parametrů $\boldsymbol{\theta}^* = (\mathbf{A}^\mathsf{T}\mathbf{A})^{-1}\mathbf{A}^\mathsf{T}\mathbf{y} \in \mathbb{R}^2$.

Algoritmus

Optimální směs zeleniny $\mathbf{x}^* \in \mathbb{R}^3$ minimalizující cenu v úloze lineárního programování spočítá simplexová metoda.

Iterační metoda

Lokální minimum $\mathbf{x}^*$ funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ aproximuje posloupnost $\mathbf{x}_1, \mathbf{x}_2, \ldots$ generovaná gradientní metodou.

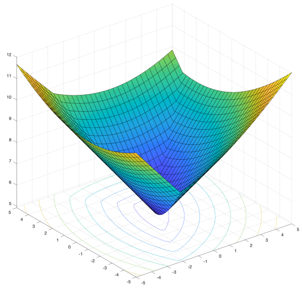
Úloha na optimální umístění

Hledáme lokaci pro heliport, z něhož dolétne helikoptéra po úsečce do nejvzdálenějšího z m bodů $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^2$ v nejkratším čase.

Optimalizační úloha

Minimalizuj $f(\mathbf{x}) = \max\limits_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|$ za podmínky $\mathbf{x} \in \mathbb{R}^2$

$$\mathbf{a}_1 = (0,0),\ \mathbf{a}_2 = (5,-1),\ \mathbf{a}_3 = (1,-4),\ \mathbf{a}_4 = (-4,3)$$



Úloha na nejmenší kruh

Nejděte nejmenší kruh obsahující m daných bodů $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^2$.

Optimalizační úloha

Min r za podmínek $\|\mathbf{x} - \mathbf{a}_i\| \leq r,\ \mathbf{x} \in \mathbb{R}^2,\ r \in \mathbb{R},\ i = 1, \ldots, m$

- Tato úloha je ekvivalentní úloze na optimální umístění
- Lze ji ekvivalentně formulovat jako úlohu, v níž jsou všechny funkce diferencovatelné:

Min y za podmínek $\|\mathbf{x} - \mathbf{a}_i\|^2 \leq y,\ \mathbf{x} \in \mathbb{R}^2,\ y \in \mathbb{R},\ i = 1, \ldots, m$

Časté prohřešky při formulaci a řešení optimalizačních úloh

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

- Není zřejmé, co je účelová funkce f
- Omezující podmínky nejsou jasně specifikovány
- Účelová funkce je zaměněna s funkcemi definujícími omezení
- Proměnné a zadané parametry úlohy nejsou rozlišeny
- Záměna nutných a postačujících podmínek optimality
- Řešení neexistuje, přesto je nalezeno
- Funkce f nemá derivaci, přesto je spočtena
- Lokální optimum je zaměněno za globální
- Maticové výrazy jsou nesmyslné (viz maticové zločiny)

Optimalizace B0B33OPT

O čem to bude

1. Aplikace lineární algebry (8 přednášek)
    - Metoda nejmenších čtverců, lineární regrese
    - PCA, ortogonální Prokrustův problém
    - Maticové rozklady: QR, spektrální, Choleského, SVD
2. Analýza a numerické metody (6 přednášek)
    - Podmínky optimality pro volné lokální extrémy
    - Iterační metody: gradientní, Newtonova, Gauss-Newtonova, Levenberg-Marquardtova
    - Omezení ve tvaru rovností, Lagrangeovy multiplikátory
3. Lineární programování (6 přednášek)
    - Konvexní polyedry
    - Simplexová metoda
    - Dualita
4. Úvod do konvexní optimalizace (3 přednášky)
    - Konvexní množiny a funkce
    - Třídy konvexních úloh

O čem to nebude a přitom by mohlo

- Numerická lineární algebra
- Výpočetní složitost optimalizačních algoritmů
- Numerické metody pro úlohy s omezeními
- Dualita pro konvexní úlohy
- KKT podmínky
- Variační počet
- Celočíselné programování
- Semidefinitní programování
- Polynomiální optimalizace

V jakých kurzech optimalizaci využijete

- *Kombinatorická optimalizace*
- *Rozpoznávání a strojové učení*
- *Robotika*
- *Julia for Optimization and Learning*
- *Umělá inteligence v robotice*
- *Optimalizace a teorie her*
- *Výpočetní teorie her*
- *Statistical Machine Learning*
- *Optimální a robustní řízení*

Výpočetní nástroje

Programovací jazyky a prostředí

- Python + NumPy + SciPy
- Matlab
- Rychle se prosazuje Julia + JuMP

Solvery (řešiče)

- SeDuMi, GLPK, Gurobi