Optimalizace

1. Optimalizační úlohy

Tomáš Kroupa   Tomáš Werner

2023 LS

Fakulta elektrotechnická
ČVUT v Praze

O čem je optimalizace?

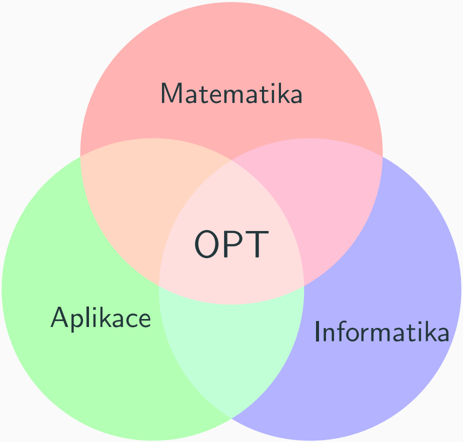
Mathematical optimization

The selection of a best element with regard to some criterion from some set of available alternatives. (*Wikipedia*)

Optimalizační úloha je zadána

1. množinou prvků, ze kterých vybíráme,
2. kritériem, které prvky ohodnocuje reálným číslem, a
3. požadavkem na minimalizaci/maximalizaci toho kritéria.

Optimalizace v kontextu



Motivace pro studium optimalizace na FEL

Základní otázky

PROČ?	AI, ML, robotika, řízení, statistika, teorie her
JAK?	Programování
CO?	Optimalizační úlohy jsou formulovány matematicky

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

O čem to bude

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

Výpočetní nástroje

Programovací jazyky a balíky

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

Solvery (řešiče)

- SeDuMi, GLPK, Gurobi, MOSEK

Úvodní definice a příklady

Úloha minimalizace

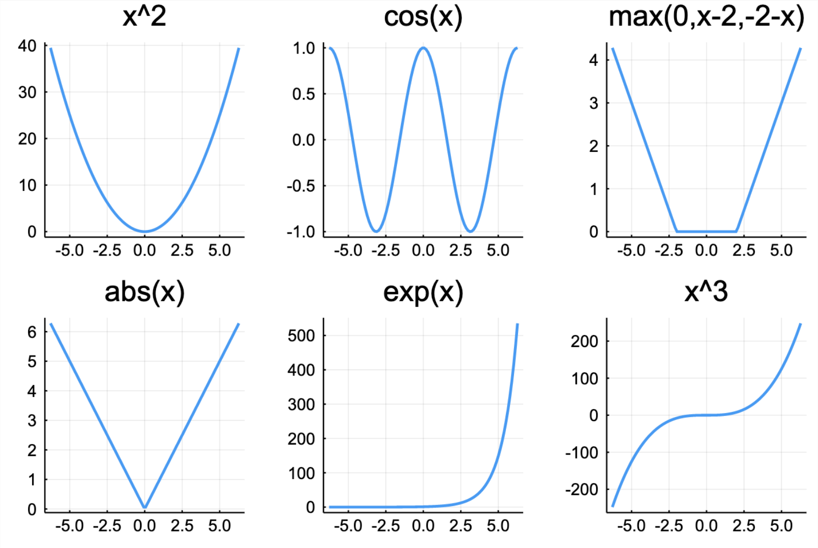
Minimalizuj $f(x)$ za podmínky $x \in X$

- Účelová funkce $f\colon Y \to \mathbb{R}$
- Množina přípustných řešení $X \subseteq Y$
- Minimum funkce f na množině X je prvek $x^* \in X$ splňující

$$f(x^*) \leq f(x) \quad \forall x \in X$$

a množinu všech minim funkce f na X značíme $\operatorname*{arg\,min}_{x \in X} f(x)$

Funkce a jejich minima



Úloha maximalizace

Maximalizuj $f(x)$ za podmínky $x \in X$

- Maximum funkce f na množině X je prvek $x^* \in X$ splňující

$$f(x^*) \geq f(x) \quad \forall x \in X$$

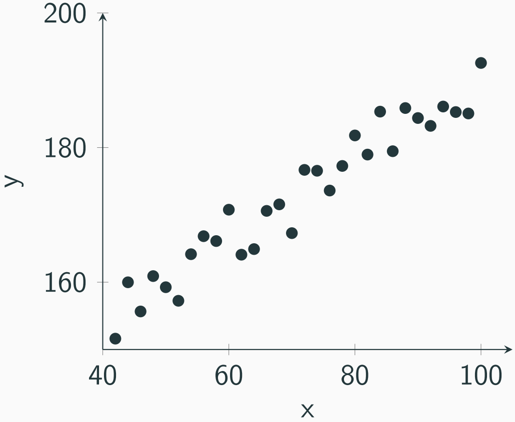
a množinu všech maxim funkce f na X značíme $\arg\max_{x \in X} f(x)$

Převod maximalizace na minimalizaci

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

Prokládáme body přímkou

Modelujeme vztah váhy x a výšky y na základě dat.



Cíl

Hledáme přímku, která co nejtěsněji proloží černé body.

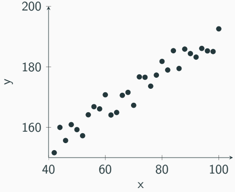
Prokládáme body přímkou – formulace modelu

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

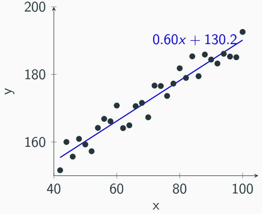
Pomocí lineární algebry lze úlohu reformulovat jako hledání řešení soustavy lineárních rovnic

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

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, 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

Hledáme optimální směs zeleniny – jak nalezneme řešení?

Vektor $(0.115, 0.027, 0)$ je řešením soustavy lineárních rovnic:

Procházíme přípustná řešení lineárního programování

$$
\begin{aligned}
x_1 &\geq 0 \\
x_2 &\geq 0 \\
x_3 &= 0 \\
35x_1 + 0.5x_2 + 0.28x_3 &\geq 0.5 \\
60x_1 + 300x_2 + 80x_3 &= 15 \\
30x_1 + 20x_2 + 10x_3 &= 4
\end{aligned}
$$

Simplexová metoda je základním algoritmem pro řešení úloh LP.

Hledáme optimální směs zeleniny – jak na to v JULIi?

[2]:
```
using JuMP, GLPK
```

[3]:
```
A = [35 .5 .28;
     60 300 80;
     30 20 10]
b = [.5, 15, 4]
c = [26, 22, 60];
```

[6]:
```
LP = Model(GLPK.Optimizer)
@variable(LP, x[1:3] >= 0)
@constraint(LP, con, A*x .>= b)
@objective(LP, Min, c'*x);
```

[7]:
```
optimize!(LP)
solution_summary(LP)
```

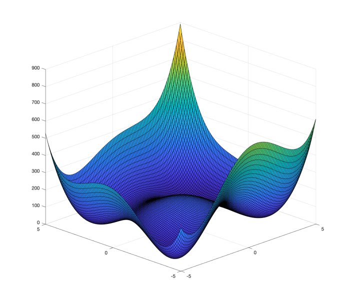
[7]:
```
* Solver : GLPK

* Status
  Termination status : OPTIMAL
  Primal status      : FEASIBLE_POINT
  Dual status        : FEASIBLE_POINT
  Message from the solver:
  "Solution is optimal"

* Candidate solution
  Objective value      : 3.59231e+00
  Objective bound      : -Inf
  Dual objective value : 3.59231e+00

* Work counters
  Solve time (sec)    : 5.19753e-05
```

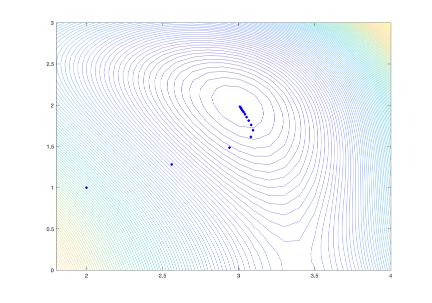
Minimalizujeme Himmelblauovu funkci

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 $\alpha = 0.01$:

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

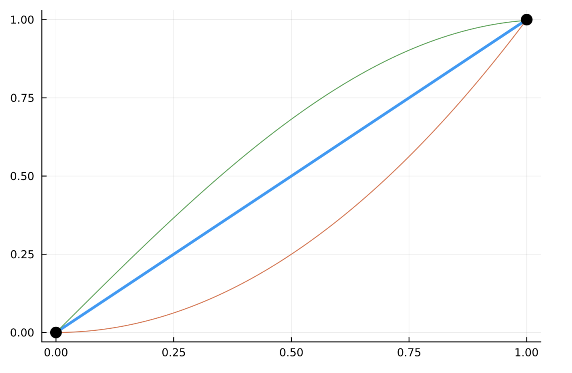




Nejkratší křivka

Cíl

Nalezněte nejkratší křivku spojující 2 body v rovině.



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

Podle typu množiny přípustných řešení X mluvíme o

- spojité optimalizaci, kde $X \subseteq \mathbb{R}^n$ je nespočetná množina,
- diskrétní optimalizaci, kde X je konečná/spočetná,
- variačním počtu, kde X obsahuje reálné funkce.

V tomto kurzu se budeme zabývat spojitou optimalizací.

Obecně o úloze spojité optimalizace

Úloha spojité optimalizace

Obecný tvar

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

Vektorový zápis:

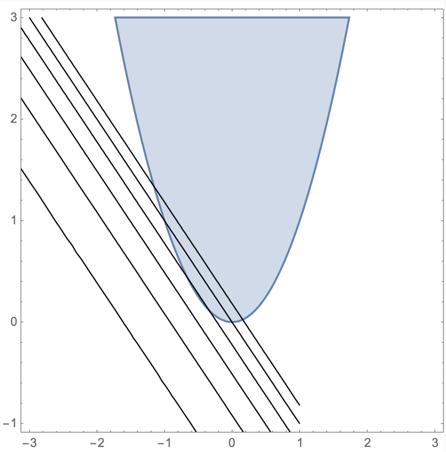
$$\min \{ f(\mathbf{x}) \mid \mathbf{g}(\mathbf{x}) \le \mathbf{0},\ \mathbf{h}(\mathbf{x}) = \mathbf{0},\ \mathbf{x} \in \mathbb{R}^n \}$$

Příklad

Řešíme úlohu

$$\min_{\mathbf{x} \in \mathbb{R}^2} e^{x_1 + x_2}$$

za podmínky $x_1^2 \leq x_2$.



- Na obrázku jsou vrstevnice účelové funkce a množina přípustných řešení
- Globální minimum $(-\frac{1}{2}, \frac{1}{4})$ lze snadno nalézt úvahou

Základní otázky

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

Je úloha přípustná?

Je množina X neprázdná?

Existuje globální/lokální minimum?

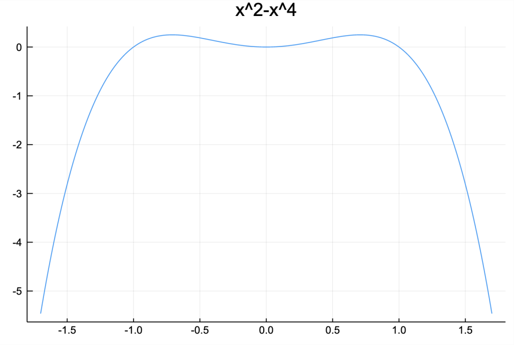
- Nabývá funkce f na X minima, neboli $\operatorname*{arg\,min}_{\mathbf{x} \in X} f(\mathbf{x}) \neq \emptyset$?
- Jak velká je množina $\operatorname*{arg\,min}_{\mathbf{x} \in X} f(\mathbf{x})$?
- Pokud $\operatorname*{arg\,min}_{\mathbf{x} \in X} f(\mathbf{x}) = \emptyset$, spokojíme se s lokálním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

Příklad

Řešíme úlohu

$$\min_{x \in \mathbb{R}} (x^2 - x^4)$$

bez omezujících podmínek.



- Globální minimum neexistuje: $\lim_{x \to \pm\infty} x^4(\frac{1}{x^2} - 1) = -\infty$
- Nutná podmínka $f'(x) = 2x - 4x^3 = 2x(1 - 2x^2) = 0$
- Lokální extrém může nastat pouze v bodech $0, \frac{1}{\sqrt{2}}, -\frac{1}{\sqrt{2}}$
- Jen bod 0 je lokální minimum, protože $f''(0) = 2$

Různé formy řešení optimalizačních úloh

Analytický tvar

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

Algoritmus

Globální minimum v úloze lineárního programování je směs zeleniny $\mathbf{x}^* \in \mathbb{R}^3$ na výstupu simplexové metody.

Iterační metoda

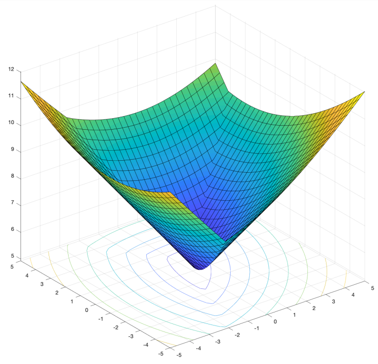
Lokální minimum $\mathbf{x}^*$ funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ aproximuje posloupnost $\mathbf{x}_1, \mathbf{x}_2, \ldots$ generovaná gradientní metodou.

Úloha na optimální umístění

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

Definice úlohy

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



Účelová funkce je konvexní, ale nehladká. Minimum existuje.

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

Ekvivalentní úlohy

- Formulujeme úlohy, které jsou ekvivalentní té předchozí
- Najděte nejmenší kruh obsahující body $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^2$:

Úloha s nehladkými omezeními

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

Kvadratické programování s kvadrat. omezeními (QCQP)

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

Kvadratické programování (QP)

Min $\rho + \|\mathbf{x}\|^2$

z.p. $\rho + 2\mathbf{a}_i^\mathsf{T}\mathbf{x} \geq \|\mathbf{a}_i\|^2,\ i = 1, \ldots, m,$ $\mathbf{x} \in \mathbb{R}^2,\ \rho \in \mathbb{R}$