Optimalizace

1. Optimalizační úlohy

Tomáš Kroupa

FEL 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*)

- Je zadána účelová funkce $f\colon X' \to \mathbb{R}$
- Matematická optimalizace spočívá v hledání minima funkce f na množině přípustných řešení $X \subseteq X'$
- Optimální řešení je prvek $\mathbf{x}^* \in X$ splňující

$$f(\mathbf{x}^*) = \min_{\mathbf{x} \in X} f(\mathbf{x})$$

a číslo $f(\mathbf{x}^*)$ je optimální hodnota úlohy

Aplikace

Nepřeberné množství úloh ve strojovém učení, rozpoznávání, informatice, statistice, fyzice, ekonomii. Namátkou:

- 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 minimalizaci/maximalizaci jistého kritéria při zadaných omezeních
2. Porovnat různé varianty zadaného problému a posoudit jejich obtížnost
3. Navrhnout vhodnou metodu řešení

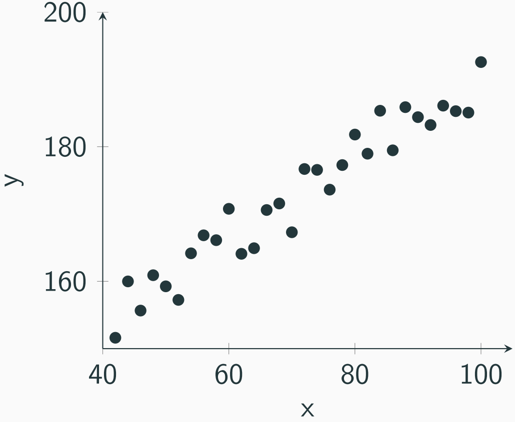
Matematické nástroje

- lineární algebra
- vícedimenzionální kalkulus

Příklady optimalizačních úloh

Prokládáme body přímkou

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

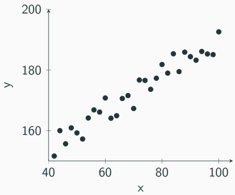


Cíl

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.

Vlivem náhodných faktorů je soustava lin. rovnic $y_i = \theta_1 + \theta_2 x_i$, $i = 1, \ldots, m$, s neznámými θ_1 a θ_2 přeurčená.

Úloha nejmenších čtverců

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

Prokládáme body přímkou – řešení

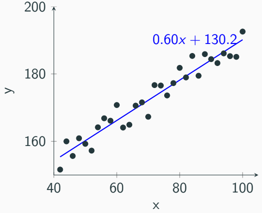
Pomocí lineární algebry lze úlohu reformulovat jako hledání řešení soustavy lineárních rovnic

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

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

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 zajistí minimální 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

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

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

Optimální komprese

Prvky množiny $\Omega = \{a, b, c, d, e\}$ se vyskytují s těmito pravděpodobnostmi:

$$(0.25, 0.25, 0.20, 0.15, 0.15)$$

Množinu Ω lze jednoznačně zakódovat pomocí 3bitových slov. Lze to udělat v průměru úsporněji?

Cíl

Hledáme binární kódování minimalizující střední délku kódu.

Optimální komprese – formulace úlohy

- Binární kód je zobrazení $C\colon \Omega \to \{0,1\}^*$
- Požadujeme, aby kód C byl jednoznačně dékodovatelný
- Střední délka kódu C je číslo $\mathbb{E}(C) = \sum_{x\in\Omega} p(x)\cdot \ell(C(x))$

Úloha diskrétní optimalizace

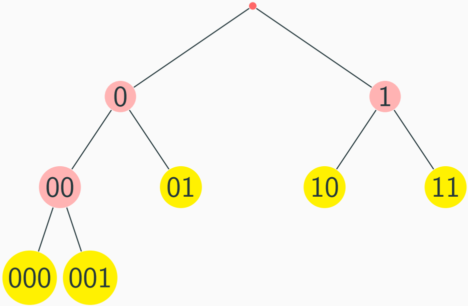
Pro zadanou konečnou množinu Ω a pravděpodobnosti $p(x)$ symbolů $x \in \Omega$, řeš úlohu

$$\min \ \mathbb{E}(C),$$

kde C je libovolný jednoznačně dekódovatelný binární kód.

Optimální komprese – řešení

a	0.25	(d,e)	0.30	(b,c)	0.45	(a,d,e)	0.55	01
b	0.25	a	0.25	(d,e)	0.30	(b,c)	0.45	10
c	0.20	b	0.25	a	0.25			11
d	0.15	c	0.20					000
e	0.15							001



Hufmannův kód má minimální střední délku a ta je rovna 2.3.

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.

Spojitá optimalizace

Úloha spojité optimalizace v obecném tvaru

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

Úsporněji lze tuto úlohu zapsat pomocí vektorové notace:

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

Účelová funkce f i omezující podmínky mají často speciální tvar.

Konvexní a nekonvexní úlohy

Konvexní množiny

Definice

Množina $X \subseteq \mathbb{R}^n$ je konvexní, pokud pro všechna $\mathbf{x}, \mathbf{y} \in X$ a libovolné $\alpha \in [0, 1]$ platí $(1 - \alpha)\mathbf{x} + \alpha\mathbf{y} \in X$.

Příklady

- úsečka, přímka, lineární podprostor
- kruh
- mnohostěn (polyedr)

Konvexní funkce

Definice

Funkce $f\colon \mathbb{R}^n \to \mathbb{R}$ je konvexní na konvexní množině $X \subseteq \mathbb{R}^n$, jestliže pro všechna $\mathbf{x}, \mathbf{y} \in X$ a každé $\alpha \in [0, 1]$ platí

$$f((1-\alpha)\mathbf{x} + \alpha\mathbf{y}) \leq (1-\alpha)f(\mathbf{x}) + \alpha f(\mathbf{y})$$

Příklady

- exponenciální funkce, absolutní hodnota, afinní funkce
- $f(x) = x^3$ na množině $[0, \infty)$
- $f(\mathbf{x}) = \max\{x_1, \dots, x_n\}$

Konvexní optimalizace

Úloha konvexní optimalizace

Minimalizuj konvexní funkci f na konvexní množině X

Příklady (viz úvod)

- úloha nejmenších čtverců
- úloha lineárního programování

Teoreticky i výpočetně představují konvexní úlohy velmi dobře zmapovanou třídu optimalizačních problémů.

Nekonvexní úloha – shlukování

- Pro zadaných m bodů $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^d$ hledáme n shluků $\mathcal{C}_j$ popsaných prototypem $\mathbf{x}_j \in \mathbb{R}^d$,

$$\mathcal{C}_j = \{\mathbf{a}_i \mid \|\mathbf{a}_i - \mathbf{x}_j\| \leq \|\mathbf{a}_i - \mathbf{x}_k\| \;\; \forall i, k\},$$

  tak, aby součet vzdáleností k prototypům byl minimální

- Minimalizujeme tak funkci

$$f(\mathbf{x}_1, \ldots, \mathbf{x}_n) = \sum_{i=1}^{m} \min_{j=1}^{n} \|\mathbf{a}_i - \mathbf{x}_j\|$$

  na množině vektorů $(\mathbf{x}_1, \ldots, \mathbf{x}_n) \in \mathbb{R}^{nd}$

Složitost

Jde o NP-těžkou úlohu, minimum je prakticky nemožné nalézt.

Nekonvexní úloha – varianta `Set-Partitioning`

Lze danou n-tici $(\alpha_1, \ldots, \alpha_n) \in \mathbb{N}^n$ rozdělit na 2 části se stejným součtem? Přesněji: existuje $\mathbf{x} \in \{-1, 1\}^n$ splňující $\sum_{i=1}^n \alpha_i x_i = 0$?

Ekvivalentní optimalizační úloha

$$\max \sum_{i=1}^{n} x_i^2 \quad \text{za podmínek} \quad \sum_{i=1}^{n} \alpha_i x_i = 0, \ \mathbf{x} \in [-1, 1]^n$$

Složitost

Jde o NP-těžkou optimalizační úlohu.

Obecně o úloze spojité optimalizace

Základní otázky

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

- Je úloha přípustná, tedy $X \neq \emptyset$?
- Nabývá funkce f na X minima, tedy $\operatorname*{argmin}_{\mathbf{x} \in X} f(\mathbf{x}) \neq \emptyset$?
- Nutné podmínky: jaké vlastnosti splňuje optimum?
- Postačující podmínky: jaké vlastnosti garantují optimalitu?
- Jak velká je množina $\operatorname*{argmin}_{\mathbf{x} \in X} f(\mathbf{x})$?
- Řešení úlohy je obtížné nalézt, stačí nám lokální minimum?

Příklad: účelová funkce $f\colon \mathbb{R} \to \mathbb{R}$

$$\min f(x) \qquad \text{za podmínky } x \in \mathbb{R}$$

Deduktivní metoda

- Předpoklady: f je diferencovatelná a $\operatorname*{argmin}_{x \in X} f(x) \neq \emptyset$
- $x^* \in \operatorname*{argmin}_{x \in X} f(x)$ najdeme mezi stacionárními body (f'(x)=0)

Induktivní metoda

- Předpoklad: f je diferencovatelná
- Pokud je f konvexní a $f'(x^*) = 0$, pak $x^* \in \operatorname*{argmin}_{x \in X} f(x)$
- Je-li f dvakrát diferencovatelná, $f'(x^*) = 0$ a $f''(x^*) > 0$, pak x^* je jen lokální minimum

Formy řešení

- Analytický tvar (lineární regrese)
- Algoritmus (úloha lineárního programování)
- Iterační metoda konvergující k (lokálnímu) minimu

Elementární úlohy a jejich řešení

Feasibility problem

Jako odměnu za realizaci projektu si mají 3 kolegové rozdělit celkem 100 tis. Kč. Práce na projektu probíhaly postupně ve 3 dvojicích. Každá dvojice se dohodla na minimální částce, kterou si přeje získat (tis. Kč): $(62, 90, 50)$. Lze dojít ke shodě?

Formulace

Označme spolupracovníky jako $1, 2, 3$ a podíl i-tého z nich jako x_i. Úlohu lze formulovat jako nalezení libovolného vektoru $\mathbf{x} = (x_1, x_2, x_3) \in \mathbb{R}^3_+$, který řeší soustavu

$$\begin{aligned} x_1 + x_2 + x_3 &= 100 \\ x_1 + x_2 &\geq 62 \\ x_1 + x_3 &\geq 90 \\ x_2 + x_3 &\geq 50 \end{aligned}$$

Feasibility problem – řešení

Původní *feasibility problem* má řešení právě tehdy, když má řešení tato optimalizační úloha:

Ekvivalentní úloha

Minimalizuj libovolnou konstantní funkci $\mathbb{R}^3 \to \mathbb{R}$ za podmínek

$$\begin{aligned} x_1 + x_2 + x_3 &= 100 \\ x_1 + x_2 &\geq 62 \\ x_1 + x_3 &\geq 90 \\ x_2 + x_3 &\geq 50 \\ (x_1, x_2, x_3) &\in \mathbb{R}^3_+ \end{aligned}$$

Úloha je nepřípustná, řešení neexistuje.

Maximalizace příjmu

Krejčí má 16 jednotek materiálu A, 11 jednotek materiálu B a 15 jednotek materiálu C. Vyrábí obleky a šaty. Oblek vyrobí z 2 jednotek materiálu A, jednotky materiálu B a jednotky materiálu C. Šaty vyrobí z jednotky materiálu A, 2 jednotek materiálu B a 3 jednotek materiálu C. Oblek se prodává za 30 a šaty za 50 (v tis. Kč). Kolik má vyrobit obleků a šatů, aby byl příjem maximální?

Úloha lineárního programování

$$
\begin{aligned}
\max \quad & 30x_1 + 50x_2 \\
\text{za podmínek} \quad & x_i \geq 0, \qquad i = 1, 2 \\
& 2x_1 + x_2 \leq 16 \\
& x_1 + 2x_2 \leq 11 \\
& x_1 + 3x_2 \leq 15
\end{aligned}
$$

Maximalizace příjmu – řešení

Úloha lineárního programování

$$\begin{aligned} \max \quad & 30x_1 + 50x_2 \\ \text{za podmínek} \quad & x_i \geq 0, & i = 1, 2 \\ & 2x_1 + x_2 \leq 16 \\ & x_1 + 2x_2 \leq 11 \\ & x_1 + 3x_2 \leq 15 \end{aligned}$$

- Řešením je vektor $(7, 2)$, optimální hodnota je 310
- Zásoby materiálů A i B jsou vyčerpány, zbydou 2 jednotky C

Optimální umístění

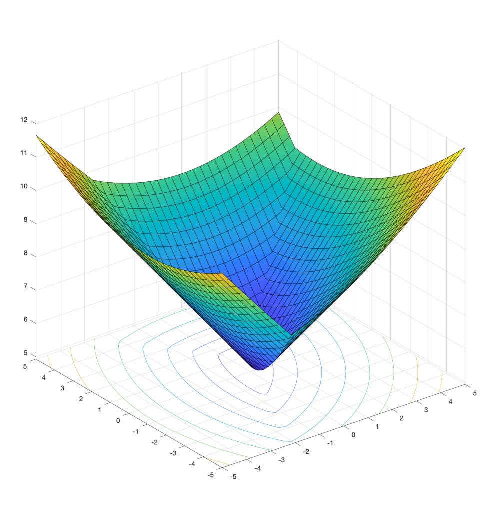
Hledáme lokaci pro heliport, z něhož dolétne helikoptéra po úsečce do nejvzdálenějšího z m bodů $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^2$ v nejkratším čase.

Optimalizační úloha

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

Optimální umístění – příklad

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



Optimální umístění – vlastnosti úlohy

Optimalizační úloha

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

- Konvexní úloha bez omezení
- Účelová funkce není hladká
- Lze ukázat, že řešení existuje pro libovolné m

Nejmenší kruh obsahující zadané body

Původní konvexní úloha

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

Úloha programování na kuželu druhého řádu

Minimalizuj funkci $g(r) = r$ za podmínek

$$\|\mathbf{x} - \mathbf{a}_i\| \leq r, \quad i = 1, \dots, m$$

$$\mathbf{x} \in \mathbb{R}^2$$

$$r \in \mathbb{R}$$

Časté prohřešky při formulaci a řešení optimalizačních úloh

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

- 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
- Lokální optimum je zaměněno za globální
- Maticové výrazy jsou nesmyslné (viz maticové zločiny)