Optimalizace

Řešení vybraných úloh. Pohled zpět.

Tomáš Kroupa Tomáš Werner

2021 LS

Fakulta elektrotechnická
ČVUT v Praze

Fermat-Weberův problém

Úloha na optimální umístění – geometrický medián

Fermat-Weberův problém

Pro zadané body $\mathbf{a}_1, \ldots, \mathbf{a}_m \in \mathbb{R}^n$ hledáme minimum funkce

$$f(\mathbf{x}) = \sum_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|, \qquad \mathbf{x} \in \mathbb{R}^n .$$

- Norma je zde eukleidovská, jsou ale i jiné možnosti
- Pro $n \geq 2$ neexistuje vzorec pro řešení pomocí konečného počtu operací, nutno použít numerické metody
- To je v kontrastu s minimalizací funkce $f(\mathbf{x}) = \sum\limits_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|^2$, kde je řešením těžiště $\frac{1}{m} \sum_{i=1}^{m} \mathbf{a}_i$

Analytické vlastnosti

- Funkce $f(\mathbf{x}) = \sum\limits_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|$ je konvexní (součet konv. funkcí)
- Gradient je $\nabla f(\mathbf{x}) = \sum\limits_{i=1}^{m} \frac{\mathbf{x}-\mathbf{a}_i}{\|\mathbf{x}-\mathbf{a}_i\|}$ pro $\mathbf{x} \neq \mathbf{a}_i$
- Pomocí konvexní analýzy lze ukázat, že minimum funkce f existuje a platí následující věta

Věta

Nechť body $\mathbf{a}_1, \ldots, \mathbf{a}_m$ neleží na společné přímce.

- Existuje jediný bod minima $\mathbf{x}^*$ funkce f
- Pokud $\mathbf{x}^* \neq \mathbf{a}_i$ splňuje $\sum\limits_{i=1}^{m} \frac{\mathbf{x}^*-\mathbf{a}_i}{\|\mathbf{x}^*-\mathbf{a}_i\|} = \mathbf{0}$, pak je bodem minima

Případ $n = 1$

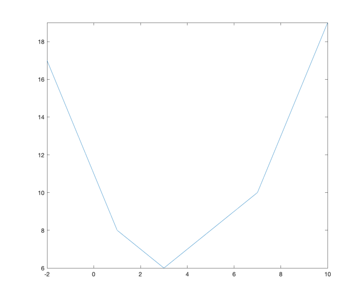
Úloha

Pro zadaná čísla $a_1, \dots, a_m \in \mathbb{R}$ hledáme minimum funkce

$$f(x) = \sum_{i=1}^{m} |x - a_i|, \qquad x \in \mathbb{R}\,.$$

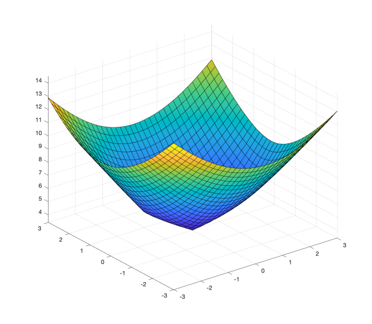
Řešením je medián čísel $a_1, \dots, a_m$.

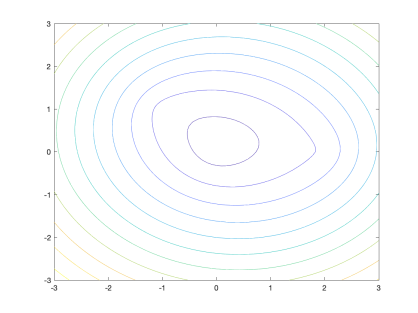
Funkce f pro $a_1 = 1$, $a_2 = 3$, $a_3 = 7$



Případ (1) pro $n = 2$ a $m = 3$

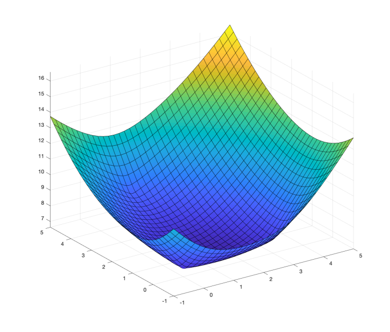
- Funkce f pro $\mathbf{a}_1 = (0, 0)$, $\mathbf{a}_2 = (2, 0)$, $\mathbf{a}_3 = (-1, 1)$
- Minima se nabývá v bodě $\mathbf{a}_1$, kde funkce nemá derivaci
- V trojúhelníku $\triangle \mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$ je úhel u vrcholu $\mathbf{a}_1$ větší než $120°$

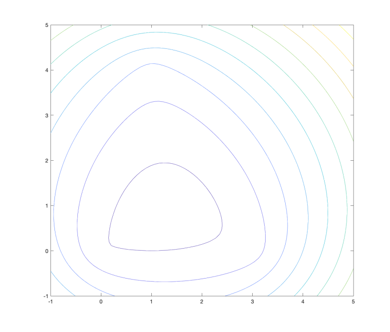




Případ (2) pro $n = 2$ a $m = 3$

- Funkce f pro $\mathbf{a}_1 = (0, 0)$, $\mathbf{a}_2 = (3, 0)$, $\mathbf{a}_3 = (1, 4)$
- Minima se nabývá v bodě uvnitř trojúhelníka $\triangle \mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$, který má všechny úhly ostré





Obecný případ pro $n = 2$ a $m = 3$

Předpoklad: body $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3 \in \mathbb{R}^2$ neleží na přímce

Trojúhelník $\triangle \mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$ má úhel větší než $120°$

Řešením je vrchol $\mathbf{a}_i$, u něhož je úhel větší než $120°$

Trojúhelník $\triangle \mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$ má úhly menší než $120°$

Řešením je Torricelliho bod trojúhelníka $\triangle \mathbf{a}_1\mathbf{a}_2\mathbf{a}_3$, tj. bod ze kterého jsou "vidět" všechny strany pod úhlem $120°$

Numerický výpočet geometrického mediánu

- Geometrický náhled na situaci v $\mathbb{R}^2$ nepomůže v tom, jak řešení Fermat-Weberova problému najít v $\mathbb{R}^n$ pro m bodů
- Iterační metoda na výpočet řešení je založena na pozorování, že podmínka $\nabla f(\mathbf{x}) = \sum_{i=1}^{m} \frac{\mathbf{x}-\mathbf{a}_i}{\|\mathbf{x}-\mathbf{a}_i\|} = \mathbf{0}$ je ekvivalentní

$$T(\mathbf{x}) = \mathbf{x}$$

kde $T(\mathbf{x}) := \frac{1}{\sum_{i=1}^{m}\|\mathbf{x}-\mathbf{a}_i\|} \sum_{i=1}^{m} \frac{\mathbf{a}_i}{\|\mathbf{x}-\mathbf{a}_i\|}$

- Tedy hledáme tzv. pevný bod zobrazení $T\colon \mathbb{R}^n \to \mathbb{R}^n$

Weiszfeldova metoda

$$T(\mathbf{x}) := \frac{1}{\sum_{i=1}^{m} \|\mathbf{x} - \mathbf{a}_i\|} \sum_{i=1}^{m} \frac{\mathbf{a}_i}{\|\mathbf{x} - \mathbf{a}_i\|}$$

Weiszfeld (1937)

1. Inicializace: $\mathbf{x}_0 \in \mathbb{R}^n$, $\mathbf{x}_0 \neq \mathbf{a}_i$, $i = 1, \ldots, m$
2. Krok: $\mathbf{x}_{k+1} := T(\mathbf{x}_k)$, $k = 0, 1, \ldots$

- Jde o gradientní metodu s délkou kroku $\left(\frac{1}{\sum_{i=1}^{m} \|\mathbf{x}_k - \mathbf{a}_i\|}\right)^{-1}$
- Předpokládáme, že body $\mathbf{a}_1, \ldots, \mathbf{a}_m$ neleží na přímce
- Posloupnost $\mathbf{x}_0, \mathbf{x}_1, \ldots$ za rozumných podmínek konverguje ke geometrickému mediánu

Nekonvexní úlohy

Nekonvexní úloha s konvexním polynomem

Optimalizační úloha

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

Ekvivalentní rozhodovací úloha `Set-Partitioning`

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

Jde o NP-těžkou optimalizační úlohu.

Nekonvexní úloha – shlukování

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

$$\mathcal{C}_j = \{\mathbf{a}_i \mid \|\mathbf{a}_i - \mathbf{x}_j\| \leq \|\mathbf{a}_i - \mathbf{x}_\ell\| \ \forall i, \ell\},$$

  tak, aby součet vzdáleností k prototypům byl minimální

- Minimalizujeme tak funkci

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

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

Jde o NP-těžkou úlohu.

Snadná nekonvexní úloha

Minimalizujeme kvadratickou formu na jednotkové sféře:

Věta (Courant–Fischer)

Nechť $\mathbf{B} \in \mathbb{R}^{m \times m}$ je symetrická matice a seřaďme její vlastní čísla vzestupně, $\lambda_1 \leq \cdots \leq \lambda_m$. Potom platí

$$\lambda_1 = \min \{\mathbf{x}^\mathsf{T} \mathbf{B} \mathbf{x} \mid \mathbf{x} \in \mathbb{R}^m,\ \|\mathbf{x}\| = 1\}$$

a minima se nabývá pro vlastní vektor $\mathbf{v}_1$ odpovídající λ_1.

Reference

T. Werner. *Optimalizace*. Elektronická skripta. FEL ČVUT, 2020.

F. J. Aragón, M. A. Goberna, M. A. López, M. M. Rodríguez. *Nonlinear optimization*. Springer, 2019.

A. Beck and S. Sabach. Weiszfeld's method: Old and new results. *Journal of Optimization Theory and Applications*, 164(1):1–40, 2015.