Lineární programování (optimalizace) a soustavy lineárních nerovností

2017

tuma@karlin.mff.cuni.cz

Příklad úlohy lineárního programování

najděte maximální hodnotu funkce $x_1 + x_2$

přes všechny vektory $\mathbf{x} = (x_1, x_2)^T \in \mathbb{R}^2$

které splňují podmínky

$$\begin{aligned} x_1 &\geq 0 \\ x_2 &\geq 0 \\ x_2 - x_1 &\leq 1 \\ x_1 + 6x_2 &\leq 15 \\ 4x_1 - x_2 &\leq 10 \quad . \end{aligned}$$

Geometrická interpretace

- K úloze si můžete nakreslit obrázek - každá z *omezujících nerovností* popisuje nějakou polorovinu.
- Body **x**, které splňují všechny omezující nerovnosti, tvoří průnik těchto polorovin. V našem případě je to konvexní pětiúhelník. Body splňující všechny omezující nerovnosti (také se jim říká omezující podmínky) jsou *přípustná řešení*.
- Mezi všemi přípustnými řešeními **x** hledáme to, které má maximální standardní skalární součin s vektorem $(1,1)^T$, neboli největší orientovanou projekci do směru tohoto vektoru. Takovému přípustnému řešení říkáme optimální řešení.
- Pokud si vše nakreslíme, vyjde nám maximální hodnota *účelové funkce* $x_1 + x_2$ rovná 5 a dosahuje se jí v optimálním bodě $(3,2)^T$, který je vrcholem pětiúhelníku všech přípustných řešení.

Česko-anglický slovníček

účelová funkce	–	cost function
omezující podmínky	–	constraints
přípustné řešení	–	feasible solution
množina přípustných řešení	–	feasible set
optimální řešení	–	optimal solution
optimální hodnota	–	optimal value

Obecná formulace úlohy lineárního programování

najděte maximální hodnotu funkce $f(\mathbf{x}) = \mathbf{c}^T\mathbf{x}$

přes všechny vektory $\mathbf{x} \in \mathbb{R}^n$

splňujícími podmínky $A\mathbf{x} \leq \mathbf{b}$,

kde $\mathbf{c} \in \mathbb{R}^n$, $\mathbf{b} \in \mathbb{R}^m$ jsou dané vektory a $A \in \mathbb{R}^{m \times n}$ je daná matice.

Pro dva vektory $\mathbf{d} = (d_1, d_2, \ldots, d_m)^T, \mathbf{e} = (e_1, e_2, \ldots, e_m)^T \in \mathbb{R}^m$ píšeme

$$\mathbf{c} \leq \mathbf{d} \ \text{ právě když } \ c_i \leq d_i \ \text{ pro každé } \ i = 1, 2, \ldots, m \ .$$

Nepřípustná úloha

Množina všech přípustných řešení je průnik uzavřených poloprostorů v $\mathbb{R}^n$. Pokud je neprázdná, nazývá se *konvexní polyedr*. Ten může být i degenerovaný, tj. mít menší dimenzi než n, v extrémním případě může jít o jediný bod.

Můžeme nějak poznat, kdy je množina přípustných řešení prázdná, tj. kdy je úloha *nepřípustná*?

Řádky v matici A označíme jako obvykle $\tilde{\mathbf{a}}_1^T, \tilde{\mathbf{a}}_2^T, \ldots, \tilde{\mathbf{a}}_m^T$.

Chceme poznat, kdy je neřešitelná soustava lineárních nerovností

$$\begin{array}{rcl} \tilde{\mathbf{a}}_1^T\mathbf{x} & \leq & b_1 \\ \tilde{\mathbf{a}}_2^T\mathbf{x} & \leq & b_2 \\ & \vdots & \\ \tilde{\mathbf{a}}_m^T\mathbf{x} & \leq & b_m \ . \end{array}$$

Jak dokážeme neřešitelnost soustavy nerovností?

Odvodíme z předpokladu řešitelnosti spor.

Jestliže nějaké $\mathbf{x} \in \mathbb{R}^n$ splňuje všechny nerovnosti, jaké další nerovnosti musí splňovat?

Očividně můžeme každou nerovnost $\tilde{\mathbf{a}}_i^T \mathbf{x} \leq b_i$ vynásobit libovolným číslem $\lambda_i \geq 0$ a dostaneme důsledek

$$\lambda_i(\tilde{\mathbf{a}}_i^T \mathbf{x}) \leq \lambda_i b_i \ .$$

Dále můžeme všechny tyto nerovnosti sečíst (přes i) a dostaneme

$$\sum_{i=1}^{m} \lambda_i(\tilde{\mathbf{a}}_i^T \mathbf{x}) \leq \sum_{i=1}^{m} \lambda_i b_i \ .$$

Odvození sporu

Levou stranu poslední nerovnosti upravíme

$$\sum_{i=1}^{m} \lambda_i(\tilde{\mathbf{a}}_i^T \mathbf{x}) = \sum_{i=1}^{m}(\lambda_i \tilde{\mathbf{a}}_i^T)\mathbf{x} = \left(\sum_{i=1}^{m} \lambda_i \tilde{\mathbf{a}}_i^T\right)\mathbf{x} = \left(\boldsymbol{\lambda}^T A\right)\mathbf{x} \ ,$$

kde $\boldsymbol{\lambda} = (\lambda_1, \lambda_2, \ldots, \lambda_m)^T$.

Také pravou stranu si přepíšeme

$$\sum_{i=1}^{m} \lambda_i b_i = \boldsymbol{\lambda}^T \mathbf{b} \ .$$

Takže každý $\mathbf{x} \in \mathbb{R}^n$ splňující původní soustavu lineárních nerovností musí splňovat také každou nerovnost

$$\left(\boldsymbol{\lambda}^T A\right)\mathbf{x} \leq \boldsymbol{\lambda}^T \mathbf{b} \ ,$$

pro libovolný vektor $\boldsymbol{\lambda} \geq 0$.

Nutná podmínka neřešitelnosti

Kdy žádné $\mathbf{x} \in \mathbb{R}^n$ poslední nerovnost nesplňuje?

Zjevně pokud pro vektor $\boldsymbol{\lambda} \geq 0$ platí, že $\boldsymbol{\lambda}^T A = \mathbf{0}_m^T$ a $\boldsymbol{\lambda}^T \mathbf{b} < 0$.

Odvodili jsme tak nutnou podmínku pro neřešitelnost soustavy lineárních nerovností:

je-li řešitelná následující soustava lineárních nerovností (a rovnic)

$$\begin{array}{rcll} \lambda_i & \geq & 0, & \qquad i = 1, \ldots, m, \\ \boldsymbol{\lambda}^T \mathbf{b} & < & 0, & \\ A^T \boldsymbol{\lambda} & = & \mathbf{0}_m, & \end{array}$$

pak je neřešitelná soustava lineárních nerovností

$$A\mathbf{x} \leq \mathbf{b} \; .$$

Věta o alternativě pro soustavy lineárních nerovností

Svět je takový, jaký má být, takže právě uvedená nutná podmínka pro neřešitelnost soustavy lineárních nerovností $A\mathbf{x} \leq \mathbf{b}$ je také postačující (důkaz je ale obtížnější).

A protože soustava lineárních nerovností $A\mathbf{x} \leq \mathbf{b}$ je buď řešitelná nebo neřešitelná, platí následující *věta o alternativě* pro soustavy lineárních nerovností.

Pro každou matici $A \in \mathbb{R}^{m\times n}$ a vektor $\mathbf{b} \in \mathbb{R}^n$ nastává právě jedna z následujících možností

- je řešitelná soustava lineárních nerovností $A\mathbf{x} \leq \mathbf{b}$,
- je řešitelná soustava lineárních nerovností

$$\begin{aligned} \lambda_i &\geq 0, \qquad i = 1, \ldots, m, \\ \boldsymbol{\lambda}^T\mathbf{b} &< 0, \\ A^T\boldsymbol{\lambda} &= \mathbf{0}_m \ . \end{aligned}$$

Příklad úlohy na lineární programování

Příklad je převzatý ze skript Jiřího Matouška.

Na základě regulací ministerstva zdravotnictví musí každé jídlo podávané v restauracích obsahovat dostatek vitamínů A, C a vlákniny. Doporučená denní množství jsou 0,5mg vitamínu A, 15mg vitamínu C a 4g vlákniny. V naší restauraci se s regulacemi chtějí vypořádat přidáváním zeleninového salátu ze syrové mrkve, nakládaných okurek a kysaného bílého zelí jako přílohy ke každému jídlu. Současně je snahou majitele zajistit tímto způsobem požadované množství vitamínů a vlákniny co nejlevněji. Potřebné údaje jsou v následující tabulce.

Tabulka

surovina	syrová mrkev	kysané zelí	nakládané okurky	požadováno na 1 porci
vitamín A [mg/kg]	35	0.5	0.5	0.5 mg
vitamín C [mg/kg]	60	300	10	15 mg
vláknina [g/kg]	30	20	10	4 g
cena[Kč/kg]	15	10	3	—

Nakládané okurky mají zjevně prošlou trvanlivost.

Označíme si x_M, x_Z a x_O množství mrkve, zelí a okurek v kilogramech na jednu porci.

Tato množství jsou přirozeně omezena podmínkami $x_M \geq 0$, $x_Z \geq 0$, $x_O \geq 0$.

Formulace pomocí lineárního programování

Požadavek ministerstva zdravotnictví na dostatečné množství vitamínu A v každé porci zapíšeme jako nerovnost

$$35x_M + 0.5x_Z + 0.5x_O \geq 0.5 \ .$$

Podobně zapíšeme další dva požadavky

$$\begin{array}{rcl} 60x_M + 300x_Z + 10x_O & \geq & 15 \ , \\ 30x_M + 20x_Z + 10x_O & \geq & 4 \ . \end{array}$$

Cena takového salátu je

$$15x_M + 10x_Z + 3x_O \ .$$

Požadavky ministerstva splníme za minimální možnou cenu vyřešením úlohy lineárního programování na následující straně.

Řešení

$$
\begin{array}{rl}
\text{minimalizujte funkci} & 15x_M + 10x_Z + 3x_O \\
\text{za podmínek} & x_M \geq 0 \\
 & x_Z \geq 0 \\
 & x_O \geq 0 \\
 & 35x_M + 0.5x_Z + 0.5x_O \geq 0.5 \\
 & 60x_M + 300x_Z + 10x_O \geq 15 \\
 & 30x_M + 20x_Z + 10x_O \geq 4 \; .
\end{array}
$$

Standardní algoritmus pro řešení takových úloh je *simplexová metoda*. Jeho autorem je George Dantzig, jeden ze zakladatelů lineárního programování. Pochází (ten algoritmus) z roku 1948.

Nejlevnějším řešením požadavků ministerstva zdravotnictví je salát za 1,40 Kč složený z 9,5 g mrkve, 38 g zelí a 290 g okurek.

Současný stav

Simplexová metoda postupně probírá vrcholy polyedru přípustných řešení, až najde to optimální.

Problém může nastat ve chvíli, kdy algoritmus musí probrat všechny vrcholy polyedru, kterých může být také 2^n v případě, kdy optimalizovaný vektor $\mathbf{x}$ má n složek (např. n-dimenzionální krychle).

Novější a teoreticky rychlejší algoritmy jsou např. *metody vnitřního bodu*.

V současnosti je řešení úlohy s desetitisíci proměnných a statisíci omezujících podmínek záležitostí vteřin na běžném stolním počítači.

Metody vnitřního bodu jsou navíc použitelné na mnohem širší třídy úloh, kdy jak omezující podmínky tak účelová funkce nejsou lineární (jsou např. kvadratické nebo obecněji konvexní).