Domácí úkol OPT 5

by a very good student

15. prosince 2019

1. Úloha Jistá výhra

15. srpna roku 2009 se v rámci čtvrtého kola fotbalové Gambrinus ligy konalo utkání mezi Spartou Praha (domácí) a Kladnem (hosté). Jistá sázková kancelář vypsala na výsledek utkání následující kurzy:

událost	1	10	0	02	2
kurz	1.27	1.02	4.7	3.09	9

První řádek tabulky označuje tyto události: 1 je výhra domácích, 10 je výhra domácích nebo remíza, 0 je remíza, 02 je remíza nebo výhra hostů, 2 je výhra hostů. Druhý řádek je kurz vypsaný na danou událost. Sázková kancelář umožňuje každému sázejícímu uzavírat sázky na libovolnou kombinaci událostí (např. lze současně vsadit na vítězství Sparty i Kladna).

V rámci své marketingové strategie nabízí sázková kancelář sázejícím jednorázový bonus ve výši 50 % z vložené částky a to až do výše 1000 korun. Bonus nelze vybrat přímo, ale lze ho použít k sázení. To znamená, že pokud například sázející vloží svých 2000 korun, přidá mu sázková kancelář bonus 1000 korun. Sázející pak může prosázet celkem 2000 + 1000 = 3000 korun.

Je spousta strategií, jak sázet. Jednou z nich je sázet tak, abychom maximalizovali jistou výhru. Jistá výhra je taková, kterou dostaneme bez ohledu na výsledek utkání, který je (nebo by alespoň měl být) neznámý. Jinými slovy jistá výhra je minimální možná výhra. Uveďme jednoduchý příklad. Předpokládejme, že vsadíme celých 3000 korun (našich 2000 plus 1000 bonus) na výhru Kladna. V tomto případě je jistá výhra 0 korun, které dostaneme pokud vyhraje Sparta nebo pokud bude remíza. Je očividné, že k tomu abychom zajistili nenulovou jistou výhru, musíme pokrýt nějakou částkou každý z možných výsledků daného utkání. První co nás napadne je vsadit 3000 korun rovnoměrně na všechny události, tj. $3000/5 = 600$ korun na každou událost vypsanou sázkovou kanceláří. V tomto případě bude hodnota jisté výhry $600 \times (1.27 + 1.02) = 1374$ korun, které dostaneme pokud vyhraje Sparta (ověřte si, že ve všech ostatních případech bude výhra vyšší). Jak vidno ani tato sázka nevypadá optimálně, protože ve výsledku můžeme přijít o $2000 - 1374 = 626$ korun.

Naším cílem je nalézt optimální sázku, tj. optimální rozložení vložených 2000 korun plus 1000 korunový bonus na jednotlivé události, tak aby jistá výhra byla maximální. Pokud úlohu správně vyřešíte, uvidíte, že optimální sázka dokonce zaručí vyšší výhru než vložených 2000 korun.

Formalizace slovního zadání úlohy

Zavedeme proměnné $x_1, ..., x_5$, které budou odpovídat množství peněz vsazených na danou událost. Jejich význam je v tabulce:

událost	1	10	0	02	2
vložené peníze [Kč]	x_1	x_2	x_3	x_4	x_5

Musí být $x_1 + ... + x_5 = 3000$ (chceme prosázet přesně 3000 korun) a $x_1, ..., x_5 \geq 0$ (sázky nemohou být záporné). Z tabulek odvodíme, jaká bude hodnota výhry v závislosti na výsledku utkání:

výsledek	výhra domácích	remíza	výhra hostů
hodnota výhry [Kč]	$1.27x_1 + 1.02x_2$	$1.02x_2 + 4.7x_3 + 3.09x_4$	$3.09x_4 + 9x_5$

Hodnota jisté (tj. minimální) výhry je tudíž rovna

$$f(x_1, ..., x_5) = \min\{1.27x_1 + 1.02x_2, 1.02x_2 + 4.7x_3 + 3.09x_4, 3.09x_4 + 9x_5\}$$

Hledání optimální sázky pak můžeme vyjádřit jako maximalizaci hodnoty jisté výhry za daných omezení:

$$\begin{gathered} \max f(x_1, ..., x_5) \\ \text{za podmínek } x_1 + ... + x_5 = 3000 \\ x_1, ..., x_5 \geq 0 \end{gathered}$$

Převod úlohy na lineární program

Vytvoříme pomocnou proměnnou y tak, že jí omezíme hledání maxima zespoda

$$\max\{f(\mathbf{x}), \mathbf{x} \in \mathbb{R}^5\} = \max\{y \mid (\mathbf{x}, y) \in \mathbb{R}^6, f(\mathbf{x}) \geq y\}$$

Pro každé optimální řešení pak musí platit, že $f(\mathbf{x}) = y$, jinak by bylo možné y zvětšit a nejednalo by se o optimální řešení.
Zároveň platí, je-li minimum nějaké množiny větší nebo rovno nějakému číslu, musí být každý prvek této množiny větší nebo roven tomuto číslu, tedy každý prvek této množiny nám vytvoří samostatnou podmínku v LP:

$$\begin{aligned} f(\mathbf{x}) &\geq y \\ \min_{i \in \{1,2,3\}} \{a_i\} &\geq y \\ \forall i \in \{1,2,3\} : a_i &\geq y \end{aligned}$$

A nakonec využijeme faktu, že

$$\max y = -\min(-y)$$

Výsledný LP poté vypadá ve standardizovaném zápisu jako

$$\begin{aligned} \min \; & -y \\ \text{za podm. } & x_1 + x_2 + x_3 = 3000 \\ & -1.27x_1 - 1.02x_2 + y \leq 0 \\ & -1.02x_2 - 4.7x_3 - 3.09x_4 + y \leq 0 \\ & -3.09x_4 - 9x_5 + y \leq 0 \\ & x_1, x_2, x_3, x_4, x_5, y \geq 0 \end{aligned}$$

Numerické řešení úlohy

Numerickým řešením dostaneme následující sázkovou tabulku:

událost	1	10	0	02	2
vložené peníze [Kč]	0	2694.6	0	0	305.4

A hodnota minimální výhry je $\min\{1.02 \cdot 2694.6, 1.02 \cdot 2694.6, 9 \cdot 305.4\} = 2748.5$ Kč.

Modifikovaná úloha a její řešení

Rovnice pro modifikovanou úlohu vypadá následovně:

$$f(x_1, x_2, x_3) = \min\{1.27x_1, 4.7x_2, 9x_3\}$$

kde je navíc podmínka $x_1, x_2, x_3 \geq 400$.
Tuto úlohu vyjádříme lineárním programem jako

$$\begin{aligned} \min \; & -y \\ \text{za podm. } & x_1 + x_2 + x_3 = 3000 \\ & -1.27x_1 + y \leq 0 \\ & -4.7x_2 + y \leq 0 \\ & -9x_3 + y \leq 0 \\ & x_1, x_2, x_3 \geq 400 \\ & y \geq 0 \end{aligned}$$

Numerickým řešením této úlohy dostaneme sázkovou tabulku

událost	1	0	2
vložené peníze [Kč]	2046.9	553.1	400

A hodnota minimální výhry je $\min\{1.27 \cdot 2046.9, 4.7 \cdot 553.1, 9 \cdot 400\} = 2599.6$ Kč.

2. Úloha Minimaxní prokládání bodů přímkou

Dáno je m bodů $(x_1, y_1), ..., (x_m, y_m)$, kde x_i, $y_i \in \mathbb{R}$. Úkolem je nalézt takovou lineární funkci

$$f_{a,b}(x) = ax + b$$

parametrizovanou čísly $a, b \in \mathbb{R}$, která nejlépe aproximuje funkční závislost odpovídající daným bodům.

Je mnoho kritérií, kterými můžeme definovat dobrou aproximaci. Konkrétní volba kritéria závisí na aplikaci. My si jako kritérium dobré aproximace zvolíme maximální absolutní odchylku definovanou jako

$$\epsilon(a, b) = \max_{i=1,...,m} |f_{a,b}(x_i) - y_i|$$

Naším cílem bude nalézt parametry a, b lineární funkce, pro které je $\epsilon(a, b)$ nejmenší.

Úloha má názornou geometrickou interpretaci, znázorněnou na obrázku: hledáme pás s minimální šířkou (měřeno svisle), jenž obepíná všechny body. Střed pásu je přímka $ax + b = 0$ a jeho šířka je $2\epsilon(a, b)$.

Převod úlohy na lineární program

Řešíme problém, kdy se snažíme minimalizovat maximální absolutní odchylku, tedy

$$\min \epsilon(a, b)$$

Přepíšeme rovnici pro maximální absolutní odchylku jako

$$\epsilon(a, b) = \max_{i=1,...,m} \max\{f_{a,b}(x_i) - y_i, -f_{a,b}(x_i) + y_i\} = \max_{i=1,...,m} \max\{ax_i + b - y_i, -ax_i - b + y_i\}$$

Což nám pro LP dá $2m$ podmínek. Opět vytvoříme novou slackovou pomocnou proměnnou, z, a problém zapíšeme jako

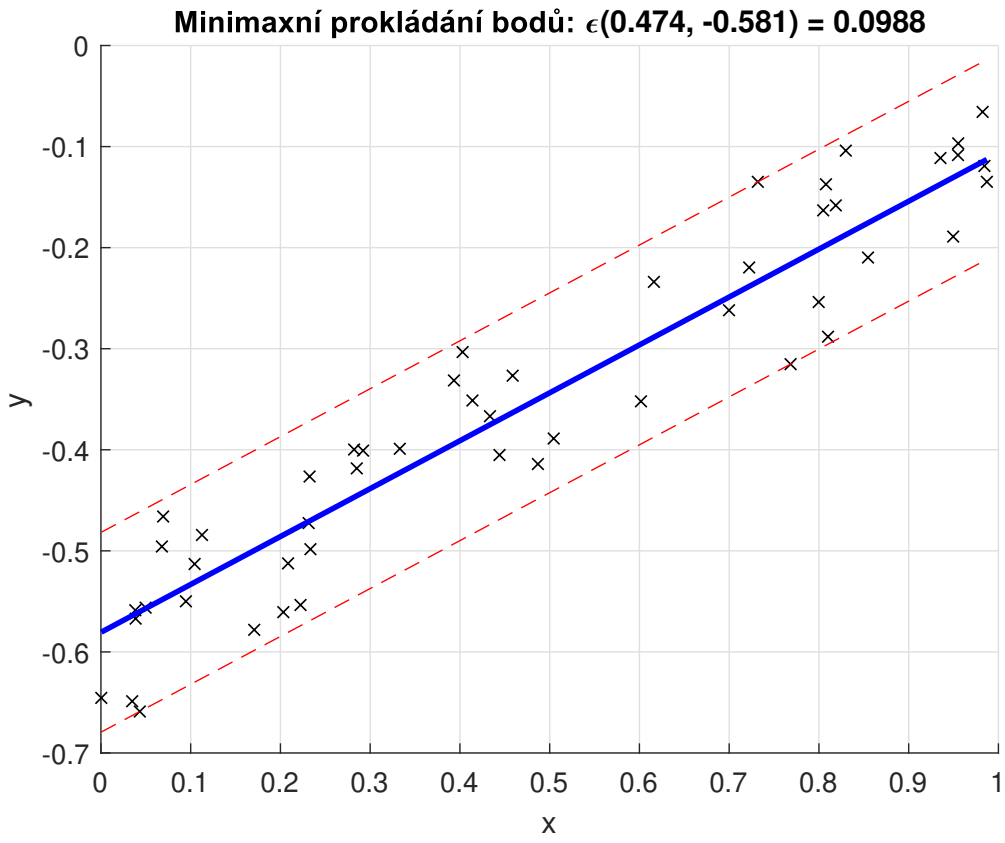
$$\min \epsilon(a, b) = \min\{z \mid \epsilon(a, b) \leq z\}$$

Zápis LP s proměnnými a, b, z, vypadá následovně

$$\begin{aligned} & \min z \\ \text{za podm. } \forall i \in \{1, ..., m\} : & \; ax_i + b - z \leq y_i \\ \forall i \in \{1, ..., m\} : & -ax_i - b - z \leq -y_i \\ & a, b, z \in \mathbb{R} \end{aligned}$$

Numerické řešení úlohy

Vyřešíme problém pro 50 zadaných bodů.
Nalezená přímka s minimální absolutní odchylkou od daných bodů je dána rovnicí $0.474x - 0.581$ a velikost této odchylky je 0.0988.



Modifikovaná úloha a její řešení

V případě, že $\mathbf{a}, \mathbf{x}_i (\forall i) \in \mathbb{R}^n$, se změní zápis rovnice pro přímku v $\mathbb{R}^n$ a maximální absolutní odchylku jen takto:

$$f_{a,b}(\mathbf{x}_i) = \mathbf{a}^T\mathbf{x}_i - y_i$$
$$\epsilon(a,b) = \max_{i=1,\dots,m} \max\{f_{a,b}(\mathbf{x}_i) - y_i, -f_{a,b}(\mathbf{x}_i) + y_i\} = \max_{i=1,\dots,m} \max\{\mathbf{a}^T\mathbf{x}_i + b - y_i, -\mathbf{a}^T\mathbf{x}_i - b + y_i\}$$

S použitím stejných kroků jako v nemodifikované úloze můžeme vyjádřit výsledný LP jako

$$\begin{aligned} & \min z \\ \text{za podm. } \forall i \in \{1,\dots,m\} : & \ \mathbf{a}^T\mathbf{x}_i + b - z \le y_i \\ \forall i \in \{1,\dots,m\} : & -\mathbf{a}^T\mathbf{x}_i - b - z \le -y_i \\ & \mathbf{a} \in \mathbb{R}^n; b, z \in \mathbb{R} \end{aligned}$$