Binárna Klasifikácia pomocou LP

Dávid Daniš, Michal Chymo, Maximilián Zivák, Adam Chrenko

December 2021

1 Popis problému

TODO

2 Dôkaz ekvivalencie

Ak $\forall x, y\ \exists a, b$ také, že:

$$a^T x - b > 0$$
$$a^T y - b < 0$$

Potom:

$$\begin{aligned} &nech\ R\ = \{(a^T x - b) \mid x \in X\} \\ &nech\ F\ = \{-(a^T y - b) \mid y \in Y\} \\ &nech\ k = \frac{1}{minimum(F \cup R)} \qquad kde\ 1\ je\ \text{ľubovolná konštanta} \end{aligned}$$

Z definície k vieme, že $k > 0$ a platí:

$$\begin{aligned} &k(a^T x - b) \geq 1 && x \in X \\ &k(a^T y - b) \leq -1 && y \in Y \end{aligned}$$

Dôkaz že predošlé dva riadky platia nechávame ako cvičenie pre čitateľa.
Z čoho vyplýva, že existujú $a' = ka$, $b' = kb$ také, že:

$$a'^T x - b' \geq 1$$

$$a'^T y - b' \leq -1$$

3 Konštantná účelová funkcia

Vyberme $\vec{a}$, b tak, že

$$\vec{a}^T \vec{x_i} - b \geq 1, \quad \forall i \in 1, 2, 3 \dots N \qquad \vec{a}^T \vec{y_i} - b \leq -1, \quad \forall i \in 1, 2, 3 \dots M$$

Toto potrebujeme dostať do tvaru úlohy LP, tak že

$$\min_{\vec{p}} \vec{c}^T \vec{p}$$

$$A_{ub} \vec{p} \leq \vec{b}_{ub}$$
$$kde \quad \vec{p} = [a_1, a_2 \dots a_n, b]^T$$

Za $\vec{c}$ si dosadíme $\vec{0}$ a teda účelovú funkciu môžeme zatial ignorovať.
Zároveň skalár **b** je niekoľkými krokmi možné "skryť"do vektorov $x_i, y_i\, \forall i \in 1, 2, 3 \ldots N = M$ a teda do matice $\mathbf{A_{ub}}$.
Do vektorov x_i, y_i pridáme na koniec matice riadok samých -1, tieto matice označíme $\mathbf{X'}, \mathbf{Y'}$. Potom tieto matice transponujeme, čiže výsledok úlohy LP bude vektor, pre ktorý platí $|\mathbf{v}| = \mathbf{n+1}$ kde $n+1$ hodnota určí hodnodu skaláru **b** a prvých n hodnôt určí koeficienty hľadaného vektora $\vec{a}$.
n je dimenzia priestoru v ktorom pracujeme

Úlohu potrebujeme dostať do štandardizovaného tvaru pre počítanie na PC: $LHS \leq RHS$.
Takže maticu $\mathbf{X'}$ prenásobíme -1 a za vektor **b** dosadiť samé -1. Tieto 2 matice $(\mathbf{X'}, \mathbf{Y'})$ spojíme do jednej nasledovne:

$$A_{ub} = \left[\frac{-X'^T}{Y'^T}\right] \qquad b_{ub} = (-1 \ldots -1)^T$$

$$-X'^T = \begin{bmatrix} -x_{1,1} & -x_{1,2} & \ldots & -x_{1,50} & 1 \\ -x_{2,1} & -x_{2,2} & \ldots & -x_{2,50} & 1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ -x_{50,1} & -x_{50,2} & \ldots & -x_{50,50} & 1 \end{bmatrix} \qquad Y'^T = \begin{bmatrix} y_{1,1} & y_{1,2} & \ldots & y_{1,50} & -1 \\ y_{2,1} & y_{2,2} & \ldots & y_{2,50} & -1 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ y_{50,1} & y_{50,2} & \ldots & y_{50,50} & -1 \end{bmatrix}$$

Úlohu v tomto tvare je možné dosadiť do vášho obľúbeného programu na riešenie lineárnej optimalizácie. (SciPy, MatLab ...).

$$A_{ub} * [a_1, a_2 \ldots a_n, b]^T \leq \vec{b}_{ub}$$

Toto riešenie porovnáme pri použití rôznych metód

4 L1 norma

LP s nulovou účelovou funkciou nájde ľubuvolnú nadrovinu. Chceli by sme ale odstrániť nesignifikantné parametre. Ak účelovú funkciu zameníme za L_1 normu $\vec{a}$, dostaneme riešenie s veľkým počtom nulových zložiek. Nový problém LP je tvaru:

$$\min_a ||a||_1$$

$$A_{ub} * \vec{p} \leq \vec{b}_{ub}$$

$$kde \quad \vec{p} = [a_1, a_2 \ldots a_n, b]^T$$

Čo je:

$$\min_a \sum_{i=1}^{n} |a_i|$$

$$A_{ub} * \vec{p} \leq \vec{b}_{ub}$$

Keďže účelová funkcia LP v štandardnom tvare musí byť lineárna, odstránime absolútnu hodnotu štandardným spôsobom pomocou novej premennej $\vec{t}$

$$\min_a \sum_{i=1}^{n} |a_i| \quad \rightarrow \quad \min_t \sum_{i=1}^{n} t_i$$

$$-t_i \leq a_i \leq t_i \quad i \in \{1, 2, ..., n\}, \quad t_i \in [0, \infty)$$

$$-t_i \leq a_i \quad a_i \leq t_i$$

$$-I_n a - I_n t \leq \vec{0}$$

$$I_n a - I_n t \leq \vec{0}$$

A teda nová úloha LP je:

$$
\begin{aligned}
\min_t \sum_{i=1}^{n} t_i &= \vec{1}_n^T * \vec{t} \\
A_{ub} * \vec{p} &\leq \vec{b}_{ub} \\
-I_n a - I_n t &\leq \vec{0}_n \\
I_n a - I_n t &\leq \vec{0}_n
\end{aligned}
$$

Na prepísanie do formy s jednou premennou a jednou maticou musíme spraviť pár úprav. Chceli by sme aby hodnoty a_i a b neovplyňovali hodnotu účelovej funkcie, čiže je nutné dosadiť za ich koeficienty nuly, zárovneň potrebujeme aby t_i(čo sú nové neznáme) všetky prispievali rovným dielom.
Maticu A je nutné rozšíriť o nulové a jednotkové matice, aby sme zachovali dimenzie, keďže chceme aby nám po násobení matice a vektoru vyšiel vektor dĺžky $k + 2n$.
Pravú stranu nerovnice musíme upraviť podobným spôsobom, tentokrát to dosiahneme rozšírením vektoru na pravej strane, nulovým vektorom dĺžky $2n$.Týmito úpravami docielimi toho, že LP nájde vektor ktorého prvých n členov je hľadaný vektor, a $n + 1$ člen je hľadaný skalár, ostatné prvky vektoru ignorujeme.

$$\min_y q^T y$$

$$q = [0_1, 0_2 \ldots 0_{n+1} | 1_1, 1_2 \ldots 1_n] \qquad y = [a_1, a_2 \ldots, a_n | b | t_1, t_2 \ldots t_n]$$

$$
\begin{bmatrix}
A_{k,n} & -\vec{1}_k & -0_{k,n} \\
I_{n,n} & -\vec{0}_k & -I_{n,n} \\
-I_{n,n} & -\vec{0}_k & -I_{n,n}
\end{bmatrix}
*
\left[\begin{array}{c} a_1 \\ \vdots \\ a_n \\ \hline b \\ \hline t_1 \\ \vdots \\ t_n \end{array}\right]
\leq
\left[\begin{array}{c} -1_1 \\ \vdots \\ -1_k \\ \hline 0_1 \\ \vdots \\ 0_{2n} \end{array}\right]
\qquad k = N + M
$$

Kde $A_{k,n}$ je A_{ub} bez posledného stĺpca.
V tejto forme je možné nerovnicu dosadiť do $< Algoritmu >$ na riešenie problémov LP.

Z vačsiny Algoritmov sa dozvieme, že viac ako polovica "parametrov(features)"je nesignifikantných. $(\leq \epsilon)$

5 Využitia Binárnej Klasifikácie

TODO

6 Bonus

TODO?