Lokacija pravaca u normiranom prostoru

1 Svojstva optimalnih medijan i centar pravaca

Neka je $I = \{1, \ldots, m\}$, $m \geq 2$ skup indeksa, $\Lambda = \{T_i(x_i, y_i) \in \mathbb{R}^2 : i \in I\}$ skup točaka u ravnini s odgovarajućim težinama $\omega_i > 0$, $i \in I$ i $W = \sum_{i \in I} \omega_i$.

Promatramo probleme:

Medijan problem: problem određivanja pravca l tako da je $f(l) = \sum_{i \in I} \omega_i d(T_i, l)$ minimalna

Centar problem: problem određivanja pravca l tako da je $g(l) = \max_{i \in I} \omega_i d(T_i, l)$ minimalan.

Svojstva koja nas zanimaju (za različite udaljenosti d):

Med 1 (Weak incidence property): postoji optimalan medijan pravac koja prolazi kroz barem dvije točke iz Λ;

Med 2 (Pseudo–halving property): za svaki optimalan medijan pravac l vrijedi

$$\sum_{T_i \in B_l^-} \omega_i \leq \frac{W}{2} \quad \text{and} \quad \sum_{T_i \in B_l^+} \omega_i \leq \frac{W}{2};$$

pri čemu su B_l^- i B_l^+ poluravnine na koje pravac l dijeli $\mathbb{R}^2$ (bez pravca l)

Cen 1 (Weak blockedness property): postoji optimalan centar pravac l koja ima svojstvo da je barem tri točke iz Λ težinski maksimalno udaljeno od l;

Cen 2 (Paralel facets property): ako je $\omega_1 = \cdots = \omega_m$, onda postoji optimalan centar pravac koji je paralelan stranici konveksne ljuske skupa točaka Λ;

Horizontalna i vertikalna udaljenost točaka $T_i(x_i, y_i)$ i $T_j(x_j, y_j)$:

$$d_{hor}(T_i, T_j) = \begin{cases} |x_i - x_j|, & y_i = y_j \\ \infty, & \text{inače} \end{cases}, \quad d_{ver}(T_i, T_j) = \begin{cases} |y_i - y_j|, & x_i = x_j, \\ \infty, & \text{inače} \end{cases}$$

Teorem 1 *Za d_{ver} vrijedi $Med\,1$, $Med\,2$, $Cen\,1$, $Cen\,2$.*

Korolar 1 *Za d_{hor} vrijedi $Med\,1$, $Med\,2$, $Cen\,1$, $Cen\,2$.*

t-udaljenost točaka

Neka je $t \in \mathbb{R}^2$ dani vektor smjera. Za točke $P, Q \in \mathbb{R}^2$ definiramo udaljenost u smjeru t s

$$d_t(P, Q) = \gamma_t(Q - P), \; \gamma_t(x) = \begin{cases} |\alpha|, & x = \alpha t \\ \infty, & \text{inače} \end{cases}$$

t-udaljenost točke do pravca

Neka je $l_{P,t}$ pravac određen vektorom smjera t i točkom P. Tada je

$$d_t(P,l) = \min_{Q\in l} d_t(P,Q) = \begin{cases} \frac{\|\overrightarrow{PQ}\|}{\|t\|}, & l \cap l_{P,t} = \{Q\} \\ 0, & l = l_{P,t} \\ \infty, & l \cap l_{P,t} = \emptyset \end{cases}$$

Primjer 1 $d_{e_1}(P,l) = d_{hor}(P,l)$, $d_{e_2}(P,l) = d_{ver}(P,l)$

Lema 1 *Neka su $p,q \in \mathbb{R}^2 \backslash \{0\}$ i D linearan operator takav da je $D(p) = q$ i $\det D \neq 0$. Tada za sve $P \in \mathbb{R}^2$ vrijedi*

$$d_q(D(P), D(l)) = d_p(P,l), \text{ gdje je } D(l) = \{D(Q) : Q \in l\}.$$

Dokaz: Najprije ćemo dokazati da je $d_q(D(P), D(Q)) = d_p(P,Q)$, $\forall P,Q \in \mathbb{R}^2$.

i) Neka je $d_p(P,Q) = |\alpha| < \infty$. Tada je $Q - P = \alpha p$ i slijedi

$$\begin{aligned} d_q(D(P), D(Q)) &= \gamma_q(D(Q) - D(P)) \\ &= \gamma_q(D(Q-P)) \\ &= \gamma_q(\alpha D(p)) \\ &= \gamma_q(\alpha q) \\ &= |\alpha| \end{aligned}$$

ii) Neka je $d_p(P,Q) = \infty$. Tada su $Q - P$ i p linearno nezavisni pa su i $D(Q) - D(P)$ i $D(p) = q$ linearno nezavisni, a onda je $d_q(D(P), D(Q)) = \infty$.

Dokažimo sada tvrdnju teorema.

$$\begin{aligned} d_q(D(P), D(l)) &= \min_{Q\in l} d_q(D(P), D(Q)) \\ &= \min_{Q\in l} d_p(P,Q) \\ &= d_p(P,l) \end{aligned}$$

□

Teorem 2 *Za sve udaljenosti d_t vrijede svojstva $Med\,1$, $Med\,2$, $Cen\,1$, $Cen\,2$*

Da bi odredili optimalni medijan ili centar pravac s udaljenošću d_t, odaberemo regularan operator D takav da je $D(t) = e_2$. Definiramo $T_i' = D(T_i)$, $i \in I$ i riješimo vertikalni problem s točkama $D(T_i)$. Ako je rješenje vertikalnog problema pravac l^*, onda je rješenje polaznog problema $D^{-1}(l^*)$.

Lema 2 *Neka je d udaljenost inducirana normom γ, te l pravac. Tada postoji $t \in \mathbb{R}^2$, $\gamma(t) = 1$ tako da vrijedi*

$$d(P,l) = d_t(P,l), \ \forall\, P \in \mathbb{R}^2.$$

Teorem 3 *Za sve udaljenosti d inducirane normom vrijede svojstva $Med\,1$, $Med\,2$, $Cen\,1$, $Cen\,2$.*

Dokaz: Dokažimo svojstvo Med 1.

Neka je l^* optimalan pravac koji ne prolazi kroz dvije točke podataka. Odaberimo t^* tako da je $d(T_i, l^*) = d_{t^*}(T_i, l^*)$. Znamo da možemo odabrati pravac l_0 koji minimizira medijan problem s udaljenošću d_{t^*} i koji prolazi kroz najmanje dvije točke podataka. Tada je

$$\begin{aligned} f(l^*) &= \sum_{i\in I} \omega_i d(T_i, l^*) \\ &= \sum_{i\in I} \omega_i d_{t^*}(T_i, l^*) \\ &\geq \sum_{i\in I} \omega_i d_{t^*}(T_i, l^0) \\ &\geq \sum_{i\in I} \omega_i d(T_i, l^0) \\ &= f(l^0) \geq f(l^*) \end{aligned}$$

pa je i l^0 također optimalni pravac i prolazi kroz dvije točke podataka. $\square$

1.1 Algoritmi za određivanje optimalnih medijan i centar pravaca

ALGORITAM 1 - optimalni medijan pravac za proizvoljnu udaljenost induciranu normom - od svih pravaca koji prolaze kroz dvije točke podataka, za optimalan medijan pravac uzeti onaj za koji je najmanja vrijednost funkcije koju minimiziramo.

Neka je $P_{jk} = \dfrac{\omega_j}{\omega_j + \omega_k} T_j + \dfrac{\omega_k}{\omega_j + \omega_k} T_k, \quad j, k \in I, \quad j \neq k.$

Teorem 4 *Za sve udaljenosti d inducirane normom postoji optimalan centar pravac l i $i, j, k \in I$ tako da l prolazi kroz kroz točke P_{ik} i P_{jk}.*

ALGORITAM 2 - optimalni centar pravac za proizvoljnu udaljenost induciranu normom - od svih pravaca koji prolaze kroz dvije točke T_{ik}, T_{jk}, za optimalan medijan pravac uzeti onaj za koji je najmanja vrijednost funkcije koju minimiziramo.

Kažemo da pravac l podupire skup $A \subseteq \mathbb{R}^2$ ako $A \cap l \neq \emptyset$ i $A \cap B_l^- = \emptyset$ ili $A \cap B_l^+ = \emptyset$.
Za stranicu e i vrh v konveksnog skupa kažem da su međusobno suprotni par ako postoje paralelni podupirujući pravci l^1, l^2 takvi da je $e \subseteq l_1$ i $v \in l^2$.

ALGORITAM 3 - optimalni netežinski centar pravac za proizvoljnu udaljenost d induciranu normom

1. Odrediti konveksnu ljusku skupa Λ, te vrhove i stranice konveksne ljuske.
2. $z^* = \infty$
3. Za sve parove međusobno suprotnih vrhova v i stranica e konveksne ljuske od Λ, pri čemu je e određen vrhovima v_1 i v_2
    1. Odrediti pravac l kroz v_1 i v_2
    2. Izračunati $d(v, l)$
    3. Ako je $d(v, l) < z^*$, $z^* := d(v, l)$ i l^* je l translatiran za $z^*/2$ prema v
4. Output: l^*

1.2 Blok norme

Blok norma je norma kod koje je pripadna jedinična kugla konveksan poligon.

Lema 3 *Neka je d_B udaljenost inducirana blok normom γ_B s osnovnim smjerovima $\{b_1, \ldots, b_G\}$ i neka je l pravac. Tada postoji $g \in \{1, \ldots, G\}$ tako da vrijedi*

$$d_B(P, l) = d_{b_g}(P, l), \quad \forall P \in \mathbb{R}^2.$$

Prethodna lema sugerira da bi se problem s blok normom mogao razdvojiti na G neovisnih potproblema i za rješenje uzeti najbolje rješenje potproblema, a ono se može dobiti pomoću vertikalne udaljenosti.

Lema 4 *Neka je α_g kut koji zatvara smjer b_g s pozitivnim dijelom x osi*

$$D^r = \begin{bmatrix} \sin\alpha_g & -\cos\alpha_g \\ \cos\alpha_g & \sin\alpha_g \end{bmatrix} \quad i \quad D^{rs} = \frac{1}{|b_g|} D^r.$$

Tada je

$$d_{b_g}(P, l) = d_{ver}(D^{rs}(P), D^{rs}(l)) = \frac{1}{|b_g|} d_{ver}(D^r(P), D^r(l)).$$

Brzi algoritam za blok norme:

1. $z^* = \infty$
2. Za g=1 to G
   1. Izračunati kut α_g
   2. $D = \begin{bmatrix} \sin\alpha_g & -\cos\alpha_g \\ \cos\alpha_g & \sin\alpha_g \end{bmatrix}$
   3. $T_i' = D(T_i)$
   4. Odrediti l_g^* koji minimizira
      - Medijan problem: $f(l) = \sum_{i \in I} \omega_i d_{ver}(T_i', l)$
      - Centar problem: $f(l) = \max_{i \in I} \omega_i d_{ver}(T_i', l)$
   5. Ako $f(l_g^*) < z^* |b_g|$, $z^* := \frac{f(l_g^*)}{|b_g|}$ i $l^* = D^{-1}(l_g^*)$

Složenost prethodnog algoritma je $O(Gm)$.

1.3 Glatke norme

Neka je γ norma i B odgovarajuća jedinična kugla. Norma γ se naziva glatka norma ako za sve $t \in \mathbb{R}^2$, $\gamma(t) = 1$ postoji točno jedan podupirujući pravac za B.
Primjeri glatkih normi su l_p, $1 < p < \infty$. Nijedna blok norma nije glatka.

Lema 5 *Neka je d udaljenost inducirana normom γ. Neka su $t, x \in \mathbb{R}^2$ takvi da je $\gamma(t) = 1$ i $d(x, l) = d_t(x, l)$. Tada postoji podupirujući pravac jedinične kugle B za normu γ koji je paralelna s pravcem l.*

Lema 6 *Ako medijan problem s vertikalnom udaljenošću ima dva optimalna medijan pravca, onda postoje i dva neparalelna optimalna medijan pravca.*

Korolar 2 *Ako medijan problem s udaljenošću d_t, $t \in \mathbb{R}^2 \backslash \{0\}$ ima dva optimalna medijan pravca, onda postoje i dva neparalelna optimalna pravca.*

Teorem 5 *Norma γ je glatka norma ako i samo ako vrijedi da svaki optimalan medijan pravac prolazi kroz dvije točke podataka.*

Dokaz: Neka je γ glatka norma, $d(x,y)=\gamma(y-x)$ i l^* optimalan medijan pravac za medijan problem s udaljenošću d. Pretpostavimo da l^* ne sadrži dvije točke podataka i odaberimo $t^* \in \mathbb{R}^2$ takav da je

$$d_{t^*}(T_i,l^*)=d(T_i,l^*), \quad \forall T_i \in \Lambda.$$

Promatrajmo medijan problem za skup Λ i udaljenost d_{t^*} i odaberimo optimalan pravac l^0 koji prolazi kroz dvije točke podataka. On je oprimalan medijan pravac i za medijan problem s udaljenošću d. Vrijedi $l^* \neq l^0$, pa prema Lemi 6 postoji i optimalan pravac l' s $s' \neq s^*$. Odaberimo t', $\gamma(t')=1$ tako da je

$$d_t^{'}(T_i,l')=d(T_i,l').\ \forall i \in I.$$

Tada je

$$d_{t^*}(T_i,l') \geq d_{t'}(T_i,l'),\ \forall i \in I.$$

Nadalje, kako je γ glatka norma,

$$d_{t^*}(T_i,l') \neq d_{t'}(T_i,l'),\ \forall i \in I,$$

jer bi inače $d_{t^*}(T_i,l')=d_{t'}(T_i,l')=d(T_i,l'),\ \forall i \in I$ značilo da u t^* postoje dva podupirujuća pravca za jediničnu kuglu u normi γ, jedan paralelan s l^*, drugi d $l^{'}$ pa γ ne bi bila glatka norma. Stoga je

$$d_{t^*}(T_i,l') > d_{t'}(T_i,l'),\ \forall i \in I.$$

Sada je

$$\begin{aligned}
f(l^*) &= \sum_{i\in I}\omega_i d(T_i,l^*)\\
&= \sum_{i\in I}\omega_i d_{t^*}(T_i,l^*)\\
&= \sum_{i\in I}\omega_i d_{t^*}(T_i,l')\\
&> \sum_{i\in I}\omega_i d_{t'}(T_i,l')\\
&= \sum_{i\in I}\omega_i d(T_i,l^{'})=f(l')
\end{aligned}$$

što je proturječno s pretposlavkom optimalnosti od l^*.
Za obrat vidi [1].

Korolar 3 *Neka je d udaljenost inducirana glatkom normom γ. Tada su svi optimalni medijan pravci halving.*

Teorem 6 *Norma γ je glatka norma ako i samo ako vrijedi da je svaki optimalan centar pravac maksimalno udaljen od tri točke podataka.*

Literatura

[1] A. Schöbel, *Locating Lines and Hyperplanes: Theory and Algorithms*, Springer Verlag, Berlin, 1999.