20/6/06

אלגוריתמים ברשתות

מבוא לתכנות ליניארי

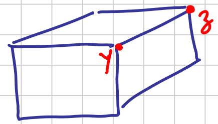
- פתרונות בסיסיים
- ניוון (degeneracy)
- קיום נק' קיצון
- אופטימליות של נק' קיצון.
- הכללה: אלימינציה בשיטת Fourier - Motzkin

הגדרה: 2 פתרונות בסיסיים $y \neq z$ נקראים **שכנים** אם $I_y \cap I_z$ מכיל $n-1$ אילוצים בת"ל, כאשר

$$I_y \triangleq \{ a_i \mid a_i y = b_i \}$$

$$.I_z \triangleq \{ a_i \mid a_i z = b_i \}$$

הערה: הבסיס "המתאים" ל- z מתקבל ע"י החלפת אילוץ אחד בבסיס "המתאים" ל- y.



צורה סטנדרטית

$$P = \{ x \mid Ax = b,\ x \geq 0 \}$$

נסמן: m - מס' עמודות A, n - מס' שורות A.

בה"כ מניחים: שורות A בת"ל (ולכן $m \leq n$).

אם x^* פתרון בסיסי אז $Ax^* = b$ ולכן $n-m$ אילוצי סימן מתקיימים בשוויון, כלומר: ב x^* יש לפחות $n-m$ רכיבים ששווים 0.

משפט: (ביחס לצורה סטנדרטית)

x^* פתרון בסיסי אמ"ם $Ax=b$ וכן יש m אינדקסים $\{B(i)\}_{i=1}^{m}$ המקיימים:

① העמודות $\{A^{B(i)}\}_{i=1}^{m}$ בת"ל.

וכן ② $i \notin \{B(1), \dots, B(m)\} \Longrightarrow x_i = 0$

הוכחה: ($\Longrightarrow$) בה"כ $\{B(i)\}_{i=1}^{m} = \{1, \dots, m\}$.

ואז נסתכל במערכת השוויונים?

$$\begin{array}{c} \begin{array}{l} m\left\{\right. \\ n-m\left\{\right. \end{array} \left(\begin{array}{ccc|ccc} A^1 & \cdots & A^m & & 0 & \\ \hline & & & 1 & & \\ & 0 & & & \ddots & \\ & & & & & 1 \end{array} \right) \cdot x = \begin{pmatrix} b \\ 0^{n-m} \end{pmatrix} \\ \qquad\qquad \underbrace{\qquad\qquad}_{m} \quad \underbrace{\qquad\quad}_{n-m} \end{array}$$

כיון ש-$A^1 \cdots A^m$ בת"ל, m הרכיבים הראשונים של
כל פתרון נקבעים באופן יחיד.
$n-m$ הרכיבים האחרונים של כל פתרון הם אפס.
דרגת המטריצה היא n, ולכן קב' האילוצים הפעילים
ביחס ל-x מכילה בסיס, כנדרש.

($\Leftarrow$) נסמן $Z \triangleq \{i \mid x_i^* = 0\}$.
כיון ש-x^* פתרון בסיסי נקבל שקב' האילוצים A
ביחד עם וקטורי היחידה $\{e^i\}_{i\in Z}$ מכילים בסיס.
ולכן למערכת השוויונים
$$\begin{pmatrix} A \\ \{e^i\}_{i\in Z} \end{pmatrix} x = \begin{pmatrix} b \\ 0^{|Z|} \end{pmatrix}$$
יש פתרון יחיד.
לכל פתרון: נקבל $i\in Z \Rightarrow x_i = 0$.

ולכל $i \notin Z$: $x_i = x_i^*$, משמע יש למערכת
$$\left(A^{B(1)} \cdots A^{B(n-|Z|)}\right)\tilde{x} = b$$
יש פתרון יחיד כאשר $\{B(1) \cdots B(n-|Z|)\} \triangleq \{1,\dots,n\} \setminus Z$
ולכן העמודות הן בת"ל, משמע $\mathrm{rank}(A) = m \geq n-|Z|$
$\Rightarrow n-m \leq |Z|$ (כנדרש בחלק ⓐ).
אם $n-m < |Z|$, אז נוסיף עמודות מ-A על חשבון
אינדקסים ב-Z. ∎

מינוח

עבור פתרון בסיסי x, הרכיבים $x_{B(1)} \cdots x_{B(m)}$
נקראים משתנים בסיסיים (או משתני בסיס).
שאר הרכיבים הם לא-בסיסיים.
העמודות $A^{B(1)} \cdots A^{B(m)}$ נקראות עמודות הבסיס,
ומהוות בסיס של $\mathbb{R}^m$.
תת-קבוצה של m עמודות בת"ל נקראת בסיס.
ונתייחס לבסיסים שמתקבלים מתת-קבוצות שונות כשונים.
המטריצה $m \times m$ המתקבלת מעמודות בסיס נקראת
מטריצת בסיס. נסמן $B = \left(A^{B(1)} \cdots A^{B(m)}\right)$ מטריצת בסיס.

בניית פתרונות בסיסיים

① מצא m עמודות בת"ל ב-A. נסמן $A^{B(1)} \cdots A^{B(m)}$.

② קבע $x_i^* = 0$ לכל $i \notin \{B(1), \dots, B(m)\}$.

③ פתור את השוויונים $\left(A^{B(1)} \cdots A^{B(m)}\right)\tilde{x} = b$

או באופן שקול:
$$\begin{pmatrix} A \\ \hline e^{\bar{B}(1)} \\ \vdots \\ e^{\bar{B}(n-m)} \end{pmatrix} x = \begin{pmatrix} b \\ 0^{n-m} \end{pmatrix}$$
אם פתרון בסיסי x מקיים $x \geq 0$, אז הוא גם פיזיבילי.

וכמו כן $x^B = \begin{pmatrix} x_{B(1)} \\ \vdots \\ x_{B(m)} \end{pmatrix}$

ואז B הפיכה ומקיימת: $B \cdot x^B = b$

ולכן $\boxed{x^B = B^{-1} \cdot b}$.

הערות

1. בסיסים שונים יכולים להשרות פתרונות בסיסיים שונים. לדוגמא, אם $b=0$, אז כל פתרון בסיסי $= 0^n$.

2. אם x^*, y^* פתרונות בסיסיים שכנים, אז אפשר למצוא בסיסים מתאימים B_{x^*}, B_{y^*} הנבדלים רק בעמודה אחת. (תזכרו להוכחה שמוצאת בסיס מתאים, בהינתן פתרון בסיסי!).

3. טענה: אם פוליהדרון הנתון בצורה סטנדרטית אינו ריק, אזי ניתן להפחית את מטריצת האילוצים כך שיתקבל תיאור של הפוליהדרון בצורה סטנדרטית ויתקיים: מס' שורות המטריצה = rank (מטריצת האילוצים)

הוכחה: $P = \{x \mid Ax=b, x \geq 0\}$.

A מטריצה $m \times n$ ונניח $k = rank(A) < m$

בה"כ: השורות $a_1, \dots, a_k$ בת"ל.

נגדיר $Q \triangleq \{x \mid \forall 1 \leq i \leq k: a_i x = b_i \ \& \ x \geq 0\}$

נראה ש $Q = P$:

כמובן $P \subseteq Q$. (למה?)

נראה כי $Q \subseteq P$. נבחר $y \in Q$ ונראה כי $y \in P$.

יהי $1+k \leq j \leq m$. כמובן a_j נפרש ע"י k השורות הראשונות:

$$a_j = \sum_{i=1}^{k} \lambda_i a_i$$

אם $x \in P$ אז:

$$b_j = a_j x = \sum_{i=1}^{k} \lambda_i a_i x = \sum \lambda_i b_i$$

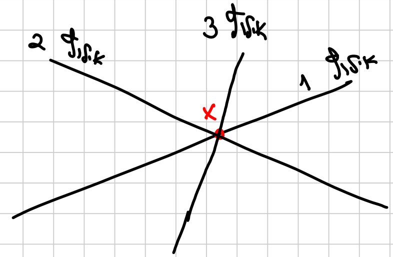
לכן

$$a_j \cdot y = \sum_{i=1}^{m} \lambda_i a_i y = \sum_{i=1}^{m} \lambda_i b_i = b_j$$

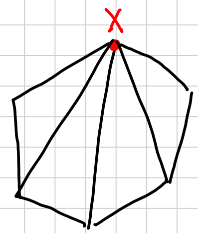
ומכאן $y \in P$, כנדרש.

נוון: פתרון בסיסי מנוון אם מס' האילוצים הפעילים ביחס ל-x גדול מ-n.

בדוגמה: $n=2$



$n=3$:



בצורה סטנדרטית: נוון $\Longleftrightarrow$ יותר מ $(n-m)$ רכיבים מתאפסים.

פתרון בסיסי נקרא מנוון אם מכיל יותר מ-$(n-m)$ אפסים.

נכון הוא תלוי בייצוג:

$$P=\{x \mid Ax=b,\ x\ge 0\}$$

$$P=\{x \mid Ax\ge b,\ Ax\le b,\ x\ge 0\}$$

האם לכל פוליהדרון יש נק' קיצון?

<u>הגדרה:</u> פוליהדרון P <u>מכיל ישר</u> אם קיים $x\in P$, וכן וקטור $d\in\mathbb{R}^n$ המקיימים:

$$\forall\lambda\in\mathbb{R}:\ x+\lambda d\in P$$

<u>משפט:</u> יהי $P=\{x\in\mathbb{R}^n : Ax\ge b\}$ פוליהדרון שאינו ריק. 3 התנאים הבאים שקולים:

1. P אינו מכיל ישר.
2. P מכיל נק' קיצון אחת לפחות.
3. $\operatorname{rank}(A)\ge n$.

<u>הוכחה:</u> $(2\Leftarrow 1)$ נראה שהעדר ישר ב-P מאפשר להפעיל פרוצדורה בעלת התכונה הבאה. בהינתן $x\in P$, אם

$$r(x)\triangleq \operatorname{rank}(\{a_i \mid a_i x=b_i\})<n$$

נמצא $\hat{x}\in P$ המקיים $r(\hat{x})>r(x)$.

בסוף נמצא באופן כזה $y\in P$ המקיים $r(y)=n$, ואותו y הוא פתרון בסיסי פיזיבילי, כנדרש.

שיטה: אם $r(x)<n$, נבחר

$$d\in\mathbb{R}^n\setminus\operatorname{span}(\{a_i \mid a_i x=b_i\})^{\perp}$$

ונסתכל על הישר $\{x+\lambda d\}_{\lambda\in\mathbb{R}}$. בסביבה של x, הישר מוכל ב-P, אבל עבור λ מספיק גדול, הישר מהדק לפחות עוד אילוץ בכדי לצאת מ-P. כלומר: קיים λ^* עבורו

$x+\lambda^* d$ נמצא על שפת P ומהדק אילוץ a_j שאינו הדוק ביחס ל-x.

כמובן $a_j\notin\operatorname{span}\{a_i \mid a_i x=b_i\}$, שכן אחרת

$$a_j(x+\lambda^* d)=\underbrace{a_j x}_{=b_j}+\lambda\cdot\underbrace{a_j d}_{=0}=b_j.$$

ולכן $r(x^*)>r(x)$, כנדרש.

$(3\Leftarrow 2)$ בנקודת קיצון, האילוצים ההדוקים מהווים בסיס, ולכן $\operatorname{rank}(A)=n$.

$(1\Leftarrow 3)$ נניח $a_1,\dots,a_n$ בת"ל ונניח בשלילה P מכיל את הישר $x+\lambda d$ עבור וקטור $d\neq 0$.

משפט: תהי $\min\{c^t x \mid x\in P\}$ תכנית לינארית ביחס לפוליהדרון P. אם P מכיל לפחות קודקוד אחד, והמינימום הנ"ל מוגדר היטב (גדול מ-$(-\infty)$), אזי המינימום מתקבל בקודקוד של P.

כלומר:

$$\left.\begin{array}{l} -\infty < \min\{c^t x : x\in P\} \\ \& \\ P \text{ מכיל קודקוד} \end{array}\right\} \Longrightarrow \text{קיים קודקוד } x^* \in P \text{ כך ש} \quad c^t x^* = \min\{c^t x \mid x\in P\}$$

<u>הוכחה</u>: יהי $\gamma \triangleq \min\{c^t x \mid x\in P\}$.

נגדיר $Q \triangleq \{x \mid x\in P \,\&\, c^t x = \gamma\}$.

P מכיל קודקוד $\Longrightarrow$ P לא מכיל ישר.

$Q \subseteq P \Longrightarrow Q$ לא מכיל ישר $\Longrightarrow Q$ מכיל קודקוד.

לפיכך: $\forall i\ \forall \lambda: \quad a_i(x+\lambda d) \geq b_i$

וזה אפשרי רק אם $\forall i: a_i d = 0$.

כלומר $d \in \text{span}(\{a_1 \cdots a_n\})^{\perp}$

אבל $a_1 \cdots a_n$ פורשים, ולכן $d=0$. סתירה. ☒

הרביע (החיובי) $\{x \mid x \geq 0\}$ לא מכיל ישר.

כל פוליהדרון בצורה הסטנדרטית מוכל ברביע החיובי.

לכן כל פוליהדרון בצורה סטנדרטית אינו מכיל ישר, לכן יש לו לפחות קודקוד אחד.

הטלות: אלימינציה בשיטת Fourier-Motzkin

הפונקציה $\Pi_K : \mathbb{R}^n \longrightarrow \mathbb{R}^K$

המוגדרת ע"י $\Pi_K(x_1,\ldots,x_n) \triangleq (x_1,\ldots,x_K)$

נקראת הטלה.

$$\Pi_K(S) \triangleq \{\Pi_K(x) \mid x\in S\}$$

<u>שאלה</u>: נתון פוליהדרון $P=\{x \mid Ax \geq b\}$.

האם $P \neq \emptyset$?

<u>רעיון לפתרון</u>: נמצא "ייצוג" של $\Pi_{n-1}(P)$.

ונענה על: האם $\Pi_{n-1}(P) \neq \emptyset$?

כמובן $P\neq\emptyset \Longleftrightarrow \Pi_{n-1}(P) \neq \emptyset$. ונמשיך ברקורסיה...

יהי x^* קודקוד של Q. נוכיח ש-x^* גם קודקוד של P.

אם x^* אינו קודקוד של P, אז הוא לא נק' קיצון.

משמע x^* בפנים הקטע $[y,z]$ שכולו ב-P.

$$c^t y,\ c^t z \geq \gamma \quad \& \quad c^t x = \gamma$$

$$\Longleftarrow \quad c^t y = c^t z = \gamma$$

$\Longleftarrow \quad y,z\in Q \quad \Longleftarrow \quad x$ לא נק' קיצון ב-Q. ☒

<u>מסקנה</u>: $\min\{c^t x \mid x\in P\} \in \mathbb{R}\cup\{-\infty\}$.

הערה: $\min\{\frac{1}{x} \mid x\geq 0\}$ אינו קיים!

ולכן המסקנה אינה לגמרי מאליה.

יתרה מזאת: $\pi_1(\pi_2(\pi_3(\cdots(\pi_{n-1}(P)\cdots) = \pi_1(P)$

לכן $P \neq \emptyset \iff \pi_1(P) \neq \emptyset$.

ועכשיו "קל" לבדוק האם $\pi_1(P)\neq\emptyset$ בהינתן ייצוג $\pi_1(P) = \{x \mid \tilde{A}x \geq \tilde{b}\}$.

איך מוצאים ייצוג $\pi_{n-1}(P) = \{x \mid A'x \geq b'\}$?

אלגוריתם האלימינציה

① נכתוב כל אי-שוויון $a_i x \geq b_i$ כאי-שוויון $a_{in}x_n \geq b_i - \bar{a}_i \bar{x}$

כאשר $\bar{x} \triangleq (x_1,\dots,x_{n-1})$, $\bar{a}_i = (a_{i1},\dots,a_{i,n-1})$

← אם $a_{in}>0$ נכתוב (*) $\boxed{x_n \geq \frac{1}{a_{in}}\cdot b_i - \bar{a}_i\bar{x}}$

$= d_i + f_i\bar{x}$, $d_i \triangleq \frac{b_i}{a_{in}} \in \mathbb{R}$, $f_i = \frac{-\bar{a}_i}{a_{in}} \in \mathbb{R}^{n-1}$

← אם $a_{in}<0$ נכתוב (**) $\boxed{d_i + f_i\bar{x} \geq x_n}$

← אם $a_{in}=0$ נכתוב (***) $\bar{a}_i\bar{x} \geq b_i$

הגדר פוליהדרון $Q \subseteq \mathbb{R}^{n-1}$ ע"י קב' האי-שוויונים

(***) $\bar{a}_i \bar{x} \geq b_i$

$\forall i \in (*)\ \forall j \in (**):\ d_j + f_j\bar{x} \geq d_i + f_i \bar{x}$

$\iff (f_j - f_i)\bar{x} \geq (d_i - d_j)$

טענה: $Q = \pi_{n-1}(P)$.

הערות: אם P מוגדר ע"י m אי-שוויונים, Q מוגדר ע"י לכל היותר $(\frac{m}{2})^2$ אי-שוויונים. ולכן זה אינו אלג' פולינומי.

מסקנה: אם P פוליהדרון, אזי $\pi_k(P)$ גם כן פוליהדרון.

תוצאה: אם $P \subseteq \mathbb{R}^{n+k}$ פוליהדרון, אזי

$$P' \triangleq \{x\in\mathbb{R}^n \mid \exists y \in \mathbb{R}^k : (x,y)\in P\}$$

הוא גם כן פוליהדרון.

תוצאה: אם $P\subseteq\mathbb{R}^n$ פוליהדרון, וכן $T:\mathbb{R}^n\to\mathbb{R}^k$ טרנספורמציה לינארית, אז $T(P)$ גם כן פוליהדרון.

הוכחה: $Q = \{(x,y) \mid x\in P,\ Tx=y\}$ הוא פוליהדרון, ו- $T(P)$ הוא ההטלה של Q.

תוצאה: הקמור של קבוצה סופית $X\subseteq\mathbb{R}^n$ הוא פוליהדרון.

הוכחה: $\text{conv-hull}(x_1 \dots x_k) = \left\{ \sum_{i=1}^{k} \lambda_i x_i \;\middle|\; \vec{\lambda} \geq 0,\ 1^t \vec{\lambda} = 1 \right\}$

אבל $P = \left\{ (\lambda_1, \dots, \lambda_k) \;\middle|\; \lambda_1 \dots \lambda_k \geq 0,\ \Sigma \lambda_i = 1 \right\}$

הוא פוליהדרון.

$T(\lambda_1 \dots \lambda_k) \triangleq \sum_{i=1}^{k} \lambda_i x_i$ היא פונק' לינארית.

ולכן $T(P)$ הוא פוליהדרון. אבל

$T(P) = \text{conv-hull}(x_1, \dots, x_k)$. $\boxtimes$