פתרון תרגיל מספר 11 ־ כלים מתמטיים

שם: מיכאל גרינבאום, **ת.ז:** 211747639

1 בינואר 2021

1. **פתרון:**

(א) **צ"ל:** $x \in \text{vertices}(\mathcal{P})$ אזי x הוא extremal

הוכחה:

תהי $x \in \text{vertices}(\mathcal{P})$, אזי קיים $c \in \mathbb{R}^n$ כך ש־ x הוא ממקסם יחיד של $c^T \cdot x$ ב־ $\mathcal{P}$.

יהי $0 \neq u \in \mathbb{R}^n$, נניח בשלילה ש־ $x + u \in \mathcal{P}$ וגם $x - u \in \mathcal{P}$.

i. אם $\langle c, u\rangle \geq 0$, נשים לב כי

$$\langle c, x+u\rangle = \langle c, x\rangle + \langle c, u\rangle \geq \langle c, x\rangle$$

וגם $x + u \in \mathcal{P}$, מהיות x הוא ממקסם אז מתקיים גם $\langle c, x+u\rangle \leq \langle c, x\rangle$ ולכן נקבל כי $\langle c, x+u\rangle = \langle c, x\rangle$, ונקבל סתירה ליחידות הממקסם של $c^T \cdot x$. (כי גם x וגם $x+u$ ממקסים אותו)

ii. אחרת, $\langle c, u\rangle < 0$, נשים לב כי

$$\langle c, x-u\rangle = \langle c, x\rangle - \langle c, u\rangle \geq \langle c, x\rangle$$

וגם $x - u \in \mathcal{P}$, מהיות x הוא ממקסם אז מתקיים גם $\langle c, x-u\rangle \leq \langle c, x\rangle$ ולכן נקבל כי $\langle c, x-u\rangle = \langle c, x\rangle$, ונקבל סתירה ליחידות הממקסם של $c^T \cdot x$. (כי גם x וגם $x-u$ ממקסים אותו)

כלומר ההנחה בשלילה לא נכונה, ולכן מתקיים $x + u \notin \mathcal{P}$ או ש־ $x - u \notin \mathcal{P}$,
כלומר x הוא extremal מההגדרה

מ.ש.ל.א.☺

(ב) **צ"ל:** $x \in \text{vertices}(\mathcal{P})$ אם"ם $\text{span}\{a_i \mid i \in I(x)\} = \mathbb{R}^n$

הוכחה:

$\Leftarrow$: נניח כי $x \in \text{vertices}(\mathcal{P})$,

נניח בשלילה ש־ $\text{span}\{a_i \mid i \in I(x)\} \neq \mathbb{R}^n$, אזי $\exists u \in \mathbb{R}^n \setminus \text{span}\{a_i \mid i \in I(x)\}$ וגם ש־ $\langle a_i, u\rangle = 0$ לכל $i \in I(x)$,

תחילה אם מתקיים $\max_{i \notin I(x)} |\langle a_i, u\rangle| = 0$, נשים לב כי $\langle a_i, u\rangle = 0$ לכל $1 \leq i \leq n$ ולכן

$$\langle a_i, x \pm \varepsilon \cdot u\rangle = \langle a_i, x\rangle \pm \varepsilon \cdot \langle a_i, u\rangle = \langle a_i, x\rangle + 0 \leq b_i$$

ולכן $x \pm \varepsilon \cdot u \in \mathcal{P}$ כי מקיימים את כל האילוצים,

אחרת נגדיר $\varepsilon = \frac{\min_{i \notin I(x)} (b_i - \langle a_i, x\rangle)}{\max_{i \notin I(x)} |\langle a_i, u\rangle|}$, נסתכל על $x \pm \varepsilon \cdot u$, נחלק למקרים:

i. אם $i \in I(x)$ נקבל כי $\langle a_i, u\rangle = 0$ ולכן מתקיים

$$\langle a_i, x \pm \varepsilon \cdot u\rangle = \langle a_i, x\rangle \pm \varepsilon \cdot \langle a_i, u\rangle = \langle a_i, x\rangle + 0 = b_i$$

ii. אחרת $i \notin I(x)$ נקבל כי

$$\begin{aligned}\langle a_i, x \pm \varepsilon \cdot u\rangle &= \langle a_i, x\rangle \pm \varepsilon \cdot \langle a_i, u\rangle = \langle a_i, x\rangle \pm \min_{i \notin I(x)} (b_i - \langle a_i, x\rangle) \cdot \frac{\langle a_i, u\rangle}{\max_{i \notin I(x)} |\langle a_i, u\rangle|} \\ &\leq \langle a_i, x\rangle \pm \min_{i \notin I(x)} (b_i - \langle a_i, x\rangle) \cdot 1 \leq \langle a_i, x\rangle \pm (b_i - \langle a_i, x\rangle) \\ &\overset{b_i \geq \langle a_i, x\rangle}{\leq} \langle a_i, x\rangle + (b_i - \langle a_i, x\rangle) = b_i\end{aligned}$$

כלומר קיבלנו כי $x \pm \varepsilon \cdot u$ מקיים את כל הדרישות של $\mathcal{P}$ ולכן $x \pm \varepsilon \cdot u \in \mathcal{P}$,

כלומר ב־ 2 המקרים מצאנו v כך ש־ $x + v \in \mathcal{P}$ וגם $x - v \in \mathcal{P}$,
מהסעיף הקודם נקבל סתירה ולכן מתקיים $\text{span}\{a_i \mid i \in I(x)\} = \mathbb{R}^n$, כנדרש.
$\Rightarrow$: נניח כי $\text{span}\{a_i \mid i \in I(x)\} = \mathbb{R}^n$,
נניח בשלילה ש־ $x \notin \text{vertices}(\mathcal{P})$, כלומר לכל c מתקיים ש־ x הוא לא ממקסם יחיד של $c^T \cdot x$ ב־ $\mathcal{P}$.
נבחר $c = \sum_{i \in I(x)} a_i$, נשים לב כי לכל $y \in \mathcal{P}$ מתקיים

$$c^T \cdot y = \left\langle \sum_{i \in I(x)} a_i, y \right\rangle = \sum_{i \in I(x)} \langle a_i, y\rangle \leq \sum_{i \in I(x)} \langle a_i, x\rangle = \left\langle \sum_{i \in I(x)} a_i, x \right\rangle = c^T \cdot x$$

כלומר $c^T \cdot y \leq c^T \cdot x$ לכל $y \in \mathcal{P}$,
מהטענה שלכל c מתקיים ש־ x הוא לא ממקסם יחיד של $c^T \cdot x$ ב־ $\mathcal{P}$ נסיק שקיים $y \in \mathcal{P}$ המקיים $c^T \cdot y = c^T \cdot x$ וגם $y \neq x$.
נשים לב כי מתקיים $\langle a_i, y\rangle = b_i$ לכל $i \in I(x)$ (בשביל שיתקיים שוויון בינהם).
מהיות $\{a_i \mid i \in I(x)\}$ פורש את $\mathbb{R}^n$, נוכל לכתוב

$$0 \neq y - x = \sum_{i \in I(x)} \langle a_i, y - x\rangle \cdot a_i$$

מהיות וזהו לא וקטור ה־ 0, $\exists j \in I(x)$ כך ש־ $\langle a_j, y - x\rangle \neq 0$,
נשים לב כי $0 = b_j - b_j = \langle a_j, y\rangle - \langle a_j, x\rangle = \langle a_j, y - x\rangle$ בסתירה לזה שקודם הראנו כי $\langle a_j, y - x\rangle \neq 0$.
לכן מתקיים כי $x \in \text{vertices}(\mathcal{P})$, כנדרש.

מ.ש.ל.ב.☺

(ג) **צ"ל:** x הוא extremal אזי $x \in \text{vertices}(\mathcal{P})$
הוכחה:
יהי x שהוא extremal, אזי לכל $u \neq 0$ מתקיים $x + u \notin \mathcal{P}$ או ש־ $x - u \notin \mathcal{P}$,
נניח בשלילה ש־ $\text{span}\{a_i \mid i \in I(x)\} \neq \mathbb{R}^n$, אזי $\exists u \in \mathbb{R}^n \backslash \text{span}\{a_i \mid i \in I(x)\}$ וגם ש־ $\langle a_i, u\rangle = 0$ לכל $i \in I(x)$,
תחילה אם מתקיים $\max_{i \notin I(x)} |\langle a_i, u\rangle| = 0$, נשים לב כי $\langle a_i, u\rangle = 0$ לכל $1 \leq i \leq n$ ולכן

$$\langle a_i, x \pm \varepsilon \cdot u\rangle = \langle a_i, x\rangle \pm \varepsilon \cdot \langle a_i, u\rangle = \langle a_i, x\rangle + 0 \leq b_i$$

ולכן $x \pm \varepsilon \cdot u \in \mathcal{P}$ כי מקיימים את כל האילוצים,
אחרת נגדיר $\varepsilon = \frac{\min_{i \notin I(x)} (b_i - \langle a_i, x\rangle)}{\max_{i \notin I(x)} |\langle a_i, u\rangle|}$, נסתכל על $x \pm \varepsilon \cdot u$, נחלק למקרים:

i. אם $i \in I(x)$ נקבל כי $\langle a_i, u\rangle = 0$ ולכן מתקיים

$$\langle a_i, x \pm \varepsilon \cdot u\rangle = \langle a_i, x\rangle \pm \varepsilon \cdot \langle a_i, u\rangle = \langle a_i, x\rangle + 0 = b_i$$

ii. אחרת $i \notin I(x)$ נקבל כי

$$\begin{aligned}\langle a_i, x \pm \varepsilon \cdot u\rangle &= \langle a_i, x\rangle \pm \varepsilon \cdot \langle a_i, u\rangle = \langle a_i, x\rangle \pm \min_{i \notin I(x)} (b_i - \langle a_i, x\rangle) \cdot \frac{\langle a_i, u\rangle}{\max_{i \notin I(x)} |\langle a_i, u\rangle|} \\ &\leq \langle a_i, x\rangle \pm \min_{i \notin I(x)} (b_i - \langle a_i, x\rangle) \cdot 1 \leq \langle a_i, x\rangle \pm (b_i - \langle a_i, x\rangle) \\ &\overset{b_i \geq \langle a_i, x\rangle}{\leq} \langle a_i, x\rangle + (b_i - \langle a_i, x\rangle) = b_i\end{aligned}$$

כלומר קיבלנו כי $x \pm \varepsilon \cdot u$ מקיים את כל הדרישות של $\mathcal{P}$ ולכן $x \pm \varepsilon \cdot u \in \mathcal{P}$ בסתירה לנתון.

כלומר ב־ 2 המקרים מצאנו v כך ש־ $x + v \in \mathcal{P}$ וגם $x - v \in \mathcal{P}$,
נקבל סתירה מהנתון ולכן מתקיים $\text{span}\left\{a_i \mid i \in I\left(x\right)\right\} = \mathbb{R}^n$, כנדרש.
עתה, מהסעיף הקודם נקבל כי $x \in \text{vertices}\left(\mathcal{P}\right)$, כנדרש.

מ.ש.ל.ג.☺

2. **צ"ל:** $\text{dimVC}\left(\mathcal{H}\right) = d + 1$

הוכחה:

תחילה נראה כי $\text{dimVC}\left(\mathcal{H}\right) \geq d + 1$,
נסתכל על $S = \{e_1, \ldots, e_n, e_1 + e_2\}$, יהיו $\varepsilon = \left(\varepsilon_1, \ldots, \varepsilon_n, \varepsilon_{n+1}\right)$ לייבלים,
נבחר $b = \varepsilon_1 + \varepsilon_2 - \varepsilon_{n+1}$, $w = \left(\varepsilon_1 - b, \varepsilon_2 - b, \varepsilon_3 - b, \ldots, \varepsilon_n - b\right)$.
נשים לב כי

$$h_{w,b}\left(e_i\right) = \text{sign}\left(\varepsilon_i - b + b\right) = \text{sign}\left(\varepsilon_i\right) = \varepsilon_i$$

וגם כי

$$\begin{aligned} h_{w,b}\left(e_1 + e_2\right) &= \text{sign}\left(\left(\varepsilon_1 - b\right) + \left(\varepsilon_2 - b\right) + b\right) = \text{sign}\left(\varepsilon_1 + \varepsilon_2 - b\right) \\ &= \text{sign}\left(\varepsilon_1 + \varepsilon_2 - \left(\varepsilon_1 + \varepsilon_2 - \varepsilon_{n+1}\right)\right) = \text{sign}\left(\varepsilon_{n+1}\right) = \varepsilon_{n+1} \end{aligned}$$

כלומר הראנו שהקבוצה S מנתצת את $\mathcal{H}$ (כי לכל ε לייבלים מצאנו $h \in \mathcal{H}$ כך שמתקיים $h\left(x_i\right) = \varepsilon_i$)
ולכן $\text{dimVC}\left(\mathcal{H}\right) \geq d + 1$,
עתה נניח בשלילה ש־ $\text{dimVC}\left(\mathcal{H}\right) > d + 1$, לכן קיימת $S \subseteq \mathbb{R}^d$ כך ש־ $|S| = d + 2$ והיא מנתצת את $\mathcal{H}$ (אם יש קבוצה נתיצה בגודל גדול מ־ $d + 2$ אז יש אחת בגודל $d + 2$).
נסמן $S = \{x_1, \ldots, x_{d+2}\}$, כמו בהוכחה של Radon's theorem, קיימים $\gamma_1, \ldots, \gamma_{d+2}$ כך שלא כולם אפסים וגם $\sum_{i=1}^{d+2} \gamma_i = 0$ וגם $\sum_{i=1}^{d+2} \gamma_i \cdot x_i = 0$.
נסמן ב־ j איבר γ_j שאינו 0, אזי נקבל $x_j = \frac{\left(-\sum_{i \neq j} \gamma_i \cdot x_i\right)}{\gamma_j}$ וגם $\gamma_j = \sum_{i \neq j} -\gamma_i$
מהיות S מנתצת את $\mathcal{H}$, לכל וקטור לייבלים ε קיימת פונקציה $h \in \mathcal{H}$ כך ש־ $h\left(x_i\right) = \varepsilon_i$.
נסתכל על $\varepsilon_i = \begin{cases} \text{sign}\left(\gamma_i\right) & i \neq j \\ -\text{sign}\left(\gamma_j\right) & i = j \end{cases}$ לכל $1 \leq i \leq d + 2$,
מההנחה קיימת $h_{w,b} \in \mathcal{H}$ כך ש־ $h_{w,b}\left(x_i\right) = \varepsilon_i$,
נשים לב כי

$$\begin{aligned} h_{w,b}\left(x_j\right) &= \text{sign}\left(w^T \cdot x_j + b\right) = \text{sign}\left(w^T \cdot \frac{\left(-\sum_{i \neq j} \gamma_i \cdot x_i\right)}{\gamma_j} + b\right) = \text{sign}\left(\sum_{i \neq j} \frac{-\gamma_i}{\gamma_j} \cdot w^T \cdot x_i + b\right) \\ &= \text{sign}\left(\sum_{i \neq j} \frac{-\gamma_i}{\gamma_j} \cdot \left(w^T \cdot x_i + b - b\right) + b\right) = \text{sign}\left(\sum_{i \neq j} \frac{-\gamma_i}{\gamma_j} \cdot \left(w^T \cdot x_i + b\right) + \frac{\sum_{i \neq j} -\gamma_i}{\gamma_j} \cdot (-b) + b\right) \\ &= \text{sign}\left(\sum_{i \neq j} \frac{-\gamma_i}{\gamma_j} \cdot \left(w^T \cdot x_i + b\right) + \frac{\gamma_j}{\gamma_j} \cdot (-b) + b\right) = \text{sign}\left(\sum_{i \neq j} \frac{-\gamma_i}{\gamma_j} \cdot \left(w^T \cdot x_i + b\right)\right) \\ &\overset{\text{sign}\left(w^T \cdot x_i + b\right) = \text{sign}\left(\gamma_i\right)}{=} \text{sign}\left(\sum_{i \neq j} \frac{-1}{\gamma_j} \cdot \left|\gamma_i \cdot w^T \cdot x_i + b\right|\right) = \text{sign}\left(\sum_{i \neq j} \frac{-1}{\gamma_j}\right) \\ &= -\text{sign}\left(\sum_{i \neq j} \frac{1}{\gamma_j}\right) = \text{sign}\left(\frac{d+1}{\gamma_j}\right) = \text{sign}\left(\gamma_j\right) \end{aligned}$$

קיבלנו כי $\text{sign}\left(\gamma_j\right) = h_{w,b}\left(x_j\right) = \varepsilon_j = -\text{sign}\left(\gamma_j\right)$, סתירה ולכן אין קבוצה נתיצה בגודל $d + 2$,
ולכן $\text{dimVC}\left(\mathcal{H}\right) \leq d + 1$ וגם הראנו כי $\text{dimVC}\left(\mathcal{H}\right) \geq d + 1$ ולכן $\boxed{\text{dimVC}\left(\mathcal{H}\right) = d + 1}$, כנדרש

מ.ש.ל.☺

3. **פתרון:**

(א) **צ"ל:** להסביר מדוע זה מספיק
הוכחה:
נוכל להוכיח באינדוקציה לכל $m \geq d+2$ יש חיתוך לא ריק כשהבסיס יהיה $d+2$ שנובע מ־ Radon's theorem שהוכח בתרגול, והצעד נובע בדיוק מהטענה שאנחנו רוצים להוכיח.

מ.ש.ל.א.☺

(ב) **צ"ל:** $\bigcap_{i=1}^{m} A_i \neq \emptyset$
הוכחה:
נסמן ב־ $x_j \in \bigcap_{i\in[m],i\neq j} A_i$ (קיים מהנחה שהחיתוך לא ריק),
נגדיר $S = \{x_1, \ldots, x_m\}$,
מהיות $m \geq d+2$, מ־ Radon's theorem שהוכח בתרגול קיימת תת קבוצה T כך ש־ $\text{conv}(T) \cap \text{conv}(S \setminus T) \neq \emptyset$.
נסמן $y \in \text{conv}(T) \cap \text{conv}(S \setminus T)$
נראה כי $y \in \bigcap_{i=1}^{m} A_i$,
יהי $1 \leq j \leq m$, נשים לב כי $\{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m\} \subseteq A_j$,
ולכן מהיות A_j קמור מתקיים כי $\text{conv}(\{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m\}) \subseteq A_j$.
עתה נחלק למקרים:

i. אם $x_j \notin T$ אזי $T \subseteq S \setminus \{x_j\} = \{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m\} \subseteq A_j$,
ולכן נקבל כי $y \in \text{conv}(T) \subseteq \text{conv}(\{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m\}) \subseteq A_j$
ii. אם $x_j \in T$ אזי $S \setminus T \subseteq S \setminus \{x_j\} = \{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m\} \subseteq A_j$,
ולכן נקבל כי $y \in \text{conv}(S \setminus T) \subseteq \text{conv}(\{x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_m\}) \subseteq A_j$

כלומר הראנו שלכל $1 \leq j \leq m$ מתקיים $y \in A_j$ ולכן $y \in \bigcap_{i=1}^{m} A_i$, כנדרש.

מ.ש.ל.ב.☺

4. **פתרון:**

(א) **צ"ל:** $m^* \leq m$
הוכחה:
נסמן את הפתרון של האופטימלי של הבעיה בשלמים ב־ $w : E \to \{0,1\}$ ואת הפתרון האופטימלי של הבעיה בשברים ב־ $w' : E \to [0,1]$.
נניח בשלילה ש־ $m^* > m$,
נגדיר $w''(e) = w(e)$ לכל $e \in E$.
נשים לב שזה פתרון חוקי וגם כי

$$\sum_{e\in E} w''(e) = \sum_{e\in E} w(e) = m < m^* = \sum_{e\in E} w'(e)$$

נשים לב שקיבלנו סתירה לאופטימליות w' כי מצאנו פתרון חוקי w'' שהוא יותר טוב.
ולכן אכן מתקיים $m^* \leq m$.

מ.ש.ל.א.☺

(ב) **צ"ל:** צריך להוריד לכל היותר $3m^*$ צלעות כדי למחוק את כל המשולשים
הוכחה:
נסמן את הפתרון של האופטימלי של הבעיה בשברים ב־ $w' : E \to [0,1]$.
נגדיר $w : E \to \{0,1\}$ באופן הבא: $w(e) = \begin{cases} 1 & w'(e) \geq \frac{1}{3} \\ 0 & else \end{cases}$ לכל $e \in E$.
תחילה נראה שנשאר עם גרף ללא משולשים.
יהי $a, b, c \in V$ משולש, אזי $w'(ab) + w'(bc) + w'(ac) \geq 1$,
לכן מתקיים כי $\max_{e\in\{ab,bc,ac\}} w'(e) \geq \frac{1}{3}$, נסמן את הצלע המקסימלית ב־ e' ונשים לב כי $w(e') = 1$, כלומר מחקנו

צלע מהמשולש.
כלומר אכן מחקנו לפחות צלע אחת בכל משולש ולכן נשאר בגרף ללא משולשים.
נשים לב כי

$$w\left(e\right)=\begin{cases}1 & w'\left(e\right)\geq\frac{1}{3}\\0 & else\end{cases}\leq\begin{cases}1 & w'\left(e\right)\geq\frac{1}{3}\\3\cdot w'\left(e\right) & else\end{cases}\leq\begin{cases}3\cdot w'\left(e\right) & w'\left(e\right)\geq\frac{1}{3}\\3\cdot w'\left(e\right) & else\end{cases}=3\cdot w'\left(e\right)$$

לכל $e\in E$.
ולכן מתקיים

$$\sum_{e\in E}w\left(e\right)\leq\sum_{e\in E}3\cdot w'\left(e\right)=3\cdot\sum_{e\in E}w'\left(e\right)=3\cdot m^{*}$$

כלומר הפתרון החוקי שמצאנו מוריד פחות מ־ $3m^{*}$ צלעות, כנדרש

מ.ש.ל.ב.☺