20.12.2020		***תכנון אלגוריתמים***
	# הרצאה 18: תכנון לינארי	
Source: Lecture notes by Eden Chlamtáč		מרצה: אורי שטמר

תזכורת – **<u>בעית כיסוי בצמתים במשקל מינימום בגרף דו-צדדי:</u>**

<u>מופע:</u> גרף דו-צדדי $G = (U, V, E)$ ופונק' משקל על הצמתים $w: (U \cup V) \to \mathbb{R}^+$

<u>פתרון חוקי:</u> כיסוי בצמתים. כלומר, קבוצת צמתים $S \subseteq (U \cup V)$ כך שלכל קשת $(u, v) \in E$ מתקיים $u \in S$ או $v \in S$ (או שניהם).

<u>יש למצוא:</u> כיסוי בצמתים S במשקל $w(S) = \sum_{u \in S} w(u)$ קטן ביותר

ראינו שניתן לנסח בעיה זו כתוכנית לנארית בשלמים:

תוכנית *A*			
min	$\sum_{u \in (U \cup V)} w(u) \cdot x_u$		(1)
$s.t.$	$x_u + x_v \geq 1$	$\forall\ (u, v) \in E$	(2)
	$x_u \geq 0$	$\forall\ u \in (U \cup V)$	(3)
	$-x_u \geq -1$	$\forall\ u \in (U \cup V)$	(4)
	$x_u \in \mathbb{Z}$	$\forall\ u \in (U \cup V)$	(5)

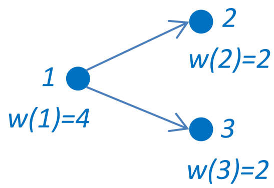
ראינו: עבור קלט מסויים (G, w) משקל פתרון אופטימלי לבעיה שלנו שווה בדיוק לערך האופטימלי של תוכנית A.

נסמן ב- OPT את <u>ערך</u> פתרון אופטימלי של תוכנית A. אז OPT הוא <u>משקל</u> כיסוי מינימלי בגרף G.

מכיוון שלא ידוע (ולא מאמינים שקיים) אלגוריתם יעיל שפותר תוכניות לנאריות בשלמים באופן כללי, התסכלנו על תוכנית לינארית רגילה (עם משתנים רציפים) דומה:

תוכנית *B*			
min	$\sum_{u \in (U \cup V)} w(u) \cdot z_u$		(6)
$s.t.$	$z_u + z_v \geq 1$	$\forall\ (u, v) \in E$	(7)
	$z_u \geq 0$	$\forall\ u \in (U \cup V)$	(8)
	$-z_u \geq -1$	$\forall\ u \in (U \cup V)$	(9)

זאת אכן תוכנית לינארית, אך היא לא מיצגת את בעיית כיסוי צמתים ממושקלת, מכיוון שהיא מאפשרת גם ערכים שבריים למשתני z_u.



למשל, בדוגמה שראינו

$$z_1 = \frac{2}{3}, \qquad z_2 = z_3 = \frac{1}{3}$$

מקיים את כל האילוצים ומביא למינימום את פונק' המטרה.
אבל, בניגוד לפתרונות שלמים, פתרון כזה לא מייצג כיסוי צמתים, שכן ההשמה $z_2 = \frac{1}{3}$ (לדוגמה) לא אומרת לנו האם יש להוסיף את צומת 2 לכיסויי או לא. באופן כללי (בגרפים דו-צדדיים כלליים), למרות השוני בין התוכניות, נראה באופן מפתיע שהן מחזירות את אותה תוצאה.

נסמן ב- $FRAC$ את <u>ערך</u> פתרון אופטימלי של תוכנית B

מה יכול להיות היחס בין $FRAC$ ל- OPT ?

כמו שאמרנו, די ברור שתמיד $FRAC \leq OPT$ כי כל פתרון חוקי לתוכנית *A* הוא גם פתרון חוקי לתוכנית *B* (כי רק מחקנו אילוצים) ולכן בפרט פתרון אופטימלי לתוכנית *A* הוא פתרון חוקי לתוכנית *B*. אבל אולי יכול להיות שלתוכנית *B* קיים פתרון אחר שמשיג ערך נמוך יותר?

<u>**משפט:**</u> לכל מופע $G = (U, V, E), w$ לבעיית כיסוי בצמתים במשקל מינימום בגרף דו-צדדי מתקיים כי

$$OPT = FRAC$$

<u>**מסקנה:**</u> ניתן לחשב את OPT בזמן פולינומי

לצורך הוכחת המשפט, נצטרך את 2 הדברים הבאים:

<u>**אבחנה 1:**</u> $FRAC \leq OPT$

<u>**טענה 2:**</u> קיים אלגוריתם אקראי שמקבל פתרון כלשהו $(\hat{z}_u)_{u \in U \cup V}$ המקיים את אילוצים $(9) - (7)$ ומחזיר וקטור ערכים (אקראיים) $(x_u^*)_{u \in U \cup V}$ כך שמתקיים:

1. הערכים $(x_u^*)_{u \in U \cup V}$ תמיד מקיימים את האילוצים $(5) - (2)$. כלומר האלגוריתם תמיד מחזיר השמה של ערכים שלמים המייצגים כיסוי צמתים חוקי
2. תוחלת משקל הכיסוי מקיימת

$$\mathbb{E}\left[\sum_{u \in U \cup V} w(u) \cdot x_u^*\right] = \sum_{u \in U \cup V} w(u) \cdot \hat{z}_u \qquad (10)$$

לפני שנוכיח את טענת עזר 2, ניזכר במשפט חשוב מאוד שלמדתם בהסתברות:

<u>**משפט (לינאריות התוחלת):**</u> לכל אוסף של משתנים אקראיים $X_1, \dots, X_n$ מתקיים:

$$\mathbb{E}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbb{E}[X_i]$$

(במידת הצורך, בסוף ההרצאה מצורפת תזכורת קצרה להוכחת משפט ליניאריות התוחלת)

<u>הוכחת טענה 2:</u>
יהי $(\hat{z}_u)$ פתרון חוקי לאילוצים $(9) - (7)$. נתבונן באלגוריתם הבא, שמקבל כקלט את ההשמה $(\hat{z}_u)$.

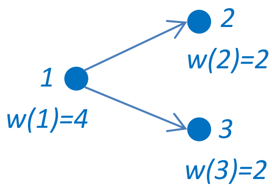
- נגריל האופן אחיד מספר ממשי $r \in [0,1]$
- לכל $u \in U$ בצע:

$$x_u^* \leftarrow \begin{cases} 1 & , \quad r \le \hat{z}_u \\ 0 & , \quad \text{אחרת} \end{cases}$$

- לכל $v \in V$ בצע:

$$x_v^* \leftarrow \begin{cases} 1 & , \quad r \ge 1 - \hat{z}_v \\ 0 & , \quad \text{אחרת} \end{cases}$$

- החזר את ההשמה $(x_u^*)_{u \in U \cup V}$.



דוגמה: עבור הקלט משמאל ועבור ההשמה $\hat{z}_1 = \frac{2}{3}, \quad \hat{z}_2 = \hat{z}_3 = \frac{1}{2}$
האלגוריתם הנ"ל יבחר השמה אקראית (x^*) באופן הבא:

ערך ההשמה	x^*	הסתברות	טווח של r
4	$x_1^* = 1, \quad x_2^* = x_3^* = 0$	$\frac{2}{3}$	$0 \le r < \frac{2}{3}$
8	$x_1^* = x_2^* = x_3^* = 1$	0	$r = \frac{2}{3}$
4	$x_1^* = 0, \quad x_2^* = x_3^* = 1$	$\frac{1}{3}$	$\frac{2}{3} < r \le 1$

צריך להוכיח: ההשמה האקראית (x^*) תמית מקיימת את האילוצים $(5) - (2)$ ותוחלת משקל הכיסוי (המקרי) שהיא מחזירה מקיימת את (10).

נשים לב שאכן בדוגמה הנ"ל שני התנאים מתקיימים – כל ההשמות הנבחרות חוקיות ותוחלת ערך הפתרון היא

$$\frac{2}{3} \cdot 4 + 0 \cdot 8 + \frac{1}{3} \cdot 4 = 4$$

שזהו גם ערך הפתרון השברי $(\hat{z})$.

נוכיח את שני התנאים בנפרד.

<u>קיום האילוצים:</u> מהגדרת האלגוריתם, באופן מיידי ברור שההשמה מקיימת את אילוצים $(3)-(5)$, כלומר כל הערכים הם 0 או 1. נתמקד באילוץ (2). תהי (u,v) קשת כלשהי בגרף. בה"כ מתקיים $u \in U$ ו- $v \in V$. נניח בשלילה שהאילוץ $x_u^* + x_v^* \geq 1$ לא מתקיים. מכיוון ששני הערכים הם 0 או 1, נובע מכך ש- $x_u^* = x_v^* = 0$. לפי הגדרת האלגוריתם, ניתן להסיק ש-

$$\hat{z}_u < r < 1 - \hat{z}_v$$

או בהעברת אגפים,

$$\hat{z}_u + \hat{z}_v < 1$$

בסתירה להנחה שההשמה $(\hat{z}_u)_{u \in U \cup V}$ מקיימת את אילוץ (7).

<u>תוחלת ערך הפתרון:</u> נשתמש במשפט לינאריות התוחלת עבור המשתנים המקריים $\{x_u^* : u \in U \cup V\}$. כדי להשתמש בלינאריות התוחלת, נראה את הטענה הבאה:

<u>טענת עזר 3:</u> לכל צומת $u \in U \cup V$ מתקיים $\mathbb{E}[x_u^*] = \hat{z}_u$

בעזרת טענת עזר 3, ומתוך לינאריות התוחלת, נובע כי

$$\mathbb{E}\left[\sum_{u \in U \cup V} w(u) \cdot x_u^*\right] \underbrace{=}_{\text{לינאריות התוחלת}} \sum_{u \in U \cup V} \mathbb{E}[w(u) \cdot x_u^*] \underbrace{=}_{w(u)\ \text{קבוע}} \sum_{u \in U \cup V} w(u) \cdot \mathbb{E}[x_u^*] \underbrace{=}_{\text{ט.ע. 3}} \sum_{u \in U \cup V} w(u) \cdot \hat{z}_u$$

וזה מה שרצינו להוכיח (תנאי (10)).

נותר להוכיח את טענת עזר 3.

נחלק לשני מקרים: צמתים ב- U וצמתים ב- V.

עבור צומת $u \in U$, לפי הגדרת האלגוריתם מתקיים $x_u^* = 1$ אם"ם $r \in [0, \hat{z}_u]$. מכיוון ש- r מתפלג אחיד בקטע $[0,1]$, זה קורה בהסתברות $\hat{z}_u$. ביתר ההסתברות מתקיים $x_u^* = 0$. לכן אנו מקבלים

$$\mathbb{E}[x_u^*] = \Pr[x_u^* = 0] \cdot 0 + \Pr[x_u^* = 1] \cdot 1 = \Pr[x_u^* = 1] = \hat{z}_u$$

כעת ניקח צומת $v \in V$. לפי הגדרת האלגוריתם מתקיים $x_v^* = 1$ אם"ם $r \in [1 - \hat{z}_v, 1]$. מכיוון ש- r מתפלג אחיד בקטע $[0,1]$, ההסתברות שזה קורה שווה לאורך הקטע $[1 - \hat{z}_v, 1]$, שהוא $\hat{z}_v$. כמו במקרה הקודם אנו מקבלים ש-

$$\mathbb{E}[x_v^*] = \Pr[x_v^* = 1] = \hat{z}_v$$

זה מסיים את הוכחת טענת עזר 3 ולכן מסיים את הוכחת טענה 2.

<u>הערה:</u> האלגוריתם שהשתמשנו בו בהוכחה של טענה 2 נקרא "אלגוריתם עיגול". זוהי טכניקה נפוצה להשוות בין פתרון לתוכנית לינארית רגילה לתוכנית לינארית בשלמים, על ידי לקיחת פתרון שברי והמרתו לפתרון שלם. בתרגול תראו דוגמה נוספת לאלגוריתם עיגול.

<u>חומר למחשבה:</u> למעשה, מתוך ההוכחה של טענה 2 ניתן להסיק מסקנה חזקה יותר: ההשמה המקרית (x_u^*) היא פתרון אופטימלי לתוכנית A בהסתברות 1. למה?

<u>הוכחת המשפט</u>

לאור אבחנה 1, מספיק להראות כי $OPT \le FRAC$. יהי $(\hat{z}_u)$ פתרון אופטימלי לתוכנית B ויהיו (x_u^*) משתנים מקריים כמו בטענה 2. נזכור שתוחלת היא ממוצע משוקלל ולכן קיימת השמה (ספציפית, לא אקראית) כלשהי $(\hat{x}_u)$ שהמשתנים המקריים יקבלו בהסתברות חיובית, כך ש-

$$\sum_{u\in U\cup V} w(u)\cdot \hat{x}_u \le \mathbb{E}\left[\sum_{u\in U\cup V} w(u)\cdot x_u^*\right]$$

לפי טענה 2, זוהי השמה חוקית המקיימת את האילוצים $(5)-(2)$ ולכן הערך האופטימלי OPT של תוכנית A יכול להיות רק קטן או שווה לערך של אותה השמה. לכן קיבלנו

$$OPT \le \sum_{u\in U\cup V} w(u)\cdot \hat{x}_u \le \mathbb{E}\left[\sum_{u\in U\cup V} w(u)\cdot x_u^*\right] = \sum_{u\in U\cup V} w(u)\cdot \hat{z}_u = FRAC$$

כאשר השוויון האחרון נובע מהאופטימליות של הפתרון $(\hat{z}_u)$ והשיוויון הקודם לו נובע מטענה 2.

מ.ש.ל.

<u>דואליות של תוכניות לינאריות – הגדרה</u>

<u>הגדרה:</u> לכל אוסף מקדמים $\{c_j : j=1,\dots,n\}$, $\{b_i : i=1,2,\dots,m\}$, $\left\{a_{ij} : \begin{matrix} i=1,2,\dots,m \\ j=1,2,\dots,n \end{matrix}\right\}$,
שתי התוכניות הלינאריות הבאות נקראות דואליות אחת לשניה (כל אחת מהתוכניות היא דואלית לשניה):

MAX	MIN
$\max \quad \sum_{j=1}^{n} c_j\cdot x_j$	$\min \quad \sum_{i=1}^{m} b_i\cdot y_i$
$\text{s.t.} \quad \sum_{j=1}^{n} a_{ij}\cdot x_j \le b_i \qquad \forall\, i=1,\dots,m$	$\text{s.t.} \quad \sum_{i=1}^{m} a_{ij}\cdot y_i \ge c_j \qquad \forall\, j=1,\dots,n$
$x_j \ge 0 \qquad \forall\, j=1,\dots,n$	$y_i \ge 0 \qquad \forall\, i=1,\dots,m$

<u>הערות:</u>

1. התוכנית הדואלית לתוכנית מקסימיזציה היא תוכנית מינימיזציה, ולהפך.
2. אם בתוכנית המקורית יש n משתנים אז בתוכנית הדואלית יש n אילוצים (לא כולל אילוצי סימן). אם בתוכנית המקורית יש m אילוצים (לא כולל אילוצי סימן) אז בתוכנית הדואלית יש m משתנים.
3. המקדמים של פונקציית המטרה בתוכנית הדואלית הם הקבועים בצד ימין של האילוצים בתוכנית המקורית, ולהפך.
4. בתוכנית הדואלית, המקדם של המשתנה ה- i באילוץ ה- j הוא בדיוק המקדם של המשתנה ה- j באילוץ ה- i בתוכנית המקורית. כלומר מטריצת המקדמים משתחלפת.

<u>דוגמה מספרית:</u> נזכר באחת התוכניות הלינאריות שראינו בהרצאה הקודמת. נקרא לה תוכנית C:

$$\begin{array}{ll} \max & 10x_1 + 20x_2 \\ \mathrm{s.t.} & x_2 \le 3 \\ & 20x_1 + 30x_2 \le 100 \\ & x_1, x_2 \ge 0 \end{array}$$

מהי התוכנית הדואלית לתוכנית C?

נכתוב את תוכנית C בצורה השקולה הבאה:

$$\begin{array}{ll} \max & \mathbf{10} \cdot x_1 + \mathbf{20} \cdot x_2 \\ \mathrm{s.t.} & 0 \cdot x_1 + 1 \cdot x_2 \le 3 \\ & 20 \cdot x_1 + 30 \cdot x_2 \le 100 \\ & x_1, x_2 \ge 0 \end{array}$$

שימו לב, כאן המקדמים של האילוצים הם:

$$\begin{array}{ll} a_{11} = 0, & a_{12} = 1 \\ a_{21} = 20, & a_{22} = 30 \end{array}$$

אזי, לפי הגדרה, התוכנית הדואלית, שנקרא לה תוכנית D, היא:

$$\begin{array}{ll} \min & 3 \cdot y_1 + 100 \cdot y_2 \\ \mathrm{s.t.} & 0 \cdot y_1 + 20 \cdot y_2 \ge \mathbf{10} \\ & 1 \cdot y_1 + 30 \cdot y_2 \ge \mathbf{20} \\ & y_1, y_2 \ge 0 \end{array}$$

<u>עוד דוגמה:</u> מה תהייה התוכנית הדואלית לתוכנית הבאה:

$$\begin{array}{ll} \min & 3 \cdot y_1 + 100 \cdot y_2 + 8 \cdot y_3 + (-1) \cdot y_4 \\ \mathrm{s.t.} & 0 \cdot y_1 + 20 \cdot y_2 + (-3) \cdot y_3 + 1 \cdot y_4 \ge \mathbf{10} \\ & 1 \cdot y_1 + 15 \cdot y_2 + 4 \cdot y_3 + (-2) \cdot y_4 \ge \mathbf{20} \\ & 0 \cdot y_1 + (-10) \cdot y_2 + 0 \cdot y_3 + 7 \cdot y_4 \ge (\mathbf{-4}) \\ & y_1, y_2, y_3, y_4 \ge 0 \end{array}$$

נשים לב, כאן מטריצת המקדמים של האילוצים היא:

$$\begin{array}{cccc} 0 & 20 & (-3) & 1 \\ 1 & 15 & 4 & (-2) \\ 0 & (-10) & 0 & 7 \end{array}$$

אזי, לפי הגדרה, התוכנית הדואלית היא:

$$\begin{aligned} \max\quad & 10\cdot x_1 + 20\cdot x_2 + (-4)\cdot x_3 \\ \text{s.t.}\quad & 0\cdot x_1 + 1\cdot x_2 + 0\cdot x_3 \le 3 \\ & 20\cdot x_1 + 15\cdot x_2 + (-10)\cdot x_3 \le 100 \\ & (-3)\cdot x_1 + 4\cdot x_2 + 0\cdot x_3 \le 8 \\ & 1\cdot x_1 + (-2)\cdot x_2 + 7\cdot x_3 \le (-1) \\ & x_1, x_2, x_3 \ge 0 \end{aligned}$$

כלומר כאן מטריצת המקדמים של האילוצים היא

$$\begin{matrix} 0 & 1 & 0 \\ 20 & 15 & (-10) \\ (-3) & 4 & 0 \\ 1 & (-2) & 7 \end{matrix}$$

משפטי דואליות

נסתכל על תוכניות לינראיות דואליות:

MAX	MIN
$\max \quad \sum_{j=1}^{n} c_j \cdot x_j$ $\text{s.t.} \quad \sum_{j=1}^{n} a_{ij}\cdot x_j \le b_i \qquad \forall\, i = 1,\dots,m$ $x_j \ge 0 \qquad \forall\, j = 1,\dots,n$	$\min \quad \sum_{i=1}^{m} b_i \cdot y_i$ $\text{s.t.} \quad \sum_{i=1}^{m} a_{ij}\cdot y_i \ge c_j \qquad \forall\, j = 1,\dots,n$ $y_i \ge 0 \qquad \forall\, i = 1,\dots,m$

נסמן את הערך האופטימלי של התוכנית הימנית כ- *MIN* ואת זה של התוכנית השמאלית כ- *MAX*.

<u>משפט (דואליות חזקה):</u> אם שתי התוכניות הנ"ל פיזביליות (יש לשתיהן פתרונות חוקיים), אזי מתקיים

$$MAX = MIN$$

לא נראה את ההוכחה של המשפט הזה בקורס שלנו. נראה רק גרסה חלשה יותר:

<u>משפט (דואליות חלשה):</u> אם שתי התוכניות הנ"ל פיזביליות (יש לשתיהן פתרונות חוקיים), אזי מתקיים

$$MAX \le MIN$$

<u>הוכחה:</u>

יהיו $(\hat{y}_i)_{i=1}^{m}, (\hat{x}_j)_{j=1}^{n}$ פתרונות חוקיים לשתי התוכניות הנ"ל (בהתאמה).

אזי מהחוקיות של $\left(\hat{x}_j\right)_{j=1}^{n}$ ומאי-השליליות של $\left(\hat{y}_i\right)_{i=1}^{m}$ נובע לכל i:

$$\sum_{j=1}^{n} a_{ij}\cdot\hat{x}_j\cdot\hat{y}_i \le b_i\cdot\hat{y}_i \qquad (11)$$

באופן דומה, מהחוקיות של $\left(\hat{y}_i\right)_{i=1}^{m}$ ומאי-השליליות של $\left(\hat{x}_j\right)_{j=1}^{n}$ נובע לכל j:

$$\sum_{i=1}^{m} a_{ij}\cdot\hat{y}_i\cdot\hat{x}_j \ge c_j\cdot\hat{x}_j \qquad (12)$$

לכן מתקיים:

$$\sum_{j=1}^{n} c_j\cdot\hat{x}_j \underbrace{\le}_{\text{לפי (12)}} \sum_{j=1}^{n}\sum_{i=1}^{m} a_{ij}\cdot\hat{y}_i\cdot\hat{x}_j \underbrace{=}_{\text{שינוי סדר סכימה}} \sum_{i=1}^{m}\sum_{j=1}^{n} a_{ij}\cdot\hat{x}_j\cdot\hat{y}_i \underbrace{\le}_{\text{לפי (11)}} \sum_{i=1}^{m} b_i\cdot\hat{y}_i$$

אי השיוויון האחרון נכון <u>לכל</u> זוג פתרונות חוקיים $\left(\hat{y}_i\right)_{i=1}^{m}, \left(\hat{x}_j\right)_{j=1}^{n}$ עבור שתי התוכניות הנ"ל.

בפרט, עבור פתרונות אופטימליים $(y_i^*)_{i=1}^{m}, (x_j^*)_{j=1}^{n}$ לשתי התוכניות הנ"ל מתקיים:

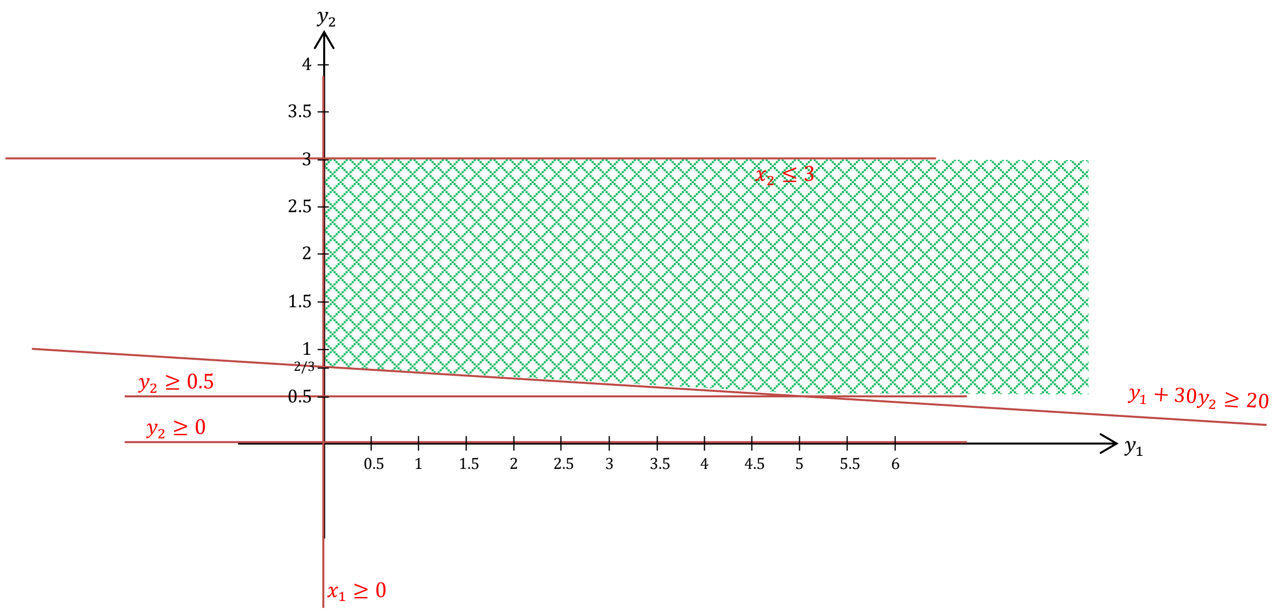
$$MAX = \sum_{j=1}^{n} c_j\cdot x_j^* \le \sum_{i=1}^{m} b_i\cdot y_i^* = MIN$$

<u>דוגמה:</u>

כפי שאמרנו, שתי התוכניות הבאות דואליות זו לזו:

תוכנית D	תוכנית C
$\min \quad 3y_1+100y_2$ $\text{s.t.} \quad y_2 \ge 1/2$ $y_1+30\cdot y_2 \ge 20$ $y_1, y_2 \ge 0$	$\max \quad 10x_1+20x_2$ $\text{s.t.} \quad x_2 \le 3$ $20x_1+30x_2 \le 100$ $x_1, x_2 \ge 0$

כפי שראינו בהרצאה הקודמת, פתרון אופטימלי (יחיד) לתוכנית C מתקבל בנקודה $x_1 = 0.5,\ x_2 = 3$ עם ערך 65



מהו הערך האופטימלי של התוכנית הזאת?
פונקציית המטרה היא

$$\min \quad 3y_1 + 100y_2$$

לכן בתחום הפיזבילי תמיד כדאי לנו (כשאפשר) להקטין את y_1 ו/או את y_2. מסקנה: פתרון אופטימלי ימצא על הקו $y_1 + 30y_2 = 20$. כשאנחנו על הקו הזה, פונקציית המטרה היא:

$$60 + 10y_2$$

כלומר פונק' המטרה יורדת עם y_2.
לכן ערך אופטימלי יתקבל בנקודה

$$y_1 = 5, \;\; y_2 = \frac{1}{2}$$

ולפתרון הזה יש ערך:

$$3y_1 + 100y_2 = 3 \cdot 5 + 100 \cdot \frac{1}{2} = 65$$

כלומר הערך של התוכנית הזאת אכן שווה לערך של התוכנית הדואלית לה.
(נזכור שלכל פתרון <u>חוקי</u> לתוכנית D יש ערך גדול שווה מ 65 ולכל פתרון <u>חוקי</u> לתוכנית C יש ערך קטן שווה 65)

תזכורת:

הוכחת משפט לינאריות התוחלת

(עבור שני משתנים מקריים X, Y המקבלים מספר סופי של ערכים)

$$\mathbb{E}[X+Y] = \sum_x \sum_y (x+y) \cdot \Pr[X=x\,, Y=y]$$

$$\underbrace{=}_{\text{פתיחת סוגריים}} \left(\sum_x x \sum_y \Pr[X=x\,, Y=y] \right) + \left(\sum_y y \sum_x \Pr[X=x\,, Y=y] \right)$$

$$\underbrace{=}_{\text{הנוסחה להסתברות שלמה}} \left(\sum_x x \cdot \Pr[X=x] \right) + \left(\sum_y y \cdot \Pr[Y=y] \right) = \mathbb{E}[X] + \mathbb{E}[Y]$$