נושא 5 - תכנון לינארי

א) הגדרה

תכנון לינארי (בקיצור LP) היא טכניקה למציאת פתרון אופטימלי של ביטוי לינארי עם משתנים ממשיים תחת אילוצים לינאריים. ישנם בעיות רבות במדעי המחשב שניתן להביע ולפתור אופטימליות שלהן באמצעות תכנון לינארי. כמו שנראה בהמשך אנו נשתמש בטכניקה זו בעיקר עבור בעיות השייכות ל-NPC, כדי למצוא להן אלגוריתם קירוב בזמן פולינומי. האופטימליות של הפתרון תלוי בהגדרת הבעיה, יכול להיות פתרון מקסימלי או מינימלי שעונה את כל האילוצים.

ב) ייצוג בעיה התכנון לינארי

בעיה התכנון לינארי מורכבת מארבעה ערכים: X, A, b, C.

X - וקטור משתנים ממשיים בגודל n, $X \in \mathbb{R}^n$, המייצג פתרון אופטימלי אותו נרצה לחשב ולהחזיר. כלומר, כל משתנה ב-X מייצג ערך שאנו מעוניינים לקבוע אותו כדי לקבל פתרון אופטימלי לבעיה.*

A - מטריצה ממשית מסדר $A_{m \times n}$. כל עמודה במטריצה זו מייצגת משתנה אחר ב-X, וכל שורה מספקת מידע על כל משתנה ב-X. המידע שנמצא ב-A משמש כדי לקבוע את האילוצים הלינאריים על וקטור X. מטריצה A ידועה כבר בהגדרת הבעיה.

b - וקטור משתנים ממשיים בגודל m, $b \in \mathbb{R}^m$. כל ערך i ב-b מתאים לשורה i ב-A, ומהווה את האילוץ הלינארי של בחירת המשתנים ב-X עבור שורה זו. בפועל נבדוק את ההכפלה $A \cdot X$ שתייצר וקטור באורך m שצריך להתאים לאילוצים ב-b. עבור בע"ת מקסימום צריך להתקיים $A \cdot X \le b$, ועבור בע"ת מינימום צריך להתקיים $A \cdot X \ge b$. b גם כן ידוע בהגדרת הבעיה.

c - וקטור משתנים ממשיים בגודל n, $c \in \mathbb{R}^n$. כל ערך i ב-c מתאים למשתנה i ב-X, ומייצג את הכמות שנרצה לקחת מ-X_i בסכימה של כל המשתנים ב-X.* המכפלה $c^T X$ תייצר לנו מספר ממשי המבטא את היעילות של הפתרון שלנו. עבור בע"ת מקסימום נחפש $c^T X$ מקסימלי, ועבור בע"ת מינימום נחפש $c^T X$ מינימלי.*

סיכום: נייצג בע"ה התכנון לינארי: $\boxed{MAX\{c^T X : A \cdot X \le b\}_{\text{בע"ת מקסימום}} \;.\; MIN\{c^T X : A \cdot X \ge b\}_{\text{בע"ת מינימום}}}$

*וקטור X_1 המקיים $A \cdot X_1 \le b$ מהווה פתרון אפשרי (Feasible) ל-LP. ואם גם לכל X_2 המקיים $A \cdot X_2 \le b$, מתקיים $c^T X_1 \ge c^T X_2$, אז X_1 הוא פתרון אופטימלי ל-LP. בע"ת תכנון לינארי בלי פתרון אפשרי (תחום ריק) נקראת "בלתי פתירה" (Infeasible). ואם פתירה אך בלי פתרון אופטימלי נקראת "לא מוגבלת" (Unbounded).

וב הבע"ת X מייצג נים ... שם עבור כל עצם ה שבחור, לכן בדרכ"כ ריש $0 \le X$.

* $c_1 \cdot x_1 + c_2 \cdot x_2 + \ldots + c_n \cdot x_n$

פונקציה $X \to c^T X$ נקראת Objective Function

* זו הסדר עבור בע"ת מקסימום.

ג) תכנון לינארי ותכנון בשלמים

$\max_{x\in\mathbb{R}^n}\{c^Tx : Ax\le b\}$

$\max_{x\in\mathbb{Z}^n}\{c^Tx : Ax\le b\}$

עד כה ראינו כיצד ניתן לייצג תכנון לינארי (Linear Programming), אמנם קיים שיטה נוספת של תכנון בשלמים (Integer Programming). ההבדל היחיד בין LP ל-IP הוא שוקטור המשתנים x ב-LP הוא בממש ($x\in\mathbb{R}^n$)*, ואילו ב-IP x בשלמים ($x\in\mathbb{Z}^n$)*. הבדל זה נראה קטן אך הוא משמעותי ביותר. ב-LP כל הבעיות הן ב-P, כלומר ידוע לנו אלגוריתמים שפותרים אותם בזמן פולינומי. (בדר"כ יעבור קבוע גדול), ואילו ב-IP רוב הבעיות הן ב-NPC (תלוי ב-A). השיטה שלנו לפתרון בעיות ב-IP, היא לפתור אותן קודם ב-LP, כך שנדע מהו הפתרון האופטימלי בממש, ואז באמצעות טכניקות באלגברה לינארית למצוא את הפתרון בשלמים שהכי קרוב. מבחינת יעילות לפתרון שמצאנו. בפרק זה נעסוק בעיקר במציאת פתרון אופטימלי ב-LP.

3) דוגמאות לתכנון לינארי

דוגמא:

	e_1	e_2	e_3
v_1	0	1	1
v_2	1	0	1
v_3	1	1	0

כל שורה v_i תכיל עבור כל קשת שיוצאת v_i כל עמודה תכיל הבני השורות של ... את המרכיבים את ... ית.

1) שידוך מקסימלי - בגרף $G=(V,E)$ עם $v_1,v_2,\dots,v_n\in V(G)$ צמתים ו-$e_1,e_2,\dots,e_m\in E(G)$ קשתות. נרצה להציג את בעיית מציאת שידוך מקסימלי (קבוצת קשתות ללא חולקות צומת) כבעיית LP.

נגדיר את $A_{n\times m}$ להיות "מטריצת השכיחות" (Incidence matrix) של G, שבה הקשתות $e_1,\dots,e_m$ הם העמודות והצמתים $v_1,\dots,v_n$ הם השורות.* בכל תא $[v_i,e_j]$ יהיה 1 אם ורק אם $v_i\in e_j$, אחרת 0.*

אנו מחפשים וקטור $x\in\{0,1\}^m$ כך ש-x_j יכיל 1 רק אם x_j שייך לשידוך, אחרת 0.

עבור כל שורה v_i אפשר לבחור רק קשת 1 לכל היותר. לכן תיה להתקיים האילוץ $A\cdot x\le 1$.

כלומר וקטור b מוגדר להיות אחדות. בנוסף כיוון שאנו רוצים לסכום את כל המשתנים ב-x כדי לחשב מספר צלעות בשידוך, נגדיר c גם כן וקטור אחדות, ונרצה למקסם את $x^T\cdot 1$.

נשים לב שבעיה זו היא ב-IP שכן כל איבר ב-x הוא 0 או 1. $\max_{x\in\mathbb{Z}^m}\{1^Tx : Ax\le 1,\ x\ge 0\}$

למרות שבעיה זו ב-IP יש לה פתרון פולינומי כמו שראינו במטלה 1, אך זהו מקרה נדיר יחסית ב-IP.

2) כיסוי צמתים מינימלי - עבור גרף $G=(V,E)$ נרצה להציג את בעיית מציאת כיסוי צמתים מינימלי VC (קבוצת צמתים שמכסה את כל הקשתות) כבעיית LP. נגדיר $A_{m\times n}$ להיות המטריצה המשוחלפת של A שהגדרנו לעיל, כלומר שב-A הנוכחי הצמתים עמודות והקשתות שורות.* אנו מחפשים וקטור $x\in\{0,1\}^n$ כך ש-x_i יכיל 1 רק אם $v_i\in VC$, אחרת 0. בכל שורה e_j חייב להיות לפחות צומת אחת שנבחרה, לכן האילוץ הוא $A\cdot x\ge 1$. אנו רוצים את המינימום של x^T1.

כל שורה e_j תכיל 1 את הצמתים המרכיבים את e_j. וכל עמודה v_i יהיה 1 בכל הקשתות הנוגעות ב-v_i.

גם כאן הבעיה היא ב-IP: $\min_{x\in\mathbb{Z}^n}\{1^Tx : A\cdot x\ge 1,\ x\ge 0\}$, להבדיל משידוך מקסימלי בעיה זו היא NPC.

בגרף דו"צ תמיד נקבל ש-LP מספר שלם, כלומר LP=IP. וגם ראינו שבגרף דו"צ יש פתרון פולינומי.

לכן נרצה לפתור ב-LP ($x\in\mathbb{R}^n$) ולמצוא קירוב בשלמים.*

(1) $\max\{c^T x : A\cdot x \le b\}$
(2) $\max\{c^T x : x \ge 0, A\cdot x \le b\}$
(3) $\max\{c^T x : x \ge 0, A\cdot x = b\}$
(4) $\min\{c^T x : A\cdot x \ge b\}$
(5) $\min\{c^T x : x \ge 0, A\cdot x \ge b\}$
(6) $\min\{c^T x : x \ge 0, A\cdot x = b\}$

ה) Equational Form

ה: גם Accepted Fo
Standard Fo

ישנם שישה דרכים להציג בעיה תכנון ליניארי. הפשוטה ביותר מביניהם לפתרון היא הצגה הנקראת Equational Form, שבה כל המשתנים ב-X הם גדולים או שווים ל-0, והאילוצים הם $A\cdot X=b$ (שורות (3) ו-(6) עצמאלה)

מסתבר שמכל בעיית LP אפשר לעבור ל-Equational Form בדרך הבאה:

- עבור אילוץ של $\le$: נוסיף משתנה $x_m \ge 0$ בדרך הבאה: $x_1+x_2\le 3 \Rightarrow x_1+x_2+x_3=3$
באמצעות משתנה נוסף זה נוכל להעביר את האילוץ ל-"=".
- עבור אילוץ של $\ge$: נכפיל את המשוואה ב-1-, כך שנקבל משוואה עם "$\le$" שאותה כבר למדנו להעביר.
- עבור משתנה X שיכול להיות שלילי: נחליף את X בשני משתנים חיוביים $y,z \ge 0$, שההפרש ביניהם שווה ל-X. $x=y-z$. נשים לב ש-$y \ge z$.
- לאחר כל הפעולות אלו ניתן שמות חדשים לכל המשתנים. נסמן את סך המשתנים ב-n. נקבל בעיית LP מוגדרת Equational Form.

עבור כל בעיית LP, השלב הראשון בדרך לפתרון יהיה לעבור לצורה זו של Equational Form.

1) פתרונות בסיס

תת מטריצה A_I: תהי בעיית ה-LP: $\max\{c^T x : x\ge 0, A\cdot x=b\}$, כאשר $A_{m\times n}$ מטריצה מדרגה m כאשר $n\ge m$. בהינתן $I\subseteq[n]$ קבוצה של אינדקסים לא בהכרח רצופים מ-1 עד n, נגדיר את A_I להיות התת-מטריצה של A המכילה את כל העמודות ב-A שהאינדקס שלהם שייך ל-I.

$$A = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

$$A_I,\ I=\{1,3\}: \begin{bmatrix} 1 & 7 \\ 2 & 8 \\ 3 & 9 \end{bmatrix}$$

וקטור בסיס: וקטור $X\in\mathbb{R}^n$ המקיים $A\cdot X=b$ נקרא "וקטור בסיס" אם קיימת קבוצה של אינדקסים B בגודל m, $B\subseteq[n]$ $|B|=m$, כך שמתקיימים שני התנאים הבאים:

(1) התת-מטריצה A_B היא לא סינגולרית (הפיכה). כלומר, כל העמודות שלה בלתי תלויות.

(2) $x_j=0, \forall j\notin B$. כל איבר x_j כאשר $j\in B$ נקרא "משתנה בסיס" וכל שאר המשתנים ב-X שהם 0 נקראים "משתנים שאינם בסיס".

נשים לב שלא כל וקטור בסיס X הוא גם פתרון אפשרי ל-LP, מפני שלא בהכרח $x\ge 0$. אם גם $x\ge 0$, אזי X יקרא "**פתרון בסיס**".

ג) משפטים על פתרונות בסיס

(1) **משפט:** כל וקטור $x\in\mathbb{R}^n$ הוא פתרון בסיס אם ורק אם כל העמודות של A_K בלתי תלויות
באשר $K=\{j\in[n] : x_j>0\}$.

הוכחה: כיוון 1 - אם x פתרון בסיס אז קיימת לו מטריצה A_B שכל עמודותיה בלתי תלויות
ו- $|B|=m$. לכן נוכל לקחת $K\subseteq B$ כלשהו והטענה מתקיימת.

כיוון 2 - נניח שהעמודות של A_K בלתי תלויות. אז נוכל להרחיב אותם עם עוד עמודות מ-A
כך שבסוף ביחד יהוו בסיס ל-A בגודל m. כל עמודות אלו הן בלתי תלויות מהגדרת בסיס
ולכן נוכל להגדיר אותן כ-B. התכונה השניה של $\forall j\notin B, x_j=0$ מתקיימת מהגדרת K.

מסקנה: כל פתרון בסיס מקושר לבסיס כלשהו של A. בנוסף, בהינתן וקטור x נוכל לבדוק האם
הוא פתרון בסיס ע"י מציאת ה-K שלו כמו שהוגדר לעיל, ובדיקה האם העמודות של A_K בלתי תלויות.

(2) **משפט:** כל קבוצת אינדקסים B בגודל m, $|B|=m$, $B\subseteq[n]$, מתאימה לכל היותר לפתרון בסיס אחד.

הוכחה: תהי קבוצה B כמתואר לעיל ויהי $x\in\mathbb{R}^n$ פתרון בסיס המתאים ל-B, כלומר $\forall j\notin B, x_j=0$.
כמה פתרונות בסיס x כאלו יש? אם אין ס"מנו שכן גם עבור 0 פתרונות המשפט מתקיים.
לכן נניח שיש פתרון ונוכיח שיש רק 1 כזה. כיוון ש-x פתרון ל-LP מתקיים: $A\cdot x=b$.
בגלל שכל אינדקס j ב-x שאינו ב-B שווה ל-0, נוכל להתעלם מכל עמודה j כזו וכן מכל איבר x_j
ונקבל לבסוף: $A_B\cdot x_B=b$ (ב-A_B יש m עמודות, ב-x נותרו m איברים ו-$b\in\mathbb{R}^m$). כעת מכיוון ש-A_B
הפיכה, שהרי מתאימה לפתרון בסיס x, למערכת המשוואות $A_B\cdot x_B=b$ יש פתרון יחיד מתכונות
מטריצה הפיכה. ולכן יש לכל היותר פתרון בסיס 1.

(3) **משפט:** מספר פתרונות הבסיס בכל בע"ת LP הוא קטן מ-$\binom{n}{m}$.

הוכחה: ב-A יש n עמודות, ולכל m עמודות מתוכן יש לכל היותר פתרון בסיס 1. לכן מספר
פתרונות הבסיס הוא לכל היותר $\binom{n}{m}$.

(4) **משפט:** עבור כל בע"ת LP $\max\{c^Tx : x\geq 0, A\cdot x=b\}$ שיש לה פתרון אופטימלי, אז לכל
פתרון x_0 קיים פתרון בסיס $\tilde{x}$ המקיים $c^T\tilde{x}\geq c^Tx$. בפרט טענה זו מתקיימת עבור x_0 אופטימלי.
לכן בכל בע"ת LP פתירה עם פתרון אופטימלי <u>קיים פתרון בסיס אופטימלי</u>.

הוכחה: יש בספר ובשקפים, בהרצאה דילג על רוב ההוכחה והסביר בקיצור.

ח) אלגוריתם לפתרון LP

מכל המשפטים מסעיף קודם נוכל לבנות אלגוריתם שמוצא פתרון אופטימלי עבור LP בדרך הבאה:

תהי $A_{m\times n}$ מטריצה מדרגה m, ותהי LP $\max\{c^Tx : x\geq 0, A\cdot x=b\}$ פתירה ולא אינסופית, לכן קיים לה פתרון בסיס אופטימלי. שלבי האלגוריתם למציאת פתרון לה הם:

- העבר את ה-LP לצורה של Equational Form.
- עבור על כל תתי הקבוצות $B\subseteq[n]$, $|B|=m$, כך ש-A_B היא לא סינגולרית (עמודות בת"ל). יש $\binom{n}{m}$ כאלו.
- עבור כל B משלב קודם מצא פתרון בסיס x המקיים $A_B\cdot x_B=b$. שמור כל פתרון כזה.
- עבור כל הפתרונות הבסיס שנמצאו משלב קודם, בחר את האחד שבו c^Tx אופטימלי.

יש מספר אלגוריתמים במבנה הזה. המוכרים מתוכם ובשימוש הרחב ביותר נקראים "סימפלקס" ו"אליפסואיד".

סימפלקס הוא אלגוריתם 'יחסית' פשוט. לא הצליחו להוכיח שרץ בזמן ריצה פולינומי. אמנם בגלל שקורה - רק במקרי קיצון שלא מייצגים בעיות אמיתיות ובגלל פשטותו, לרוב משתמשים בו. אליפסואיד כן ניתן להוכיח שרץ בזמן פולינומי אך עבור קבוע מאוד גדול. בנוסף נחשב למורכב יותר ולכן פחות בשימוש.

סימפלקס תמ"ז / מעשי למעשה / ... נדירים / עולם כנראה / ... נדירים / יותר.

ט) גיאומטריה של תכנון לינארי

כל מרחב הפתרונות של LP ניתן לתאר כפוליהדרון קמור (ציור במצגת). כל פתרונות הבסיס נמצאים על הצלעות והקודקודים של הפוליהדרון. בדרך כלל פתרון בסיס אופטימלי יהיה על הקודקודים אך גם יכול להימצא על הצלעות.

י) הלמה של גאוס

בסעיף ח הצגנו אלגוריתם לפתרון בעיית LP, אמנם כאמור לעיל לא תמיד לבעיית LP יש פתרון. לפני שנריץ את האלגוריתם נרצה לבדוק האם בכלל יש לבעיה פתרון. בסעיף זה נלמד מתי למשוואה המטריציונית $A\cdot x=b$ יש פתרון באמצעות "הלמה של גאוס", ובסעיף הבא נלמד מתי ל-Equational Form $x\geq 0$, $A\cdot x=b$ יש פתרון באמצעות "הלמה של פרקש".

<u>למה של גאוס:</u> המשוואה המטריציונית $A\cdot x=b$ פתירה אם ורק אם לכל וקטור y השייך לגרעין השמאלי של A, y^T אורתוגונלי ל-b, כלומר $y^Tb=0$. הגרעין השמאלי של $A_{m\times n}$ הם כל הוקטורים שאם נכפיל איתם משמאל את A נקבל 0. $\text{left ker}(A)=\{y\in\mathbb{R}^m : y^TA=0\}$.

הוכחה: כיוון 1 - נניח שקיים $x\in\mathbb{R}^n$ כך ש-$A\cdot x=b$. ויהי $y\in\mathbb{R}^m$ המקיים $y^TA=0$, אזי מתקיים:

$$y^Tb=y^T(A\cdot x)=(y^TA)x=0\cdot x=0 \Rightarrow \text{מש"ל}$$

כיוון 2 - נניח כי לכל וקטור $y\in R^m$ המקיים $y^TA=0$ גם מתקיים $y^Tb=0$. מרחב העמודה של A, המסומן ב-$R(A)$ ניתן להגדירו בשני דרכים:

(1) $R(A)=\{z\in R^m: \exists x\in R^n,\ A\cdot x=z\}$. (2) $R(A)=\{z\in R^m: \exists B\in R^{1\times m},\ B\cdot z=0\}$.

מספיק להוכיח כי $b\in R(A)$, שכן אז יש x כך ש-$A\cdot x=b$. נבחן שני מקרים:

- אם $R(A)=\{0\}$ הלבב, אזי בהכרח $A=0$. כתוצאה מכך כל $y\in R^m$ בגרעין השמאלי של A. לכן לפי הנחה גם $y^Tb=0$ לכל $y\in R^m$. מכאן שבהכרח $b=0$, ואזי $b\in R(A)$ ו-$A\cdot x=0$ פתירה.
- כאשר $R(A)\neq\{0\}$. נוכיח $b\in R(A)$ באמצעות ההגדרה השנייה של $R(A)$. נגדיר מטריצה $B\in R^{1\times m}$ שבה כל שורה היא וקטור אחר מהגרעין השמאלי של A, שבמקרה זה אינו ריק. נשים לב כי אכן $B\cdot b=0$ שהרי לכל y הגרעין השמאלי של A $y^Tb=0$. לכן $b\in R(A)$. מ.ש.ל.

נוכל להציג את הלמה של גאוס בדרך נוספת. תמ"ז רק אחד משני המשפטים מתקיים:

(1) למשוואה המטריציאלית $A\cdot x=b$ יש פתרון.

(2) למשוואות $y^TA=0$ וגם $y^Tb=-1$ יש פתרון. לא בדוקא -1 אלא כל מספר ממשי שונה מ-0.

יא) <u>הלמה של פרקש</u>

<u>למה של גאוס:</u> תהי מטריצה $A_{m\times n}$ ויהי וקטור $b\in R^m$. תמ"ז רק אחד משני המשפטים מתקיים:

(1) למשוואה המטריציאלית $A\cdot x=b$ וגם $x\geq 0$ יש פתרון.

(2) למשוואות $y^TA\geq 0$ וגם $y^Tb<0$ יש פתרון. כלומר קיים $y\in R^m$ שבהכפלתו משמאל מפריד בין A ל-b.

הוכחה: נוכיח קודם ששני המשפטים לא יכולים להתקיים במקביל. נניח בשלילה ששניהם מתקיימים. נסתכל על הביטוי $y^TA\cdot x$. מצד אחד נקבל שהוא שלילי, אך מצד שני נקבל שהוא גדול שווה ל-0. זהו בוודאי סתירה!

$$\left.\begin{array}{l} y^T(A\cdot x)=y^Tb<0 \\ (\underset{\geq 0}{y^TA})\cdot \underset{\geq 0}{x}\geq 0 \end{array}\right\}\text{סתירה!}$$

כדי לסיים את ההוכחה צריך להוכיח שאם (1) לא מתקיים אזי (2) מתקיים, ובכך להוכיח שלא יכול להיות ששניהם לא מתקיימים במקביל. ההוכחה בעמוד 11 ובספר. בשיעור לא הוסבר טוב.

יב) <u>דואליות בתכנון ליניארי</u>

דואליות בתכנון ליניארי היא מנגנון שעוזר לנו להוכיח מתי פתרון לבעיית LP פתירה הוא אופטימלי או כמה רחוק הוא מלהיות אופטימלי. בהמשך נלמד שיטה באמצעות דואליות כדי למצוא פתרון מספק לבעיית LP. הרעיון בדואליות הוא שעבור בעיית LP כלשהי נבנה בעיית LP נוספת הפוכה לה, שפתרון לה מהווה חסם לבעיה הראשית. לבעיה הראשית נקרא Primal LP ולבעיה שמהווה חסם נקרא Dual LP, ובקיצור P ו-D.

אם P בע"ת מקסימום אז D בע"ת מינימום ומהווה חסם עליון ל-P. ואם P בע"ת מינימום אז D בע"ת מקסימום ומהווה חסם תחתון ל-P. אם נמצא פתרון ב-P שהוא גם פתרון אופטימלי ל-D, אז פתרון זה בהכרח אופטימלי גם ב-P.

דואליות בין שתי בעיות שאחת רק חסם של השניה נקראת "דואליות חלשה", ואילו דואליות בין שני בעיות שהמקסימום שווה למינימום נקראת "דואליות חזקה". דוגמאות לדואליות חזקה:

- משפט קניג: בגרף דו-צדדי שידוך מקסימלי שווה לכיסוי צמתים מינימלי.
- משפט מנגר: המספר המקסימלי של מסלולים זרים בין A ל-B שווה למספר מינימלי של צמתים מפרידים ביניהם

יג) אלגוריתם למציאת בעיה דואלית

עבור בעיה P נרצה למצוא את ה-D המתאים לה. לשם כך נצטרך קודם להגדיר כיצד נראה אילוץ כזה ב-P. למשוואה המטריצאלית המייצגת את האילוצים ב-P יש 3 אפשרויות $A\cdot x \ge/=/\le b$. נסמן את השורות של A כך $a_1,a_2,\dots,a_m$. כעת נוכל להגדיר כל אילוץ $i\in[m]$ בדרך הבאה: $a_i \ge/=/\le b_i$. נייצר את D לפי השלבים הבאים:

1) אם P בע"ת MAX אז D בע"ת MIN, ואם P בע"ת MIN אז D בע"ת MAX.

2) כל אילוץ $i\in[m]$ ב-P הוא משתנה דואלי y_i בוקטור הפתרונות של D. האילוץ על כל משתנה y_i ב-D הוא בהתאם לאילוץ המתאים לו ב-P, לפי העקרונות הבאים:

$a_i\cdot x < b_i \Rightarrow y_i \ge 0$ $\qquad$ $a_i\cdot x \ge b_i \Rightarrow y_i \le 0$ $\qquad$ $a_i\cdot x = b_i \Rightarrow y_i\in\mathbb{R}$

בדר"כ כל האילוצים הם תחת אותו סימן, לכן העקרונות אלו נכונים לוקטור הפתרונות y כולו.

3) הערך אותו אנו מעוניינים שיהיה מקסימום או מינימום (בהתאם ל-D) הוא y^Tb.

4) כל משתנה x_j ב-P כאשר $j\in[n]$ הוא אילוץ j ב-D. הסימן בכל אילוץ j ב-D הוא בהתאם לסימן באילוץ על המשתנה x_j המתאים לו ב-P. בצד השני של המשוואה יהיה האיבר ה-j בוקטור c.

$$x_j\in\mathbb{R} \Rightarrow y_1\cdot a_{1j}+y_2\cdot a_{2j}+\dots+y_m\cdot a_{mj}=c_j$$

$$x_j\ge 0 \Rightarrow y_1\cdot a_{1j}+y_2\cdot a_{2j}+\dots+y_m\cdot a_{mj}\ge c_j$$

$$x_j\le 0 \Rightarrow y_1\cdot a_{1j}+y_2\cdot a_{2j}+\dots+y_m\cdot a_{mj}\le c_j$$

במילים אחרות, כל אילוץ j ב-D הוא מכפלת העמודה ה-j ב-A בוקטור הפתרונות y של D, כאשר בצד השני האיבר ה-j ב-c. נוכל לסכם זאת: $A^Ty \ge/=/\le c$.

דוגמאות: ① $MAX\{c^Tx: x\ge 0, A\cdot x=b\} \Rightarrow MIN\{y^Tb: A^Ty\ge c,\ y\in\mathbb{R}\}$

② $MIN\{c^Tx: A\cdot x\ge b\} \Rightarrow MAX\{y^Tb: y\le 0,\ A^Ty=c\}$

במצגת 12 ובספר יש הסבר עם דוגמאות מספריות מדוע אלגוריתם זה נכון. לא נוכיח זאת כאן, אך נאמר כי הרעיון המרכזי הוא למצוא וקטור y שאם נכפלו בכל האילוצים ב-P, אזי חיבור כל אילוצים אלו, מהצורה $\sum_{i=1}^{m} y_i(a_i\cdot x \ge/=/\le b_i)$, יתן לנו את המשוואה: $c^Tx \ge/=/\le y^Tb$. לכן y^Tb מהווה חסם ל-P.

דואליות חזקה

עבור בע"ה P פתירה, אם האלגוריתם לעיל החזיר בע"ה D פתירה, אזי בין P ל-D מתקיים קשר של דואליות חזקה. דואליות חזקה אומרת שבהינתן שני פתרונות x^* ו-y^* ל-P ו-D בהתאמה, אזי מתקיים $c^Tx^*=y^{*T}b$.

3) מצב אופטימליות - רווח הדואליות

באמצעות דואליות נוכל להגדיר כמה אנו קרובים לפתרון אופטימלי באמצעות המושג הבא: בהינתן בע"ות P ו-D ושני פתרונות להם x ו-y בהתאמה, נגדיר את הערך $|c^Tx - y^Tb|$ להיות "רווח הדואליות". ערך זה מהווה מדד לאופטימליות של x ו-y. כאשר רווח הדואליות הוא 0, x ו-y הם פתרונות אופטימליים. לעומת זאת בכל שמתקיים א"ס כך הפתרונות אלו הן פתיח ופתיח אופטימליים.

Complementary Slackness

$Ax=b,\ x\ge 0$

נגדיר את רווח הדואליות בצורה שונה. בדרך כלל הבע"ה P תהיה בצורה של Equational Form, שהרי זהו אחד השלבים באלגוריתם לפתרון LP. נרצה להעביר גם את D להיות בצורה של Equational Form באותו אופן כמו שעשינו בסעיף ה', על ידי הוספה של משתני וקטור העתרונות y, הנקראים Slack Variables. משתנים אלו ייוצגו בוקטור $s\ge 0$. נקבל בע"ה חדשה $D'=\cancel{\min}\max\{b^Ty : A^Ty+s=c,\ s\ge 0\}$. כעת נגדיר את רווח הדואליות בדרך הבאה:

$$c^Tx - y^Tb \underset{\text{תכונות } T}{=} c^Tx - b^Ty \underset{Ax=b}{=} c^Tx - (Ax)^Ty \underset{\text{פתיחת סוגריים}}{=} c^Tx - x^TA^Ty \underset{\text{גורם משותף}}{=} x^T(c - A^Ty) \underset{c-A^Ty=s}{=} x^Ts$$

כעת נוכל להגדיר את משפט Complementary Slackness המציג 3 תנאים שקולים מתי שני פתרונות x ו-y ל-P ו-D' הם אופטימליים.

<u>משפט</u>: יהי x, (y,s) שני פתרונות אפשריים ל-P ו-D', הטענות הבאות שקולות:

(1) x ו-(y,s) הם פתרונות אופטימליים ל-P ו-D' בהתאמה.

(2) $x^Ts=0$.

(3) $x_j\cdot s_j=0\quad \forall j\in[n]$, כלומר לא רק חיבור האיברים שווה ל-0 אלא כל איבר שווה ל-0.

10) Primal-Dual Method

עבור בעיית IP עד כה למדנו שיש להעבירה לבעיית LP, לפתור אותה באמצעות האלגוריתם מסע"ל יא (כמו סימפלקס ואליפסואיד), ואז לחזור חזרה ל-IP באמצעות אלגוריתם קירוב.
הבעיה בפתרון בעיה ב-IP בדרך זו היא שאנו מתייחסים לכל בעיה באותה צורה ומפעילים עליה אלגוריתם לפתרון LP שזמן הריצה שלו די גדול. אנו לא מתחשבים בסוג והתכונות של הבעיה הספציפית איתה אנו מתמודדים, באמצעותם אנו יכולים לשפר את זמן הריצה. בנוסף, המעבר חזרה ל-IP יכול להיות גם כן מסובך ואף כן לדרוש זמן ריצה גדול.

במקום זאת שיטת Primal-Dual אומרת שעבור כל בעיית IP, ננסה אלגוריתם ייעודי לבעיה זו שמשתמש בשני התכונות ל-P ו-D, וכל הזמן נמדוד את "הרווח" הדואלי שלהם וננסה לקרב אותם ל-0. (נק"ק תנאי complementary slackness)
אלגוריתם זה מתחשב בכל התכונות והאילוצים של הבעיה ובכך משפר את זמן הריצה.
האלגוריתם לעולם לא יקרא לסימפלקס או אלגוריתם אחר שפותר LP מפני שהוא לא מנסה למצוא פתרון אופטימלי של ה-LP. נדגים את השימוש בשיטה על מסלול הכי קצר (דייקסטרה).

דוגמאות להגדרת בעיות ב-LP

1) MAX-FLOW

נתון גרף G, צומת מקור s וצומת יעד t. המטרה להזרים כמה שיותר מ-s ל-t כאשר לכל קשת $e\in E(G)$ יש משקל מקסימלי $w(e)$ שיכול לעבור בה. גם הזרימה עוברת בכיוון הנגדי המשקל שלילי.

הגדרת הבעיה בתכנון לינארי:

- לכל קשת $(u,v)\in E(G)$ נגדיר משתנה $x_{u,v}$ שיהיה הכמות המועברת מ-u ל-v.
- לכל קשת e נגדיר אילוץ שכמות הזרימה מוגבלת בין $w(e)$ ל-$-w(e)$.
- לכל צומת פרט ל-s ו-t נגדיר אילוץ שסכום הזרימה שנכנס שווה לזה שיוצא.
- מטרה למקסם את כמות המשתנים שיוצאת מ-s השווה לנכנסת ל-t.

LP

$$\text{MAX} \sum_{v\in N_G(s)} x_{s,v}$$

$$\text{s.t}\quad -w(e)\le x_e\le w(e)\quad \forall e\in E(G)$$

$$\sum_{v\in N_G(u)} x_{u,v}=0\quad \forall u\in V(G)\setminus\{s,t\}$$

$$x_{v,u}=-x_{u,v}$$

2) רגרסיה לינארית

בהינתן n נקודות במישור $(x_1,y_1),\ldots,(x_n,y_n)$ צריך למצוא 'ישר' $y=ax+b$ שהכי קרוב לכל הנקודות, הקרבה מחושבת באמצעות סכום הסטיות של כל נקודה מהישר בערכי y. $\sum_{i=1}^{n}|ax_i+b-y_i|$.

- לכל נקודה נגדיר משתנה x_i שיחשב את הסטיה.
- לכל נקודה נגדיר אילוץ שהמשתנה x_i גדול מהסטיה.

כיוון שהסטיה מחושבת בערך מוחלט נצטרך שני אילוצים אחד במקרה שהסטיה שלילית והשני במקרה חיובית.

- המטרה למנמם את סכום כל המשתנים (סטיות).

כמות ≠ מידת קודות הודות שמצאות ישר.

$$\text{MIN} \sum_{i=1}^{n} x_i$$

$$\text{s.t}\quad x_i\ge ax_i+b-y_i\quad \forall i\in[n]\quad a,b\in\mathbb{R}$$

$$x_i\ge -(ax_i+b-y_i)\quad \forall i\in[n]\quad a,b\in\mathbb{R}$$

3) שידוך ממושקל מקסימלי בגרף דו"צ

נתון גרף דו-צדדי $G=(A\cup B,E)$ ממושקל. צריך למצוא שידוך מושלם m שממקסם את משקל הקשתות.

- לכל קשת e נגדיר משתנה x_e שיהיה 1 אם e בשידוך ו-0 אחרת.
- כל צומת $u\in V(G)$ צריכה להופיע בשידוך פעם אחת, כלומר סכום המשתנים של הקשתות היוצאות מ-u שווה ל-1.

בנוסף לכל קשת e מתקיים $x_e\in\{0,1\}$. ההגדרת LP לה שונה.

- המטרה למקסם את סכום משקל הקשתות שמופיעות בשידוך.

הפתרון ב-LP כמובן גדול מהפתרון ב-IP כי שוקל יותר אפשרויות. בנוסף הפתרון ב-LP הוא פולינומיאלי באמצעות אלגוריתמים שלמדנו.

$$\text{MAX} \sum_{e\in E(G)} x_e\cdot w_e$$

$$\text{s.t}\quad \sum_{v\in N_G(u)} x_{u,v}=1\quad \forall u\in V(G)$$

$$\underset{IP}{x_e\in\{0,1\}}\ \ /\ \ \underset{LP}{0\le x_e\le 1}\quad \forall e\in E(G)$$

בס"ד

<u>משפט:</u> מן הפתרון LP נוכל לבנות פתרון IP בזמן פולינומי.

הוכחה: יהי x^* פתרון אופטימלי. אם כל המשתנים שלמים סיימנו. לכן יש x_e^* שהוא $0<x_e^*<1$.
אמנם בגלל האילוץ, אם $e=(u,v)$ אזי ל-u חייבת להיות עוד קשת (u,a) ששייכת ל-x^* כך
שתשלים ל-1. אמנם אז גם ל-a יש קשת נוספת כזאת. כך נמשיך עד שנקבל מעגל. כיוון
שהגרף דו"צ המעגל זוגי. נסמן קשתות מעגל $e_1,e_2,\dots,e_k$. לכל הקשתות הזוגיות נוסיף ערך m
כך שלא יעברו את 1 ולכל הקשתות האי-זוגיות נחסר m כך שלא יהיו קטנים מ-0.
סך המשקל לא היקטן ועדיין עונה על כל האילוצים עבור כל צומת כי הוספנו והחסרנו
אותו ערך. נחזור על התהליך עד שישארו ב-x^* רק קשתות שהן 1 והן הפתרון ב-IP.

זו max IS
וזה בגלל זוגיים
וכי בדיקה
ומגיעות Z

מטריצה Totally Unimodular

היא מטריצה שלכל תת מטריצה ריבועית שלה הדטרמיננטה היא ב-$\{-1,0,1\}$. כיוון שכל איבר יחיד הוא תת מטריצה ריבועית, איברי המטריצה הנ"ל מורכבים מ-$\{-1,0,1\}$.
נסמן מטריצה כזו $M_{m\times n}(\{-1,0,1\})$. התכונות של מטריצה כזו הן:

(1) אם לבעיית LP מטריצת האילוצים שלה A היא Totally Unimodular, $A\in M_{m\times n}(\{-1,0,1\})$.
וגם $b\in\mathbb{Z}^m$ אז הפתרון האופטימלי של הבעיה הוא בשלמים.

(2) אם לבעיית LP מטריצת האילוצים שלה A היא Totally Unimodular, $A\in M_{m\times n}(\{-1,0,1\})$,
וגם $b\in\mathbb{Z}^m$, $c\in\mathbb{Z}^n$. אז לבעיית ה-LP ולבעיה הדואלית שלה יש אותו פתרון אופטימלי, שהוא גם בשלמים.

משפט: מטריצה $A_{n\times m}$ היא מטריצת השכנויות של גרף דו-צדדי $G=(X\cup Y,E)$ אם ורק אם A היא מטריצה Totally Unimodular.

הוכחה: צ"ל כל תת מטריצה ריבועית $Q_{k\times k}$ של A היא עם דטרמיננטה $\{-1,0,1\}$. נוכיח באינדוקציה על גודל מטריצה Q.

בסיס - עבור $k=1$ הדטרמיננטה.

צעד - נניח עבור $k-1$ ונוכיח עבור k. נחלק לשלושה מקרים:

- אם ב-Q יש עמודת 0, אזי דטרמיננטה 0.
- אם ב-Q יש עמודה עם 1 בודד. אזי לפי כלל חישוב דטרמיננטה, הדטרמיננטה של $Q_{k\times k}$ שווה לתת מטריצה ריבועית של A בגודל $(k-1)\times(k-1)$, ולפי הנחת האינדוקציה הטענה מתקיימת.
- אם ב-Q כל העמודות שני 1, (לא יכול להיות יותר היות וכל עמודה היא קשת). אזי עבור כל עמודה השורה שבה יש 1 שייכת ל-X והשנייה שייכת ל-Y, היות ומיקום האחדות הוא קצוות הקשת שמייצגת העמודה, ו-A היא מטריצת שכנויות של גרף דו"צ. אזי, מספר השורות ב-Q השייכות לצמתים מ-X שווה למספר השורות השייכות ל-Y. ולכן השורות ב-Q הן תלויות ומכאן דטרמיננטה 0. מ.ש.ל.

מסקנה - הוכחה נוספת למשפט קניג (König)

יהי $G=(X\cup Y,E)$. צ"ל כ"ס צמתים מינימלי $\tau(G)$ שווה לשידוך מקסימלי $\nu(G)$. ידוע כי הבעיה הדואלית של תוח' min VC בתבנית LP היא max Matching בתבנית LP. מכיוון שלהן גרף דו"צ, לפי משפט קודם הפתרון האופטימלי שלהם שווה והוא בשלמים. לכן $\tau(G) = \nu(G)$.