חקר ביצועים- גיל אריאל תשפ"ד

רגב יחזקאל אימרה

July 20, 2024

לשאלות/ הצעות שיפורים מוזמנים לשלוח מייל ל-$regevel2006@gmail.com$.

הרצאה 1.

תכנון ליניארי.

תכנון ליניארי זה בעיית אופטימיזציה עם אילוצים.
בהינתן $f, c_i : \mathbb{R}^n \to \mathbb{R}$ פונקציות המטרה שלנו היא למצוא

$$\begin{aligned} &\min_{x \in \mathbb{R}} f(x) \\ &\text{s.t.} c_i(x) = 0_{(i \in E)} \\ &c_i(x) \geq 0_{(i \in I)} \end{aligned}$$

לאילוצים מהסוג $c_i(x) = 0$ נקרא **אילוצים פעילים** ולאילוצים מהסוג $c_i(x) \geq 0$ נקרא **אילוצים לא פעילים**.

בעיות ליניאריות.

חשוב לומר כי f וכל הפונקציו c_i הן פונקציות ליניאריות, כלומר עבור איזשהו $c \in \mathbb{R}^n$ נוכל לכתוב

$$\begin{aligned} f(x) &= < c, x > + F \\ c_i(x) &= < a_i, x > - b_i \end{aligned}$$

כאשר לרוב $F = 0$ ומתקיים $i \in E \cup I, a_i \in \mathbb{R}^n, b \in \mathbb{R}$.

הרעיון הכללי:

המינימום שאנחנו מחפשים מתקבל רק בקצוות התחום.
נוכיח בהמשך: מספיק להסתכל רק בפינות של התחום שלנו למציאת המינימום.
דוגמה:

$$\begin{aligned} \max\ z = &x_1 + 2x_2 - x_3 \\ &\text{s.t.} 2x_1 + x_2 + x_3 \leq 14 \\ &4x_1 + 2x_2 + 3x_3 \leq 28 \\ &2x_1 + 5x_2 + 5x_3 \leq 30 \\ &x_1, x_2, x_3 \geq 0 \end{aligned}$$

פתרון אפשרי: היא נקודה שמקיימת את האילוצים.
פתרון אופטימלי: פתרון אפשרי מקסימלי.
נמיר את הבעיה לצורה סטנדרטית:

$$\begin{aligned} \max z &= x_1 + 2x_2 - x_3 \\ \text{s.t.} S_1 &= 14 - 2x_1 - x_2 - x_3 \\ S_2 &= 28 - 4x_1 - 2x_2 - 3x_3 \\ S_3 &= 30 - 2x_1 - 5x_2 - 5x_3 \\ x_1, x_2, x_3, &S_1, S_2, S_3 \geq 0 \end{aligned}$$

כעת ברור ש$(x_1,x_2,x_3)=(0,0,0)$ פתרון אפשרי, אבל אפשר להגדיל את x_1 עד שS_1 או S_2 או S_3 יתאפסו.
לכן נשווה את x_2,x_3 ל-0 ונראה עמה אפשר להגדיל את x_1:

$$\begin{aligned}
S_1 &: 14-2x_1\underbrace{-x_2-x_3}_{=0}\geq 0\\
&x_1\leq 7\\
S_1 &: 28-4x_1\underbrace{-2x_2-3x_3}_{=0}\geq 0\\
&x_1\leq 7\\
S_1 &: 30-2x_1\underbrace{-5x_2-5x_3}_{=0}\geq 0\\
&x_1\leq 15
\end{aligned}$$

ולכן $x_1\in[0,7]$. נחלץ את x_1 מ-S_1:

$$x_1=7-\frac{1}{2}x_2-\frac{1}{2}x_3-\frac{1}{2}S_1$$

ונציב את זה בכל מקום שאנחנו רואים x_1. לכן הבעיה שלנו תהיה

$$\begin{aligned}
\max z &= 7+\frac{3}{2}x_2-\frac{3}{2}x_3-\frac{1}{2}S_1\\
\text{s.t.} x_1 &= 7-\frac{1}{2}x_2-\frac{1}{2}x_3-\frac{1}{2}S_1\\
S_2 &= -x_3-2S_1\\
S_3 &= 16-4x_2-4x_3+S_1\\
&x_1,x_2,x_3,S_1,S_2,S_3\geq 0
\end{aligned}$$

אפשר להגדיל את x_2 עד שx_1 או S_2 או S_3 יתאפסו:

$$\begin{aligned}
x_1 &: \ldots x_2\leq 14\\
S_2 &: x_2 \text{ לא משפיע}\\
S_3 &: \ldots x_2\leq 4
\end{aligned}$$

ונחלץ את x_2 מ-S_3:

$$x_2=4-\frac{1}{4}S_3-x_3+\frac{1}{4}S_1$$

ונחליף את x_2 בכל מקום:

$$\begin{aligned}
\max z &= 7+\frac{3}{2}x_2-\frac{3}{2}x_3-\frac{1}{2}S_1\\
\text{s.t.} x_1 &= 5+\ldots\\
x_2 &= 4+\ldots\\
S_2 &= \ldots\\
&x_1,x_2,x_3,S_1,S_2,S_3\geq 0
\end{aligned}$$

נתקענו, מצאנו כי $z=13$.
נוכיח כי $z=13$ הוא המקסימום:

$$\begin{aligned}
\max\ z = & x_1+2x_2-x_3\\
\text{s.t.} & (1)\ 2x_1+x_2+x_3\leq 14\\
& (2)\ 4x_1+2x_2+3x_3\leq 28\\
& (3)\ 2x_1+5x_2+5x_3\leq 30\\
& x_1,x_2,x_3\geq 0
\end{aligned}$$

נתבונן ב-$0\cdot(1)+\frac{1}{16}\cdot(2)+\frac{3}{8}\cdot(3)$:
אחרי המון דם יזע ודמעות של פתיחת סוגריים וכינוס איברים נקבל

$$x_1+2x_2+\frac{33}{16}x_3\leq 13$$

אבל! מתקיים

$$z = x_1 + 2x_2 - x_3 \le 13x_1 + 2x_2 + \frac{33}{16}x_3 \le 13$$

לכן $13 \le z$ אבל $z \le 13$ כלומר $z = 13$, כנדרש.

■

הרצאה 2.

הגדרה: **בעיית תכנון ליניארי** עם n משתנים $(x^1,\dots,x^n)$ היא מהצורה

$$\min\setminus\max z = c_0 + x^i c_i$$
$$\text{s.t.} x^i a_i^j \underbrace{\square}_{\geq\setminus\leq\setminus=} b_j$$

כאשר $i\in 1,\dots,m, j\in 1,\dots,n$ וגם c_i,b_i,a_{ij} פרמטרים.

צורה סטנדרטית.

הגדרה: בעיה ליניארית נקראת **בצורה סטנדרטית** אם היא בצורה:

$$\max z = x^i c_i$$
$$\text{s.t.} x^j a_j^i = b_i$$
$$x_i \geq 0$$
$$i\in 1,\dots,m$$

ובצורה מטריציונית נכתוב

$$\max z = \bar{c}\cdot\bar{x}$$
$$\text{s.t.} A\bar{x} = \bar{b}$$
$$\bar{x}\geq 0$$

כאשר $\bar{x} = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_m \end{pmatrix}, \bar{c} = \begin{pmatrix} c_1 \\ c_2 \\ \vdots \\ c_m \end{pmatrix}, \bar{b} = \begin{pmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{pmatrix}$ וגם A מטריצה מגודל $m\times n$.

הערה: $\forall x_i : x_i \geq 0 \iff \bar{x}\geq 0$, כנ"ל לגבי $\leq$ או $\geq$.

בה"כ אנחנו יכולים להניח שהשורות של A בת"ל וגם כי $m\leq n$. לכן, $\text{Rank}(A)=m$

טענה: כל בעיה ליניארית ניתנת לכתיבה בצורה הסטנדרטית.

הוכחה:

- c_0: לא משנה את מיקום המינימום/ מקסימום.
- $\min$: נכפיל את z ב-1 ונקבל בעיית מקסימום.
- אילוצי אי שוויון: אם יש לנו אילוץ מהצורה $a_j^i x_i \leq b_j$ נגדיר משתנה חדש $S_j := b_j - (a_j^i x_i)$, נוסיף את S_j לרשימת האילוצים ונוסיף אילוץ $S_j\geq 0$.
  אם יש לי $a_j^i x_i \geq b_j$ שוב נגדיר משתנה חדש $S_j := a_j^i x_i - b_j$, נוסיף את S_j לרשימת האילוצים ונוסיף אילוץ $S_j \geq 0$
- אם x_i כלשהו לא חסום להיות אי שלילי, נחליף אותו בגרסה חדשה $x_i = y_{i1} - y_{i2}$ ונוסיף לאילוצים $y_{i1}, y_{i2}\geq 0$.

■

צורה קנונית:

בעיית תכנון ליניארי היא **בצורה קנונית** אם היא מהצורה

$$\max z = \bar{c}\cdot\bar{x}$$
$$\text{s.t.} A\bar{x} \leq \bar{b}$$
$$\bar{x}\geq 0$$

ראינו: קנוונית $\Leftarrow$ סטנדרטית.

האם סטנדרטית $\Leftarrow$ קנונית? כן! בהינתן בעיה

$$\max z = \bar{c}\cdot\bar{x}$$
$$\text{s.t.} A\bar{x} = \bar{b}$$
$$\bar{x}\geq 0$$

נוסיף את האילוצים

$$\begin{aligned}\max z &= \bar{c}\cdot\bar{x}\\ \text{s.t.} A\bar{x} &\le \bar{b}\\ -A\bar{x} &\le -\bar{b}\\ \bar{x} &\ge 0\end{aligned}$$

והרי לנו אותה בעיית תכנון ליניארית בצורתה הקנונית.

תכונות:

אילוץ מהצורה $a_i \cdot x_i \le b_i$ חוצה את המרחב $\mathbb{R}^n$ לחלק אחד מצידו של היפר מישור עם נורמל a_i.

הגדרה אוסף הנקודות בו מתקיימים כל האילוצים נקרא **פוליטופ**.

הגדרה קבוצה $s \subseteq \mathbb{R}^n$ נקראת **קמורה** אם לכל 2 נקודות בS, הקבוצה S מכילה את הקטע המחבר ביניהן.

$$\forall x,y \in S, \lambda \in [0,1] : \ell = \{z \in \mathbb{R}^n | z = \lambda x + (1-\lambda)y\} \subseteq S$$

הגדרה נקודה $x \in \mathbb{R}^n$ נקראת **אפשרית** אם היא מקיימת את כל האילוצים.

טענה: אוסף כל הנקודות האפשריות היא קבוצה קמורה.

הוכחה: יהיו $x, y \in \mathbb{R}^n$ נקודות אפשריות, אז במקרה של צורה סטנדרטית x ו-y נמצאים על אותו ההיפר מישור ולכן גם הקטע המחבר ביניהם נמצא על ההיפר מישור. במקרה של צורה קנונית x ו-y נמצאים על אותו צד של ההיפר מישור ולכן גם הקטע המחבר ביניהם נמצא על ההיפר מישור.

מסקנה: קבוצת הנקודות האפשריות היא פוליטופ קמור.

הערה: השפה של פוליטופ קמור ממימד n הם פוליטופים קמורים ממימד $n-1$.

מושגים:

- כל וקטור $x \in \mathbb{R}^n$ נקרא פתרון.
- פתרון המקיים את כל האילוצים נקרא פתרון אפשרי.
- אוסף כל הפתרונות האפשריים נקרא קבוצה אפשרית.
- פתרון אפשרי בו מתקבל ה-max נקרא פתרון אופטימלי.
- פתרון בסיסי הוא פתרון אפשרי שבו לכל היותר m משתנים שונים מ0.
- נקודת קצה (או נקודת קודקוד) של S היא נקודה שאינה נקודה פנימית על קטע ישר בתוך S (כולל השפה).

משפט: נקודה x היא פתרון בסיסי $\Longleftrightarrow$ היא קודקוד.

הוכחה: $\Leftarrow$ נניח שx פתרון בסיסי. נניח שמצאנו $y, z \in S$ כך ש$x = wy + (1-w)z$, צ"ל $y = z$.

נניח בשלילה כי $z \neq y$, נתבונן באיברי x:

אם $x_i = 0$ נקבל $x_i = \underbrace{w}_{>0}\underbrace{y_i}_{\ge 0} + \underbrace{(1-w)}_{>0}\underbrace{z_i}_{\ge 0} = 0$ ולכן $y_i = z_i = 0$. נתבונן ב$y - z$.

יש לפחות $n - m$ מקומות עם 0, כלומר לכל היותר m מקומות עם 0. בנוסף,

$$A(y-z) = Ay - Az = b - b = 0$$

לכן מצאנו קומבינציה ליניארית של $k \le m$ עמודות של A שתלויות ליניארית, בסתירה לכך ש$\text{Rank}(A) = m$. לפיכך $y = z$.

$\Rightarrow$ נניח בשלילה שx נקודת קצה ונקודה אפשרית שאינה בסיסית. נסמן ב$i_1, \ldots, i_r$ את הרכיבים של x השונים מ0.

כיוון ש$\text{Rank}(A) = m$, כל r עמודות של A הן תלויות ליניארית. לכן, קיימים סקלרים $d_1, \ldots, d_r$ שאינם כולם 0 כך ש:

$$d_1 a_{i_1} + \cdots + d_r a_{i_r} = 0$$

כאשר a_{i_k} הוא העמודה הi_k של A.

נשלים את $d = (d_1, \ldots, d_r)^t$ לווקטור ב$\mathbb{R}^n$ על ידי הוספת אפסים, ובפרט אם $x_i = 0$ אז $d_i = 0$. קיבלנו $Ad = 0$.

נבחר

$$\begin{aligned}\varepsilon_1 &= \min\{|x_i| : x_i \neq 0\}\\ \varepsilon_2 &= \max\{|d_i| : d_i \neq 0\}\\ \varepsilon &= \tfrac{1}{2}\tfrac{\varepsilon_2}{\varepsilon_1}\end{aligned}$$

ואז $y^{\pm} = x_i \pm \varepsilon$. אם $x_i = 0$ אז $d_i = 0$ כלומר $y_i^{\pm} = 0$

אם $x \neq 0$ אז $\frac{\varepsilon_2}{2} < \frac{|x_i|}{2} \leq |x_i| - \varepsilon|d_i| \leq |x_i \pm \varepsilon| = |y^{\pm}|$. לכן,

$$Ay^{\pm} = A(x \pm \varepsilon d) = Ax \pm \varepsilon Ad = 0$$

כלומר מצאנו שתי נקודות אפשריות שונות y^+, y^-. כיוון ש$x = \frac{1}{2}(y^+ + y^-)$ אזי x לא נקודת קצה, בסתירה.

■

משפט: אם x^* היא פתרון אופטימלי אז היא נקודת קצה אז היא נקודת קצה או נמצאת על שפה של S הנוצרת מנק' קצה שהם פתרונות אופטימלים.

משפט: כל פתרון מקומי הוא פתרון גלובאלי.

תרגול 1.

בעיות תכנון.

בעיית תכנון מורכבת מ-3 רכיבים:

1. פונקציית מטרה- פונקציה שנרצה למקסם/ למזער.
2. משתני החלטה- המשתנים שעבורת נחפש ערכים שמביאים למקסום/ מזעור של פונקציית המטרה.
3. אילוצים-אילוצים שקיימים עבור משתני ההחלטה.

הגדרה: בעיות תכנות ליניאריות הן בעיות תכנון שפונקציית המטרה והאילוצים הם ליניארים.

מטרה: למצוא פתרון **אופטימלי** לבעיה כלומר סט של ערכים עבור משתני החלטה שגם מקיים את האילוצים וגם מביא למקסימום/ מינימום של פונקציית המטרה.

דוגמה: לדניאל יש חברה ליצור ספרי מתמטיקה. החברה מייצר שני סוגי ספרים:

1. מתמטיקה בדידה.
2. אלגברה ליניארית.

ברשת החברה 3 מפעלים:

1. מפעל לייצור הדפים עבור 2 הספרים.
2. מפעל לכריכה ועריכה של ספרי בדידה.
3. מפעל לכריכה ועריכה של ספרי ליניארית.

להלן טבלה:

בהנחה שכל ספר שנייצר יימכר בשוק, כמה ספרי בדידה וספרי ליניארית יש לייצר כדי להרוויח את הרווח המקסימלי תחת אילוצי הזמן?

פתרון:

שלב ראשון: ננסח את הבעיה כבעיית תכנון:

משתני ההחלטה- כמות הספרים שנייצר של בדידה = x_1, כמות הספרים שנייצר של ליניארית = x_2.

פונקציית מטרה- $z = 3x_1 + 5x_2$ הרווח הכללי.

אילוצים:

עבור מפעל א: $3x_1 + 2x_2 \leq 18$, עבור מפעל ב: $x_1 \leq 4$, עבור מפעל ג: $2x_2 \leq 12$. אילוצי אי שליליות של משתני ההחלטה: $x_1, x_2 \geq 0$.

הבעיה:

$$\begin{aligned} &\max z = 3x_1 + 5x_2 \\ &\text{s.t.} x_1 \leq 4 \\ &\quad 2x_2 \leq 12 \\ &\quad 3x_1 + 2x_2 \leq 18 \\ &\quad x_1, x_2 \geq 0 \end{aligned}$$

אפשר לפתור את זה לפי גרף ולראות איפה מתקבלת הנקודה הכי גבוהה, אבל מעל 2 מימדים זה כבר לא עובד ☹.

נרצה אלגוריתם שפותר לנו את הבעיה עבור כל כמות של משתני החלטה.

סימפלקס:

האלגוריתם מתחיל בנקודת קצה אחת ועובר לנקודת קצה אחרת בצורה שממקסמת/ ממזערת את פונקצית המטרה עד שלא ניתן לשפר.

שלבי האלגוריתם:

1. הנחה- הבעיה:

$$\begin{aligned} \min \setminus \max z = &\sum_{i=1}^{n} c_i x_i \\ \text{s.t.} A\bar{x} &\leq \bar{b} \end{aligned}$$

2. מעבירים את הבעיה לצורה סטנדרטית:
   - נהפוך את פונקציית המטרה ל-max.
   - נהפוך את כל האילוצים להיות אילוצי שוויון (נגדיר משתני חוסר).
   - נהפוך את כל המשתנים להיות גדולים שווים ל-0.

3. מתחילים מפתרון בסיסי בו כל ה-xים שלי שווים ל0- ומשתני החוסר יהיו שונים מ-0.
4. מכניסים לבסיס את המשתנה שהמקדם שלו בפונקצית המטרה הוא השלילי הגדול ביותר.
5. מוציאים מהבסיס את המשתנה בעל היחס הקטן ביותר בין ערכו הנוכחי לבין מקדם המשתנה הנכנס (מתעלמים מיחסים שליליים).
6. אחרי הכנסת המשתנה לבסיס מאפסים את שאר הערכים בעמודה שלו ומביאים את ערך עמודת המשתנה הנכנס ל1 ומבצעים דירוג מטריצה.
7. חוזרים על שלב 4 ו5 עד שאין מקדמים שליליים בשורת z.

הרצאה 3.

צורה סטנדרטית:

$$\max z = c^T x$$
$$\text{s.t.} Ax = b$$
$$x \geq 0$$

צורה קנונית:

$$\max z = c^T x$$
$$\text{s.t.} Ax \leq b$$
$$x \geq 0$$

משפט: אם x^* היא פתרון אופטימלי אז היא נקודת קצה או נמצאת על שפה הנוצרת מנקודות קצה שהם פתרונות אופטימלים (ואז כל נקודה על השפה היא נקודה אופטימלית).

מסקנה: מספיק לחפש מקסימום בנקודות קצה. כמה נקודות קצה יש? 2^{m+n}.

מסקנה: לכל בעיית תכנות ליניארי אחת ורק אחת מהאפשרויות הבאות מתקבלות:

1. אין נקודות אפשריות.
2. המקסימום הוא אינסוף.
3. יש פתרון יחיד והוא מתקבל בקודקוד.

טענה: כל מינימום מקומי מבין הפינות הוא מינימום גלובאלי.

עקרון הצמידות.

בבעיות מקסימום, אני רוצה לחסום מלמעלה את z שלי כדי להראות שאכן הגעתי למקסימום. איך אני אעשה את זה?

נחפש את החסם העליון הכי טוב, כלומר הכי קטן. מותר לקחת כח קומבינציה ליניארית

באופן כללי: יש לנו בעיית תכנות ליניארי

$$P$$
$$\max z = c^T x$$
$$\text{s.t.} Ax \leq b$$
$$x \geq 0$$

אז יש לנו את הבעיה הצמודה שלה

$$D$$
$$\min w = \lambda^T b$$
$$\text{s.t.} A^T \lambda \geq c$$
$$\lambda \geq 0$$

כיוון שכל נקודה אפשרית של P היא חסם תחתון וכל נקודה של D היא חסם עליון, נקבל את תכונת הצמידות החלשה: עם x אפשרית ל-P ו-y אפשרית ל-D אז $z(x) \leq w(y)$.

נשאלת השאלה כמה הדוקים החסמים.

הגדרה: אוסף נקודות ב$\mathbb{R}^n$ הסגור תחת כפל בסקלר חיובי נקרא **חרוט** (חרוט הוא קמור).

משפט: יהי K חרוט ב$\mathbb{R}^n$ כאשר K מהצורה $K = \{By + Cw | y > 0\}$ כאשר $C \in M_{n,p}, B \in M_{n,p}, w \in \mathbb{R}^D, y \in \mathbb{R}^m$, אזי:

יהי $g \in \mathbb{R}^n$ אזי אחד ורק אחד משני התנאים הבאים מתקיימים:

1. $g \in K$
2. קיים וקטור $d \in \mathbb{R}^n$ כך ש: $g \cdot d < 0, B^T d \geq 0, C^T d = 0$

כלומר: נתבונן במישור העובר דרך הראשית עם נורמל d. המישור מפריד בין החרוט K ובין g.

תכונת הצמידות החזקה: אם P סופית אז גם D סופית והפתרונות האופטימלים שלהם שווים.

מסקנות:

1. אם P סופית אז D סופית.
2. אם P לא חסומה אז אין פתרון לD.
3. אם D לא חסומה אז לP אין פתרון.
4. אם לP אין פתרון אז D לא חסומה או שאין לD פתרון.
5. אם לD אין פתרון אז p לא חסומה או שאין לP פתרון.

הרצאה 4.

תזכורת: תכונת הצמידות: בהינתן בעיה P

$$\max z = c^T x$$
$$\text{s.t.} Ax \le B$$
$$x \ge 0$$

קיימת הבעיה הדואלית D

$$\min w = b^T \lambda$$
$$\text{s.t.} A^T \lambda \ge c$$
$$\lambda \ge 0$$

תכונת הצמידות החלשה:

אם x נקודה אפשרית בP ו-y נקודה אפשרית בD אזי $z(x) \le w(y)$

תכונת הצמידות החזקה:

אם P סופית אז D סופית והפתרונות האופטימלים שווים.

למת פארקש:

יהי חרוט $K = \{By + Cw | y \ge 0\} \subseteq \mathbb{R}^n$ כאשר $C \in M_{n,p}, B \in M_{n,m}, w \in \mathbb{R}^p, y \in \mathbb{R}^m$. יהי $g \in \mathbb{R}^n$, אז אחד ורק אחד מהתנאים הבאים מתקיימים:

א $g \in K$.

או

ב קיים וקטור $d \in \mathbb{R}^n$ כך ש

$$g \cdot d < 0$$
$$B^T d \ge 0$$
$$C^T d = 0$$

כלומר ניתן למצוא היפר מישור d שנמצא בין g וK.

משפט 5.1. נתונה מערכת ליניארית

$$\begin{cases} Ax = b \\ x \ge 0 \end{cases}$$

אז אחת ורק אחת מהאפשרויות הבאות מתקיימות:

א קיים x כך ש$Ax = b, x \ge 0$.

או

ב קיים y כך ש$A^T y \ge 0$ וגם$b^T y < 0$.

הוכחה: למת פארקש עם $B = A, B = 0, g = b$.

משפט 5.2. נתונה מערכת ליניארית קנונית

$$Ax \le b$$

אז אחת ורק אחת מהאפשרויות הבאות מתקיימות:

א קיים x כך ש$Ax \le b$

או

ב קיים y כך ש $b^T Y = 0, Y \ge 0, A^T y < 0$.

הוכחה: למת פארקש עם $B = I, C = A, g = b$.

תרגול 2.

תזכורת:

$$\begin{aligned}\max Z &= c^T x\\ \textbf{s.t.}\, Ax &\le b\\ x &\ge 0\end{aligned}$$

אלגוריתם big M:

1. נהפוך לצורה סטנדרטית.
2. עבור אילוצי $'='$: נוסיף משתנה מלאכותי R_i.
עבור אילוצי $'\geq'$: מחסיר משתנה עודף S_i ומוסיף משתנה מלאכותי R_i.
3. מחסיר לפונקציית המטרה את הסכום $\sum_{i=1}^{k} MR_i$.
4. לבטא את ה-R_i בפונקציית המטרה לפי המשתנים x_i, S_i לפי האילוצים של המשתנים המלאכותיים.
5. לפתור רגיל לפי סימפלקס.

הרצאה 5.

משפט 5.1. אם P סופית אז D הבעיה הצמודה לה סופית.

הרצאה 6.

אופטימיזציה עם אילוצים.

אילוצי שוויון.

המקרה הפשוט: אילוצי שיוויון $\Leftarrow$ כופלי לגראנז'.

$$\star\begin{cases}\min_x & f(\overline{x})\\ \text{s.t.} & \overline{g}(\overline{x})=\overline{0}\end{cases}$$

כאשר

$$f:\mathbb{R}^n\to\mathbb{R}$$
$$g:\mathbb{R}^n\to\mathbb{R}^k$$

נתחיל עם אילוץ יחיד $k=1$.

דוגמה:

$$\min_{x,y} x+y$$
$$\text{s.t.} x^2+y^2-1=0$$

לפי לגראנז', נגדיר

$$L(\overline{x},\lambda)=f(\overline{x})-\lambda\cdot g(\overline{x})$$

ואז הפתרון של $\star$ הוא נקודת אוכף של L. לכן:

$$\nabla_{\overline{x},\lambda}L(\overline{x},\lambda)=0$$

היתרון: אין אילוץ.

החיסרון: הגדלנו את המימד.

בנקודת מינימום/ מקסימום, האילוץ חייב להשיק לקווי הגובה. כיוון מאונך לקווי הגובה של f יהיה ∇f.

מגדירים

$$D_{\hat{p}}f(x)=\lim_{h\to 0}\frac{f(\bar{x}+\hat{p}h)-f(\bar{x})}{h}=\nabla f\cdot\hat{p}$$

קווי גובה הם בכיוונים $\hat{p}$ המקיימים $\nabla f\cdot\hat{p}$. יש $n-1$ וקטורים $\hat{p_1},\hat{p_2},\ldots,\hat{p}_{n-1}$ ב"ת שמקיימים $\nabla f\cdot\hat{p}=0$ כלומר ∇f מאונך לכל המשיקים במשטח הגובה כלומר ∇f מאונך למשטח שלנו. האילוץ הוא $g(\bar{x})=0$, ∇g מאונך לאילוץ $\bar{g}(\bar{x})=0$ כלומר $\nabla g||\nabla f$ כלומר $\nabla f=\lambda\cdot\nabla g$ כאשר λ סקלר. זה בדיוק האלגוריתם של כופלי לגראנז'.

אילוצי אי שוויון.

דוגמה:

$$\begin{cases}\min & x_1+x_2\\ \text{s.t.} & 2-x_1^2-x_2^2\geq 0\end{cases}$$

הגדרה נקודה המקיימת את כל האילוצים תיקרא **נקודה אפשרית**.

הגדרה בעיית מינימציה היא מהצורה:

$$\star\begin{cases}\min_{x\in\mathbb{R}^n} & f(\bar{x})\\ \text{s.t.} & c_i(\bar{x})=0, i\in E\\ & c_i(\bar{x})\geq 0, i\in I\end{cases}$$

כאשר $f:\mathbb{R}^n\to\mathbb{R}$ וגם $|E|<\infty$ כמות אילוצי השוויון ו-$|I|<\infty$ כמות אילוצי האי שוויון.

הגדרה קבוצת הנקודת האפשריות הינה

$$\Omega=\{x\in\mathbb{R}^n|\forall i\in E, c_i(x)=0\wedge\forall i\in I, c_i(x)>0\}$$

הגדרה נאמר ש-x^* הוא **פתרון גלובאלי** של $\star$ אם $\forall x\in\Omega, f(x)\geq f(x^*)$.

הגדרה נאמר ש-x^* הוא **פתרון לוקאלי** של $\star$ אם קיימת ε-סביבה של x^* כך ש:

$$\forall x\in\Omega\cap B_\varepsilon(x^*): f(x^*)\leq f(x)$$

בנוסף, נניח שכל הפונקציות f,c_i גזירות ברציפות. $\partial\Omega$ לא חייבת להיות גזירה.

לפעמים אפשר לקחת בעיה לא גזירה ולהפוך אותה לגזירה.

דוגמה :

$$\min_{x\in\mathbb{R}} \max\left\{x^2, x\right\}$$

ניסוח שקול וגזיר ∞ פעמים :

$$\begin{aligned} &\min_{x,t} t \\ &\text{s.t.}\, t - x \geq 0 \\ &\quad t - x^2 \geq 0 \end{aligned}$$

הגדרה קבוצת האילוצים הפעילים בנקודה אפשרית $x \in \Omega$ הוא אוסף האינדקסים של האילוצים בהם מתקיים שוויון.

$$A(x) = E \cup \{i \in I | c_i(x) = 0\}$$

הגדרה נאמר שתנאי **LICQ** מתקיים אם הגרדיאנטים של האילוצים הפעילים ב-x הם בלתי תלויים ליניארית.

תנאי הכרחי מסדר ראשון kkt

$$\star \begin{cases} \min_{x\in\mathbb{R}^n} & f(\bar{x}) \\ \textbf{s.t.} & c_i(\bar{x}) = 0, i \in E \\ & c_i(\bar{x}) \geq 0, i \in I \end{cases}$$

הגדרה הלגראנז'יאן להיות

$$L(x, \lambda) = f(x) - \Sigma\lambda_i c_i(x)$$

אם x^* מינימום לוקאלי של $\star$ ו-f ו-c_i גזירות ברציפות ותנאי $LICQ$ מתקיים ב-x^* אזי קיים וקטור כופלי לגראנז' $\bar{\lambda}^\# \in \mathbb{R}^{|I|+|E|}$ כך ש :

1. $\nabla_x L(x^*, \lambda^\#) = 0$.
2. $\forall i \in E, c_i(x^*) = 0$.
3. $\forall i \in I, c_i(x) \geq 0$.
4. $\forall i \in I, \lambda_i^\# \geq 0$.
5. $\forall i \in E \cup I, \lambda^\# c_i(x^*) = 0$.

תרגול 3.

מקרים חריגים בסימפלקס:

מקרה 1- ללא פתרון אפשרי.

איך תיראה הטבלה שלנו? לדוגמה

	Z	X_1	X_2	a_1	b
Z	1	4	2	1	7
a_1	2	3	3	5	2
X_1	4	-1	2	3	

כאשר a_1 הוא משתנה מלאכותי. שורת פונקצית המטרה היא ללא מקדמים שליליים (כלומר הגענו לסיום האלגוריתם) אבל יש משתנה מלאכותי בבסיס עם ערך ששונה מ-0, לכן הגענו למצב של מקרה ללא פתרון אפשרי.

מקרה 2- אינסוף פתרונות.

$$\begin{aligned} &\max z = y + x \\ &\text{s.t.} y + x \leq 1 \end{aligned}$$

קו הגובה $y + x = 1$ יתן לנו את המקסימום, ויש עליו אינסוף פתרונות!

איך זה יראה בטבלה שלנו?

	Z	X_1	X_2	S_1	b
Z	1	2	0	1	7
X_1	2	3	5	2	4
S_1	2	1	8	0	2

כרגע הפתרון האופטימלי הוא $(X_1, X_2) = (4, 0)$. נכניס את X_2 לבסיס:

	Z	X_1	X_2	S_1	b
Z	1	2	0	1	7
X_1	$\frac{3}{4}$	$\frac{11}{4}$	0	2	$\frac{3}{4}$
X_2	$\frac{1}{4}$	$\frac{1}{8}$	1	0	$\frac{1}{4}$

הפתרון החדש הוא $(X_1, X_2) = \left(\frac{3}{4}, \frac{1}{4}\right)$. אם בשורת פונקציית המטרה קיים משתנה שלא בבסיס עם מקדם 0 בשורת פונקציית המטרה ושאפשר להכניס אותו לבסיס, אזי נקבל שיש לנו אינסוף פתרונות.

מקרה -3 פתרון לא חסום.

לדוגמה

$$\begin{aligned} &\max z = y + x \\ &\text{s.t.} y + x \geq 1 \end{aligned}$$

בסימפלקס:

	Z	X_1	X_2	S_1	b
Z	1	2	-4	1	7
X_1	1	3	-3	1	2
S_1	4	2	1	0	-2

אני יכול להגדיל את X_2 אבל מצד שני אני לא יכול להכניס אותו לבסיס שלי, לכן יש לי פתרון לא חסום.

הבעיה הדואלית.

יש לי בעיית סימפלקס.

$$\begin{aligned} &\max Z = 3x_1 + 5x_2 \\ &\text{s.t.} x_1 \leq 4 && (1) \\ &x_2 \leq 6 && (2) \\ &3x_1 + 2x_2 \leq 18 && (3) \\ &x_1, x_2 \geq 0 \end{aligned}$$

ואני רוצה לחסום את Z. נתבונן ב-$(2)\cdot 3+(1)+(3)$:

$$3x_1+2x_2+x_1+3x_2=3x_1+5x_2\le 18+4+3\cdot 6=40$$

כלומר

$$4x_1+5x_2\le 40$$

ובגלל ש-$3x_1+5x_2\le 4x_1+5x_2$ נקבל

$$Z\le 40$$

ואנחנו יכולים לשפר את החסם על Z על ידי $(2)\cdot 3+(3)$ ונקבל

$$Z=3x_1+5x_2\le 18+3\cdot 6=36$$

כלומר

$$Z\le 36$$

לכן המרנו את הבעיה שלנו ללמצוא חסם הדוק ל-Z לפי המקדמים של משתני ההחלטה, והיא תיראה כך:

$$\begin{aligned}\min W&=4y_1+6y_2+18y_3\\ \text{s.t.}y_1+3y_3&\ge 3\\ y_2+2y_3&\ge 5\\ y_1,y_2,y_3&\ge 0\end{aligned}$$

ובאופן כללי:

הבעיה הראשונית שלי תהיה

$$\begin{aligned}\max Z&=\overrightarrow{c}\cdot\overrightarrow{x}\\ \text{s.t.}A\overrightarrow{x}&\le\overrightarrow{b}\\ \overrightarrow{x}&\ge 0\end{aligned}$$

והבעיה הדואלית שלי תהיה

$$\begin{aligned}\min W&=\overrightarrow{b}^T\cdot\overrightarrow{y}\\ \text{s.t.}A^T\cdot\overrightarrow{y}&\ge\overrightarrow{c}^T\\ \overrightarrow{y}&\ge 0\end{aligned}$$

כאשר $\overrightarrow{c}$ וקטור המקדמים של $\overrightarrow{x}$ בשורת Z, $\overrightarrow{b}$ וקטור הסקלרים באילוצים, A מטריצת המקדמים.

דוגמה:

$$\begin{aligned}\max Z&=4x_1+3x_3+5x_3\\ \text{s.t.}x_1&\le 5\\ 2x_1+x_2&\le 7\\ x_3&\le -8\\ x_1,x_2,x_3&\ge 0\end{aligned}$$

נגדיר את מטריצת המקדמים להיות

$$\begin{pmatrix} x_1 & x_2 & x_3 & | & b\\ 1 & 0 & 0 & | & 5\\ 2 & 1 & 0 & | & 7\\ 0 & 0 & 1 & | & -8\\ - & - & - & + & -\\ 4 & 3 & 5 & | & 1\end{pmatrix}$$

נשחלף ונקבל

$$\begin{pmatrix} y_1 & y_2 & y_3 & | & \\ 1 & 2 & 0 & | & 0\\ 0 & 1 & 0 & | & 3\\ 0 & 0 & 1 & | & 5\\ - & - & - & + & -\\ 4 & 7 & -8 & | & 1\end{pmatrix}$$

כלומר הבעיה הדואלית תהיה

$$\begin{aligned} &\min W = 5y_1 + 7y_2 - 8y_3 \\ &\text{s.t.} y_1 + 2y_2 \geq 4 \\ &\quad y_2 \geq 3 \\ &\quad y_3 \geq 5 \\ &\quad y_1, y_2, y_3 \geq 0 \end{aligned}$$

הערה : הבעיה הדואלית לבעיה הדואלית הינה הבעיה המקורית.

הערה : ערך הפונקציה של הבעיה הפרימלית עבור פתרון אופטימלי שווה לערך הפונקציה של הבעיה הדואלית עבור הפתרון האופטימלי שלה כלומר הפתרון של בעיה אחת היא הפתרון של הבעיה הדואלית שלה.

שימושים :

1. עבור בעיה שצריך לפתור בשיטת $Big - M$ ניתן לעבור לבעיה הדואלית ולפתור אותה בצורה יותר קלה.
2. אם מספר האילוצים m מאוד גדול ממספר המשתנים n (כלומר $m >> n$) אם נעבור לבעיה הדואלית נקבל בעיה עם n אילוצים ו-m משתנים, וזה יותר קל!

דוגמה :

$$\begin{aligned} &\min W = 20y_1 + 10y_2 \\ &\text{s.t.} 5y_1 + y_2 \geq 6 \\ &\quad 2y_1 + 2y_2 \geq 8 \\ &\quad y_1, y_2 \geq 0 \end{aligned}$$

לכאורה נצטרך לעשות פה $Big - M$, אבל אם נעבור לבעיה הדואלית נראה שזה יומר לנו לבעיית סימפלקס רגילה!

הבעיה היא min וכל האילוצים הם $\geq$ לכן ניתן לעבור לבעיה הדואלית ישירות.

$$\begin{pmatrix} 5 & 1 & | & 6 \\ 2 & 2 & | & 8 \\ 1 & 4 & | & 7 \\ - & - & + & - \\ 20 & 10 & | & 1 \end{pmatrix}$$

ונשחלף :

$$\begin{pmatrix} 5 & 2 & 1 & | & 20 \\ 1 & 2 & 4 & | & 10 \\ - & - & - & + & - \\ 6 & 8 & 7 & | & 1 \end{pmatrix}$$

ולכן הבעיה הדואלית תהיה

$$\begin{aligned} &\max Z = 6x_1 + 8x_2 + 7x_3 \\ &\text{s.t.} 5x_1 + 2x_2 + x_3 \leq 20 \\ &\quad x_1 + 2x_2 + 4x_3 \leq 10 \\ &\quad x_1, x_2, x_3 \geq 0 \end{aligned}$$

ונפתור את זה בעזרת סימפלקס רגיל :

	Z	x_1	x_2	x_3	S_1	S_2	b
Z	1	-6	-8	-7	0	0	0
S_1	0	5	2	1	1	0	20
S_2	0	1	2	4	0	1	10

המשתנה הנכנס הוא x_2 והמשתנה שיוצא הוא S_2. נעדכן את הטבלה :

	Z	x_1	x_2	x_3	S_1	S_2	b
Z	1	-2	0	9	0	4	40
S_1	0	4	0	-3	1	-1	10
x_2	0	$\frac{1}{2}$	1	2	0	$\frac{1}{2}$	5

המשתנה הנכנס הוא x_1 והמשתנה שיוצא הוא S_1. נעדכן את הטבלה :

	Z	x_1	x_2	x_3	S_1	S_2	b
Z	1	0	0	$\frac{15}{2}$	$\frac{1}{2}$	$\frac{7}{2}$	45
x_1	0	1	0	$\frac{-3}{4}$	$\frac{1}{4}$	$-\frac{1}{4}$	$\frac{5}{2}$
x_2	0	0	1	$\frac{19}{8}$	$-\frac{1}{8}$	$\frac{5}{8}$	$\frac{15}{4}$

וכעת אי אפשר להכניס לבסיס משתנים כלומר הגעתי לפתרון האופטימלי

$$x_1 = 2.5, x_2 = 3.75 \Rightarrow Z = 45$$

האם זה הפתרון של הבעיה המקורית? נחשב את המינימום בנקודה הזאת. המקדמים יהיו המקדמים של משתני החוסר S_i, כלומר

$$S_1 = y_1 = 0.5, S_2 = y_2 = 3.5 \Rightarrow W = 45$$

ובגלל שאנחנו יודעים שהפתרון של הבעיה הפרימלית והפתרון של הבעיה הדואלית שווים $\Longleftrightarrow$ זה הפתרון האופטימלי, נקבל כי $W = 45$ הוא הפתרון האופטימלי.

הערה: אם יש אילוץ שוויון = מפרקים אותו לשני אילוצים (פעם $\leq$ ופעם $\geq$ ואת האילוץ של $\leq$ מכפילים ב-1 כדי לעבור בחזרה מהבעיה הדואלית).

אופטימיזציה עם אילוצים.

עכשיו יש לנו בעיה לא ליניארית:

$$\begin{aligned} &\min_{x_1,\ldots,x_n} f(x_1,...,x_n) \\ &\quad \text{s.t.} g_1(x_1,...,x_n) \geq=\leq 0 \\ &\qquad \vdots \\ &\qquad g_m(x_1,...,x_n) \geq=\leq 0 \end{aligned}$$

השיטה: כופלי לגראנז'.

מסתכלים על הפונקציה

$$L = (\overrightarrow{x}, \overrightarrow{\lambda}) = f(\overrightarrow{x}) - \sum_{i=1}^{m} \lambda_i \cdot g(\overrightarrow{x})$$

ולזה קוראים הלגראנז'יאן.

אלגוריתם לפתירת בעיה לא ליניארית:

1. מנסחים את הלגראנז'יאן.
2. דורשים את הדרישות הבאות:
   (א) $\nabla_{\overrightarrow{x}} L(\overrightarrow{x}, \overrightarrow{\lambda}) = 0$
   (ב) עבור אילוצי שיוויון:

$$\forall i \in E : g_i(\overrightarrow{x}) = 0$$

   (ג) עבור אילוצי אי שיוויון:

$$\forall t \in I : \lambda_t \cdot g_t(\overrightarrow{x}) = 0$$

3. פותרים אם מערכת המשוואות. אחרי שפותרים את מערכת המשוואות יש נקודות חשודות, בודקים את הערכים שלהן בפונקציה המקורית ולוקחים את הנקודה עם הערך המינימלי.
4. משרטטים את האילוצים. אם האילוצים יוצרים תחום סגור וחסום ו-f רציפה אז לפי ויירשטראס קיים ל-f מינימום גלובאלי ואכן הנקודה המינימלית שקיבלנו חייבת להיות המינימום הגלובלי.

תרגיל: מצאו מינימום מוחלט לבעיה הבאה:

$$\begin{aligned} &\min f(x,y) = 2x^2 + 9y^2 - 24x + 4 \\ &\text{s.t.} x + y \geq 0 \\ &\quad x^2 + y^2 = 25 \end{aligned}$$

פתרון:

התחום האפשרי חסום וסגור ולכן קיים לו מינימום מקומי.

הלגראנז'יאן יהיה

$$L(x,y,\lambda_1,\lambda_2) = 2x^2 + 9y^2 - 24x + 4 - \lambda_1(x+y) - \lambda_2(x^2+y^2-25)$$

נקבל את מערכת המשוואות

$$\begin{aligned} L_x =& 4x - 24 - \lambda_1 + 2\lambda_2 x = 0 && (1)\\ L_y =& 18y - \lambda_1 - 2\lambda_2 y = 0 && (2)\\ & x^2 + y^2 - 25 = 0 && (3)\\ & \lambda_1 (x+y) = 0 && (4) \end{aligned}$$

ונפתור את מערכת המשוואות: נפצל את (4) למקרים:

1. אם $x + y = 0$ אחרי הרבה חישובים נקבל $y = \pm\sqrt{12.5}$ כלומר נקבל את הנקודות $\left(\sqrt{12.5}, -\sqrt{12.5}\right)$ ואת $\left(-\sqrt{12.5}, \sqrt{12.5}\right)$.
2. אם $\lambda_1 = 0$ נקבל ממשוואה (2) את המשוואה $2y(9-\lambda_2) = 0$. שוב נפצל למקרים:
   (א) אם $y = 0$ נקבל $x = 5$ כלומר קיבלנו את הנקודה $(5,0)$.
   (ב) אם $\lambda_2 = 9$ נקבל $x = -\frac{12}{7}$ ואז $y = 4.696$ כלומר קיבלנו את הנקודה $\left(-\frac{12}{7}, 4.696\right)$.

נציב את הנקודות לקביעת המינימום:

(x,y)	$f(x,y)$
$\left(\sqrt{12.5}, -\sqrt{12.5}\right)$	56.647
$\left(-\sqrt{12.5}, \sqrt{12.5}\right)$	226.675
$(5,0)$	-66
$\left(-\frac{12}{7}, 4.696\right)$	249.49

כלומר נקודת המינימום שלנו היא $(5,0)$ והערך בה הוא -66.

כיוון שהתחום הנוצר מהאילוצים שלנו הוא קומפקטי, נקבל ממשפט ויירשטראס כי $(5,0)$ הינה נקודת מינימום גלובלי.

הרצאה 7.

אופטימיזציה:

דוגמה: מספר המבקרים בסניף דואר. נעריך את קבצ המבקרים בסניף הדואר. נסמן ב-t_i את זמני ההגעה כאשר $t_0 = 0$.

הגדרה: הזמנים בין אירועי ההגעה נקראים **זמני ההמתנה**.

נגדיר $\tau_i = t_i - t_{i-1}$.

הנחה: זמני ההמתנה בלתי תלויים, חסרי זיכרון, שווי התפלגות ורציפים.

טענה: ההתפלגות האקספוננציאלית היא רציפה, אי שלילית וחסרת זיכרון.

אפשר לומר ש-$t_i \sim \text{Exp}\left(\lambda\right)$. אנחנו רוצים לאמוד את λ.

נחשב את זמן ההמתנה הממוצע ונסמנו $\langle\tau\rangle = \frac{\Sigma\tau_i}{N}$ כאשר N מספר הדגימות.

ידוע $\tau_1, \ldots, \tau_n$ ורוצים למצוא את λ. כלומר אנחנו יודעים לחשב את $P\left(\tau_1, \ldots, \tau_n|\lambda\right)$ אבל אנחנו רוצים לחשב את $P\left(\lambda|\tau_1, \ldots, \tau_n\right)$.

לפי חוק בייס אפשר לכתוב

$$P\left(\lambda|\tau_1, \ldots, \tau_n\right) = \frac{P\left(\tau_1, \ldots, \tau_n|\lambda\right) \cdot P\left(\lambda\right)}{P(\text{data})}$$

ולכל אחד מהחלקים קוראים

$$\begin{aligned} P\left(\lambda|\tau_1, \ldots, \tau_n\right) &- \text{posterior} \\ P\left(\text{data}|\lambda\right) &- \text{likelihood} \\ P\left(\lambda\right) &- \text{prior} \\ P(\text{data}) &- \text{evidence} \end{aligned}$$

כאשר

$$P(\text{data}) = \int P\left(\text{data}|\lambda\right) \cdot P\left(\lambda\right) d\lambda$$

ידע על λ: חיובי. לכן ניקח

$$P\left(\lambda\right) = \begin{cases} 1 & \lambda \geq 0 \\ 0 & \lambda < 0 \end{cases}$$

אבל זה לא התפלגות כי אי אפשר לנרמל את $P\left(\lambda\right)$, אבל זה לא בעיה כי אם אני רוצה למצוא מקסימום ל-posterior זה בעצם למצוא מקסימום ל-$\frac{\text{likelihood}\cdot 1}{\text{קבוע}}$ שזה למצוא מקסימום ל-likelihood.

הנחה: נשתמש גם ב-$\langle\tau\rangle$. הגיוני ש-$\lambda \sim \frac{1}{\langle\tau\rangle}$ כאשר $\lambda \geq 0$. לכן

$$\lambda \sim \text{Exp}\left(\langle\tau\rangle\right)$$

לכן

$$\begin{aligned} P\left(\text{data}|\lambda\right) &= \lambda^N e^{-\lambda N\langle\tau\rangle} \\ P\left(\lambda\right) &= f\left(\lambda\right) = \langle\tau\rangle\, e^{-\lambda\langle\tau\rangle} \end{aligned}$$

כלומר

$$\begin{aligned} P(\text{data}) &= \int_0^\infty P\left(\text{data}|\lambda\right) \cdot P\left(\lambda\right) d\lambda \\ &= \langle\tau\rangle \int_0^\infty \lambda^N \cdot e^{-\lambda(N+1)\langle\tau\rangle} d\lambda \\ &= \frac{\langle\tau\rangle\, N!}{(N+1)^N\,\langle\tau\rangle^N} \end{aligned}$$

לכן

$$P\left(\lambda|\text{data}\right) = C\lambda^N e^{-\lambda N\langle\tau\rangle}$$

ואם נסמן

$$\begin{aligned} \alpha &= N+1 \\ \beta &= N\langle\tau\rangle \end{aligned}$$

נקבל

$$X\sim\Gamma(\alpha,\beta)$$

כלומר משתנה מקרי שמתפלג גאמה

$$f(X)=\frac{\beta^{\alpha}}{\Gamma(\alpha)}X^{\alpha-1}e^{-\beta X}$$

ואז כדי לקחת את ה-λ המתאים שלי אני יכול לקחת כמה דברים :

$$\begin{aligned}\max&=\frac{1}{\langle\tau\rangle}\\ \text{mean}&=\left(1+\frac{1}{N}\right)\frac{1}{\langle\tau\rangle}\\ \text{חציון}&=\text{אין נוסחא סגורה}\end{aligned}$$

צנזורה:

המטרה : להעריך את הסבלנות של אנשים שמתקשרים לפני שהם מנתקים.

הנחה : חסרי זיכרון : זמן ההמתנה עד שמנתקים מתפלג אקספוננציאלי.

הבעיה : אם עונים למתקשרים, זה הורס את הדאטה, אבל יודעים שהסבלנות > זמן ההמתנה.

יש שתי שיטות למצוא את זה :

1. נזרוק את הדאטה המצונזר.
2. ניקח את הזמן של הצינזור.

בשניהם הזמנים שמחשבים קצרים מהאמת.

נניח שזמן ההמתנה חסר זיכרון : נניח שהוא מתפלג אקספוננציאלית עם קצב μ, לכן אנחנו רוצים למצוא את λ ואת μ.

הדאטה :

נסמן

$$\delta_i=\begin{cases}0 & \text{ענו למתקשר}\\ 1 & \text{ניתק}\end{cases}$$

ואז יש לנו

$$\begin{gathered}(\tau_1,\delta_1)\\ (\tau_2,\delta_2)\\ (\tau_3,\delta_3)\\ \vdots\end{gathered}$$

לכל מתקשר i נגריל 2 משתנים מקריים :

$$\begin{aligned}x_i&:\text{סבלנות}\\ y_i&:\text{מתי יענו לי}\end{aligned}$$

הנחנו שהם ב"ת :

$$x_i\sim\text{Exp}\left(\lambda\right),y_i\sim\text{Exp}\left(\mu\right)$$

לכן

$$\begin{aligned}F_0(E)&=P\left(\tau_i\leq t\wedge\delta_i=0\right)=P\left(y_i\leq t\wedge y_i\leq x_i\right)\\ F_1(t)&=P\left(\tau_i\leq t\wedge\delta_i=1\right)=P\left(y_i\leq t\wedge x_i\leq y_i\right)\end{aligned}$$

לכן נחשב את הנראות :

$$L\left(\lambda,\mu\right)=P\left(\tau_1,\delta_1,\ldots,\tau_N,\delta_N|\lambda,\mu\right)=P\left(\tau_1,\delta_1|\lambda,\mu\right)\cdots P\left(\tau_N,\delta_N|\lambda,\mu\right)=\left(\prod_{\{j|\delta_j=0\}}F_0'\left(\tau_j\right)\right)\cdot\left(\prod_{\{j|\delta_j=1\}}F_1'\left(\tau_j\right)\right)$$

לכן

$$\log\left(L\left(\lambda,\mu\right)\right)=\sum_{\{j|\delta_j=0\}}F_0'\left(\tau_j\right)+\sum_{\{j|\delta_j=1\}}F_1'\left(\tau_j\right)$$

נמצא מקסימום :

$$\frac{\partial L}{\partial \lambda} = 0$$
$$\frac{\partial L}{\partial \mu} = 0$$

הרבה חישובים יוצא

$$\log\left(L\left(\lambda,\mu\right)\right) = \sum_{\{j|\delta_j=0\}} \left[\ln\lambda - \lambda\tau_j + \ln(1-e^{-\mu\tau_j})\right] + \sum_{\{j|\delta_j=1\}} \left[\ln\mu - \mu\tau_j + \ln(1-e^{-\lambda\tau_j})\right]$$

ולזה נמצא מקסימום נומרית.

תרגול 4.

בעיית אופטימיזציה כללית עם אילוצים:

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} f(x) \\ &\text{s.t.} c_i(x) = 0 && i \in E \\ &c_i(x) \geq 0 && i \in I \end{aligned}$$

כאשר:

- E קבוצת האינדקסים לאילוצי שוויון.
- I קבוצת האינדקסים לאילוצי אי שוויון.

הקבוצה הפיזבילית (קבוצת הנקודות האפשריות) היא

$$\Omega = \{x|\forall i \in E, c_i(x) = 0 \wedge \forall i \in I, c_i(x) \geq 0\}$$

כלומר קבוצת כל הנקודות שעומדות באילוצים.

קבוצת האילוצים הפעילה הינה

$$A(x) = E \cup \{i \in I : c_i(x) = 0\}$$

תהי $x \in \Omega$. נאמר שתנאי משפט $LICQ$ מתקיים ב-x אם הגראדיאנטים של האילוצים הפעילים ב-x הם בלתי תלויים ליניארית.

משפט (תנאי kkt):

אם נגדיר את הלגראנז'יאן להיות

$$L(x, \lambda) = f(x) - \sum_{i\in E\cup I} \lambda_i c_i(x)$$

נניח ש-$\hat{x}$ הוא פתרון לוקאלי ונניח ש-f ו-c_i גזירות ברציפות ושתנאי $LICQ$ מתקיימים ב-$\hat{x}$, אזי קיים וקטור כופלי לגראנז' $\hat{\lambda} \in \mathbb{R}^{|I|+|E|}$ כך ש:

1. $\nabla_x L(\hat{x}, \hat{\lambda}) = 0$.
2. $\forall i \in E, c_i(\hat{x}) = 0$.
3. $\forall i \in I, c_i(\hat{x}) \geq 0$.
4. $\forall i \in I \cup E, \hat{\lambda} \geq 0$.
5. $\forall i \in I \cup E, \hat{\lambda} \cdot c_i(\hat{x}) = 0$.

אלגוריתם לפתרון של בעיית אופטימיזציה עם תנאי kkt:

1. מציירים את האילוצים.
2. עוברים על כל סט אפשרי של אילוצים)מגדירים אותם להיות האילוצים הפעילים).
   הערה: עבור האילוצים הלא פעילים, תסמנו עבורם כופלי לגראנז' ותציבו בהם 0 (לפי תנאי 5).
3. מצאו נקודות אפשריות על בסיס האילוצים הפעילים.
4. תפתרו את מערכת המשוואות ותוודאו שהתנאים מתקיימים עבור הפתרונות שתמצאו.

דוגמה:

$$\begin{aligned} &\min_{x\in\mathbb{R}^n} \frac{1}{2}(x-2)^2 + \frac{1}{2}\left(y - \frac{1}{2}\right)^2 \\ &\text{s.t.} \frac{1}{x+1} - y - \frac{1}{4} \geq 0 && (1) \\ &x \geq 0 && (2) \\ &y \geq 0 && (3) \end{aligned}$$

נצייר את האילוצים:

מקרה I: $A(x) = \{\}$. כופלי הלגראנז' המתאימים לאילוצים שווים ל-0: $\lambda_1 = \lambda_2 = \lambda_3 = 0$.

$$\nabla_x L(\hat{x}, \hat{\lambda}) = \nabla f = \left(x - 2, y - \frac{1}{2}\right) = 0$$

מקבלים שפתרון שמקיים את תנאי (1) הוא $x = 2, y = \frac{1}{2}$. אבל אילוץ (1) לא מתקיים ולכן זה לא פתרון אפשרי.

מקרה II: $A(x)=\{1,2,3\}$ לא קיים פתרון שמפעיל את כל האילוצים ולכן ניתן לעבור למקרה הבא.

מקרה III: $A(x)=\{1,2\}$. פתרון שמפכיל את האילוצים הוא $(0,0.75)$. עבור אילוץ (3), $\lambda_3=0$.

$$\begin{aligned}L&=f-\lambda_1c_1-\lambda_2c_2\\ \nabla L&=\nabla f-\lambda_1\nabla c_1-\lambda_2\nabla c_2\\ &=\begin{pmatrix}x-2\\ y-\frac{1}{2}\end{pmatrix}-\lambda_1\begin{pmatrix}-(x+1)^{-2}\\ -1\end{pmatrix}-\lambda_2\begin{pmatrix}1\\ 0\end{pmatrix}\end{aligned}$$

נציב $(0,0.75)$:

$$\begin{pmatrix}x-2\\ \frac{3}{4}-\frac{1}{2}\end{pmatrix}-\lambda_1\begin{pmatrix}-1\\ -1\end{pmatrix}-\lambda_2\begin{pmatrix}1\\ 0\end{pmatrix}$$

כלומר נקבל את מערכת המשוואות הבאה:

$$\begin{cases}-2+\lambda_1-\lambda_2=0\\ \frac{1}{4}+\lambda_1=0\end{cases}$$

שפתרונה הוא

$$\lambda_1=-\frac{1}{4},\lambda_2=-2\frac{1}{4}$$

והפתרונות האלו לא מקיימים את התנאי הרביעי ולכן זה לא פתרון שיכול להיות. נמשיך לפתרון הבא.

מקרה IV: $A(x)=\{1,3\}$ ופתרון שמפעיל את האילוצים הוא $(3,0)$. אילוץ 2 לא פעיל ולכן $\lambda_2=0$.

$$\nabla L=\begin{pmatrix}-2\\ -\frac{1}{2}\end{pmatrix}-\lambda_1\begin{pmatrix}-(x+1)^{-2}\\ -1\end{pmatrix}-\lambda_3\begin{pmatrix}0\\ 1\end{pmatrix}$$

נציב את הנקודה $(3,0)$:

$$\begin{pmatrix}1\\ -\frac{1}{2}\end{pmatrix}-\lambda_1\begin{pmatrix}-\frac{1}{16}\\ -1\end{pmatrix}-\lambda_3\begin{pmatrix}0\\ 1\end{pmatrix}$$

ונקבל את המערכת

$$\begin{cases}1+\frac{\lambda_1}{16}=0\\ -\frac{1}{2}+\lambda_1-\lambda_3=0\end{cases}$$

ונקבל $\lambda_1=-16$ כלומר שלילי ולכן לא מקיים את האילוץ הרביעי. נמשיך למקרה הבא:

מקרה V: $A(x)=\{2,3\}$ ופתרון שמפעיל את האילוצים הוא $(0,0)$. אילוץ 1 לא פעיל ולכן $\lambda_1=0$.

$$\nabla L=\begin{pmatrix}x-2\\ y-\frac{1}{2}\end{pmatrix}-\lambda_2\begin{pmatrix}1\\ 0\end{pmatrix}-\lambda_3\begin{pmatrix}0\\ 1\end{pmatrix}$$

נציב את הנקודה $(0,0)$:

$$\begin{pmatrix}-2\\ -\frac{1}{2}\end{pmatrix}-\lambda_2\begin{pmatrix}1\\ 0\end{pmatrix}-\lambda_3\begin{pmatrix}0\\ 1\end{pmatrix}$$

ונקבל את המערכת

$$\begin{cases}-2-\lambda_2=0\\ -\frac{1}{2}-\lambda_3=0\end{cases}$$

ונקבל $\lambda_2=-2$ כלומר שלילי ולכן לא מקיים את האילוץ הרביעי. נמשיך למקרה הבא:

מקרה VI: $A(x)=\{1\}$. אילוצים 2 ו-3 לא פעילים ולכן $\lambda_2=\lambda_3=0$.

$$\nabla L=\begin{pmatrix}-2\\ -\frac{1}{2}\end{pmatrix}-\lambda_1\begin{pmatrix}-(x+1)^{-2}\\ -1\end{pmatrix}=0$$

משוואה שלישית תהיה המשוואה של האילוץ הפעיל

$$\frac{1}{x+1}-y-\frac{1}{4}=0$$

כלומר קיבלנו 3 משוואות ב-3 נעלמים:

$$\begin{cases}x-2+\frac{\lambda_1}{(x+1)^2}=0\\ y-\frac{1}{2}+\lambda_1=0\\ \frac{1}{x+1}-y-\frac{1}{4}=0\end{cases}$$

שפתרונה הוא

$$\begin{aligned} x =& 1.949 \\ y =& 0.411 \\ \lambda_1 =& 0.089 \end{aligned}$$

מקרה VII : $A(x) = \{2\}$. הפתרון הוא מהצורה $(0, y)$ ובנוסף $\lambda_1 = \lambda_3 = 0$.

$$\nabla L = \begin{pmatrix} x-2 \\ y-\frac{1}{2} \end{pmatrix} - \lambda_2 \begin{pmatrix} 1 \\ 0 \end{pmatrix} = 0$$

וממערכת המשוואות נקבל $\lambda_2 = -2$ כלומר תנאי המשפט לא מתקיימים. נעבור למקרה האחרון :

מקרה VIII : $A(x) = \{3\}$. הפתרון הוא מהצורה $(x, 0)$ ובנוסף $\lambda_1 = \lambda_2 = 0$.

$$\nabla L = \begin{pmatrix} x-2 \\ y-\frac{1}{2} \end{pmatrix} - \lambda_3 \begin{pmatrix} 0 \\ 1 \end{pmatrix} = 0$$

וממערכת המשוואות נקבל $\lambda_3 = -\frac{1}{2}$ כלומר תנאי המשפט לא מתקיימים.

נשארנו עם נקודה אחת חשודה, שהיא $(1.94, 0.089)$. בגלל ש-f היא פונקציה קמורה נקבל שהנקודה החשודה חייבת להיות מינימום (כיוון שיש רק נקודה אחת חשודה, ו-f חייבת להיות בעלת נקודת מינימום מטעמי קמירות).