Assignment 4-Optimization

שאלה 1:

נתונה בעיית האופטימיזציה הבאה:

$$\max\{x_1x_2 + x_2x_3 + x_1x_3\} \quad s.t \quad \{x_1 + x_2 + x_3 = 3\}$$

א. נמצא את הנק' הקריטית בשיטת כופלי לגרנז':

נסמן $f(x) = x_1x_2 + x_2x_3 + x_1x_3$, $C^{eq}(x) = x_1 + x_2 + x_3 - 3 = 0$

נגדיר את פונקצית הלגרנז'יאן:

$$L(x,\lambda) = f(x) + \lambda C^{eq}(x) = x_1x_2 + x_2x_3 + x_1x_3 + \lambda(x_1 + x_2 + x_3 - 3)$$

כעת נגזור ונשווה ל-0:

$$\nabla L_x(x,\lambda) = \nabla_x f(x) + \lambda \nabla_x C^{eq}(x) = \begin{bmatrix} x_2 + x_3 + \lambda \\ x_1 + x_3 + \lambda \\ x_2 + x_1 + \lambda \\ x_1 + x_2 + x_3 - 3 \end{bmatrix} = 0$$

נפתור את המערכת משוואות ונקבל:

$$\lambda = -2, x_1 = x_2 = x_3 = 1$$

כלומר הנק' החשודה היא (1,1,1)

ב. נראה שהנק (1,1,1) היא נק' מקסימום, ע"י כך שנראה:

$$\forall y \neq 0 : J^{eq}y = 0 \to y^T \nabla_x^2 L(x^*, \lambda^*) y < 0$$

יהי $J^{eq}y = 0$, $y \neq 0$. נשים לב כי $J^{eq} = [\nabla C^{eq}(x)] = [1,1,1]$

אז מתקיים: $J^{eq}y = 0 \leftrightarrow [1,1,1] * \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = 0 \leftrightarrow y_1 + y_2 + y_3 = 0$

נחשב את $\nabla_x^2 L(x^*, \lambda^*)$: $\begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix}$

נקבל כי

$$y^T \nabla_x^2 L(x^*, \lambda^*) y = [y_1, y_2, y_3] \begin{bmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} = [y_2 + y_3, y_1 + y_3, y_1 + y_2] \begin{bmatrix} y_1 \\ y_2 \\ y_3 \end{bmatrix} =$$

$$y_1y_2 + y_1y_3 + y_1y_2 + y_3y_2 + y_1y_3 + y_2y_3 = 2(y_1y_2 + y_1y_3 + y_2y_3)$$

נציב $y_1 = -y_2 - y_3$ ונקבל:

$$y^T \nabla_x^2 L(x^*, \lambda^*) y = 2(-y_2^2 - y_2y_3 - y_2y_3 - y_3^2 + y_2y_3) = -2(y_2^2 + y_2y_3 + y_3^2)$$

$$= -(y_2^2 + (y_2^2 + 2y_2y_3 + y_3^2) + y_3^2) = -((y_2 + y_3)^2 + y_2^2 + y_3^2)$$

נשים לב כי לא ייתכן ש$y_2 = y_3 = 0$ שכן זה יגרור ש $y_1 = -y_2 - y_3 = 0$ וידוע כי $y \neq 0$.

כמן כן, כל הביטויים בתוך הסוגריים הם בחזקה זוגית ולכן חיוביים,

סה"כ קיבלנו כי $y^T \nabla_x^2 L(x^*, \lambda^*) y = -((y_2 + y_3)^2 + y_2^2 + y_3^2) < 0$

שאלה 2:

נתונה בעיית האופטימיזציה הבאה:

$$\min\{(x_1 + x_2)^2 - 10(x_1 + x_2)\} \quad s.t \quad \{3x_1 + x_2 = 6,\ x_1^2 + x_2^2 \leq 5, -x_1 \leq 0\}$$

נסמן: $C_1^{eq}(x) = 3x_1 + x_2 - 6 = 0$

$C_2^{ieq}(x) = x_1^2 + x_2^2 - 5 \leq 0,$ $C_3^{ieq}(x) = -x_1 \leq 0$,

נגדיר את פונקציית הלגרנז'יאן:

$L_x(x,\lambda_1^{eq},\lambda_2^{ieq},\lambda_3^{ieq}) = f(x) + (\lambda_1^{eq})C_1^{eq}(x) + (\lambda_2^{ieq})C_2^{ieq} + (\lambda_3^{ieq})C_3^{ieq} =$
$(x_1+x_2)^2 - 10(x_1+x_2) + \lambda_1^{eq}(3x_1+x_2-6) + \lambda_2^{ieq}(x_1^2+x_2^2-5) + \lambda_3^{ieq}(-x_1)$

כעת נדרוש שיתקיימו התנאים הבאים:

1) $\nabla L_x(x,\lambda_1^{eq},\lambda_2^{ieq},\lambda_3^{ieq}) = 0$
2) $C_1^{eq} = 0$
3) $C_2^{ieq}, C_3^{ieq} \le 0$
4) $\lambda_2^{ieq}, \lambda_3^{ieq} \ge 0$

שלב 1: תחילה נחשב את $\nabla L_x(x,\lambda_1^{eq})$:

$$\nabla L_x(x,\lambda_1^{eq}) = \nabla_x f(x) + \lambda_1^{eq}\nabla_x C_1^{eq}(x) = \begin{bmatrix} 2x_1+2x_2-10 \\ 2x_2+2x_1-10 \end{bmatrix} + \lambda_1^{eq}\begin{bmatrix} 3 \\ 1 \end{bmatrix}$$
$$= \begin{bmatrix} 2x_1+2x_2-10+3\lambda_1^{eq} \\ 2x_2+2x_1-10+\lambda_1^{eq} \\ 3x_1+x_2-6 \end{bmatrix} = 0$$

נפתור את המערכת משוואות ונקבל: $\lambda_1^{eq} = 0, x_1 = 0.5, x_2 = 4.5$

נבדוק האם הנק (0.5,4.5) מקיימת את האילוצי אי-שוויון:

$(0.5)^2 + (4.5)^2 = 20.5 > 5$ ולכן לא מקיימים את האילוץ הראשון.

אז נכניס לפונקציית הלגרנז'יאן את האילוץ השני ונחשב את

$$\nabla L_x(x,\lambda_1^{eq},\lambda_2^{ieq}) = \nabla_x f(x) + \lambda_1^{eq}\nabla_x C_1^{eq}(x) + \lambda_2^{ieq}\nabla_x C_2^{ieq}(x) =$$
$$\begin{bmatrix} 2x_1+2x_2-10 \\ 2x_2+2x_1-10 \end{bmatrix} + \lambda_1^{eq}\begin{bmatrix} 3 \\ 1 \end{bmatrix} + \lambda_2^{ieq}\begin{bmatrix} 2x_1 \\ 2x_2 \end{bmatrix} = \begin{bmatrix} 2x_1+2x_2-10+3\lambda_1+2\lambda_2 x_1 \\ 2x_1+2x_2-10+\lambda_1+2\lambda_2 x_2 \\ 3x_1+x_2-6 \\ x_1^2+x_2^2-5 \end{bmatrix} = 0$$

נפתור את המערכת משוואות ונקבל 2 נקודות חשודות:

הראשונה: $x_1 = 2.174, x_2 = -0.522, \lambda_1 = 4.826, \lambda_2 = -1.79$

נפסול אותה מאחר ו- $\lambda_2 = -1.79 < 0$ וזה לא מקיים את תנאי KKT.

השנייה: $x_1 = 1.4258, x_2 = 1.722, \lambda_1 = 0.9897, \lambda_2 = 0.293$

נשים לב כי הנקודה גם עומדת באילוצי האי-שוויון השני מאחר ו $-x_1 = -1.4258 \le 0$

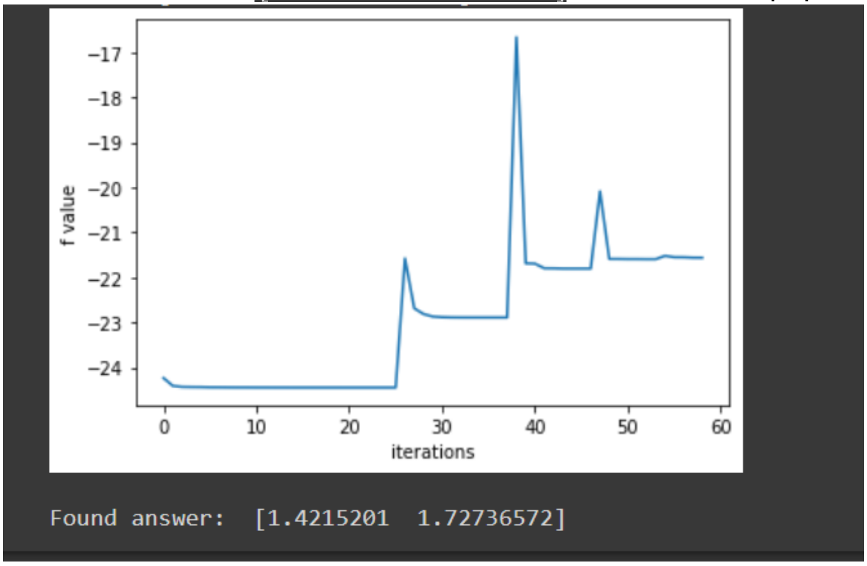
ולכן נקבל כי זוהי נק' חשודה.

נחשב את ההסיאן $\nabla L_x = \begin{bmatrix} 2+2\lambda_2 & 2 \\ 2 & 2+2\lambda_2 \end{bmatrix}$

נשים לב כי $\nabla^2 L_x = \begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} + \begin{bmatrix} 2\lambda_2 & 0 \\ 0 & 2\lambda_2 \end{bmatrix}$ מאחר ו $\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} \ge 0$ וגם
$0 < \begin{bmatrix} 0.586 & 0 \\ 0 & 0.586 \end{bmatrix} = \begin{bmatrix} 2\lambda_2 & 0 \\ 0 & 2\lambda_2 \end{bmatrix}$ ולכן $\nabla^2 L_x > 0$.
מהגדרת חיוביות, לכל $y \neq 0$, נובע כי $y^T\nabla^2 L_x y > 0$ ולכן הנק (1.4258,1.722) היא נקודת מינימום.

ג. $\min\left((x_1+x_2)^2 - 10(x_1+x_2)\right) + \mu((3x_1+x_2-6)^2 + (\max\{0, x_1^2+x_2^2-5\})^2 + (\max\{0,-x_1\})^2)$

ד. מצורף קוד: תוצאה סופית: $x^* = [1.4215201\ \ 1.72736572]$



<u>שאלה 3:</u>

א. נתונה הבעיה:

$$\min\left\{\frac{1}{2}x^THx - x^Tg\right\}\ \ s.t\ \ a \le x \le b$$

נסמן: $f(x) = \frac{1}{2}hx^2 - gx$ נגזור את f ונקבל: $f'(x) = hx - g$.

כלומר $f'(x) = hx - g = 0 \leftrightarrow x = \frac{g}{h}$.

נבדוק שהנק' החשודה $x = \frac{g}{h}$ היא נק' מינימום ע"י גזירה שנייה: $f''(x) = h$. נשים לה כי המטריצה H היא *spd* ולכן סה"כ נקבל כי $f''(x) = h > 0$ ולכן הנק' החשודה x היא נק מינימום.

נשים לב שמאחר שהפונקציה מוגדרת בתחום *[a,b]* אז נצטרך להתייחס למקרה שבו הנק החשודה x חורגת מהגבולות, ולכן היא תקבל את ערכן (שכן הנק' צריכה להישאר בתחום אך גם להיות מינימלית)

אז נקבל את הפתרון
$$\mathrm{x}^* = \begin{cases} \frac{g}{h}, & a \le \frac{g}{h} \le b \\ b, & \frac{g}{h} > b \\ a, & \frac{g}{h} < a \end{cases}$$

ב. נסמן: $f(x) = \frac{1}{2}x^THx - x^Tg$.

אז $f(x) = \frac{1}{2}x^THx - x^Tg = \frac{1}{2}\sum_{i=1}^{n}\left(\sum_{j=1}^{n}x_jH_{ji}\right)x_i - \sum_{i=1}^{n}x_ig_i = \sum_{i=1}^{n}\left(\sum_{j=1}^{n}x_jH_{ji} - g_i\right)x_i$

במקרה של *Coordinate Descent* בודקים לגבי x_i מסויים ולכן נוכל להוריד את הלולאה החיצונית ונקבל: $f(x_i) = \left(\frac{1}{2}\sum_{j=1}^{n}x_jH_{ji} - g_i\right)x_i$

מאחר וH סימטרית, אז נקבל כי

$$\forall i,j: H_{ij} = H_{ji} \rightarrow f(x_i) = \left(\sum_{j=1, j\ne i}^{n}x_jH_{ji} + \frac{1}{2}x_iH_{ii} - g_i\right)x_i =$$

$$\frac{1}{2}x_i^2 H_{ii} + \left(\sum_{j=1, j\neq i}^{n} x_j H_{ji} - g_i\right) x_i$$

אז נתאים את מה שקיבלנו לנוסחה מסעיף ב, ונקבל: $h = H_{ii}, g = \left(\sum_{j=1, j\neq i}^{n} x_j H_{ji} - g_i\right)$

נמצא את הפתרון עבור *x_i* : $f'(x_i) = H_{ii}x_i + \left(\sum_{j=1, j\neq i}^{n} x_j H_{ji} - g_i\right)$

אז נקבל: $f'(x_i) = 0 \leftrightarrow H_{ii}x_i + \left(\sum_{j=1, j\neq i}^{n} x_j H_{ji} - g_i\right) = 0 \leftrightarrow x_i = -\frac{\left(\sum_{j=1, j\neq i}^{n} x_j H_{ji} - g_i\right)}{H_{ii}}$

כלומר $x_i^* = -\frac{\left(\sum_{j=1, j\neq i}^{n} x_j H_{ji} - g_i\right)}{H_{ii}}$

ג. מצורף קוד

ד. מצורף קוד.

התוצאה שהתקבלה היא: *[5,3.880,0.825,1.741,5.889]*