הטכניון, מכון טכנולוגי לישראל
הפקולטה להנדסת חשמל

שיטות חישוביות באופטימיזציה

046197

סמסטר אביב תשפ"א 2021

מגיש:

עידו יחזקאל	204397368

תרגיל 2-

תהי $f:\mathbb{R}\to\mathbb{R}$ קמורה וגזירה ברציפות.

בעית אופטימיזציה

$$\min_x \{f(x) \mid x\geq 0\}$$

צ"ל בעזרת תנאי KKT את הטענות הבאות ולהסביר אותן.

1. אם $f'(0)>0$ אזי $x=0$ פתרון לבעיה.

הוכחה-

ראשית נשים לב כי הבעיה קמורה, פונקציית מטרה קמורה ותחום קמור. בנוסף מתקיים תנאי slater ולכן KKT הכרחיים (לדוגמא עבור $x=1$).

נחשב את הלגרנז'יאן:

$$L(x,\mu)= f(x) - \mu x$$

stationarity $\frac{\partial L}{\partial x}(x,\mu) = f'(x) - \mu = 0$

$\Rightarrow f'(x) = \mu.$

C-S: $\mu\cdot x=0$

נסתכל על הנקודה $(0,f'(0))$ ונראה היא מקיימת KKT

C-S: $\mu\cdot x\big|_{x=0} = \mu\cdot 0 = 0$ $\quad\mu= f'(0)\geq 0$ $\quad x=0\geq 0$

dual feasibility $\quad$ primal

ולכן הנקודה $\mu^*=f'(0)$; $x^*=0$ מקיימת את תנאי KKT

ומהווה פתרון לבעיה.

הפירוש הגאומטרי הוא שבגלל שהפונקציה קמורה

היא מעל המשיקה לכולה. ניתן לראות זאת מאי-שוויון הגרדיאנט:

סך $y>0$

$$f(y) \geq f(0) + y \cdot f'(0) \geq f(0)$$

ולכן $x=0$ מהווה נקודת מינימום בתחום.

2. אם $f'(0) \leq 0$ אזי כל נקודה $x \geq 0$ המקיימת $f'(x)=0$

<u>היא פתרון:</u>

נסתכל על נקודה מהצורה $(x, \mu=0)$

Stationary: $\mu=0\geq 0$ dual $x \geq 0$ primal   C-S: $0 \cdot x_{\geq 0} = 0$

$$\frac{\partial L}{\partial x} = f'(x) - \mu = 0$$

$f'(x)=0 \Rightarrow \mu=0$ ולכן כל נקודה כזאת עם $\mu=0$ מקיימת

KKT.

הפירוש הגאומטרי הוא ש $f(x)$ יורדת בתחום

ולכן אנו מחפשים נקודה שבה $f(x)$ עולה אחריה

והיא נקודת מינימום. כלומר מחפשים משהו

מהצורה:  ובגלל שקמורה אם קיים

מינימום לוקלי הוא גלובלי.

תרגיל 2:

נתונה בעיית האופטימיזציה הבאה:

$$(P) \quad \min_{x} \; x^{T} Q x$$

$$\text{s.t.} \quad x^{T} Q x + b^{T} x \geq 1$$

$$C^{T} x \geq 0$$

כאשר $Q \succ 0 \in \mathbb{R}^{n \times n}$ וכן $b, c \in \mathbb{R}^{n}$

1. הראה כי P לא קמורה.

פתרון:

$f(x)$ פונקציית מטרה היא קמורה מכיוון ש- $Q \succ 0$
וגם האילוץ האפיני ולכן נתמקד באילוץ:

$$x^{T} Q x + b^{T} x \geq 1 \Rightarrow -x^{T} Q x - b^{T} x + 1 \leq 0$$

עבור המקרה החד מימדי:

$$Q = 1 \; ; \; b = 0 \Rightarrow -x^{2} + 1 \leq 0$$

$$(1-x)(1+x) \leq 0$$

$x \leq -1 \cup x \geq 1$ איחוד בין 2 חצי מרחב

לא קמורה לדוגמא לפי הגדרה:

$$\frac{1}{2}(-1) + \left(1 - \frac{1}{2}\right)(1) \quad -\frac{1}{2} + \frac{1}{2} = 0 \notin x \leq -1 \cup x \geq 1$$

ולכן אילוץ זה לא קמור והבעיה כולה
לא קמורה.

② מציאת הבעיה הדואלית:

ראשית נסדר את הבעיה:

$$\min \quad x^TQx$$

$$\text{s.t} \quad -x^TQx - b^Tx + 1 \le 0$$

$$-c^Tx \le 0$$

נכתוב את הלגרנג'יאן:

$$L(x, \lambda_1, \lambda_2) = x^TQx - \lambda_1 x^TQx - \lambda_1 b^T x + \lambda_1 - \lambda_2 c^T x$$

$$= (1-\lambda_1)\, x^TQx - (\lambda_1 b + \lambda_2 c)^T x + \lambda_1$$

<u>נפריד למקרים:</u>

עבור $0 \le \lambda_1 < 1$, $\lambda_2 > 0$ הלגרנג'יאן קמור ונקודת

מינימום היא:

$$\frac{\partial L(x, \lambda_1, \lambda_2)}{\partial x} = 2(1-\lambda_1)Qx - (\lambda_1 b + \lambda_2 c) = 0$$

$$x = Q^{-1} \frac{1}{2(1-\lambda_1)} (\lambda_1 b + \lambda_2 c)$$

$$q(\lambda_1, \lambda_2) = \cancel{(1-\lambda_1)} \frac{1}{\cancel{2(1-\lambda_1)}} (\lambda_1 b + \lambda_2 c)^T Q^{-1} Q Q^{-1} \frac{(\lambda_1 b + \lambda_2 c)}{2(1-\lambda_1)}$$

$$- (\lambda_1 b + \lambda_2 c)^T \cdot \frac{Q^{-1}}{2(1-\lambda_1)} (\lambda_1 b + \lambda_2 c) + \lambda_1$$

$$= -\frac{(\lambda_1 b + \lambda_2 c)^T Q^{-1} (\lambda_1 b + \lambda_2 c)}{4(1-\lambda_1)} + \lambda_1$$

עבור $\lambda_1 = 1$

$$L(\lambda_1,\lambda_2) = -(b+\lambda_2 c)^T x + 1$$

אם קיים עבורו $-(b+\lambda_2 c)_i < 0$ ניתן לבחור x כך ש- $L(x,\lambda_2) \longrightarrow -\infty$

ולכן אם $-(b+\lambda_2 c)x \geq 0$ הערך המינימלי לה $x=0$ ולכן הפונקציה הדואלית היא $\lambda_1 = 1$ ומכיוון שאין הגבלה על x נדרש שהביטוי יתאפס כלומר

$$\lambda_2 c = -b.$$

<u>אחרת</u> הלגרנז'יאן יכול להתבדר לאינסוף:

$$q(\lambda_1,\lambda_2) = \begin{cases} \lambda_1 - \dfrac{(\lambda_1 b+\lambda_2 c)^T Q^{-1}(\lambda_1 b+\lambda_2 c)}{4(1-\lambda_1)} & 0\leq\lambda_1<1,\ \lambda_2\geq 0 \\ 1 & \lambda_1=1,\ \lambda_2 c=-b \\ -\infty & \text{else.} \end{cases}$$

$$\max\ q(\lambda_1,\lambda_2)$$

s.t.

$$0\leq\lambda_1\leq 1,\ \lambda_2\geq 0 \quad || \quad \lambda_1=1,\ \lambda_2 c=-b$$

③ נסמן $\beta = b^T Q^{-1} b$ ג״ כי ערך אופטימלי

של הבעיה חסום מלמעלה על ידי:

$$\frac{2+\beta-\sqrt{\beta(\beta+4)}}{2}.$$

<u>פתרון:</u>

לפי דואליות חלשה:

$$p^* \geq q(\lambda_1, \lambda_2) \qquad \forall (\lambda_1, \lambda_2) \in dom(q)$$

ולכן עבור $\lambda_2 = 0$; $0 \leq \lambda_1 < 1$ (נמצא ערך מקסימלי של $q(\lambda_1, 0)$:

$$q(\lambda_1, 0) = \frac{-\lambda_1 b^T Q^{-1} \lambda_1 b}{4(1-\lambda_1)} + \lambda_1 = \frac{\beta}{4} \cdot \frac{-\lambda_1^2}{(1-\lambda_1)} + \lambda_1$$

$$\max q(\lambda_1, 0) \Rightarrow \frac{\partial q(\lambda_1)}{\partial \lambda_1} = \frac{\beta}{4} \cdot \left(\frac{-2\lambda_1 (1-\lambda_1) + \lambda_1^2 \cdot -(1)}{(1-\lambda_1)^2} \right) + 1$$

$$\frac{\beta}{4} \left(\frac{-2\lambda_1 + 2\lambda_1^2 - \lambda_1^2}{(1-\lambda_1)^2} \right) + 1 = \frac{\beta}{4} \left(\frac{\lambda_1^2 - 2\lambda_1}{(1-\lambda_1)^2} \right) + 1$$

$$\cancel{\frac{\beta}{4} \lambda_1 \left(\frac{\lambda_1 - 2}{(1-\lambda_1)^2} \right) + \lambda_1 = 0} \qquad \cancel{\frac{\frac{\beta}{4}\lambda_1 - \frac{\beta}{2} + \lambda_1 (1-\lambda_1)^2}{(1-\lambda_1)^2} = 0}$$

$$\cancel{\lambda_1 \left(\frac{\beta}{4} \frac{\lambda_1 - 2}{(1-\lambda_1)^2} + \lambda_1 \right) = 0}$$

$$\frac{\lambda_1^2\beta - 2\beta\lambda_1 + 4(1-\lambda_1)^2}{4(1-\lambda_1)^2} = 0$$

$$\frac{\beta\lambda_1^2 - 2\beta\lambda_1 + 4 - 8\lambda_1 + 4\lambda_1^2}{4(1-\lambda_1)^2}$$

$\Leftrightarrow$

$$(1+\beta)\lambda_1^2 - 2(\beta+4)\lambda_1 + 4 = 0$$

$$\lambda_{1_{1,2}} = \frac{2(\beta+4) \pm \sqrt{4(\beta+4)^2 - 4\cdot 4(4+\beta)}}{2(\beta+4)}$$

$$= \frac{2(\beta+4) \pm \sqrt{(\beta+4)(4\beta+16-16)}}{2(\beta+4)}$$

$$= 1 \pm \frac{\sqrt{(\beta+4)\beta}}{\beta+4} = 1 \pm \sqrt{\frac{\beta}{\beta+4}}$$

מכיוון שאנו מחפשים פתרון בתחום $0 \le \lambda_1 < 1$

אז $\lambda_1 = 1 - \sqrt{\frac{\beta}{\beta+4}}$

ונציב ב- $q(\lambda_1)$

$$q\left(1 - \sqrt{\frac{\beta}{\beta+4}}\right) = \frac{\beta}{4} \cdot \frac{-\left(1 - \frac{\sqrt{\beta}}{\sqrt{\beta+4}}\right)^2}{\left(1 - 1 + \sqrt{\frac{\beta}{\beta+4}}\right)} + 1 - \sqrt{\frac{\beta}{\beta+4}}$$

$$= -\frac{\beta}{4}\,\frac{\left(1-2\sqrt{\frac{\beta}{\beta+4}}+\frac{\beta}{\beta+4}\right)}{\frac{\sqrt{\beta}}{\sqrt{\beta+4}}} + 1 - \sqrt{\frac{\beta}{\beta+4}}$$

$$= -\frac{1}{4}\cdot\sqrt{\beta+4}\,\sqrt{\beta}\left(1-2\sqrt{\frac{\beta}{\beta+4}}+\frac{\beta}{\beta+4}\right)+1-\sqrt{\frac{\beta}{\beta+4}}$$

$$= -\frac{\sqrt{\beta(\beta+4)}}{4}+\frac{\beta}{2}-\frac{\beta\sqrt{\beta}}{4\sqrt{\beta+4}}+1-\sqrt{\frac{\beta}{\beta+4}}$$

$$= \frac{-\sqrt{\beta}(\beta+4)+2\beta\sqrt{\beta+4}-\beta\sqrt{\beta}+4\sqrt{\beta+4}-4\sqrt{\beta}}{4\sqrt{\beta+4}}$$

$$= \frac{-\sqrt{\beta}(\beta+4+\beta+4)+(2\beta+4)\sqrt{\beta+4}}{4\sqrt{\beta+4}}$$

$$= \frac{-\sqrt{\beta}\,2(\beta+4)+2(\beta+2)\sqrt{\beta+4}}{4\sqrt{\beta+4}}$$

$$= \frac{\beta+2-\sqrt{\beta(\beta+4)}}{2}$$

ולכן עבור $\lambda_2=0$ ו- $\lambda_1=1-\sqrt{\frac{\beta}{\beta+4}}<1$ כך ש-$(\lambda_1,\lambda_2)\in$
מתקיים $q=\frac{\beta+2-\sqrt{\beta(\beta+4)}}{2}$ ולפי דואליות חלשה

$$p^* \geq q \quad ∎$$

תרגיל 3:

יהי $0 \prec Q \in \mathbb{R}^{n\times n}$ ויהיו $a,b\in\mathbb{R}^n$ וקטורים
ו- $0<c\in\mathbb{R}$, בעיית אופטימיזציה:

$$\min_{x} \frac{x^TQx}{x^Ta}$$

$$\text{s.t.} \quad x^Tb=c$$
$$x^Ta>0$$

נתון כי הבעיה פיסבילית ומקיימת תנאי Slater

כמו כן:

$$\alpha=a^TQ^{-1}a$$
$$\beta=b^TQ^{-1}b$$
$$\gamma=a^TQ^{-1}b=b^TQ^{-1}a$$
$$\alpha\beta>\gamma^2$$

הוכח כי הבעיה קמורה ומצא ערך אופטימלי.

פתרון: נוסיף משתנה עזר $y=x^Ta$ ונקבל מחדש:

$$\min_{x,y} \frac{x^TQx}{y}$$

$$x^Tb=c \quad y=x^Ta$$
$$y>0$$

נראה קמירות:

נסמן $g(x)=x^TQx$ והיא קמורה כי $Q \succ 0$

מתרגיל ראינו כי עבור $y>0$ מתקיים ש-

$$h(x,y)=y\cdot g\left(\frac{x}{y}\right)=y\cdot\frac{x^T}{y}Q\frac{x}{y}=\frac{x^TQx}{y}$$

קמורה. ולכן פונקציית מטרה קמורה ואילוצים אפיניים.

ונפתור על ידי שנניח ש-y קבוע ונמצא את הלגרנז'יאן

$$L(x,\mu_1,\mu_2)=\frac{1}{y}x^TQx+\mu_1(x^Tb-c)+\mu_2(x^Ta-y)$$

$$=\frac{1}{y}x^TQx+x^T(\mu_1 b+\mu_2 a)-\mu_1 c-\mu_2 y$$

זוהי פונקציה ריבועית ב-x והיא קמורה ולכן נשווה את הגרדיאנט ל-0.

$$\frac{\partial L}{\partial x}=\frac{2}{y}Qx+(\mu_1 b+\mu_2 a)=0$$

$$x^*=-\frac{1}{2}y\,Q^{-1}(\mu_1 b+\mu_2 a).$$

אילוצי השוויון:

$$x^{*T} b = c$$

$$\Rightarrow -\frac{\ell}{2} y \left(\mu_1 b^T + \mu_2 a^T\right) Q^{-1} b = c$$

$$\mu_1 b^T Q^{-1} b + \mu_2 a^T Q^{-1} b = -\frac{2c}{y}$$

$$\Rightarrow \mu_1 \beta + \mu_2 \gamma = \frac{-2c}{y} \quad (1)$$

$$x^{*T} a = y$$

$$-\frac{\ell}{2} y \left(\mu_1 b^T + \mu_2 a^T\right) Q^{-1} a = y \quad /: y > 0$$

$$\mu_1 b^T Q^{-1} a + \mu_2 a^T Q^{-1} a = -2$$

$$\mu_1 \gamma + \mu_2 \alpha = -2 \quad (2)$$

$$\Rightarrow \mu_2 = \frac{-2 - \mu_1 \gamma}{\alpha}$$

$$\mu_1 \beta \; \frac{-2\gamma - \mu_1 \gamma^2}{\alpha} = -\frac{2c}{y} \quad / \cdot \alpha y$$

$$\mu_1 \alpha \beta y - 2\gamma y - \mu_1 \gamma^2 y = -2c\alpha$$

$$\mu_1 y (\alpha\beta - \gamma^2) = 2\gamma y - 2c\alpha$$

$$\mu_1 = \frac{2\gamma y - 2c\alpha}{y(\alpha\beta - \gamma^2)}$$

$$\Rightarrow \mu_2 = \frac{-2 - \dfrac{\gamma(\alpha\gamma y - 2c\alpha)}{y(\alpha\beta - \gamma^2)}}{\alpha}$$

$$= \frac{-2y\cancel{\alpha}\beta \cancel{+2y\gamma^2} \cancel{-2\gamma^2 y} + 2c\alpha\gamma}{\cancel{\alpha} y(\alpha\beta - \gamma^2)}$$

$$= \frac{-2y\beta + 2c\gamma}{y(\alpha\beta - \gamma^2)}$$

$$g(y) = \frac{x^{*T} Q x^{*}}{y} = \frac{-\frac{\ell}{2}\cancel{y}(\mu_1 b^T + \mu_2 a^T) Q^{-1} Q \left(-\frac{\ell}{2} y Q^{-1}(\mu_1 b + \mu_2 a)\right)}{\cancel{y}}$$

$$+\frac{\ell}{4} y(\mu_1 b^T + \mu_2 a^T) Q^{-1}(\mu_1 b + \mu_2 a)$$

$$= \frac{\ell}{4} y\left(\mu_1^2 b^T Q^{-1} b + \mu_1\mu_2 a^T Q^{-1} b + \mu_1\mu_2 b^T Q^{-1} a + \mu_2^2 a^T Q^{-1} a\right)$$

$$= \frac{\ell}{4} y\left(\mu_1^2 \beta + \mu_1\mu_2\gamma + \mu_1\mu_2\gamma + \mu_2^2 \alpha\right)$$

$$= \frac{\ell}{4} y\Big(\mu_1 \underbrace{(\mu_1\beta + \mu_2\gamma)}_{-\frac{2c}{y}} + \mu_2 \underbrace{(\mu_1\gamma + \mu_2\alpha)}_{-2}\Big)$$

על פי אילוצי שוויון:

$$= \frac{\ell}{4} y \cdot \left(\mu_1 \cdot \frac{-2c}{y} - 2\mu_2\right) = -\frac{\ell}{2}(\mu_1 c + \mu_2 y)$$

$$g(y) = -\frac{l}{2}\left(\mu_1 c + \mu_2 y\right)$$

$$= -\frac{l}{2} \cdot \left(\frac{(2\gamma y - 2c\alpha)c}{y(\alpha\beta - \gamma^2)} + \frac{y(-2y\beta + 2c\gamma)}{y(\alpha\beta - \gamma^2)} \right)$$

$$= -\frac{l}{2} \cdot \left(\frac{2\gamma c y - 2c^2\alpha - 2y^2\beta + 2yc\gamma}{y(\alpha\beta - \gamma^2)} \right)$$

$$= \frac{y^2\beta - 2c\gamma y + c^2\alpha}{y(\alpha\beta - \gamma^2)}$$

כעת נפתור את בעיית האופטימיזציה

$$\min_y g(y)$$
$$y > 0$$

נשים לב כי $y \in \mathbb{R}$ ולכן נגזור:

$$g'(y) \frac{(2y\beta - 2c\gamma)y\cancel{(\alpha\beta - \gamma^2)} - (y^2\beta - 2c\gamma y + c^2\alpha)\cancel{(\alpha\beta - \gamma^2)}}{y^2(\alpha\beta - \gamma^2)^{\cancel{2}}} = 0$$

$$\frac{2y^2\beta - 2c\gamma y - y^2\beta + 2c\gamma y - c^2\alpha}{y^2(\alpha\beta - \gamma^2)} = 0$$

$$\Rightarrow y^2\beta - c^2\alpha = 0 \quad \Rightarrow y^2 = \frac{c^2\alpha}{\beta}$$

$$y>0 \Rightarrow y^* = c\sqrt{\frac{\alpha}{\beta}}$$

סיכום:

השתמשתי במשתנה עזר $y=x^Ta$ ופתרתי
את הבעיה המקורית בעזרתו
לאחר מכן ניסחתי בעיה חדשה שבה $g(y)$
מוגדרת תחת אילוץ ש- $y>0$

וקיבלנו כי $y^*=c\sqrt{\frac{\alpha}{\beta}}$

$$\Rightarrow x^{*T}a = c\sqrt{\frac{\alpha}{\beta}}$$

תרגיל 4-

נתונה בעיית האופטימיזציה:

$$\min_{x} \left\{\frac{1}{3}x^3 - ax\right\}$$

כאשר $x, a > 0$, a פרמטר נתון.

1. צ"ל הפתרון הינו $x^* = \sqrt{a}$

הוכחה:

נגזור ונשווה ל-0:

$$f'(x) = \frac{3}{3}x^2 - a \quad ; \quad f'(x) = 0 \Rightarrow x^2 - a = 0$$

$$x = \pm\sqrt{a}$$

נבחר כי $x>0$ ולכן $x = \sqrt{a}$

$$f''(x) = 2x\Big|_{\sqrt{a}} = 2\sqrt{a} > 0$$

ולכן מינימום.

2. מצא ביטוי לאיטרציה ה-k כתלות באיטרציה ה-k-1.

$$x_k = x_{k-1} - f''(x_{k-1})^{-1} f'(x_{k-1})$$

$$= x_{k-1} - \frac{1}{2x_{k-1}} \cdot (x_{k-1}^2 - a)$$

$$= \frac{2x_{k-1}^2 - x_{k-1}^2 + a}{2x_{k-1}} = \frac{x_{k-1}^2 + a}{2x_{k-1}}$$

3. הוכח כי אם $x_{k-1} > \sqrt{a}$ אז $x_{k-1} > x_k > \sqrt{a}$

הוכחה:

$$x_k = \frac{x_{k-1}^2 + a}{2x_{k-1}} < x_{k-1} \quad / \cdot 2x_{k-1}$$

$$\Leftrightarrow$$

$$x_{k-1}^2 + a < 2x_{k-1}^2$$

$$\Leftrightarrow \quad a < x_{k-1}^2$$

$$\Leftrightarrow \quad \sqrt{a} < x_{k-1}$$

ולכן אם $x_{k-1} > \sqrt{a}$ אז $x_k < x_{k-1}$

$$\frac{x_{k-1}^2 + a}{2x_{k-1}} > \sqrt{a} \Leftrightarrow x_{k-1}^2 + a > 2x_{k-1}\cdot\sqrt{a}$$

$$\Leftrightarrow x_{k-1}^2 - 2x_{k-1}\sqrt{a} + a > 0$$

$$\Leftrightarrow (x_{k-1} - \sqrt{a})^2 > 0$$

ולכן ביטוי אמת ולכן מתקיים.

---

4. נניח כי $x_k = (1+\delta_k)\sqrt{a}$

$$\delta_k = \frac{\delta_{k-1}^2}{2(1+\delta_{k-1})}$$

$$(1+\delta_k)\cancel{\sqrt{a}} = \frac{(1+\delta_{k-1})^2\cancel{\sqrt{a}^2} + \cancel{a}}{2(1+\delta_{k-1})\cancel{\sqrt{a}}}$$

$$1+\delta_k = \frac{1+2\delta_{k-1} + \delta_{k-1}^2 + 1}{2(1+\delta_{k-1})} = \frac{2(1+\delta_{k-1}) + \delta_{k-1}^2}{2(1+\delta_{k-1})}$$

$$1+\delta_k = 1+\frac{\delta_{k-1}^2}{2(1+\delta_{k-1})} \Leftrightarrow \delta_k = \frac{\delta_{k-1}^2}{2(1+\delta_{k-1})} \blacksquare$$

⑤ הוכח כי $\delta_k \le \frac{\delta_{k-1}}{2}$

$$\frac{\delta_{k-1}^2}{2(1+\delta_{k-1})} \le \frac{\delta_{k-1}}{2}$$

$$\Leftrightarrow \delta_{k-1}^2 \le \delta_{k-1}+\delta_{k-1}^2$$

$$\Leftrightarrow \delta_{k-1} \ge 0$$

ושים לב כי $\delta_{k-1} \ge 0$ מהנתון ש- $X_{k-1} > \sqrt{a}$

$$(1+\delta_{k-1})\sqrt{a} > \sqrt{a} \Leftrightarrow 1+\delta_{k-1} > 1 \Leftrightarrow \delta_{k-1} > 0.$$

$$X_k = (1+\delta_k)\sqrt{a} \le \left(1+\frac{\delta_{k-1}}{2}\right)\sqrt{a} \le \dots \le \left(1+\frac{\delta_0}{2^k}\right)\sqrt{a}$$

$$\xrightarrow[k\to\infty]{} (1+0)\sqrt{a} = \sqrt{a}.$$

6. הוכח כי $\delta_k \le \frac{\delta_{k-1}^2}{2}$

הוכחה -

$$\frac{\delta_{k-1}^2}{2(1+\delta_{k-1})} \le \frac{\delta_{k-1}^2}{2} \Leftrightarrow \frac{1}{1+\delta_{k-1}} \le 1 \Leftrightarrow \delta_{k-1} \ge 0$$

7. נניח כי $a=2$ וניחוש ראשוני הוא $x_0=2$.

עבור כל אחד מהביטויים לשגיאה הכפלית (1), (2)

ו-(3) חשבו כמה איטרציות דרושות על מנת

לקבל דיוק של 50 ספרות לאחר הנקודה.

$$X_0 = 2 = \sqrt{2}(\delta_0+1)$$

$$\sqrt{2} = \delta_0 + 1$$

$$\Rightarrow \delta_0 = \sqrt{2}-1$$

הדרישה על X_k:

$$X_k - \sqrt{a} \le 10^{-50}$$

$$\Rightarrow \sqrt{a}(\delta_k+1) - \sqrt{a} \le 10^{-50}$$

$$\sqrt{a}\,\delta_k \le 10^{-50}$$

$$\delta_k \le \frac{10^{-50}}{\sqrt{a}} = \frac{10^{-50}}{\sqrt{2}}$$

עבור (2):

$$\delta_k \le \frac{\delta_{k-1}}{2} \le \frac{\delta_{k-2}}{4} \le \dots \le \frac{\delta_0}{2^k}$$

$$\frac{\delta_0}{2^k} \le \frac{10^{-50}}{\sqrt{2}} \iff \frac{\sqrt{2}(\sqrt{2}-1)}{10^{-50}} \le 2^k \qquad k \ge 165.3$$

$$k = 166.$$

לטור (3):

$$\delta_k \le \frac{\delta_{k-1}^2}{2} \le \frac{\left(\frac{\delta_{k-2}^2}{2}\right)^2}{2} = \frac{\delta_{k-2}^{2^2}}{2^3} \le \ldots \le 2^{-2^k+1} \cdot \delta_0^{2^k}.$$

$$2^{-2^k+1} \cdot \delta_0^{2^k} \le \frac{10^{-50}}{\sqrt{2}} \quad / \log_2$$

$$\left(-2^k+1+2^k\log(\delta_0)\right) \le -50\log(10) - \log(\sqrt{2})$$

$$\left(-2^k+2^k\log(\delta_0)\right) \le -50\log(10) - \frac{1}{2} - 1$$

$$2^k \ge \frac{50\log(10) - 1.5}{1 - \log(\sqrt{2}-1)}$$

$$2^k \ge 72.46$$

$$k \ge 6.2$$

$$\boxed{k = 7}$$

קוד מצורף.

סעיף 7:

קוד:

```python
import numpy as np

a = np.sqrt(2)


def _stop_condition(current_delta):
    global a
    return current_delta <= (1e-50) / a


def _next_delta_1(current_delta):
    return (current_delta ** 2) / (2 * (1 + current_delta))


def _next_delta_2(current_delta):
    return current_delta / 2


def _next_delta_3(current_delta):
    return (current_delta ** 2) / 2


def _evaluations(update_function, delta_0):
    k = 0
    current_delta = delta_0
    while not _stop_condition(current_delta):
        k += 1
        current_delta = update_function(current_delta)
    return k


if __name__ == '__main__':
    delta_0 = a - 1
    print("1st expression reached bound in {} iterations".format(_evaluations(_next_delta_1, delta_0)))
    print("2nd expression reached bound in {} iterations".format(_evaluations(_next_delta_2, delta_0)))
    print("3rd expression reached bound in {} iterations".format(_evaluations(_next_delta_3, delta_0)))
```

פלט, מספר איטרציות לכל חסם:

```
1st expression reached bound in 7 iterations
2nd expression reached bound in 166 iterations
3rd expression reached bound in 7 iterations
```

תרגיל 5-

$$(P_1)\quad \min_{x,y} \{x^2+y^2+5x-10y \mid |x-y/2-2|\le 1\}$$

$$(P_2)\quad \min_{x,y} \{(x-3)^2+y^2+3x-2y \mid (x-1)^2+(y-2)\le 9;\ x\ge 2,\ y\ge 0\}$$

א. הראו כי הבעיות קמורות.

הוכחה:

עבור P_1: תחום הבעיה האילוצים הוא קמור כי הוא חיתוך בין 2 חצאי מרחב.

$$f_1(x,y)=x^2+y^2+5x-10y$$

$$\nabla f_1(x,y)=\begin{bmatrix}\nabla_x f_1(x,y)\\ \nabla_y f_1(x,y)\end{bmatrix}=\begin{bmatrix}2x+5\\ 2y-10\end{bmatrix}$$

$$\nabla^2 f_1(x,y)=\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix}\succ 0$$

בגלל ערכים עצמיים חיוביים.

$$\Rightarrow \nabla^2 f_1(x,y)\succ 0$$

ולכן הבעיה P_1 קמורה.

הבעיה (ק$_2$):

תחום האילוץ קמור כי הוא חיתוך בין חצאי מרחב ומעגל, כל אחד מהתחומים קמור ולכן בפרט החיתוך.

$$f_2(x,y)=(x-3)^2+y^2+3x-2y$$

$$\nabla f_2(x,y)=\begin{bmatrix}2(x-3)+3\\ 2y-2\end{bmatrix}$$

$$\nabla^2 f_2(x,y)=\begin{bmatrix}2 & 0\\ 0 & 2\end{bmatrix}$$

שכן ההסיאן מוגדר להיות PD ולכן פ. מטרה קמורה ולכן כל הבעיה קמורה.

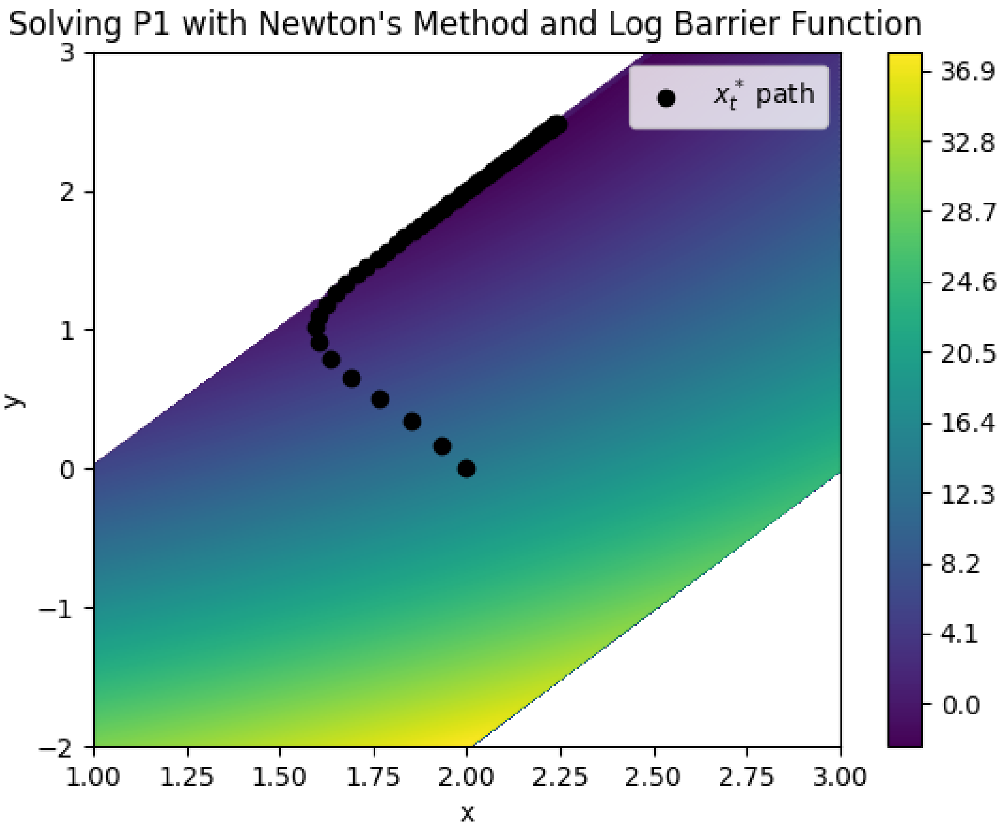
שאלה 5 סעיפים 2,3:

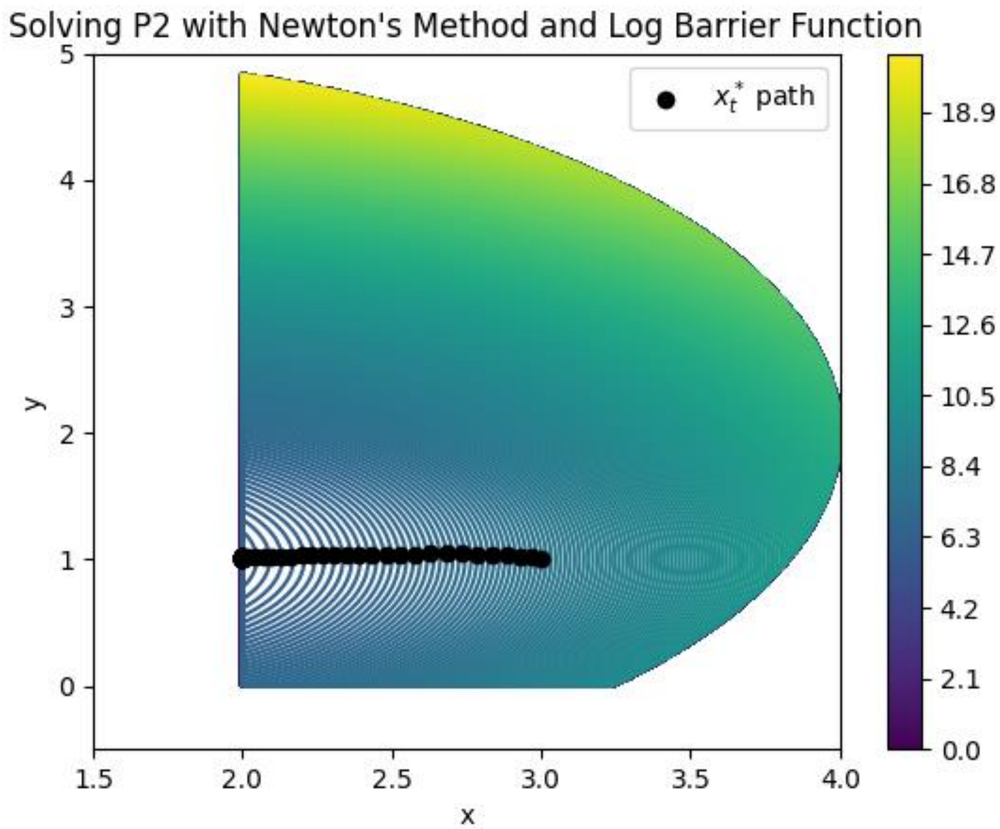
קוד:

```python
import numpy as np
from matplotlib import pyplot as plt


class MinimizationProblem:
    def __init__(self, name: str, number_constraints):
        self.name: str = name
        self.number_constraints: int = number_constraints
        self.target_func = None
        self.dLdx = self.dLdy = self.ddLdxdy = self.ddLddx = self.ddLddy = None
        self.constrains = None
        self.starts_vector = None

    def gradient(self, x, y, t):
        return np.array([self.dLdx(x, y, t), self.dLdy(x, y, t)])

    def hessian(self, x, y, t):
        hess = np.array([[self.ddLddx(x, y, t), self.ddLdxdy(x, y, t)], [self.ddLdxdy(x, y, t), self.ddLddy(x, y, t)]])
        assert hess.shape == (2, 2)
        return hess


def solve_inner(current_x_vec, t, p:MinimizationProblem, num_iter=5):
    x_vec = current_x_vec
    for i in range(num_iter):
        x_vec = x_vec - 0.01 * np.linalg.inv(p.hessian(x_vec[0], x_vec[1], t)) @ p.gradient(x_vec[0], x_vec[1], t)
    return x_vec


def solve_outer(mu, t_0, epsilon, p:MinimizationProblem):
    x_vecs = [p.starts_vector]
    t = t_0
    while p.number_constraints / t > epsilon:
        next_vec = solve_inner(x_vecs[-1], t, p)
        x_vecs.append(next_vec)
        t *= mu
    return x_vecs
```

```python
def init_p1():
    p_1_problem = MinimizationProblem("P1", 2)
    p_1_problem.constrains = lambda x, y: abs(x - 0.5 * y - 2) <= 1
    p_1_problem.target_func = lambda x, y: (x ** 2) + (y ** 2) + 5 * x - 10 * y
    p_1_problem.dLdx = lambda x, y, t: t * (2 * x + 5) - 1 / (x - 0.5 * y - 3) - 1 / (x - 0.5 * y - 1)
    p_1_problem.ddLddx = lambda x, y, t: 2 * t + 1 / ((x - 0.5 * y - 3) ** 2) + 1 / ((x - 0.5 * y - 1) ** 2)
    p_1_problem.dLdy = lambda x, y, t: t * (2 * y - 10) + 0.5 / (x - 0.5 * y - 3) + 0.5 / (x - 0.5 * y - 1)
    p_1_problem.ddLddy = lambda x, y, t: 2 * t + 0.25 / ((x - 0.5 * y - 3) ** 2) + 0.25 / ((x - 0.5 * y - 1) ** 2)
    p_1_problem.ddLdxdy = lambda x, y, t: -0.5 / ((x - 0.5 * y - 3) ** 2) - 0.5 / ((x - 0.5 * y - 1) ** 2)
    p_1_problem.starts_vector = np.array([2, 0])
    return p_1_problem


def init_p2():
    p_2_problem = MinimizationProblem("P2", 3)
    p_2_problem.constrains = lambda x, y: ((x - 1) ** 2 + (y - 2) ** 2 <= 9) * (x >= 2) * (y >= 0)
    p_2_problem.target_func = lambda x, y: ((x - 3) ** 2) + (y ** 2) + 3 * x - 2 * y
    p_2_problem.dLdx = lambda x, y, t: t * (2 * x - 3) + 2 * (x - 1) / (-((x - 1) ** 2) - ((y - 2) ** 2) + 9) - 1 / (x - 2)
    p_2_problem.dLdy = lambda x, y, t: t * (2 * y - 2) + 2 * (y - 2) / (-((x - 1) ** 2) - ((y - 2) ** 2) + 9) - 1 / y
    p_2_problem.ddLdxdy = lambda x, y, t: 4 * (x - 1) * (y - 2) / ((-((x - 1) ** 2) - ((y - 2) ** 2) + 9) ** 2)
    p_2_problem.ddLddx = lambda x, y, t: 2 * t + 2 / (-((x - 1) ** 2) - ((y - 2) ** 2) + 9) + 4 * ((x - 1) ** 2) / (
            (-((x - 1) ** 2) - ((y - 2) ** 2) + 9) ** 2) \
                                         + 1 / ((x - 2) ** 2)
    p_2_problem.ddLddy = lambda x, y, t: 2 * t + 2 / (-((x - 1) ** 2) - ((y - 2) ** 2) + 9) + 4 * ((y - 2) ** 2) / (
            (-((x - 1) ** 2) - ((y - 2) ** 2) + 9) ** 2) \
                                         + 1 / (y ** 2)
    p_2_problem.starts_vector = np.array([3, 1])

    return p_2_problem


if __name__ == '__main__':
    p_1 = init_p1()
    p_2 = init_p2()

    mu = 1.5
    t0 = 1
    epsilon = 1e-10
    n = 500
    x_values_p1 = solve_outer(mu, t0, epsilon, p_1)
    x_values_p2 = solve_outer(mu, t0, epsilon, p_2)
    plot_cont(p_1, xlims=[1, 3], ylims=[-2, 3], x_values=x_values_p1, fig_path="p1.png")
    plot_cont(p_2, xlims=[1.5, 4], ylims=[-0.5, 5], x_values=x_values_p2, fig_path="p2.png")
```

פלט של הגרפים:


Solving P1 with Newton's Method and Log Barrier Function
x_t^* path
y
x
1.00
1.25
1.50
1.75
2.00
2.25
2.50
2.75
3.00
3
2
1
0
−1
−2
36.9
32.8
28.7
24.6
20.5
16.4
12.3
8.2
4.1
0.0


Solving P2 with Newton's Method and Log Barrier Function
x_t^* path
y
x
1.5
2.0
2.5
3.0
3.5
4.0
5
4
3
2
1
0
18.9
16.8
14.7
12.6
10.5
8.4
6.3
4.2
2.1
0.0