HW4 גרימוביץ' יקטרינה 094197 314092826

98

תרגיל 1

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

$$\min_x \begin{cases} f(x) \\ g(x) = -x \end{cases}$$

$\nabla g(x) = -1$ לכל x

תהיה $f: R \to R$ פונקציה קמורה וגזירה ברציפות.
יש להוכיח בעזרת תנאי KKT את הטענות הבאות:
ולתת פירוש גאומטרי.

(1) אם $f'(0) > 0$ אז $x=0$ הוא פתרון לבעיית האופטימיזציה
נגדיר לגרנז'יאן $\mathcal{L}(x,\lambda) = f(x) + \lambda g(x)$ (תקף לשני הסעיפים)
כאשר $\lambda \geq 0$. נגזור אותו $\nabla \mathcal{L}(x,\lambda) = f'(x) - \lambda$
נניח כי $x_0 = 0$ היא נקודה פיזיבילית מקיימת KKT, לכן
היא חייבת לקיים:
1) $\nabla \mathcal{L}(0,\lambda) = f'(0) - \lambda = 0 \to f'(0) = \lambda > 0$ מתקיים
2) $\lambda g(0) = \lambda \cdot (-0) = 0$ מתקיים
3) $g(0) = -0 \leq 0$ מתקיים

לכן הוכחנו כי נקודה $x_0 = 0$ היא אכן נקודת KKT לכן היא פתרון האופטימלי לבעיה.
נשים לב כי מתוך הנתון כי $f(x)$ היא פונקציה קמורה, אז בהכרח מ $x=0$ ומעלה
כי חייבת להיות לא יורדת (ו-$x=0$ הוא מינימום גלובלי ~~ו~~ יחיד)
אם נניח כי קיימת נקודה $x_0 < x_1$ כלשהי שהשיפוע מה אפס ~~...~~ אז נקבל נקודת
קיצון נוספת מקומית (מקסימום) ← נקבל כי f' קעורה ← סותר את הנתון.
לכן אכן השיפוע עולה ← f היא מונוטונית עולה ולכן נקודת $x=0$ היא מינימום

(2) אם $f'(0) \leq 0$ אז לכל נקודת $x \geq 0$ מקיימת $f'(x) = 0$ הוא פתרון לבעיית אופטימיזציה.
נגדיר כי קיים $x^* \geq 0$ כך ש $f'(x^*) = 0$ שהוא מקיים תנאי KKT, נבדוק:
כאשר $\lambda \geq 0$
1) $\nabla \mathcal{L}(x^*) = f'(x^*) - \lambda = 0 \to \lambda = 0$ מתקיים עבור $\lambda = 0$
2) $\lambda g(x^*) = 0 \cdot g(x^*) = 0 \cdot -x^* = 0$ מתקיים
3) $g(x^*) = -x^* \leq 0$ מתקיים

לכן $x^* \geq 0$ כך ש $f'(x^*) = 0$ ה"ה אכן נקודת KKT ולכן הוא פתרון
אופטימלי לבעיה.
אם כי f לא עולה ($f'(0) \leq 0$) בנק' $x=0$, לכן באיזשהי נקודה x^*
היא מקבלת מינימום, וגם כאן מקמירות נקבל נק' מינימום אחת.

HW4 גרמוביץ' יקטרינה 046197 314097826

תרגיל 2

נתונה בעיית אופטימיזציה:

$$(P)\quad \min_x \begin{cases} x^TQx \\ x^TQx + b^Tx \geq 1 \\ c^Tx \geq 0 \end{cases}$$

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

נתון להניח כי (P) פיזיבילית לכן קיים לה פתרון פיזיבילי בתחום האילוץ.

① נראה כי בעיית אופטימיזציה לא מהווה קמורה.

כדי שבעיית אופטימיזציה תהיה קמורה, גם פ' מטרה וגם תחום האילוץ חייבים להיות קמורות. נרשום בצורה אחרת את הבעיה:

$$(P)\quad \min_x \begin{cases} f(x) = x^TQx \\ g_1(x) = 1 - x^TQx - b^Tx \\ g_2(x) = -c^Tx \end{cases}$$

לפי הדרכה לשאלה, נבחר n ונראה עבורו כי בעיית אופטימיזציה לא מהווה קמורה.

נקח לדוגמה $n=1$: $c=0$, $b=0$, $Q=a\succ 0$, לכן תחום האילוץ יהיה:

$ax^2-1\geq 0 \quad ax^2 \geq 1 \Rightarrow x \geq \frac{1}{\sqrt{a}}$ או $x \leq -\frac{1}{\sqrt{a}}$

מתקבל תחום האילוץ שאינו רציף לכן אם נקח 2 נקודות הקצה של התחום $x_1=\frac{1}{\sqrt{a}}$, $x_2=-\frac{1}{\sqrt{a}}$, אז לכל $\lambda\in[0,1]$, כל נקודה z שהיא: $z=\lambda x_1+(1-\lambda)x_2$ לא תהיה שייכת לתחום האילוץ ולכן האילוץ לא קמור ולכן בעיית אופטימיזציה לא קמורה.

עבור $n=2$: עבור $b=0$, $c=0$, $Q=I\succ 0$, ו- $x=\begin{bmatrix}x\\y\end{bmatrix}$ נקבל כי בעיית אופטימיזציה היא:

$$\min \begin{cases} x^2+y^2 \\ x^2+y^2 \geq 0 \end{cases}$$

ניתן לראות כי תחום האילוץ הוא קבוצה משלימה של כדור $B(0,1)$ ומכיון שראינו בתרגול כי כדור הוא קבוצה קמורה, לכן ה- $x^2+y^2\geq 0$ הוא לא קמורה.

ניתן לראות זאת גם אם נבחר וקטור מפורש. נניח ש- $Q=I$

נבחר $x=\begin{bmatrix}0\\1\end{bmatrix}$, $y=\begin{bmatrix}1\\0\end{bmatrix}$, $b=[\frac{1}{4},\frac{1}{4}]^T$, $c=0$

x מקיימת אילוץ: $[0\ 1]\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}0\\1\end{bmatrix}+[\frac{1}{4}\ \frac{1}{4}]\begin{bmatrix}0\\1\end{bmatrix}=1\frac{1}{4}\geq 1$

y מקיימת אילוץ: $[1\ 0]\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}1\\0\end{bmatrix}+[\frac{1}{4}\ \frac{1}{4}]\begin{bmatrix}1\\0\end{bmatrix}=1\frac{1}{4}\geq 0$

אם נמצא $\lambda\in[0,1]$, למשל $\lambda=\frac{1}{2}$ נקבל כי $z=\lambda x+(1-\lambda)y=[\frac{1}{2},\frac{1}{2}]^T$

נבדוק האם z מקיים אילוץ:

$[\frac{1}{2}\ \frac{1}{2}]\begin{bmatrix}1&0\\0&1\end{bmatrix}\begin{bmatrix}1/2\\1/2\end{bmatrix}+[\frac{1}{4}\ \frac{1}{4}]\begin{bmatrix}1/2\\1/2\end{bmatrix}=\frac{1}{4}+\frac{1}{4}+\frac{1}{8}+\frac{1}{8}=\frac{3}{4}<1$

האילוץ לא מקיים, ולכן בעיית אופטימיזציה גם כאן לא מהווה קמורה.

② מצאו בעיה דואלית.

מקודם רשמתי את (P) כפי נוח לרשום לגרנז'יאן. עבור $\lambda_i\geq 0$: $\mathcal{L}(x,\lambda)=f(x)+\sum_i^n \lambda_i g_i(x)$

$$\mathcal{L}(x,\lambda)=x^TQx+\lambda_1(1-x^TQx-b^Tx)+\lambda_2(-c^Tx)=(1-\lambda_1)x^TQx-(\lambda_1 b^T+\lambda_2 c^T)x+\lambda_1$$

פ' מטרה הדואלית נתונה ע"י $q(\lambda)=\inf_{x\in C}\mathcal{L}(x,\lambda)$

נפריד את הבאה למקרים כי יש פה תלות ב- x עבור תוצאות ה- inf:
ראשית אם נגזור את הלגרנז'יאן נקבל:
על מנת למצוא inf: נשווה לאפס: $\nabla\mathcal{L}(x,\lambda)=2(1-\lambda_1)Qx-(\lambda_1 b^T+\lambda_2 c^T)=0$

עבור $\lambda_1<1$, $\lambda_2\geq 0$ מתוך לגרנז'יאן נקבל פונקציה ריבועית שהיא קמורה. לכל x המינימום מתקבל מתוך $\nabla\mathcal{L}(x,\lambda)$ ומתוך הנחה כי Q היא PD והפיכה.
לכן: $x^*=\frac{1}{2(1-\lambda_1)}Q^{-1}(\lambda_1 b^T+\lambda_2 c^T)$

$$g(\lambda)=(1-\lambda)\cdot\left[\frac{1}{2(1-\lambda_1)}Q^{-1}(\lambda_1 b^T+\lambda_2 c^T)\right]^TQ\frac{1}{2(1-\lambda_1)}Q^{-1}(\lambda_1 b^T+\lambda_2 c^T)-(\lambda_1 b^T+\lambda_2 c^T)\cdot\frac{1}{2(1-\lambda_1)}Q^{-1}(\lambda_1 b^T+\lambda_2 c^T)+\lambda_1=$$

$$=\frac{1}{4(1-\lambda_1)}(\lambda_1 b^T+\lambda_2 c^T)^TQ^{-1}QQ^{-1}(\lambda_1 b^T+\lambda_2 c^T)-\frac{(\lambda_1 b^T+\lambda_2 c^T)Q^{-1}(\lambda_1 b^T+\lambda_2 c^T)}{2(1-\lambda_1)}+\lambda_1=$$

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

עבור $\lambda_1=1$, $\lambda_2\geq 0$ והבל פונקציה אפינית, לכן נרצה לבחור x שמתאים $g(x)\to-\infty$
כלומר נרצה למצוא $\inf\{-(\lambda_1 b^T+\lambda_2 c^T)\}$ כלומר $g(\lambda)=\lambda_1$
אם הביטוי $-\lambda_1 b^T+\lambda_2 c^T=0$

314092826 046197 גרימופיל יקטרינה HW4

תרגיל 2 המשך

עבור $1<\lambda_1$, $\lambda_2 \geq 0$ מתקבלת פ' ריבועית עם מקדמים שליליים ← פ' קערה שמקבלת מינימום ב $-\infty$.

לכן אם נסכם: $0<\lambda_1<1$, $\lambda_2>0$

נסמן $t=\lambda_1 b^* + \lambda_2 c^*$

$$g(\lambda_1,\lambda_2) = \begin{cases} -\frac{1}{4}\frac{1}{(1-\lambda_1)} t^T Q^{-1} t + \lambda_1 & \lambda_2>0,\ 0<\lambda_1<1 \\ \lambda_1 & t=0,\ \lambda_1=1 \\ \infty & \text{else} \end{cases}$$

לכן הבעיה הדואלית היא:

$$\max_{\lambda_1,\lambda_2} g(\lambda_1,\lambda_2)$$
$$\text{s.t. } \operatorname{dom} g = \{\lambda_1,\lambda_2 \mid (\lambda_2>0,\ 1>\lambda_1\geq 0) \cup (\lambda_1=1,\ t=0)\}$$

סעיף 3 נסמן $\beta = b^TQ^{-1}b$ הראו כי הערך האופטימלי של הבעיה חסום מלמטה ע"י $\frac{1}{2}\left(2+\beta-\sqrt{\beta(\beta+4)}\right)$

כפי שראינו בסעיף הקודם ומהתאוריה, לכל λ_1,λ_2 שהם מקיימים את התחום החיובי מתקיים $p^* \geq g(\lambda_1\lambda_2)$

נבחר כי $\lambda_2=0$ כפי שהוצגו בהדרכה לתרגיל, לכן $t=\lambda_1 b^*$, בנוסף נסמן $b^TQ^{-1}b$ $\beta=b^TQ^{-1}b$

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

נגזור:

$$g'(\lambda_1,0) = 1 - \frac{1}{4}\frac{1}{(1-\lambda_1)^2}\beta(2\lambda_1-\lambda_1^2) = 0$$
$$4(1-\lambda_1)^2 - \beta(2\lambda_1-\lambda_1^2) = 0$$
$$4-8\lambda_1+4\lambda_1^2-2\beta\lambda_1+\beta\lambda_1^2=0$$
$$(4+\beta)\lambda_1^2-(8+2\beta)\lambda_1+4=0$$

$$\lambda_{1,2} = \frac{(8+2\beta)\pm\sqrt{(8+2\beta)^2-16(4+\beta)}}{2(4+\beta)} = \frac{2(4+\beta)\pm\sqrt{(4+\beta)^2 2^2-16(4+\beta)}}{2(4+\beta)} = 1\pm\frac{\sqrt{(4+\beta)\beta}}{4+\beta}$$

נקבל: כי $\lambda_1<1$

$$\lambda_1 = 1-\frac{\sqrt{\beta}}{\sqrt{\beta+4}} \quad \Leftarrow \quad \lambda_{1,2} = 1\pm\frac{\sqrt{\beta}}{\sqrt{4+\beta}}$$

נציב λ_1

מסעיף הקודם מצאנו כי $x^* = \frac{1}{2(1-\lambda_1)}\cdot Q^{-1}(\lambda_1 b^* + \lambda_2 c^*)$

לכן במקרה כי $\lambda_2=0$ נקבל:

$$x^* = \frac{\lambda_1}{2(1-\lambda_1)}Q^{-1}b =$$
$$= x^* = \frac{1}{2\sqrt{\beta}}\left(\sqrt{4+\beta}-\sqrt{\beta}\right)Q^{-1}b$$

נשים לב כי אכן קיימת נק' מינימום בתחום כי $\left.\frac{\partial g}{\partial \lambda_1}\right|_{\lambda_1\to 1} = -\infty$, $\left.\frac{\partial g}{\partial \lambda_1}\right|_{\lambda\to 0} > 0$

נציב ב g:

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

לצורך נוחות נסמן $k=\frac{\sqrt{\beta}}{\sqrt{\beta+4}}$

$$= -\frac{1}{4}(1-k)^2\frac{1}{k} + 1 - k = \frac{4(k-k^2)-(1-k)^2\beta}{4k} =$$
$$= \frac{\cancel{4k-4k^2-\beta+2\beta k-\beta k^2}}{\cancel{4k}} = \frac{(1-k)(4k-(1-k)\beta)}{4k} =$$

$(a=\sqrt{\beta+4})$

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

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

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

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

$$= \frac{1}{4\sqrt{\beta+4}}\cdot\sqrt{\beta+4}\left(4+2\beta-2\sqrt{\beta(\beta+4)}\right) = \frac{1}{2}\left[2+\beta-\sqrt{\beta(\beta+4)}\right] = D^*$$

מדואליות $D^* \leq P^*$ שהוא החסם מלמטה של p^*.

4 הוכיחו כי בעיית אופטימיזציה במסגרת אינה קמורה.

נראה כי קיים $r>0$ כך ש $\|x-a\|_Q^2 \geq r \iff x^TQx+b^Tx \geq 1$

כאשר עבור $Q\succ 0$ ניתן לרשום $\|v\|_Q^2 = v^TQv$

נפתח:

$$\|x-a\|_Q^2 \geq r \Rightarrow (x-a)^TQ(x-a) \geq r$$
$$x^TQx - 2a^TQx + a^TQa \geq r$$

מסימטריות של Q:

$$x^TQx + (-2Qa)^Tx \geq r - a^TQa = 1$$

נדרוש כי $b=-2Qa$, $r-a^TQa=1$

מתוך תנאי זה נקבל כי $a=-\frac{1}{2}Q^{-1}b$ ולכן

$$r = 1+a^TQa = 1+\left(-\frac{1}{2}Q^{-1}b\right)^TQ\left(-\frac{1}{2}Q^{-1}b\right) = 1+\frac{1}{4}b^TQ^{-1}b.$$

נשים לב כי $Q^{-1} \succ 0$ לכן לפי הגדרה $b^TQ^{-1}b > 0$ ולכן גם $r > 0$
כלומר האילוץ מוגדר ע"י כדור $\left(B^c_{\|\cdot\|_2}(a,r)\right)$ כלומר מחוצה לו.
כעת נתייחס לאילוץ השני שהוא היפרמישור $c^Tx \geq 0$
עבור $c=0$ נקבל כי התחום האילוץ הוא, תחום החיצוני של הכדור ← לפי תרגול 2 ראינו כי תחום לא קמור. ← בעיית אופטימיזציה במקרה זה לא קמור.
עבור $c \neq 0$ נקבל כי חיתוך (בין האילוצים הוא) $x^TQx + b^Tx - 1 = c^Tx$
באותו אופן כמו מקודם נוכל להגדיר את התחום כמו אליפסואיד במשוואת השלמה לריבוע ופיתוח של נוסחה. נקבל גם כאן תחום שהוא מחוץ לאליפסואיד שבגלל אם ניקח $x=0$ הוא יכול להיות קומבינציה של x_1, x_2 על שפת האליפסואיד כך שעבור $\lambda \in [0,1]$ נקבל $\lambda x_1 + (1-\lambda)x_2 = 0$
לכן תחום האילוץ בהכרח איננו קמור
ולכן הבעיה בהכרח איננה קמורה

314092826 046197 ארימוביץ' יקטרינה HW 4

תרגיל 3

תהי $Q\succ 0$ ב-$R^{n\times n}$, ויהיו $a,b\in R^n$ וקטורים נתונים ויהי $0<c\in R$

נתונה בעיית האופטימיזציה:

$$(P)\quad \min_x \begin{cases} \frac{x^TQx}{x^Ta} \\ x^Tb=c \\ x^Ta>0 \end{cases}$$

נתון כי הבעיה פיזיבילית ומקיימת תנאי slater

כמו כן, מתקיים $\alpha\beta>\gamma^2$ כאשר: $\alpha=a^TQ^{-1}a$, $\beta=b^TQ^{-1}b$

$\gamma=a^TQ^{-1}b=b^TQ^{-1}a$

הראו כי הבעיה קמורה ומצאו את הערך האופטימלי שלה באמצעות תנאי KKT.

הדרכה: משתנה עזר $y=x^Ta$. לקבל בעיה בשני משתנים, להניח כי y קבוע. לפתור בעיה ב-x עם אילוץ שוויון. ולהראות כי קיימים $\mu_1,\mu_2\in R$ כך שהערך האופטימלי (אם y)

$$g(y)=\frac{x^{*T}Qx^*}{y}=\frac12(\mu_1c+\mu_2y)=\frac{\beta y^2-2c\gamma y+c^2\alpha}{y(\alpha\beta-\gamma^2)}$$

לאחר מכן לפתור בעיית אופטימיזציה $\min\{g(y)\,|\,y>0\}$ לקבלת ערך האופטימום של (P).

מתוך הגדרת הבעיה ומתוך $y=x^Ta$ ניתן לרשום כי

$$(P')\quad \min_x \begin{cases} \frac{x^TQx}{y} \\ h_1(x)=c-x^Tb \\ y=x^Ta\ ,\ y>0 \Rightarrow h_2(x)=y-x^Ta \end{cases}$$

נשים לב מכיוון כי תחום האילוץ ניתן קמור כי מוגדר ע"י פ' אפיניות (היפרמישור)

פ' מטרה, תחת הנחה כי $\underset{y>0}{y=const}$ הינה גם קמורה (תרגול 3, תרגיל 1, מקרה 3)

(בנוסף ידוע שכפל ע קבוע אז מכפלה של פ' קמורה בקבוע גם קמורה)

לכן בעיית האופטימיזציה (P) הינה קמורה.

נרצה למצוא ערך האופטימלי. לצורך כך נרשום את הלגרנז'יאן שלה

$$L(x,\mu_1,\mu_2)=\frac1y x^TQx+\mu_1(c-x^Tb)+\mu_2(y-x^Ta)=\frac1y x^TQx+\mu_1c-\mu_1x^Tb-\mu_2x^Ta+\mu_2y$$

לפי הנתון הבעיה מקיימת תנאי Slator לכן קיום תנאי KKT הינו חובה כדי ש-x^* יהיה הערך האופטימלי.

נגזור את הלגרנז'יאן:

$$\nabla_x L(x,\mu_1,\mu_2)=2\frac1y Qx-\mu_1b+\mu_2a=0$$

היות ו-Q מטריצה PD לכן הפיכה, נוכל לבודד את x:

$$2\frac1y Qx=\mu_1b+\mu_2a \Rightarrow x^*=Q^{-1}\frac{y(\mu_1b+\mu_2a)}{2}$$

מבעיה IS אין אילוצי אי-שוויון, לכן אין צורך ב complementary slakness. הנקודה שהתקבלה היא נק' KKT ולכן היא הפתרון האופטימלי. נציב אותה בחזרה ב-

נמצא כעת את $g(y)$:

$$g(y)=\frac{x^{*T}Qx^*}{y}=\frac1y\cdot\left(Q^{-1}\frac{y(\mu_1b+\mu_2a)}{2}\right)^TQx^*=\frac12(\mu_1b+\mu_2a)^TQ^{-1^T}Qx^*=$$

מתוך הגדרה של Q נקבל כי $Q^{-1}Q=I \Leftarrow Q^{-1}=Q^{-1^T}$

$$g(y)=\frac12(\mu_1b+\mu_2a)^Tx^*=\frac12\mu_1b^Tx^*+\frac12\mu_2a^Tx^*=\underline{\frac12(\mu_1c+\mu_2y)}$$

כיון שבעית אופטימיזציה היא קמורה, מקיימת תנאי Slator ובנוסף אין x כזה שמביא פתרון הבעיה ל-$-\infty$ (כי $Q\succ0$, $y>0$, אז $\frac1y x^TQx>0$) לכן מתקיימת פה שוויון הדוקה. ($D^*=P^*$).

כעת נמצא את μ_1,μ_2. נציב את x^* בתנאים מאלצים:

$x^Tb=c$

$$\hookrightarrow c=x^{*T}b=\left(Q^{-1}\frac{y(\mu_1b+\mu_2a)}{2}\right)^Tb=\frac y2(\mu_1b+\mu_2a)^TQ^{-1}b=\frac y2(\mu_1b^TQ^{-1}b+\mu_2a^TQ^{-1}b)$$

$$\underline{c=\frac y2(\mu_1\beta+\mu_2\gamma)}$$

$x^Ta=y$

$$\hookrightarrow y=x^{*T}a=\left(Q^{-1}\frac{y(\mu_1b+\mu_2a)}{2}\right)^Ta=\frac y2(\mu_1b+\mu_2a)^TQ^{-1}a=\frac y2(\mu_1b^TQ^{-1}a+\mu_2a^TQ^{-1}a)=$$

$$y=\frac y2(\mu_1\gamma+\mu_2\alpha) \Rightarrow \underline{2=\mu_1\gamma+\mu_2\alpha}$$

~~נציב כעת בביטוי של $g(y)$~~:

$\cancel{g(y)=\frac12(\mu_1c+\mu_2y)=\frac12}$

נבודד את μ_1,μ_2 קודם.

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

$$\frac{2c}{y}=\mu_1\beta+\frac{2-\mu_1\gamma}{\alpha}\gamma$$

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

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

נציב כעת בתוך $g(y)$ ונקבל:

$$g(y)=\frac12(\mu_1c+\mu_2y)=\frac12\left[\frac{2(c\alpha-\gamma y)}{y(\alpha\beta-\gamma^2)}c+\frac{2(\beta y-c\alpha)}{y(\alpha\beta-\gamma^2)}y\right]=$$

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

314092826 גרימוביץ' יקטרינה HW4

תרגיל 3 המשך

כעת נפתור בעיית אופטימיזציה חדשה $\min\{g(y) \mid y>0\}$, $g(y)=\frac{\beta y^2-2c\gamma y+\alpha c^2}{y(\alpha\beta-\gamma^2)}$

נגזור את $g(y)$ לפי y ונשווה לאפס כדי למצוא נקודות קיצון:

$$g(y)=\frac{\beta}{(\alpha\beta-\gamma^2)}y-\frac{2c\gamma}{(\alpha\beta-\gamma^2)}+\frac{\alpha c^2}{(\alpha\beta-\gamma^2)}\frac{1}{y}$$

$$g'(y)=\frac{\beta}{\alpha\beta-\gamma^2}-\frac{\alpha c^2}{(\alpha\beta-\gamma^2)}\frac{1}{y^2}=0 \Rightarrow y^2\beta=\alpha c^2 \Rightarrow y^2=\frac{\alpha c^2}{\beta} \rightarrow y=\pm c\sqrt{\frac{\alpha}{\beta}}$$

לפי ההגדרה של הבעיה $y>0$, לכן ניקח רק פתרון חיובי

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

נבדוק את הנגזרת השנייה על מנת לבדוק שהתקבל min:

$$g''(y)=\frac{\alpha c^2}{(\alpha\beta-\gamma^2)}\frac{2}{y^3}>0 \Leftarrow \begin{cases} \text{מתקבל מתוך } \alpha>0 \\ y>0 \\ \text{ומה } \alpha\beta>\gamma^2 \end{cases}$$

לכן אכן מדובר ב y^* שהוא מינימום.

נציב כעת בחזרה את y^*:

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

$$=\frac{\sqrt{\beta\alpha}\,c}{(\alpha\beta-\gamma^2)}-\frac{2c\gamma}{(\alpha\beta-\gamma^2)}+\frac{\sqrt{\beta\alpha}\,c}{(\alpha\beta-\gamma^2)}=\frac{2c(\sqrt{\alpha\beta}-\gamma)}{\alpha\beta-\gamma^2}=\frac{2c}{\sqrt{\alpha\beta}+\gamma}=D^*$$

זהו הערך האופטימלי של הבעיה הדואלית אך מתוך דואליות חזקה זהו גם ערך האופטימלי של בעיה מקורית $p^*=D^*$

314092826 046197 יקטרינה גרימובסקי HW4

תרגיל 4

נתונה בעיית אופטימיזציה $\min_x \left\{\frac{1}{3}x^3 - ax\right\}$ $a>0,\ x>0$

① הראו כי פתרון הבעיה הוא $x^* = \sqrt{a}$

(נסמן) פ' $f(x) = \frac{1}{3}x^3 - ax$. נשים לב כי פ' זו היא פולינום מספר פעמים שגזירה, פעמיים ברציפות בתחום הנתון $x>0$. נגזור ונקבל: $f'(x) = x^2 - a = 0$

$x = \pm\sqrt{a}$

היות ומדובר ב $x>0$, לכן $x^* = \sqrt{a}$ הוא פתרון יחיד של בעיה בתחום הנתון.

(נבדוק שהוא אכן מינימום: $f''(x) = 2x$ עבור $f''(-\sqrt{a}) < 0$ לכן הפ' מקבלת כאן max מקס'

$f''(\sqrt{a}) > 0$ לכן פ' מקבלת פה מינימום.

② מעוניינים לפתור את בעיית אופטימיזציה בשיטת ניוטון. מצאו קישור לאיטרציה k לפתרון הבעיה כתלות באיטרציה k-1 ופרמטר a. הניחו גודל הצעד קבוע $t=1$

אלגוריתם ניוטון מתבסס על כך ש $d_{k-1} = -\left(\nabla^2 f(x_{k-1})\right)^{-1}\nabla f(x_{k-1})$

$x_k = x_{k-1} + t_k d_{k-1}$

מתוך סעיף הקודם: $\nabla f(x_k) = x_k^2 - a$ באופן דומה: $\nabla f(x_{k-1}) = x_{k-1}^2 - a$ נציב ב d_{k-1}:

$\nabla^2 f(x_k) = 2x_k$ $\nabla^2 f(x_{k-1}) = 2x_{k-1}$

$$d_{k-1} = -\frac{x_{k-1}^2 - a}{2x_{k-1}} \Rightarrow x_k = x_{k-1} + 1\left(-\frac{x_{k-1}^2 - a}{2x_{k-1}}\right) = \frac{2x_{k-1}^2 + (-x_{k-1}^2 + a)}{2x_{k-1}} = \frac{x_{k-1}^2 + a}{2x_{k-1}} = \frac{1}{2}x_{k-1} + \frac{a}{2x_{k-1}}$$

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

מכיוון שמדובר ב $\sqrt{a}$ שהוא הפתרון של הבעיה. לכן כל איטרציה x_k תהיה קטנה ממנו / פה שאיפה להתכנס לפתרון. לכן כיוון d_k הוא כיוון הירידה, לכן $x_{k-1} > x_k$

ניתן לראות זאת גם כי כיוון $\frac{x_{k-1}^2 - a}{2x_{k-1}} > 0$ כי $x_{k-1} > 0$ וגם $x_{k-1}^2 > a$ (כי $x_{k-1}^2 > \sqrt{a}^2$)

$$x_k = \frac{x_{k-1}}{2} + \frac{a}{2x_{k-1}} \underbrace{<}_{\{a < x_{k-1}^2\}} \frac{x_{k-1}}{2} + \frac{x_{k-1}^2}{2x_{k-1}} = x_{k-1}$$

נרצה להראות כי $x_k > \sqrt{a}$ נניח כי

$$x_k = \frac{x_{k-1}}{2} + \frac{a}{2x_{k-1}} > \sqrt{a}$$

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

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

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

לכן הביטוי $x_k > \sqrt{a}$ הוא נכון כי הביטוי שקיבלנו הוא נכון לפי הנתון ⇐ לכן $x_{k-1} > x_k > \sqrt{a}$ מש"ל

④ הניחו כי $x_k = (1+\delta_k)\sqrt{a}$, כלומר הגודל δ_k מסמן את השגיאה הכפלית בחישוב $\sqrt{a}$

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

מתוך סעיף ② $x_k = \frac{x_{k-1}^2 + a}{2x_{k-1}}$ נציב את $x_k = (1+\delta_k)\sqrt{a}$:

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

$:\sqrt{a}$

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

⑤ הוכיחו כי $\delta_k \le \frac{\delta_{k-1}}{2}$ הסיקו כי אלגוריתם מתכנס ל-$\sqrt{a}$.

נראה מסעיף ④. מתוכו ניתן להסיק כי $\delta_k > 0$ לכן:

$\left(x_k = (1+\delta_k)\sqrt{a} \Rightarrow \delta_k = \frac{x_k - \sqrt{a}}{\sqrt{a}} > 0\right)$

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

נשים לב כי עבור $k\to\infty$ נקבל $\delta_k \to 0$ כי זה מתקבל

מתוך $x_k = (1+\delta_k)\sqrt{a} \Leftarrow$ $x_k = \lim_{\delta_k\to 0}(1+\delta_k)\sqrt{a} = \sqrt{a}$

מתוך $\delta_k \le \frac{1}{2}\delta_{k-1}$ גם כל פעם השגיאה קטנה בחצי לפחות

$$\delta_k \le \frac{1}{2}\delta_{k-1} \le \frac{1}{2^2}\delta_{k-2} \le \frac{1}{2^3}\delta_{k-3} \dots \le \frac{1}{2^k}\delta_0$$

כלומר אם, עבור k גדול השגיאה מתכנסת לאפס

314092826 046197 גרימוביץ' יקטרינה HW4

תרגיל 4 המשך

(6) הוכיחו כי $\delta_k \le \frac{\delta_{k-1}^2}{2}$

ראינו בסעיפים הקודמים כי ההתכנסות לא תלויה בתנאי התחלה ולכן ניתן להניח גם כי $\sqrt{a} < x_0$ מסעיף 3 ולכן לכל k, $x_k > \sqrt{a}$ ואז לכל k $\delta_k > 0$ ולכן:

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

(7) נניח כי $a=2$ וכי ניחוש הראשוני $x_0 = 2$ עבור כל אחד מהביטויים חשבו כמה איטרציות נדרשות על מנת לקבל דיוק של 30 ספרות אחרי נקודה.

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

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

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

מתוך ריצת קוד מטלב נקבל:

(1) - מתכנס תוך 7 איטרציות

(2) - מתכנס תוך ~165 איטרציות

(3) - מתכנס תוך 7 איטרציות

תרגיל 5

נתונות 2 בעיות אופטימיזציה:

$(P_1)\quad \min_{x,y}\left\{ x^2+y^2+5x-10y \;\middle|\; |x-\tfrac{y}{2}-2| \le 1 \right\}$

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

① הראו כי בעיות אופטימיזציה קמורות

נתחיל עם בעיה P_1: נרשום בצורה אחרת:

$|x-\tfrac{y}{2}-2| \le 1$

$x-\tfrac{y}{2}-2 \ge -1 \qquad x-\tfrac{y}{2}-2 \le 1$

$y \le 2x-1 \qquad -x+\tfrac{y}{2}+1 \le 0 \qquad x-\tfrac{y}{2}-3 \le 0 \qquad y \ge 2x-6$

$$(P_1')\quad \min_{x,y}\begin{cases} f(x,y) = x^2+y^2+5x-10y \\ g_1(x,y) = x-\frac{y}{2}-3 \le 0 \\ g_2(x,y) = -x+\frac{y}{2}+1 \le 0 \end{cases}$$

מטרה היא פ' גזירה ברציפות פעמיים, לכן נוכל לרשום את ההסיאן שלה כי:

$$\nabla^2 f(x,y) = \begin{bmatrix} \frac{\partial^2 f}{\partial x^2} & \frac{\partial^2 f}{\partial y \partial x} \\ \frac{\partial^2 f}{\partial x \partial y} & \frac{\partial^2 f}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} = 2\cdot I^{2\times 2} \rightarrow$$

ההסיאן הוא PD
לכן לפי תנאי לקמירות
מסדר השני $f_1(x,y)$ היא קמורה.

כעת נבדוק גם את האילוצים: שניהם, גם $g_1(x,y)$ וגם $g_2(x,y)$ הם ~~אילוצים~~ אפיניים (פ' לינאריות מגדירות הישר מישור) לכן פ' קמורות ולכן חיתוך שלהם הוא גם קמור ולכן כל הבעיה האופטימיזציה קמורה.

עבור בעיה P_2: נרשום בצורה אחרת:

$g_i \le 0 \quad \Rightarrow i=1,2,3$

$$(P_2')\quad \min_{x,y}\begin{cases} h(x,y) = (x-3)^2+y^2+3x-2y \\ g_1(x,y) = 2-x \\ g_2(x,y) = -y \\ g_3(x,y) = (x-1)^2+(y-2)^2-9 \end{cases}$$

פ' מטרה היא פ' גזירה פעמיים ברציפות, לכן נוכל לרשום את ההסיאן שלה כי:

$$\nabla^2 h(x,y) = \begin{bmatrix} \frac{\partial^2 h}{\partial x^2} & \frac{\partial^2 h}{\partial y \partial x} \\ \frac{\partial^2 h}{\partial x \partial y} & \frac{\partial^2 h}{\partial y^2} \end{bmatrix} = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} \rightarrow$$

בדומה לסעיף הקודם,
ההסיאן הוא PD
לכן לפי תנאי לקמירות מסדר שני $h(x,y)$ היא פ' קמורה

כעת נבדוק את האילוצים: $g_1(x,y)$ ו- $g_2(x,y)$ הן פ' אפיניות ולכן קמורות.

$$\nabla^2 g_3(x,y) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix}$$

בדומה לסעיף הקודם,
ההסיאן הוא PD לכן אילוץ $g_3(x,y)$ הוא קמור
(פ' גזירה ברציפות פעמיים עם ההסיאן PD מקיימת תנאי השני לקמירות).
לכן חיתוך של האילוצים הוא גם קמור ולכן כל הבעיה האופטימיזציה קמורה.

② (P_1'): נגדיר עבור בעיית אופטימיזציה זו בעיה שקולה עבור $t>0$:

$$(P_{1t}')\ :\ \min\left\{ x^2+y^2+5x-10y-\frac{1}{t}\left(\ln\left(-x+\frac{y}{2}+3\right)\right)-\frac{1}{t}\ln\left(x-\frac{y}{2}-1\right)\right\}$$

$$(P_{1t}')\ :\ \min\left\{ x^2+y^2+5x-10y-\frac{1}{t}\left(\ln\left(-x+\frac{y}{2}+3\right)\left(x-\frac{y}{2}-1\right)\right)\right\}$$

על מנת להפעיל barrier נגזור את פ' המטרה החדשה (נסמנה ב- F)

$$\nabla F_1(x,y,t) = \left[\, 2x+5+\frac{2x-y-4}{t\left(-x+\frac{y}{2}+3\right)\left(x-\frac{y}{2}-1\right)}\ ,\ 2y-10-\frac{-x+\frac{y}{2}+2}{t\left(-x+\frac{y}{2}+3\right)\left(x-\frac{y}{2}-1\right)} \right]^T$$

$d_t = -\left(\nabla^2 F_1\right)^{-1}\nabla F_1$

נצטרך חישוב d בשיטת ניוטון, ונראה את לחשב $\nabla^2 F_1$
הנגזרות עבור נגזרות השניות חושבו באמצעות מטריצים סימבוליים syms במטלב עבור שני הסעיפים.

(P_2'): נגדיר עבור בעיית אופטימיזציה זו בעיה שקולה עבור $t>0$:

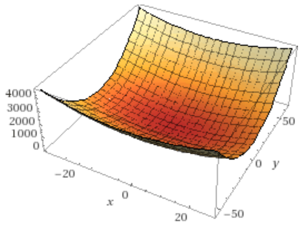
$$(P_{2t}')\ :\ \min\left\{ x^2+y^2+5x-10y-\frac{1}{t}\left(\ln\left(9-(x-1)^2-(y-2)^2\right)-\ln(x-2)-\ln(y)\right)\right\}$$

$$\nabla F_2(x,y,t) = \left[\, (2x-3)+\frac{2(x-1)}{t\left[9-(x-1)^2-(y-2)^2\right]}-\frac{1}{t[x-2]}\ ,\ 2y-2+\frac{1}{t}\frac{2(y-2)}{\left[9-(x-1)^2-(y-2)^2\right]}-\frac{1}{ty} \right]$$

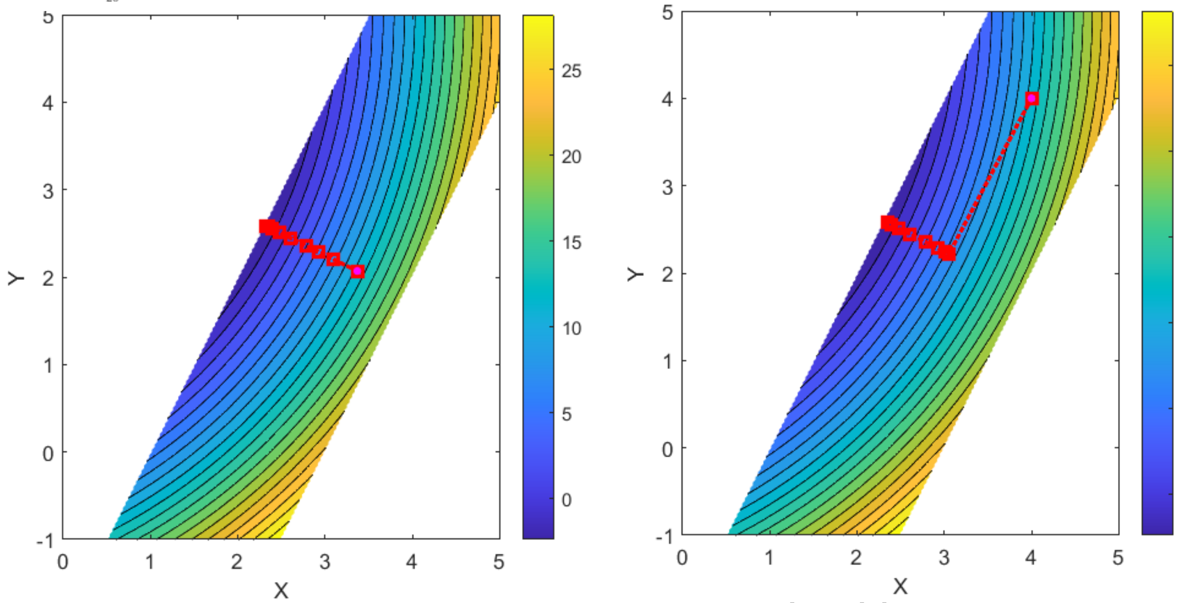
$d_t = -\left(\nabla^2 F_2\right)^{-1}\nabla F_2$

בדומה לסעיף הקודם, גם כאן נחשב את d בשיטת ניוטון,

סעיף 2

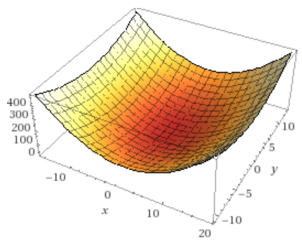


כפי שיתן לראות לפי `contourf` הפתרון של האלגוריטם מתכנס לכיוון הפתרון האנליטי של הבעיה (ראה איור בצד)$min(f(x,y)) = 31.25\ at\ [-2.5,5]$ תחת מגבלות האילוצים הנתונים בבעיה. נשים לב כי עבור כל ניחוש התחלתי אנחנו מתכנסים לאותה נקודת המינימום $[2.3, 2.6]$ של פ' מטרה תחת אילוצים.

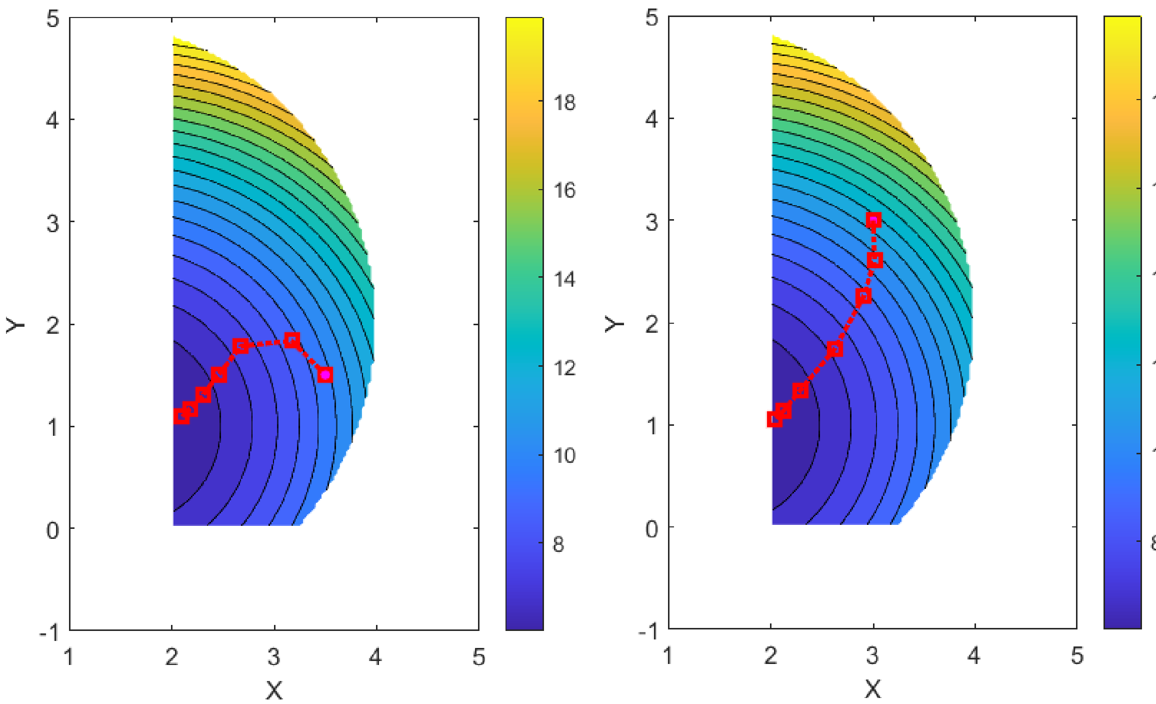


באיור מטה מוצגים שני מסלולים של התכנסות לפתרון.

סעיף 3



כפי שיתן לראות לפי `contourf` הפתרון של האלגוריטם מתכנס לכיוון הפתרון האנליטי של הבעיה (ראה איור בצד $(min(h(x,y)) = \frac{23}{4}\ at\ [1.5, 1]$ תחת מגבלות האילוצים הנתונים בבעיה. נשים לב כי עבור כל ניחוש התחלתי אנחנו מתכנסים לאותה נקודת המינימום $[2, 1]$של פ' מטרה תחת אילוצים.



Grad(P1)	[x + 1/(y/2 - x + 1) + 1/(y/2 - x + 3) + 5, 2*y - 1/(2*(y/2 - x + 1)) - 1/(2*(y/2 - x + 3)) - 10*2]
Hesan(P1)	[(y/2 - x + 1)^2 + 1/(y/2 - x + 3)^2 + 2, 1/(4*(y/2 - x + 1)^2) + 1/(4*(y/2 - x + 3)^2) + 2/1] [(2*(y/2 - x + 1)^2) - 1/(2*(y/2 - x + 3)^2), 1/(4*(y/2 - x + 1)^2) + 1/(4*(y/2 - x + 3)^2) + 2/1 -]
Grad(P2)	[x - 1/(x - 2) - (2*x - 2)/((x - 1)^2 + (y - 2)^2 - 9) - 3, 2*y - (2*y - 4)/((x - 1)^2 + (y - 2)^2 - 9) - 1/y - 2*2]
Hesan(P2)	(2*x - 2)^2/((x - 1)^2 + (y - 2)^2 - 9)^2 - 2/((x - 1)^2 + (y - 2)^2 - 9) + 1/(x - 2)^2 + 2, ((2*x -] [2)*(2*y - 4))/((x - 1)^2 + (y - 2)^2 - 9)^2 ((2*x - 2)*(2*y - 4))/((x - 1)^2 + (y - 2)^2 - 9)^2, (2*y - 4)^2/((x - 1)^2 + (y - 2)^2 -] [9)^2 - 2/((x - 1)^2 + (y - 2)^2 - 9) + 1/y^2 + 2