314092826 046197 גרימוביץ' יקטרינה HW2

הגשה באיחור באישור כפיר

תרגיל 1

תהי $f:\mathbb{R}^n\to\mathbb{R}$ הוכיחו כי f קמורה אם ורק אם $\phi(t)=f(x+td)$ קמורה לכל $x,d\in\mathbb{R}^n$

$\Leftarrow$ נניח קודם כל כי f קמורה ונוכיח כי $\phi(t)=f(x+td)$ קמורה לכל $x,d\in\mathbb{R}^n$

יהי $\lambda\in[0,1]$, ובנוסף נניח כי קיים $t_1,t_2\in\mathbb{R}$ אז מתקיים:

$$\underline{\phi(\lambda t_1+(1-\lambda)t_2)} = f(x+(\lambda t_1+(1-\lambda)t_2)d) =$$
$$= f(\lambda x+(1-\lambda)x + d\lambda t_1 + d(1-\lambda)t_2) =$$
$$= f(\lambda(x+dt_1)+(1-\lambda)(x+dt_2)) \underset{f \text{ מקמירות של}}{\leq} \lambda f(x+dt_1)+(1-\lambda)f(x+dt_2) = \underline{\lambda\phi(t_1)+(1-\lambda)\phi(t_2)}$$

מש"ל

כנדרש נוכיח כיוון השני לכל $x,d\in\mathbb{R}^n$

$\Rightarrow$ נניח כי $\phi(t)=f(x+td)$ קמורה. ונוכיח כי f קמורה..

יהי $\lambda\in[0,1]$ ובנוסף $z,y\in dom(f)$, כך שנגדיר את $y=\cancel{x}$, $d=z-y$

$$\underline{f(\lambda z+(1-\lambda)y)} = f(y+\lambda(z-y)) = f(x+\lambda d) = \phi(\lambda) = \phi(\lambda\cdot 1+(1-\lambda)\cdot 0) \underset{\text{מקמירות}}{\leq} \lambda\phi(1)+(1-\lambda)\phi(0) =$$
$$= \lambda f(x+d)+(1-\lambda)f(x) \overset{\text{נציב חזרה}}{=} \lambda f(y+z-y)+(1-\lambda)f(y) = \underline{\lambda f(z)+(1-\lambda)f(y)}$$

מש"ל

תרגיל 2

יהיו $Q,R\in\mathbb{R}^{n\times n}$ כך ש $Q,R\succ 0$ נתבונן ב: $g(x)=f(x,x)=(\frac{1}{2}x^TQx)(\frac{1}{2}x^TRx)$, $f(x,y)=(\frac{1}{2}x^TQx)(\frac{1}{2}y^TRy)$
מוגדרת לכל $x,y\in\mathbb{R}^n$

1. חשבו גרדיאנט והסיאן של $g(x)$ נעזר כאן בכלל השרשרת:

$$\nabla g(x) = \frac{1}{4}\frac{\partial[(x^TQx)(x^TRx)]}{\partial x} =$$
$$= \frac{1}{4}\left[(x^TQx)\cdot\frac{\partial x^TRx}{\partial x} + \frac{\partial(x^TQx)}{\partial x}(x^TRx)\right] = \frac{1}{4}\left[x^TQx\cdot 2Rx + 2Qx\, x^TRx\right] = \frac{1}{2}\left[(x^TQx)Rx+(x^TRx)Qx\right]$$

$$\nabla^2 g(x) = \frac{1}{2}\frac{\partial}{\partial x}\left[x^TQx\cdot Rx + x^TRx\cdot Qx\right] = \frac{1}{2}\left[x^TQx\frac{\partial Rx}{\partial x} + \frac{\partial x^TQx}{\partial x}x^TR + x^TRx\frac{\partial Qx}{\partial x} + \frac{\partial x^TRx}{\partial x}x^TQ\right] =$$
$$= \frac{1}{2}x^TQx\cdot R + Qxx^TR + \frac{1}{2}x^TRxQ + Rxx^TQ$$

2. נתון כי $g(x)$ אינה קמורה. האם בהכרח $f(x,y)$ אינה קמורה?
נשים לב כי אם $g(x)$ היא לא קמורה, לכן קיים $\lambda\in[0,1]$ עבורו יתקיים (עבור $x_1,x_2\in dom(g)$) כי
$g(\lambda x_1+(1-\lambda)x_2) > \lambda g(x_1)+(1-\lambda)x_2$
מכיוון שאם נגדיר כי $y=x$, לכן מתקיים גם: $f(\lambda x_1+(1-\lambda)x_2,\lambda x_1+1-\lambda(x_2)) > \lambda \underset{g''(x_1)}{f(x_1,x_1)}+(1-\lambda)\underset{g''(x_2)}{f(x_2,x_2)}$ ואז
לכן סתירה הפרכים האלה לתנאי קמירות
גם נסתר לכן f גם לא קמורה.

3. הוכיחו כי אם $R=\alpha Q$, $\alpha>0$ אז $g(x)$ קמורה לכל $x\in\mathbb{R}^n$
מספיק זה ניתן להוכיח בתנאי השני לקמירות (לבדיקת קמירות). נציב את $R=\alpha Q$ בהסיאן. צריך להראות כי $\nabla^2 g(x)\geq 0$

$$\nabla^2 g(x) = \frac{1}{2}\left[\frac{1}{2}x^TQx\alpha Q + Qxx^T\alpha Q + \frac{1}{2}x^T\alpha Qx Q + \alpha Qxx^TQ\right] = \alpha\left[x^TQxQ + 2Qxx^TQ\right] \circledast$$
$$\alpha\left[(x^TQx)Q + 2Qx(Qx)^T\right]$$

ניתן להגדיר טענות עזר: א) אם יש שתי מטריצות $A,Q\succeq 0$ אז גם $AQA\succeq 0$ [ניתן להוכיח בקלות:
כי לכל $x\in\mathbb{R}^n$ אם נסמן $Ax=v$ נקבל $\underbrace{x^TA}_{(Ax)^T}QAx = vQv\geq 0$]

ב) אם נניח שיש $v\in\mathbb{R}^n$ אז $vv^T\succeq 0$ [מתקיים כיוון שאם נגדיר $y\in\mathbb{R}^n$
$y^Txx^Ty = \underbrace{(y^Tx)}_{\alpha=}\underbrace{(x^Ty)}_{\alpha=} = \alpha^2\geq 0 = \|y^Tx\|_2^2$

מכאן ניתן להסיק כי $(Qx)(Qx)^T\succeq 0$

ג) בנוסף אם יש $\beta>0$ וגם $Q\succeq 0$ אז גם $\beta Q\succeq 0$ (נובע ישירות מהגדרה) (מכפלה ב.)

ד) בנוסף גם ניתן להסיק לגבי הסכום של מטריצות. אם יש $A,Q\in\mathbb{R}^{n\times n}$ שהן PSD אז
גם $A+Q$ היא PSD [ניתן להראות זאת לכל $x\in\mathbb{R}^n$, $x^T(A+Q)x =$
$= x^TAx + x^TQx \geq 0$ (לפי הגדרה $x^TAx\geq 0$, $x^TQx\geq 0$)

לכן מתוך שילוב של ג והגדרה 3 מתרגול 1 נקבל כי $(x^TQx)Q\geq 0$
ומתוך ד כי סכום הביטויים PSD הוא גם PSD
מתוך $\circledast$

לכן $\nabla^2 g(x)\succeq 0$ ולכן $g(x)$ מקיים תנאי השני לקמירות $\Leftarrow$ $g(x)$ קמורה

314092826 046197 HW 2 גרימנביץ' יקטרינה.

תרגיל 2 - המשך

4. נמצא כעת $Q = \begin{bmatrix} 3 & -2 \\ -2 & 3 \end{bmatrix}$, $R = \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$ הראנו $Q, R \succ 0$ והאם $g(x)$ קמורה עבור הבחירה הנ"ל?

נבדוק det:

$\det(R) = 4\cdot 4 - 3\cdot 3 = 5 > 0$ $tr R = 4+4 = 8$

$\det(Q) = 3\cdot 3 - (-2)(-2) = 5 > 0$ $tr Q = 6$

נמצא ע"ע של שתי המטריצות:

$R: \quad \det|R - I\lambda| = \begin{vmatrix} 4-\lambda & 3 \\ 3 & 4-\lambda \end{vmatrix} = (4-\lambda)^2 - 9 = 0 \rightarrow \lambda_{R1} = 1, \ \lambda_{R2} = 7$

$Q: \quad \det|Q - I\lambda| = \begin{vmatrix} 3-\lambda & -2 \\ -2 & 3-\lambda \end{vmatrix} = (3-\lambda)^2 - 4 = 0 \rightarrow \lambda_{Q1} = 1, \ \lambda_{Q2} = 5$

נקבל כי בשני המקרים ע"ע חיוביים

לכן R,Q הן PD

נבדוק האם $g(x)$ היא קמורה ע"י תנאי לקמירות. $\nabla^2 g(x) \succeq 0$

נקח את הביטוי שהתקבל בסעיף 1. ניתן לראות כי 2 איברים הם מתקיימים כמו בסעיף הקודם, כלומר

$$\nabla^2 g(x) = \underbrace{\tfrac{1}{2} x^T Q x R}_{\text{PD לפי סעיף 3}} + \underbrace{Q x x^T R}_{\text{כי נקבל}} + \underbrace{\tfrac{1}{2} x^T R x Q}_{\text{PD לפי סעיף 3}} + \underbrace{R x x^T Q}_{\text{כי לבד}}$$

נפתח את שני הביטויים הנוספים:

$$R x x^T Q = (Rx)(x^T Q) = \begin{pmatrix} 4x_1 + 3x_2 \\ 3x_1 + 4x_2 \end{pmatrix} \cdot (3x_1 - 2x_2 \ , \ -2x_1 + 3x_2) =$$

$$= \begin{bmatrix} 12x_1^2 + x_1x_2 - 6x_2^2 & -8x_1^2 + 6x_1x_2 + 9x_2^2 \\ 9x_1^2 + 6x_1x_2 - 8x_2^2 & -6x_1^2 + x_1x_2 + 12x_2^2 \end{bmatrix}$$

נשים לב כי: $(Rx)(x^TQ) = ((Qx)(x^TR))^T$

לכן

(1) $R x x^T Q + Q x x^T R = \begin{bmatrix} 2(12x_1^2 + x_1x_2 - 6x_2^2) & x_1^2 + 12x_1x_2 + x_2^2 \\ x_1^2 + 12x_1x_2 + x_2^2 & 2(12x_2^2 + x_1x_2 - 6x_1^2) \end{bmatrix}$

בנוסף אם נציב בביטויים הנותרים, נקבל:

(2) $\frac{1}{2}(x^T Q x) R = \frac{1}{2}[3x_1^2 + 3x_2^2 - 4x_1x_2]\begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$

(3) $\frac{1}{2}(x^T R x) Q = \frac{1}{2}[4x_1^2 + 4x_2^2 + 6x_1x_2]\begin{bmatrix} 3 & -2 \\ -2 & 3 \end{bmatrix}$

אם נבחר איזשהו וקטור x, כדי ש $g(x)$ תהיה קמורה, אז לכל $x \in \mathbb{R}^2$ חייב להתקיים כי (1)+(2)+(3) הוא PD

נבחר וקטור $x = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$ ונציב לביטויים האלה:

(1) $= \begin{bmatrix} 24 & 1 \\ 1 & -12 \end{bmatrix}$ (2) $= \frac{1}{2}\cdot 3 \cdot \begin{bmatrix} 4 & 3 \\ 3 & 4 \end{bmatrix}$ (3) $= \frac{1}{2}\cdot 4 \cdot \begin{bmatrix} 3 & -2 \\ -2 & 3 \end{bmatrix}$

$$\nabla^2 g(x) = \begin{bmatrix} 36 & 1.5 \\ 1.5 & 0 \end{bmatrix} \Rightarrow \det\begin{vmatrix} 36-\lambda & 1.5 \\ 1.5 & -\lambda \end{vmatrix} = -(36-\lambda)\lambda - 1.5^2 = 0 \rightarrow \lambda_1 = 36.06, \ \lambda_2 = -0.06$$

$(tr(\nabla^2 g(x)) = 36)$

לכן יש ע"ע שהם שליליים

לכן $\nabla^2 g(x) \not\succeq 0$

ולכן $g(x)$ איננה קמורה.

314092826 | 046197 | גרימנפיל יקטרינ | HW2

תרגיל 3

הוכיחו כי f הבאה קמורה עבור $x>0 \in \mathbb{R}^n$:

$g(x) = -\prod_{i=1}^{n} x_i^{p_i}$, $p_i \geq 0$, $\sum_i^n p_i = 1$

קודם כל נשים לב כי אם קיים $p_i = 1$ נקבל כי $g(x) = x_i$ שעבורו $\nabla^2 g(x) = 0$

אחרת נבדוק את המקרה בו אין $p_i = 1$:

נבדוק תנאי לקמירות מספר שני:

ניעזר בטענת עזר. נראה כי $-\frac{1}{g(x)}$ - קמורה.

$-\frac{1}{g(x)} = -\frac{1}{-\prod_{i=1}^{n} x_i^{p_i}} = \frac{1}{\prod_{i=1}^{n} x_i^{p_i}} = \prod_{i=1}^{n} x_i^{-p_i} =$

$= e^{\ln\left(\prod_{i=1}^{n} x_i^{-p_i}\right)} = e^{-\sum_{i=1}^{n} p_i \ln(x_i)}$

נסמן כי $f(x) = \sum_{i=1}^{n} p_i \ln(x_i)$, $h(x) = e^{-x}$

נבדוק, ראשית $f(x)$ קעורה. ו $h(x)$ קמורה ולא עולה, לכן לפי משפט על ההרכבה של פ' קמורות: $\left(h(f(x)) = -\frac{1}{g(x)}\right)$

g convex if h convex and non~~decreasing~~ increasing and f(x) concave

$f(x)$: לפי פלט עזר פ' $\ln(x)$ קעורה כי $(\ln x)' = \frac{1}{x}$, $(\ln x)'' = -\frac{1}{x^2} < 0$ ← קעורה.

לכן לכל $1 \leq i \leq n$ $\ln(x_i)$ פ' $f(x)$ היא הרכבה של פ' קעורות לכן ← קעורה. נשים לב כי

בנוסף, $p_i \geq 0$ לכן $p_i \ln(x_i)$ גם קעורה ולכן גם $\sum p_i \ln(x_i)$ קעורה.

$h(x)$: קמורה ולא עולה - מתקיים לפי גזירה: $h'(x) = -e^{-x} < 0$ ← h פ' יורדת = לא עולה

$h''(x) = e^{-x} > 0$ ← קמורה לכל x

לכן נקבל כי $h(f(x))$ היא קמורה ולכן גם $g(x)$ קמורה.

טענת עזר: תהיה $f: \mathbb{R}^n \to \mathbb{R}$ כך ש $f(x) > 0$ לכל $x \in \mathbb{R}^n$

אז אם $-\frac{1}{f(x)}$ - קמורה, אז גם $f(x)$ קמורה.

הוכחה: נגדיר $h(x) = -\frac{1}{x}$ $h: \mathbb{R}^- \to \mathbb{R}$

$g(x)$ - קמורה. $g(x) = \frac{1}{f(x)}$ $g: \mathbb{R}^n \to \mathbb{R}$

$\mathbb{R}/\mathbb{R}^+$ מאחר שהפונקציה מוגדרת עליהם

מתקיים כי הפונקציה שלנו: $f(x) = h(g(x))$

לכן $h(x)$ חייב פ' מונוטונית- לא יורדת וקמורה.

לפי תנאי השני לגזירות: $h'(x) = \frac{1}{x^2}$, $h''(x) = -\frac{2}{x^3}$

כלומר עבור $x < 0$ נקבל כי $h''(x) \geq 0$ וגם כי $h'(x) > 0$ אז היא עולה

בנוסף מתוך כך ש $f(x) > 0$ ו $g(x) < 0$ הם משפרים על אותו תחום לכן לפי תכונה משמרת קמירות מס 6 מהתרגולים, נקבל כי $f(x)$ קמורה.

תרגיל 4

f פ' (לא בהכרח גזירה) מוגדרת בקטע $I\subset\mathbb{R}$, $x\neq y$ $\quad \Delta f(x,y)=\frac{f(x)-f(y)}{x-y}.$ (קירוב לנגזרת)

1. צריך להוכיח כי $\Delta f(x,y)=\Delta f(y,x)$

$$\Delta f(x,y)=\frac{f(x)-f(y)}{x-y}=\frac{-(f(y)-f(x))}{-(y-x)}=\frac{f(y)-f(x)}{y-x}=\Delta f(y,x)$$

2. הוכיחו כי f קמורה אמ"מ לכל $x_1<x_2<x_3$ $(x_i\in I)$ מתקיים $(x_3-x_2)f(x_1)+(x_1-x_3)f(x_2)+(x_2-x_1)f(x_3)\geq 0$ ⊛

⇐ נניח כי f קמורה. נוכיח כי ⊛ מתקיים.

מכיוון ש f קמורה, לכן קיים $\lambda\in[0,1]$. לפי $x_1<x_2<x_3$, נקודה x_2 נמצאת בין x_1,x_3 לכן נגדיר אותה כ:

$$x_2=\lambda x_1+(1-\lambda)x_3 \Rightarrow x_2=\lambda x_1+x_3-\lambda x_3 \Rightarrow \lambda=\frac{x_2-x_3}{x_1-x_3} \Rightarrow (1-\lambda)=1-\frac{x_2-x_3}{x_1-x_3}=\frac{x_1-x_2}{x_1-x_3}$$

נבטא את λ.

מתוך הגדרת קמירות של f:

$$f(x_2)=f(\lambda x_1+(1-\lambda)x_3)\leq\lambda f(x_1)+(1-\lambda)f(x_3)$$

נעביר אגפים!

$$\lambda f(x_1)+(1-\lambda)f(x_3)-f(x_2)\geq 0$$

נציב את λ:

$$\underbrace{\frac{x_2-x_3}{x_1-x_3}}_{=\frac{x_3-x_2}{x_3-x_1}}f(x_1)+\underbrace{\frac{x_1-x_2}{x_1-x_3}}_{=\frac{x_2-x_1}{x_3-x_1}}f(x_3)-f(x_2)\geq 0$$

(x_1-x_3) ביטוי שלילי לכן נרצה להכפיל ב x_3-x_1 שהוא ביטוי חיובי, לכן נכפול גם סימן במונה של λ, $(1-\lambda)$

נכפיל ב (x_3-x_1)

$$(x_3-x_2)f(x_1)+(x_2-x_1)f(x_3)-(x_3-x_1)f(x_2)\geq 0$$
$$(x_3-x_2)f(x_1)+(x_2-x_1)f(x_3)+(x_1-x_3)f(x_2)\geq 0$$

כנדרש.

⇒ נניח כי ביטוי ⊛ מתקיים. נוכיח כי f קמורה. נגדיר $y_1,y_3\in I$ ו-y_2 והיעזר במשפט לקמירות נניח כי קיים $\lambda\in[0,1]$ שעבורו נוכיח כי: $f(\lambda y_1+(1-\lambda)y_3)\leq\lambda f(y_1)+(1-\lambda)f(y_3)$

נגדיר בדומה למקרה הקודם, $y_3=x_3$, $y_1=x_1$, $x_2=y_2=\lambda y_1+(1-\lambda)y_3$ שזה תואם לדרישה $x_1<x_2<x_3$

$$\Rightarrow \lambda=\frac{y_3-y_2}{y_3-y_1} \quad (1-\lambda)=\frac{y_2-y_1}{y_3-y_1}$$

נציב בביטוי:

$$(y_3-y_2)f(y_1)+(y_1-y_3)f(y_2)+(y_2-y_1)f(y_3)\geq 0$$
$$(y_3-y_1)f(y_2)\leq(y_3-y_2)f(y_1)+(y_2-y_1)f(y_3) \qquad :(y_3-y_1)$$
$$f(y_2)\leq\frac{y_3-y_2}{y_3-y_1}f(y_1)+\frac{y_2-y_1}{y_3-y_1}f(y_3)$$
$$f(y_2)\leq\lambda f(y_1)+(1-\lambda)f(y_3)$$
$$f(\lambda y_1+(1-\lambda)y_3)\leq\lambda f(y_1)+(1-\lambda)f(y_3)$$

נחליף משתנים בחזרה: לכן גם

$$f(\lambda x_1+(1-\lambda)x_3)\leq\lambda f(x_1)+(1-\lambda)f(x_3)$$

משפט לקמירות מתקיים, לכן f קמורה.

3. הוכיחו כי אם f קמורה, ו $x_1<x_2<x_3$ מתקיים $\frac{f(x_2)-f(x_1)}{x_2-x_1}\overset{①}{\leq}\frac{f(x_3)-f(x_1)}{x_3-x_1}\overset{②}{\leq}\frac{f(x_3)-f(x_2)}{x_3-x_2}$.

הביטוי הזה בעצם נובע מתוך סעיף הקודם, נסדר מחדש

$$(x_2-x_1)f(x_3)\geq-(x_3-x_2)f(x_1)-(x_1-x_3)f(x_2)$$

נוסיף $[+x_1f(x_1)]$

$$(x_2-x_1)f(x_3)-x_2f(x_1)\geq-x_3f(x_1)+(x_3-x_1)f(x_2)$$
$$(x_2-x_1)f(x_3)-(x_2-x_1)f(x_1)\geq(x_3-x_1)f(x_2)-(x_3-x_1)f(x_1)$$
$$(x_2-x_1)[f(x_3)-f(x_1)]\geq(x_3-x_1)[f(x_2)-f(x_1)]$$

היות ו $x_2-x_1>0$ וגם $x_3-x_1>0$ לכן אפשר לחלק בהם בלי שינוי סימן של אי שוויון:

$$\frac{f(x_3)-f(x_1)}{(x_3-x_1)}\geq\frac{f(x_2)-f(x_1)}{(x_2-x_1)}$$ ① כנדרש

ביטוי כאן שוויון השני ② ניתן לקבל על ידי הוספת $x_3f(x_3)$ לשני האגפים של אי שוויון וסידור ונקבל גם

$$(x_3-x_1)(f(x_3)-f(x_2))\geq(x_3-x_2)(f(x_3)-f(x_1))$$
$$\frac{f(x_3)-f(x_2)}{x_3-x_2}\geq\frac{f(x_3)-f(x_1)}{x_3-x_1}.$$

HW2 גרמנוביץ יונתן 046197 314092826

תרגיל 4 המשך

4) נדרש להוכיח כי $a\le b\le c\le d\le e$: בסעיף זה ניתן להיעזר בהגדרת השיפועים מתוך הנתון:

$a=\Delta f(x,y)=\frac{f(x)-f(y)}{x-y} \overset{\text{לפי 1.}}{\underset{\text{שקול}}{=}} \frac{f(y)-f(x)}{y-x}.$

$b=\Delta f(x,z)=\frac{f(x)-f(z)}{x-z}$

$c=\Delta f(y,z)=\frac{f(y)-f(z)}{y-z}$

$d=\Delta f(y,w)=\frac{f(y)-f(w)}{y-w}$

$e=\Delta f(z,w)=\frac{f(z)-f(w)}{z-w}$

באנלוגיה לסעיף 3 : עבור $x<y<z$ ($x<x_2<x_3$) נקבל:

$\frac{f(y)-f(x)}{y-x}\le\frac{f(z)-f(x)}{z-x} \Rightarrow \underline{\underline{a\le b}}$

עבור $y<z<w$ נקבל: $\frac{f(z)-f(y)}{z-y}\le\frac{f(w)-f(y)}{w-y} \Rightarrow \underline{\underline{c\le d}}$

ומתוך התנאי השני שהבאנו:

$\frac{f(z)-f(x)}{z-x}\le\frac{f(z)-f(y)}{z-y} \Rightarrow \underline{\underline{b\le c}}$

$\frac{f(w)-f(y)}{w-y}\le\frac{f(w)-f(z)}{w-z} \Rightarrow \underline{\underline{d\le e}}$

נאחד את כל התנאים ונקבל: $a\le b\le c\le d\le e$

5) הוכיחו כי f קמורה $\underline{\underline{\text{אמ"מ}}}$ $g_y(x)=\Delta f(x,y)$ לא יורדת ב-x לכל $y\ne x\in I$

$\Leftarrow$ $\underline{\text{אם } f \text{ קמורה}}$. צ"ל כי $g_y(x)=\Delta f(x,y)$ לא יורדת לכל $y\ne x\in I$

אם f קמורה, אז קיים $x_1,x_2\in I$ כך ש $w_1,w_2\ne y$

נפריד למקרים: 1 $y<w_1<w_2$ ((כאשר $y=x_1, w_1=x_2, w_2=x_3$)) אז:

$$\frac{f(w_1)-f(y)}{w_1-y}\le\frac{f(w_2)-f(y)}{w_2-y}\le\frac{f(w_2)-f(w_1)}{w_2-w_1}$$

$$\Delta f(w_1,y)\le\Delta f(w_2,y)$$

$$g_y(w_1)\le g_y(w_2)$$

2 $w_1<y<w_2$ ($x_1=w_1,\ y=x_2,\ w_2=x_3$) אז:

$$\frac{f(y)-f(w_1)}{y-w_1}\le\frac{f(w_2)-f(w_1)}{w_2-w_1}\le\frac{f(w_2)-f(y)}{w_2-y}.$$

$$\Delta f(w_1,y)\le\Delta f(w_2,w_1)\le\Delta f(w_2,y)$$

$$g_y(w_1)\le g_{w_1}(w_2)\le g_y(w_2)$$

3 $w_1<w_2<y$ ($x_1=w_1,\ x_2=w_2,\ x_3=y$ (כאשר)) אז:

$$\frac{f(w_2)-f(w_1)}{w_2-w_1}\le\frac{f(y)-f(w_1)}{y-w_1}\le\frac{f(y)-f(w_2)}{y-w_2}$$

$$\Delta f(w_2,w_1)\le\Delta f(w_1,y)\le\Delta f(w_2,y)$$

$$g_y(w_1)\le g_y(w_2)$$

$\Rightarrow$ אם $g_y(x)=\Delta f(x,y)$ לא יורדת לכל $x\ne y\in I$ אז f קמורה:

יהיו $x_1,x_3\in I$ ($x_1<x_3$) כך שאם נבחר $x_2=\lambda x_1+(1-\lambda)x_3$ עבור $\lambda\in[0,1]$ (כמובן מתקיים כי $x_1<x_2<x_3$)

לפי ההגדרה $g_y(x)=\Delta f(x,y)$ לא יורדת עבור x, אז נבחר $y=x_2$:

$$g_{x_2}(x_1)\le g_{x_2}(x_3)$$

$$\frac{f(x_1)-f(x_2)}{x_1-x_2}\le\frac{f(x_3)-f(x_2)}{x_3-x_2}.$$ אז

נשים לב כי $x_1-x_2<0$ וגם $x_3-x_2>0$ לכן אם נכפול:

$$(f(x_1)-f(x_2))(x_3-x_2)\ge(x_1-x_2)(f(x_3)-f(x_2))$$

$$(x_3-x_2)f(x_1)-(x_3-x_2)f(x_2)\ge(x_1-x_2)f(x_3)-(x_1-x_2)f(x_2)$$

$$(x_3-x_2)f(x_1)+(x_2-x_1)f(x_3)\ge(-x_1+x_2+x_3-x_2)f(x_2)$$

נציב x_2 מקיום הנ"ל: $(x_3-x_2)f(x_1)+(x_2-x_1)f(x_3)\ge(x_3-x_1)f(x_2)$

$$(x_3-(\lambda x_1+(1-\lambda)x_3))f(x_1)+(\lambda x_1+(1-\lambda)x_3-x_1)f(x_3)\ge(x_3-x_1)f(x_2)$$

$$(x_3-x_1)\lambda f(x_1)+(x_3-x_1)(1-\lambda)f(x_3)\ge(x_3-x_1)f(x_2)$$

נשים לב כי x_3-x_1 גם נוכל לחלק בו את התנאי ונקבל:

$$\lambda f(x_1)+(1-\lambda)f(x_3)\ge f(\lambda x_1+(1-\lambda)x_3)$$

סוף 5

314092826   046197   ג'רימנוביץ' יקטרינה   HW2

תרגיל 5

צריך לקבוע האם המודלים קמורים:
על מנת להראות שבעיה קמורה נראה כי גם פ' המטרה קמורה וגם תחום האילוץ

① $\min_x \{ x^2+x^4 \mid |x|-|x-3| \le 2 \}$

נסמן $f_1(x)=x^2+x^4$ וגם $C_1=\{x \mid |x|-|x-3|\le 2\}$

ניתן להראות כי f_1 הינה קמורה בשני דרכים
1) היות ו-$f_1(x)$ היא פ' גזירה פעמיים ברציפות לכל x, לכן:
$f_1'(x)=2x+4x^3$
$f_1''(x)=2+12x^2 \ge 0 \Rightarrow f_1(x)$ convex

2) כמו כן, ראינו כי x^p היא פ' קמורה, לכן סכום של פ' קמורות הוא גם קמור (לפי פלט ספר)
לכן $f_1(x)$ קמורה

נבדוק את C_1: עבור $x<0$: $-x+(x-3)\le 2$
$-3\le 2$ תמיד מתקיים

תחום האילוץ היא בפער חסום בין $|x|$ ל-$|3-x|$

עבור $0\le x\le 3$: $x+(x-3)\le 2$
$2x\le 5 \to x\le 2.5$
לכן $x\in C_1$ אם $0\le x\le 2.5$

עבור $x>3$: $x-x+3\le 2 \to 3\le 2 \Rightarrow x\notin C_1$ בתחום הזה.

לכן אם קיבלנו
תחום $-\infty\le x\le 2.5$ בתחום הזה C_1 תהיה קמורה
(ניתן להראות כי $C_1=\mathrm{epi}\{f(x)\}$ תחום קמור עבור $x\le 2.5$
לכן תחום כולו גם קמור)

② $\min_{(x,y)} \{ x^2y^2 \mid x+y\ge 4,\ x^2+y^2\le 10 \}$

נסמן $f_2(x,y)=x^2y^2$ פ' זו גזירה פעמיים ברציפות, לכן:

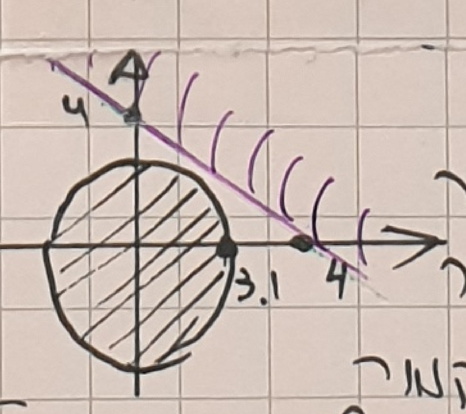
$$\nabla^2 f_2(x,y)=\begin{bmatrix} \frac{\partial^2 f_2}{\partial x^2} & \frac{\partial^2 f_2}{\partial y\partial x} \\ \frac{\partial^2 f_2}{\partial x\partial y} & \frac{\partial^2 f_2}{\partial y^2} \end{bmatrix}=\begin{bmatrix} 2y^2 & 4xy \\ 4xy & 2x^2 \end{bmatrix}$$

כעת נבדוק האם ההסיאן הוא $\nabla^2 f_2 \succeq 0$.
נבחר נקודה $(x_0,y_0)=(2,2)$ עבורה מתקיים

$$\nabla^2 f_2(2,2)=2\begin{bmatrix} 4 & 8 \\ 8 & 4 \end{bmatrix} \Rightarrow \mathrm{tr}=16$$
$$\det=16-64<0$$

$\begin{vmatrix} 4-\lambda & 8 \\ 8 & 4-\lambda \end{vmatrix}=0$
$(4-\lambda)^2=64$
$4-\lambda=\pm 8$
$\lambda=-4,\ \lambda=12$

לכן בהכרח קיימים ע"ע שליליים
ולכן $\nabla^2 f_2(x,y)\not\succeq 0$
ולכן f_2 לא קמורה



תחום האילוץ אם נחלקו לתחומים אז גרפית ניתן לראות כי מדובר במעגל כלל רדיוס 3.1 וחצי מישור כל אחד מהאילוצים האלה הוא קמור בנפרד (תרגול מספר 2 תרגיל 1 + דף ספר קמירות)

לכן בעיה ② לא קמורה.

③ $\min_{x,y} \left\{ \sqrt{1+x^2}-\sqrt{1-y^2} \;\middle|\; \frac{\sqrt{(x-1)^2+y^2}}{\sqrt{x^2+(y-2)^2}} \le \frac{1}{2},\ |y|<1 \right\}$

נסמן $C_3=\left\{x,y \;\middle|\; \frac{\sqrt{(x-1)^2+y^2}}{\sqrt{x^2+(y-2)^2}}\le \frac{1}{2},\ |y|<1\right\}$, $f_3=\sqrt{1+x^2}-\sqrt{1-y^2}$

נראה ראשית שהתחום קמור: היות והביטוי במכנה חיובי (רשום)
$\sqrt{(x-1)^2+y^2} \le \frac{1}{2}\sqrt{x^2+(y-2)^2}$
מתוך תרגול 2, תרגיל 3: $\|x-a\|_2 \le \mu\|x-b\|_2$
$\left\|\begin{bmatrix} x \\ y \end{bmatrix}-\begin{bmatrix} 1 \\ 0 \end{bmatrix}\right\|_2 \le \frac{1}{2}\left\|\begin{bmatrix} x \\ y \end{bmatrix}-\begin{bmatrix} 0 \\ 2 \end{bmatrix}\right\|_2$
עבור $1>\mu>0$ המישור יהיה מוגדר כפנים של הכדור ולכן קמור.

$|y|<1$ תחום קמור (ניתן להוכיח לפי הגדרת קמורות - נניח $y_1,y_2\in C_3$) שעבור כל $\lambda\in[0,1]$ מתקיים:
$|\lambda y_1+(1-\lambda)y_2| \le |\lambda||y_1|+|(1-\lambda)||y_2| = \lambda|y_1|+(1-\lambda)|y_2| < \lambda+(1-\lambda)=1$
(כי y_1,y_2 בתוך התחום)

לכן כחיתוך בין 2 קמוצות נקבל כי C_3 הוא קמור.

פ' מטרה f_3:
ראינו בתרגול כי עבור פ' $g(x)=\sqrt{1+x^2}$
קיימת פ' במוצה $h(y)=-\sqrt{1-y^2}$ שהיא תהיה קמורה לפי הגדרה שהוכחנו. נוכיח רגע כי $g(x)$ קמורה.
פ' גזירה ברציפות פעמיים, לכן:
$g'(x)=\frac{x}{\sqrt{1+x^2}}$
$g''(x)=\frac{1}{(x^2+1)^{3/2}}>0 \Leftrightarrow$ תמיד הביטוי מוגדר חיובית $\Rightarrow g(x)$ קמורה
לפי הגדרה, סכום של פ' קמורות הוא קמור $\Leftarrow$
בעית אופטימיזציה זו הוא קמורה

④ $\min_{x,y,z} \{ 2x+3y+4z \mid xyz\ge 1,\ \sqrt{xy}\ge z,\ x>0,\ y>0,\ z>0 \}$

פונקציית מטרה היא חיבור לינארי של 3 פונקציות אפיניות, לכן היא קמורה

תחום האילוץ

$x>0,\ y>0,\ z>0$ - הינם אילוצים אפיניים $\Leftarrow$ קמורים

$x\cdot y\cdot z\ge 1$, - כפי שראינו מתרגיל 3, ניתן לרשום את הבעיה כ:
$x^{\frac{1}{3}}y^{\frac{1}{3}}z^{\frac{1}{3}}>1^{1/3}$
$\underbrace{-x^{1/3}y^{1/3}z^{1/3}<-1^{1/3}}_{\text{sublevel set}}$ $\Rightarrow$ הקבוצה הזאת קמורה: תת קבוצה של $-x^{1/3}y^{1/3}z^{1/3}$ קמור.

$\sqrt{xy}\ge z$ - נגדיר $f_5(x,y,z)=z-\sqrt{xy}$, כאשר $f_5\le 0$
z היא קבוצה אפינית ולכן קמורה.
$-\sqrt{xy}$ - בדומה לתחום הקודם ניתן לרשום כ $-\sqrt{xy}=-x^{\frac{1}{2}}y^{\frac{1}{2}}z^{0}$ שהוא לפי תרגיל 3 גם כן קמור.

לכן הבעיה קמורה (גם תחום וגם פונקציית מטרה)

HW2 יקטרינה גרימוביץ' 094197 314092826

תרגיל 6

1) יהיה $0\le q_1 \le q_2 \le \dots \le q_n$ מספרים נתונים. הוכיחו כי $\sum_{i=1}^n q_i s_i \le q_n$ כאשר $s_i \ge 0$ לכל i מתקיים $\sum_{i}^{n} s_i = 1$

(מצאו $(s_i)_{i=1}^n$ עבור האי שוויון שמתקבל בשוויון)

הוכחה: היות ו-q_n הוא הערך המרבי של סדרת המספרים הנתונה, לכן:

$$\sum_{i=1}^n q_i s_i \le \sum_{i=1}^n q_n s_i = q_n \sum_{i=1}^n s_i = q_n$$

ניתן לראות כי $s_i = \begin{cases} 0 & i\ne n \\ n & i=n \end{cases}$ אי שוויון מתקיים בשוויון

2) תהי Q סא"ס ב-$R^{n\times n}$ נניח כי הערכים העצמיים של Q מסודרים בסדר עולה $0\le q_1\le q_2 \le \dots \le q_n$

הוכיחו כי $q_n = \max_x \{x^TQx \mid x^Tx=1\}$

היות ומטריצה Q סא"ס ב-$R^{n\times n}$ לכן היא סימטרית. לפי משפט הפירוק הספקטרלי, קיימת מטריצה יוניטרית וממשית $\underline{U} \in R^{n\times n}$ ומטריצה אלכסונית $\Lambda = diag(q_1 \dots q_n)$ כך שמתקיים $\underline{Q} = \underline{U}^T \underline{\Lambda} \underline{U}$

נסמן $\underline{U}\,\underline{x} = \underline{u}$ לכן: $x^TQx = x^TU^T\Lambda Ux = (\underline{U}\underline{x})^T\Lambda(\underline{U}\underline{x}) = \underline{u}^T\Lambda \underline{u}$ $\quad$ $\underline{U}^T\underline{u} = \underline{x}$

בנוסף ידוע כי מטריצה יוניטרית לא משנה את ערכה של הנורמה, לכן

$$x^Tx = u^T \underbrace{UU^T}_{=I} u = u^Tu = 1$$

מכאן, מכיוון ש-$q_n \dots q_1$ מסודרים בסדר עולה ואם סא"ס לכן הם כולם אי שליליים, ולכן:

$$q_n^* = \max_u \{u^T\Lambda u \mid u^Tu = 1\} = \max_u \{\sum q_i u_i^2 \mid \sum_i^n u_i^2 = 1\} = \max_{s_i}\{\sum_i^n q_i s_i \mid \sum_i^n s_i = 1\} \underset{\text{לפי סעיף 1}}{\le} q_n$$

(נסמן $s_i = u_i^2$)

לכן זה גם מתקיים עבור בעיה מקורית.

3) הוכיחו כי אם $x\ne 0$ אז $q_n \ge \frac{x^TQx}{x^Tx}$

בעצם אנחנו מתבקשים לפתור בעיה: $\max_x \left\{\frac{x^TQx}{x^Tx}\right\}$

נגדיר משתנה עזר $y = \frac{x}{\|x\|_2}$ ונציבה בבעיה: $\quad$ ($x = y\|x\|$)

(מותר לעשות החלפה הזאת כי $x\ne 0$ לפי נתוני הבעיה)

$$\frac{x^TQx}{x^Tx} = \frac{x^TQx}{\|x\|_2^2} = \frac{x^T}{\|x\|_2} Q \frac{x}{\|x\|_2} = y^TQy$$

בנוסף: נשים לב כי האילוץ לבעיה: $y^Ty = \|y\|_2^2 = \frac{\|x\|_2^2}{\|x\|_2^2} = 1$

כלומר קיבלנו בעצם בעיה שקולה לבעיה מסעיף הקודם $g(y) = \max_y\{y^TQy \mid y^Ty = 1\}$

בסעיף הקודם ראינו כי בעיה מקיימת כי $q_n^* \le q_n$ לכן גם כאן: $y^TQy = \frac{x^TQx}{x^Tx} \le q_n$