jonobarel@gmail.com

אופטימיזציה של פונקציה בלי אילוצים

תהיה $A \subset \mathbb{R}^n$, $f:A \to \mathbb{R}$ אזי נאמר שבנק' $x_0 \in A$ יש מקס' גלובאלי ל-f אם $f(x_0) \geq f(x)$ לכל $x \in A$.

נאמר שבנק' A יש מקסימום לוקאלי ל-f אם קיימת $\delta > 0$ כל שלכל $|x - x_0| < \delta$ מתקיים $f(x_0) \geq f(x)$.

x_0 תקרא נק' קריטית של f אם לכל $1 \leq i \leq n$ $f_{x_i}(x_0) = 0$.

טענה

תהיה $A \subset \mathbb{R}^n$, $f:A \to \mathbb{R}$ אז אם x_0 היא נקודת אקסטרמום של f וקיימת ל-f כל הנגזרות בנק' x_0 אזי x_0 היא נק' קריטית של f.

דוגמא

$$f(x,y) = x \cdot y$$
$$f_x(x,y) = y \quad f_y(x,y) = x$$
$$f_x = 0 \Rightarrow y = 0$$
$$f_y = 0 \Rightarrow x = 0$$

← הנקודה הקריטית היא $(x,y) = (0,0)$ היא לא נק' מקסימום.

משפטי אקסטרמום ב- $\mathbb{R}^2$

טענה: תהיה $f:\mathbb{R}^2 \to \mathbb{R}$ אז תנאי מספיק לכך ש- (x_0, y_0) היא נק' מקסימום לוקאלי עבור f היא:

1. $f_x(x_0, y_0) = 0 \quad f_y(x_0, y_0) = 0$
2. $$f_{xx}(x_0, y_0) < 0$$ $$f_{xx}(x_0, y_0) f_{yy}(x_0, y_0) - f_{xy}(x_0, y_0) f_{yx}(x_0, y_0) > 0$$

תנאי הכרחי לכך ש- (x_0, y_0) תהיה מקסימום לוקאלי הוא אי שיוויון חלש ב - (2)

jonobarel@gmail.com

<u>טענה</u>

תהיה $f:\mathbb{R}^2\rightarrow\mathbb{R}$ אז תנאי מספיק לכך ש- $(x_\cdot, y_\cdot)$ היא נק' מינימום לוקאלי עבור f הוא:

1. $f_x(x_0, y_0)=0 \quad f_y(x_0, y_0)=0$

2. $$\begin{gathered} f_{xx}(x_0, y_0)>0 \\ f_{xx}(x_0, y_0) f_{yy}(x_0, y_0)-f_{xy}(x_0, y_0) f_{yx}(x_0, y_0)>0 \end{gathered}$$

תנאי הכרחי לכך ש- (x_0, y_0) תהיה מינימום לוקאלי הוא אי שיוויון חלש ב – (2)

<u>דוגמא</u>

$$\begin{gathered} f(x, y)=x^2+y^2-2x-2y \\ f_x=2x-2=0\Rightarrow x=1 \\ f_y=2\mathrm{y}-2=0\Rightarrow y=1 \\ f_{xx}=2>0 \\ f_{xx}\cdot f_{yy}-f_{xy}\cdot f_{yx}=2\cdot 2-0=4>0 \end{gathered}$$

ולכן הנקודה $(x, y)=(1,1)$ היא נק' מינימום.

<u>**משפטי אקסטרמום ב-**</u> $\mathbb{R}^n$

<u>טענה</u>

תהיה $f:A\rightarrow\mathbb{R}$, $A\subset\mathbb{R}^n$ אזי תנאי מספיק לכך ש- $x_0\in A$ היא נק' מקסימום לוקאלי עבור f הוא:

1. $\nabla f(x)=0$ (וקטור הנגזרות החלקיות)

2. $(-1)^i \Delta_i(x_0)>0, 1\leq i\leq n$

$$\Delta_i(x_0)=\begin{vmatrix} f_{x_1x_1} & \cdots & f_{x_1x_i} \\ \vdots & & \vdots \\ f_{x_ix_1} & \cdots & f_{x_ix_i} \end{vmatrix}(x_0)$$

jonobarel@gmail.com

תנאי הכרחי למקסימום הוא אי-שוויון חלש ב-(2)

תהיה $f:A\to\mathbb{R}$, $A\subset\mathbb{R}^n$ אזי תנאי מספיק לכך ש- $x_0\in A$ היא נק' מינימום לוקאלי עבור f הוא:

1. $\nabla f(x)=0$
2. $\Delta_i(x_0)>0\,,1\le i\le n$

$$\Delta_i(x_0)=\begin{vmatrix} f_{x_1x_1} & \cdots & f_{x_1x_i} \\ \vdots & & \vdots \\ f_{x_ix_1} & \cdots & f_{x_ix_i} \end{vmatrix}(x_0)$$

תנאי הכרחי לכך ש- x_0 תהיה נק' מינימום לוקאלי הוא אי-שוויון חלש ב-(2)

מסקנה

אם $f:A\to\mathbb{R}$ היא פונק' קעורה (קמורה) אז כל נק' קריטית היא נק' מקסימום (מינימום) לוקאלי.

אם A היא קבוצה קמורה אז נק' זו היא מקסימום (מינימום) גלובאלי.

דוגמא

$$f:\mathbb{R}^3\to\mathbb{R}$$
$$f(x,y,z)=x^3+xy+2y^2+xz+z^2$$

מהן נק' האקסטרמום?

נחפש את הנק' הקריטיות.

$$(1)\,f_x=3x^2+y+z=0$$
$$(2)\,f_y=x+4y=0$$
$$(3)\,f_z=x+2z=0$$
$$(4):(2)-(3)\Rightarrow z=2y$$
$$(5):(2)\Rightarrow x=-4y$$

jonobarel@gmail.com

נציב את (4)+(5) ב-(1)

$$\begin{gathered}3(-4y)^2+y+2y=0\\ \Rightarrow y(48y+3)=0\\ \Rightarrow y=0 \text{ or } y=\frac{-1}{16}\\ \Rightarrow (y=0,x=0,z=0),\ (y=\frac{-1}{16},x=\frac{1}{4},z=\frac{-1}{8})\end{gathered}$$

$$\begin{pmatrix} f_{xx} & f_{xy} & f_{xz}\\ f_{yx} & f_{yy} & f_{yz}\\ f_{zx} & f_{zy} & f_{zz}\end{pmatrix}=\begin{pmatrix} 6x & 1 & 1\\ 1 & 4 & 0\\ 1 & 0 & 2\end{pmatrix}$$

א. $(x,y,z)=(0,0,0)$

$$\begin{gathered}\Delta_1=f_{xx}=6x=0\\ \Delta_2=\begin{vmatrix}0 & 1\\ 1 & 4\end{vmatrix}=0-1=1<0\end{gathered}$$

(כזכור: $\Delta_2>0$ מינימום, $(-1)^2\Delta_2>0$ מקסימום)

לכן הנק' שלנו היא לא מינימום ולא מקסימום.

ב. $(x,y,z)=(\frac{1}{4},-\frac{1}{16},-\frac{1}{8})$

$$\begin{gathered}\Delta_1=f_{xx}=6x=1.5>0\\ \Delta_2=\begin{vmatrix}1.5 & 1\\ 1 & 4\end{vmatrix}=6-1=5>0\\ \Delta_3=\begin{vmatrix}1.5 & 1 & 1\\ 1 & 4 & 0\\ 1 & 0 & 2\end{vmatrix}=1.5\begin{vmatrix}4 & 0\\ 0 & 2\end{vmatrix}-1\begin{vmatrix}1 & 0\\ 1 & 2\end{vmatrix}+1\begin{vmatrix}1 & 4\\ 1 & 0\end{vmatrix}=6>0\\ \Delta_1>0,\ \Delta_2>0,\ \Delta_3>0\end{gathered}$$

הנק' היא נקודת מינימום לוקאלי.

<u>דוגמא</u>

jonobarel@gmail.com

$$g(x,y,z)=e^{-(x^2+y^2+z^2)}$$
$$g_x=-2x\cdot e^{-x^2+y^2+z^2}=0\Rightarrow x=0$$
$$g_y=-2y\cdot e^{-x^2+y^2+z^2}=0\Rightarrow y=0$$
$$g_z=-2z\cdot e^{-x^2+y^2+z^2}=0\Rightarrow z=0$$
$$g_{xx}=e^{-x^2+y^2+z^2}\cdot(-2)+e^{-x^2+y^2+z^2}\cdot(4x^2)$$
$$g_{xy}=4xy\cdot e^{-x^2+y^2+z^2}$$
$$g_{xz}=4xz\cdot e^{-x^2+y^2+z^2}$$
$$g_{xx}=e^{-x^2+y^2+z^2}\cdot(4x^2-2)=e^{-x^2+y^2+z^2}(-2)<0$$

$$e^{-x^2+y^2+z^2}=1\,(x=0,y=0,z=0)$$
$$\begin{vmatrix} g_{xx} & g_{xy} \\ g_{yx} & g_{yy} \end{vmatrix}=1\cdot\begin{vmatrix} 4x^2-2 & 4xy_{(\ =0)} \\ 4xy_{(\ =0)} & 4y^2-2 \end{vmatrix}=4>0$$
$$\begin{vmatrix} g_{xx} & g_{xy} & g_{xz} \\ g_{yx} & g_{yy} & g_{yz} \\ g_{zx} & g_{zy} & g_{zz} \end{vmatrix}=\begin{vmatrix} -2 & 0 & 0 \\ 0 & -2 & 0 \\ 0 & 0 & -2 \end{vmatrix}=-8<0$$

$(0,0,0)$ נק' מקסימום.

<u>מקסימיזציה עם אילוצים</u>

$$\max_{x\in\mathbb{R}^n} f(x)$$
$$(1)\, s.t. \begin{Bmatrix} g_1(x)=b_1 \\ \vdots \\ g_m(x)=b_m \end{Bmatrix}, m<n$$

$g_1,\dots,g_m, f$ ב"ת ומוגדרות על $\mathbb{R}^n$

קבוצת כל הנקודות המקיימות את היאלוצים (1) תקרא קבוצת האילוצים ותסומן על-ידי

$$C=\{x\in\mathbb{R}^n : g_i(x)=b_i, i=1,\dots,m\}$$

$\bar{x}\in\mathbb{R}^n$ תהיה נק' מקסימום לוקאלית מאולצת אם לכל x כך ש- $|x-\bar{x}|<\delta$ מתקיים $f(\bar{x})\geq f(x)$ והיא תיקרא נק' מקסימום גלובאלית מאולצת אם $f(\bar{x})\geq f(x)$ לכל $x\in C$

<u>טענה</u>

נניח שפונק' המטרה f ופונק' האילוצים g_j , $j=1,\dots,m$ דיפרנציאבילית ו- bar x in C היא נק' מקסימום לוקאלית מאולצת, אזי קיימים מס' $\lambda_j\in\mathbb{R}$, $j=1,\dots,m$ כל ש-

$$\frac{\partial f}{\partial x_i}(\bar{x})=\sum_{j=1}^{m}\lambda_j\frac{\partial g_j}{\partial x_i}(\bar{x})$$

כלומר,

$$\begin{pmatrix} f_{x_1} \\ \vdots \\ f_{x_n} \end{pmatrix}=\lambda_1\begin{pmatrix} \frac{\partial g_1}{\partial x_1} \\ \vdots \\ \frac{\partial g_1}{\partial x_n} \end{pmatrix}+\lambda_2\begin{pmatrix} \frac{\partial g_2}{\partial x_1} \\ \vdots \\ \frac{\partial g_2}{\partial x_n} \end{pmatrix}+\cdots+\lambda_m\begin{pmatrix} \frac{\partial g_m}{\partial x_1} \\ \vdots \\ \frac{\partial g_m}{\partial x_n} \end{pmatrix}$$

<u>דוגמא</u>

$$\max_{x,y,z\in\mathbb{R}_+} \sqrt{x}+\sqrt{y}+\sqrt{z}$$
$$\text{s.t.}\quad x+y+z=100$$
$$f_x=\lambda g_x \Rightarrow \frac{1}{2\sqrt{x}}=\lambda$$
$$f_y=\lambda g_y \Rightarrow \frac{1}{2\sqrt{y}}=\lambda$$
$$f_z=\lambda g_z \Rightarrow \frac{1}{2\sqrt{z}}=\lambda$$
$$\Rightarrow x=y=z=33\frac{1}{3}$$

<u>דוגמא 2</u>

$$\max_{x,y,z\in\mathbb{R}} \sqrt{x}+\sqrt{y}+\sqrt{z}$$
$$\text{s.t.}\quad g=x+y+z=100$$
$$h=x-y-z=0$$
$$f_x=\lambda_1 g_x+\lambda_2 h_x \Rightarrow \frac{1}{2\sqrt{x}}=\lambda_1+\lambda_2$$
$$f_y=\lambda_1 g_y+\lambda_2 h_y \Rightarrow \frac{1}{2\sqrt{y}}=\lambda_1-\lambda_2$$
$$f_z=\lambda_1 g_z+\lambda_2 h_z \Rightarrow \frac{1}{2\sqrt{z}}=\lambda_1-\lambda_2$$
$$\Rightarrow y=z$$
$$\Rightarrow x+2\mathrm{y}=100,\, x+2\mathrm{y}=0$$
$$\Rightarrow 2\mathrm{x}=100$$
$$x=50,\, y=z=25$$

<u>אילוצי אי-שיוויון</u>

$$\max_{x\in\mathbb{R}^n} f(x)$$
$$\text{s.t.}\quad \begin{matrix} g_1(x)=b_1 \\ \vdots \\ g_m(x)=b_m \\ h_1(x)\le c_1 \\ \vdots \\ h_k(x)\le c_k \end{matrix} \qquad m+k\le n$$

jonobarel@gmail.com

טענה:

נניח ש- $\bar{x}\in C$ היא נק' מקסימום לוקאלית של הבעיה, אזי קיימים מספר $\lambda_1\in\mathbb{R}$, $i=1,\dots,m$ ו- $\alpha_s\in\mathbb{R}_+$, $s=1,\dots k$ כך ש-

$$\begin{aligned}&(1)\frac{\partial f}{\partial x_i}(\bar{x})=\sum_{j=1}^{m}\lambda_j\frac{\partial g_j(\bar{x})}{\partial x_i}+\sum_{s=1}^{k}\alpha_s\frac{\partial h_s(\bar{x})}{\partial x_i}\\&(2)\qquad \alpha_s\cdot(h_s(\bar{x})-c_s)=0,\quad s=1,\dots,k\end{aligned}$$

לגבי (2): יש שני מצבים. אם האי-שיוויון כובל, ואז $h_s=c_s$, אז הביטוי הוא 0, ואין אילוץ על α .

אבל אם זה לא כובל, אז α חייב להיות 0, וזה עושה את החיים הרבה יותר קלים.

הבעיה היא שצריך "לנחש" או להניח לגבי אילוצי האי-שיוויון, אילו מהם כובלים ואילו לא. עבור ההנחה הזאת צריך לנסות לפתור, ואז אם מגיעים לסתירה, אז ההנחה לא נכונה.

צריך לפתור עבור הנחות שונות עד שאנחנו מצליחים למצוא סט הנחות (כובל, לא כובל) שלא מגיע לסתירה.