אופטימיזציה - חלק ד'

① הסבירו את החסרונות בשיטת gradient decent ותנו דוגמא הדורשת מספר גדול של איטרציות.

פתרון:

1. האלגוריתם צריך לצעוד על כל הדרכים לפני שמעדכנים יכול לגרום לאלגוריתם להיות מאוד איטי (הרבה מאוד איטי)

2. אפשר להיתקע במינימום לוקלי (יכול לקרות בגלל גודל הצעד העדכון, או בחירת נק' התחלה)

3. צעד העדכון יכול להשפיע על התכנסות האלגוריתם (יכול לגרום פספוס של המינימום גם)
אם ניקח צעד גדול מידי אפשר לפספס את הנק'.
אם ניקח צעד קטן מידי אפשר להיתקע במינימום לוקלי.

דוגמא מההרצאה שדורשת מספר רב של איטרציות:

$$f(x) = x_1^2 + \gamma x_2^2$$

$$x_0 = \frac{(\gamma, 1)}{\|(\gamma, 1)\|_2}$$

$$x_k = \frac{1}{\|(\gamma, 1)\|_2}\left(\frac{\gamma-1}{\gamma+1}\right)^k (\gamma, (-1)^k)$$

ניקח $\gamma = 3$ ונקבל:

$$f(x) = x_1^2 + 3x_2^2$$ (המינימום ב-(0,0))

$$x_0 = (3, 1)\cdot\frac{1}{\sqrt{10}}$$

לפי כלל העדכון x_k נקבל

$$x_1 = \frac{1}{\sqrt{10}}\cdot\frac{2}{4}(3,-1) = \frac{1}{2\sqrt{10}}(3,-1)$$

$$x_2 = \frac{1}{\sqrt{10}}\left(\frac{1}{2}\right)^2\cdot(3,1)$$

$$x_3 = \frac{1}{\sqrt{10}}\left(\frac{1}{2}\right)^3\cdot(3,-1)$$

$$\vdots$$

$$x_k = \frac{1}{\sqrt{10}}\cdot\left(\frac{1}{2}\right)^k(3,(-1)^k)$$

יידרש הרבה איטרציות בשביל להתקרב לנק' מינימום של f.

Scanned with CamScanner

אופטימיזציה - חלק א'

2 הגדרה:

החרוט הנוצר ע"י קבוצת וקטורים $x_1,\dots,x_k \in R^n$ הוא

$$\{x = \alpha_1 x_1 + \dots + \alpha_k x_k \mid \alpha_1,\dots,\alpha_k \ge 0\}$$

הוכח כי

א. לא כל החרוטים קמורים.

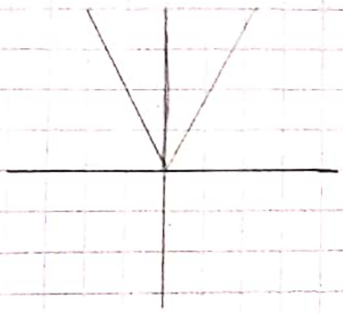
פתרון:

הגדרה של חרוט מההרצאה:

קבוצה K ב-R^n הסגורה תחת כפל בסקלר חיובי ממש נקראת חרוט

כלומר כל נק' ב-K כפול סקלר חיובי α נשאר ב-K ($\alpha x \in K$)

בשביל להראות שלא כל החרוטים קמורים, נבחר חרוט ו-2 וקטורים שנמצאים בחרוט אבל הצירוף לא בחרוט.



ניקח את $y=2x$, $y=-2x$

נגדיר את הקבוצה

$K = \{x \mid x = \alpha(1,2)$ או $x = \alpha(-1,2)\}$

הישר $y=2x$ הישר $y=-2x$

ניקח $\alpha_1=\frac{1}{3}$, $\alpha_2=\frac{2}{3}$ ואת הוקטורים $(1,2),(-1,2)$

$$\frac{1}{3}(1,2)+\frac{2}{3}(-1,2) = \left(\frac{1}{3},\frac{2}{3}\right)+\left(-\frac{2}{3},\frac{4}{3}\right)$$

$$= \left(-\frac{1}{3},2\right)$$

הוקטור $(-\frac{1}{3},2)$ לא נמצא בקבוצה K

מצאנו חרוט שאינו קמור

בס"ד

(2)

ב. כל החרוטים הנוצרים ע"י קב' וקטורים הם קמורים.

פתרון: ניקח קב' וקטורים $\{x_1,\dots,x_n\}$ $K = \{x = \alpha_1 x_1 + \dots + \alpha_n x_n \mid \alpha_1,\dots,\alpha_n \geq 0\}$

וצריך להראות ש-K קמור

יהי $k_1,k_2 \in K$ ו-$\beta_1,\beta_2 \geq 0$! צריך להראות שהצירוף הלינארי שלהם ב-K

כלומר $\beta_1 k_1 + \beta_2 k_2 \in K$

לקחנו את k_1,k_2 מהקבוצה K לכן קיים צירוף לינארי של $x_1,\dots,x_n$

עבור כל אחד מהם: $k_1 = \alpha_1 x_1 + \dots + \alpha_n x_n$

$k_2 = \gamma_1 x_1 + \dots + \gamma_n x_n$

נחשב את $\beta_1 k_1 + \beta_2 k_2$

$$\beta_1 k_1 + \beta_2 k_2 = \beta_1(\alpha_1 x_1 + \dots + \alpha_n x_n) + \beta_2(\gamma_1 x_1 + \dots + \gamma_n x_n) =$$

גורם משותף

$$= x_1(\beta_1\alpha_1 + \beta_2\gamma_1) + x_2(\beta_1\alpha_2 + \beta_2\gamma_2) + \dots + x_n(\beta_1\alpha_n + \beta_2\gamma_n)$$

קיבלנו צירוף לינארי של $x_1,\dots,x_n$ עם מקדמים חיוביים ולכן $\beta_1 k_1 + \beta_2 k_2 \geq 0$

לכן $\beta_1 k_1 + \beta_2 k_2 \in K$ החרוט קמור.

③ נתבונן בבעיית האופטימיזציה הבאה

$$\min_x \ (x_1-3/2)^2+(x_2-\tfrac{1}{2})^4$$

s.t.

$$1-x_1-x_2\geq 0$$
$$1-x_1+x_2\geq 0$$
$$1+x_1-x_2\geq 0$$
$$1+x_1+x_2\geq 0$$

הראו שבנק' $x^*=(1,0)$ מתקיימים תנאי KKT ותנאי ההכרח מסדר שני. בדקו האם התנאים המספיקים מסדר שני מתקיימים.

פתרון:

תנאי KKT:

$$\min_x \ f(x)$$

s.t.

$$c_i(x)=0 \qquad i\in E$$
$$c_i(x)\geq 0 \qquad i\in I$$

נגדיר לגרנז'יאן $L=f(x)-\sum_{i\in E\cup I}\lambda_i c_i(x)$

נניח x^* מינימום לוקלי, f, c_i גזירות ברציפות ותנאי LICQ מתקיים ב-x^*

אז קיים וקטור כופלי לגרנז' λ^* ק"ש

א. $\nabla_x L(x^*,\lambda^*)=0$

ב. $c_i(x^*)=0 \quad \forall i\in E$

ג. $c_i(x^*)\geq 0 \quad \forall i\in I$

ד. $\lambda_i^*\geq 0 \quad \forall i\in I$

ה. $\lambda_i^* c_i(x^*)=0 \quad \forall i\in E\cup I$

נבדוק שהתנאים מתקיימים:

הגרדיאנט של f

$$\nabla f(x)=\begin{pmatrix}2(x_1-\frac{3}{2})\\ 4(x_2-\frac{1}{2})^3\end{pmatrix}$$

$$\nabla f(x^*)=\nabla f((1,0))=\begin{pmatrix}-1\\ -\frac{1}{2}\end{pmatrix}$$

בס"ד

3

האילוצים C_1 ו C_2 הפעילים

$C_1: \quad 1-1-0=0$

$C_2: \quad 1-1+0=0$

$C_3: \quad 1+1-0=2$

$C_4: \quad 1+1+0=2$

כעת נבדוק ה- LICQ

נחשב גרדיאנטים לאילוצים הפעילים

$$\nabla_x C_1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \qquad \nabla_x C_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

הן בת"ל!

נמצא וקטור כופלי לגרנז' שמקיים את תנאי א'-ה'

האילוצים C_3 ו C_4 אינם פעילים ולכן $\lambda_4^* = \lambda_3^* = 0$.

הלגרנז'יאן הוא $L = f(x) - \sum \lambda_i c_i(x)$

$$\nabla_x L = \nabla_x f(x) - \sum \lambda_i \nabla_x c_i(x)$$

$\nabla_x f(x)$:

$$\nabla_x f(x) = 2\left(x_1 - \tfrac{3}{2}\right) + 4\left(x_2 - \tfrac{1}{2}\right)^3$$

$$\nabla_x f(x) = \begin{pmatrix} 2\left(x_1 - \frac{3}{2}\right) \\ 4\left(x_2 - \frac{1}{2}\right)^3 \end{pmatrix}$$

$\nabla_x C_i$:

$$\nabla_x C_1 = \begin{pmatrix} -1 \\ -1 \end{pmatrix} \qquad \nabla_x C_2 = \begin{pmatrix} -1 \\ 1 \end{pmatrix}$$

נציב x^* ב $\nabla_x L$ ונפתור $\nabla_x L(x^*) = 0$

$$\nabla_x L(x^*) = \begin{pmatrix} 2\left(1 - \frac{3}{2}\right) \\ 4\left(0 - \frac{1}{2}\right)^3 \end{pmatrix} - \lambda_1^* \begin{pmatrix} -1 \\ -1 \end{pmatrix} - \lambda_2^* \begin{pmatrix} -1 \\ 1 \end{pmatrix} =$$

2-3

$$= \begin{pmatrix} -1 \\ -\frac{1}{2} \end{pmatrix} - \lambda_1^* \begin{pmatrix} -1 \\ -1 \end{pmatrix} - \lambda_2^* \begin{pmatrix} -1 \\ 1 \end{pmatrix} = 0$$

$$-1 + \lambda_1^* + \lambda_2^* = 0$$

$$-\frac{1}{2} + \lambda_1^* - \lambda_2^* = 0 \Rightarrow \lambda_2^* = -\frac{1}{2} + \lambda_1^*$$

3

$$-1+\lambda_1^*+\lambda_2^*=0$$

$$\lambda_2^*=-\frac{1}{2}+\lambda_1^*$$

$$\Downarrow$$

$$-1+\lambda_1^*-\frac{1}{2}+\lambda_1^*=0$$

$$-\frac{3}{2}+2\lambda_1^*=0$$

$$2\lambda_1^*=\frac{3}{2}$$

אי שלילי $\lambda_1^*=\frac{3}{4}$, $\lambda_2^*=-\frac{1}{2}+\frac{3}{4}=\frac{1}{4}$

סה"כ וקטור כופל לגרנז' $\begin{pmatrix}\frac{3}{4}\\ \frac{1}{4}\\ 0\\ 0\end{pmatrix}$

נראה שתנאי ההכרחי מסדר שני מתקיים:

הגדרה משפט: אם (x^*,λ^*) פתרון לוקלי ! ו-LICQ תקף, אז:

$$\forall w\in T(x^*,\lambda^*),\quad w^T L_{xx}(x^*,\lambda^*)w\geq 0$$

כאשר KKT מתקיים ב-(x^*,λ^*) והחרוט הקריטי מוגדר כ-

$$T(x^*,\lambda^*)=\left\{ w \;\middle|\; \begin{array}{ll} \nabla c_i(x^*)w=0 & i\in E \\ \nabla c_i(x^*)w=0 & i\in I\cap A(x^*)\wedge \lambda_i^*>0 \\ \nabla c_i(x^*)w\geq 0 & i\in I\cap A(x^*)\wedge \lambda_i^*=0 \end{array}\right\}$$

במקרה הזה $E=\emptyset$

נמצא את ה-w המקיימים

$$\begin{pmatrix}-1\\-1\end{pmatrix}\cdot w=0 \qquad \begin{pmatrix}-1\\1\end{pmatrix}w=0$$

$$\begin{pmatrix}-1\\-1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=0 \qquad \begin{pmatrix}-1\\1\end{pmatrix}\begin{pmatrix}x\\y\end{pmatrix}=0$$

$$-x-y=0 \qquad -x+y=0 \Rightarrow x=y$$

$$\Leftarrow$$

$$-x-x=0$$

$$-2x=0$$

$$x=0$$

$$y=0$$

בס"ד

3

כצ

$\begin{pmatrix}0\\0\end{pmatrix}^T L_{xx}(x^*,\lambda^*)\begin{pmatrix}0\\0\end{pmatrix}$ אכן גדול שווה 0

עבור תנאים מספיקים מסדר שני:

משפט: הרצאה

(x^*,λ^*) רק בה תנאי KKT מתקיים ובנוסף

$$z^T L_{xx}(x^*,\lambda^*) z > 0$$

כאשר z היא מטריצה שעמודותיה בסיס לגרעין של $G(x^*,\lambda^*)$

אז x^* הוא מינימום לוקאלי.

ב-w היחיד שלנו בתחום הקריטי הוא $\begin{pmatrix}0\\0\end{pmatrix}$

לכן מתקיים באופן ריק.

Scanned with CamScanner

אופטימיזציה - חלק ג'

4. פתרו את בע"ת התכנון הליניארי הבאה

(זוהי הבעיה הצמודה לזו שפתרנו בכיתה)

$$\min z = 14\lambda_1 + 28\lambda_2 + 30\lambda_3$$

s.t

$$2\lambda_1 + 4\lambda_2 + 2\lambda_3 \geq 1$$

$$\lambda_1 + 2\lambda_2 + 5\lambda_3 \geq 2$$

$$\lambda_1 + 3\lambda_2 + 5\lambda_3 \geq -1$$

$$\lambda_1, \lambda_2, \lambda_3 \geq 0$$

פתרון: את הבעיה הצמודה פתרנו בכיתה וקיבלנו $z=13$, $(5,4,0)$

נחזור לזה, ולפי משפט, הפתרון לבעיה הזאת ולבעיה המקורית הינו זהה.

קודם נהפוך את הבעיה לבעיה דואלית ע"י נעביר אותה ל slack form

נהפוך את ה min ל-Max ונכפול ב-1- את z

נבצע טרנספוס, נהפוך את הא"י שוויונים ונוסיף משתני slack.

נקבל:

$$\text{Max } z = x_1 + 2x_2 - x_3$$

s.t

$$2x_1 + x_2 + x_3 + S_1 = 14$$

$$4x_1 + 2x_2 + 3x_3 + S_2 = 28$$

$$2x_1 + 5x_2 + 5x_3 + S_3 = 30$$

$$x_1, x_2, x_3 \geq 0$$

$$S_1, S_2, S_3 \geq 0$$

(4)

נשתמש באלגוריתם simplex

	x_1	x_2	x_3	slack			
Z	-1	-2	1	0	0	0	0
	2	1	1	1	0	0	14
	4	2	3	0	1	0	28
	2	5	5	0	0	1	30

כעת נרצה להכניס את x_2 (כי מקדם z הגדול ביותר במינוס) עם שורה 3 (כי $\frac{30}{5} < \frac{14}{1}, \frac{28}{2}$)

ונהפוך אותו ל-1.

	x_1	x_2	x_3	s_1	s_2	s_3	
Z	-1	-2	1	0	0	0	0
	2	1	1	1	0	0	14
	4	2	3	0	1	0	28
	$\frac{2}{5}$	1	1	0	0	$\frac{1}{5}$	6

$\frac{1}{5}R_3$

Z	$-\frac{1}{5}$	0	3	0	0	$\frac{2}{5}$	12
	$1\frac{3}{5}$	0	0	1	0	$-\frac{1}{5}$	8
	$\frac{16}{5}$	0	1	0	1	$-\frac{2}{5}$	16
	$\frac{2}{5}$	1	1	0	0	$\frac{1}{5}$	6

$R_0 - 2R_3$
$R_1 - R_3$
$R_2 - 2R_3$

כעת נרצה להכניס את x_1 (כי $-\frac{1}{5}$ המקדם הגדול ביותר במינוס)

עם שורה 1 (כי $\frac{8}{1\frac{3}{5}} < \frac{16}{16/5}, \frac{6}{2/5}$)

Z	$-1/5$	0	3	0	0	$2/5$	12
	1	0	0	$5/8$	0	$-1/8$	5
	$\frac{16}{5}$	0	1	0	1	$-\frac{2}{5}$	16
	$\frac{2}{5}$	1	1	0	0	$1/5$	6

$\frac{1}{1\frac{3}{5}}R_1$

בס"ד

אופטימיזציה - חלק א'

4

	x_1	x_2	x_3	s_1	s_2	s_3	
Z	0	0	$3\frac{1}{5}$	$\frac{1}{8}$	0	$\frac{3}{8}$	13
	1	0	0	$\frac{5}{8}$	0	$-\frac{1}{8}$	5
	0	0	1	-2	1	0	0
	0	1	1	$-\frac{1}{4}$	0	$\frac{1}{4}$	4

$R_0 + \frac{1}{5}R_1$

$R_2 - \frac{16}{5}R_1$

$R_3 - \frac{2}{5}R_1$

כאן נגמר האלגוריתם (אין מקדמים שליליים בשורה Z)

ידוע שהמשתנים שנכנסו לבסיס הם X_1, X_2, S_2 (כי יצא אפס בשורה Z)

כמו שציינו בתחילת התרגיל יצא $Z=13$ בנק' $(5, 4, 0)$