מודלים לא ליניאריים בחקר ביצועים --- תרגיל בית 8

26 במאי 2024

הנחיות להגשה

- יש להגיש שני קבצים בלבד:
  - קובץ PDF עם תשובות לשאלות וצילום מסך של קוד ופלטים.
  - קובץ ZIP המכיל את כל קובצי הקוד אותם נדרשתם לכתוב.

שאלה 1
בכל אחד מהסעיפים הבאים הראו שהבעיה היא בעיית תכנות קמור, כיתבו קוד CVXPY הפותר אותה ורישמו את הפתרון וערך פונקציית המטרה המתקבלים.

א.

$$\begin{aligned}
\underset{\mathbf{x}\in\mathbb{R}^3}{\text{minimize}} \quad & x^2+2xy+2y^2+z^2+3x-4y \\
\text{subject to} \quad & \sqrt{2x^2+xy+4y^2+4}+\frac{(x-y+z+1)^2}{x+y} \le 6 \\
& x,y,z \ge 1
\end{aligned}$$

ב.

$$\begin{aligned}
\underset{\mathbf{x}\in\mathbb{R}^2}{\text{maximize}} \quad & \sqrt{2y+5}-x^2-2xy-y^2-2x-3y \\
\text{subject to} \quad & \frac{x^2}{x+y}+\left(\frac{x^2}{y}+1\right)^8 \le 100 \\
& x+y \ge 4 \\
& y \ge 1
\end{aligned}$$

שאלה 2
בשביל למצוא את המיקום המיטבי לפתיחת מחסן משלוחים באמצעות רחפנים, מחלקים את העיר לנקודות ביקוש. המיקום של נקודה $i=1,2,\ldots,m$ הוא $\mathbf{a}^i\in\mathbb{R}^2$ ומספר ההזמנות השבועי הממוצע בכל נקודה הוא $p^i\in\mathbb{R}$. תהי $g_i(\mathbf{x})=\alpha\left\|\mathbf{a}^i-\mathbf{x}\right\|$ פונקציה המראה את זמן הטיסה הממוצע בדקות של הרחפן מנקודה $\mathbf{x}$ ל-$\mathbf{a}^i$, כך ש-$\alpha>0$ קבוע כלשהו (אנו מתעלמים מזמן הטיסה בחזרה). תהי $\gamma>0$ עלות דקת טיסה. נניח כי כל רחפן יכול לשאת הזמנה אחת בלבד בכל רגע ונניח כי יש מספר בלתי מוגבל של רחפנים.

א. כיתבו בעיית תכנות קמור המוצאת המיקום המיטבי למיקום המחסן הממזערת את עלות הטיסות השבועיות.

ב. נניח כי אנו מעוניינים לפצות לקוחות על כל משלוח שמתעכב. נניח שעבור כל דקת איחור מעבר לזמן η_1, ניתנת הנחה של μ_1 שקלים לדקה. בנוסף, עבור כל דקת איחור מעבר לזמן $\eta_2 > \eta_1$ ניתנת הנחה של $\mu_2 > \mu_1$ שקלים. כיתבו בעיית תכנות קמור שממזערת את זמן הטיסה הכולל ואת עלויות הפיצויים.

ג. הריצו את הסקריפט Q2.ipynb ליצירת נתוני הבעיה. פיתרו את שני המודלים באמצעות CVXPY ושרטטו את נקודות הביקוש ואת שני מיקומי המחסן המתקבלים בגרף אחד. צרפו גם את שני הפתרונות המיטביים שמתקבלים ואת ערכי פונקציית המטרה בנקודות המיטביות.

שאלה 3

נתונות נקודות $\{(\mathbf{x}_i, y_i)\}_{i=1}^m$ כך ש־$\mathbf{x}_i \in \mathbb{R}^n$ ו־$y_i \in \{1, -1\}$. נרצה למצוא משטח ריבועי מפריד מהצורה $\mathbf{x}^T\mathbf{P}\mathbf{x} + \mathbf{q}^T\mathbf{x} + r = 0$ כך שאם $y_i = 1$ אז $\mathbf{x}_i^T\mathbf{P}\mathbf{x}_i + \mathbf{q}^T\mathbf{x}_i + r \geq 0$ ואם $y_i = -1$ אז $\mathbf{x}_i^T\mathbf{P}\mathbf{x}_i + \mathbf{q}^T\mathbf{x}_i + r \leq 0$, עבור $\mathbf{P} \in \mathbb{R}^{n\times n}$ סימטרית, $\mathbf{q} \in \mathbb{R}^n$ ו־$r \in \mathbb{R}$. נסחו בעיית תכנות קמור שמוצאת את המקדמים $\mathbf{P}, \mathbf{q}, r$ כך שהנורמה הספקטרלית $\|\mathbf{P}\|_2$ היא מזערית. בתשובתכם, כיתבו במפורש את משתני ההחלטה והסבירו מדוע זו בעיית תכנות קמור.

שאלה 4

א. ממשו פונקציית python המופעלת באופן הבא:

```
x, fs = grad_proj(f, gf, proj, t, x0, eps)
```

כך שהיא מממשת את שיטת היטל הגרדיאנט עם גודל צעד קבוע עבור הבעיה

$$\begin{array}{ll} \min\limits_{\mathbf{x}\in\mathbb{R}^n} & f(\mathbf{x}) \\ \text{s.t.} & \mathbf{x} \in C \end{array}$$

כאשר C קבוצה קמורה ו־f גזירה ברציפות. הקלטים, בהתאמה הם פונקציית המטרה, הגרדיאנט שלה, פונקציית היטל על C (מקבלת וקטור $\mathbf{x}$ ומחזירה את $P_C(\mathbf{x})$), גודל צעד, נקודת התחלה ורמת דיוק. הפלטים הם פתרון הבעיה ומערך מגודל מספר האיטרציות עם ערכי פונקציית המטרה. אין צורך לבצע בדיקת תקינות קלט. השתמשו בתנאי העצירה $\|\mathbf{x}^{k+1} - \mathbf{x}^k\| \leq \varepsilon$.

ב. מיצאו את ההיטל האורתוגונלי של נקודה $\mathbf{p} \in \mathbb{R}^2$ על הישר $l \equiv \{\binom{x}{y} : x + y - 1 = 0\}$ ב־$\mathbb{R}^2$. אין צורך לספק הוכחה מתמטית.

הכוונה: מיצאו את משוואת הישר האנך ל־l עליו נמצאות שתי הנקודות $\mathbf{p}$ ו־$P_l(\mathbf{p})$.

ג. ממשו פונקציית python המופעלת באופן הבא:

```
proj = proj_section_b()
```

שמחזירה פונקציית היטל. הפונקציה המוחזרת proj מקבלת נקודה $\mathbf{p} \in \mathbb{R}^2$ ומחזירה את ההיטל שלה על הישר מהסעיף הקודם. כלומר, proj(p) היא $P_l(\mathbf{p})$. אין צורך לבצע בדיקת תקינות קלט.

ד. הסבירו מדוע ניתן להפעיל את שיטת היטל הגרדיאנט על הבעיה

$$\begin{array}{ll} \min\limits_{\mathbf{x}\in\mathbb{R}^2} & \|\mathbf{x}\|^2 \\ \text{s.t.} & \mathbf{x} \in l \end{array}$$

והסבירו מדוע מובטחת התכנסות השיטה ומדוע קיים לה פתרון מיטבי. מיצאו את הפתרון המיטבי בעזרת הפונקציות שכתבתם בסעיפים הקודמים, גודל צעד קבוע מיטבי $t^k \equiv \frac{1}{2}$, נקודת ההתחלה $(100,100)$ ו־$\varepsilon = 10^{-8}$. תוך כמה איטרציות התכנסה השיטה ולאיזו נקודה?

שאלה 5

בבעיית איכון מיקום של עצם מסוים (למשל, מקום נפילת רקטה על־סמך גורמי תצפית שונים) יש m מיקומים של חיישנים $\mathcal{A} \equiv \{\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_m\} \subset \mathbb{R}^n$ ומרחקים מוערכים $d_1, d_2, \ldots, d_m \geq 0$ בין החיישנים לבין העצם. כלומר, אם נסמן את מיקום העצם ב־$\mathbf{x} \in \mathbb{R}^n$, אז אנו יודעים כי $\|\mathbf{x} - \mathbf{a}_i\| \approx d_i$ לכל $1, 2, \ldots, m$. לכן, דרך להעריך את מיקום העצם $\mathbf{x} \in \mathbb{R}^n$ היא לפתור את בעיית האופטימיזציה הבאה:

$$(SL) \qquad \min_{\mathbf{x} \in \mathbb{R}^n} \sum_{i=1}^{m} \left(\|\mathbf{x} - \mathbf{a}_i\| - d_i\right)^2$$

בעייה זו היא לא גזירה ולכן לא ניתן להפעיל עליה שיטות מבוססות גרדיאנט. בנוסף, בעיה זו היא גם לא קמורה. בשאלה זו נמצא ניסוח גזיר ולא קמור שקול של הבעיה עליו נוכל להפעיל את שיטת היטל הגרדיאנט.

א. הראו שהבעיה

$$(SL2) \qquad \begin{aligned} \min_{\mathbf{x}, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m \in \mathbb{R}^n} & \left\{ f\left(\mathbf{x}, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m\right) \equiv \sum_{i=1}^{m} \left(\|\mathbf{x} - \mathbf{a}_i\|^2 - 2 d_i \mathbf{u}_i^T \left(\mathbf{x} - \mathbf{a}_i\right) + \|\mathbf{a}_i\|^2 \right) \right\} \\ \text{s.t. } & \|\mathbf{u}_i\| \leq 1, \ \forall i = 1, 2, \ldots, m \end{aligned}$$

שקולה לבעיה (SL) במובן ש־$\mathbf{x} \in \mathbb{R}^n$ הוא פתרון מיטבי של (SL) אם"ם קיימים $\mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m \in \mathbb{R}^n$ כך ש־$(\mathbf{x}, \mathbf{u}_1, \mathbf{u}_2, \ldots, \mathbf{u}_m)$ פתרון מיטבי של $(SL2)$.

הכוונה: תוכלו להשתמש באי־שוויון קושי־שוורץ.

ב. הראו שפונקציות המטרה בשתי הבעיות (SL) ו־$(SL2)$ לא קמורות.

הכוונה: מספיק להראות זאת עבור המקרה שבו $n = m = 1$.

ג. נגדיר את המטריצות

$$\mathbf{D} = \begin{bmatrix} d_1 \mathbf{I}_n & d_2 \mathbf{I}_n & \cdots & d_m \mathbf{I}_n \end{bmatrix} \in \mathbb{R}^{n \times nm}, \qquad \mathbf{P} = \begin{pmatrix} m\mathbf{I}_n & -\mathbf{D} \\ -\mathbf{D}^T & \mathbf{0}_{nm \times nm} \end{pmatrix} \in \mathbb{R}^{(n+nm)\times(n+nm)}.$$

הראו שקבוע ליפשיץ של הגרדיאנט של f הוא $L_{\nabla f} = 2\|\mathbf{P}\|$.

ד. הסתמכו על העובדה שמכפלה קרטזית משמרת קמירות. הסבירו מדוע ניתן להפעיל את שיטת היטל הגרדיאנט על בעיית $(SL2)$, מדוע מובטחת התכנסות השיטה ומדוע קיים פתרון מיטבי. האם ניתן לומר שהפתרון הוא אכן נקודת מינימום גלובלי? (כלומר, האם הפלט של שיטת היטל הגרדיאנט אכן פותר את המודל). הסבירו.