פרויקט - אופטימיזציה קמורה (67731)

אדם שטרסנר, ת"ז: 208837260, תואר ראשון

1 הגדרת הבעיה

הצגת הבעיה:

$$\min_{K\succeq 0:K\ is\ k-band} Tr(SK) - log(|K|)$$

עבור $S \in S^n_+$.
נתבונן בכל המטריצות $K \succ 0$ (זו בפרט מטריצה PSD לכן מטריצה כזו היא חוקית).
אם כך, נוכל להשתמש בפירוק צ'ולסקי עבור המטריצה K: קיימת מטריצה $L \in M_{n\times n}$ כך ש: $K = LL^T$ ו־L מקיימת את התנאים הבאים:

- L משולשית תחתונה
- הפיכה
- ערכי האלכסון חיוביים
- L היא $k-band$ (לא קשור לפירוק צ'ולסקי אלא בגלל האילוץ ש־K היא $k-band$).

טענה: אם L היא $k-band$, אזי K היא $k-band$.
הוכחה:
נניח כי L היא $k-band$, אזי עבור $i > j$ כך ש־$i - j > k$ מתקיים:

$$\begin{aligned}\left(LL^T\right)_{ij} &= \sum_m L_{im}\cdot\left(L_{mj}\right)^T\\ &= \sum_m L_{im}\cdot L_{jm}\end{aligned}$$

נחלק למקרים:
(i) עבור $1 \le m \le j$, נקבל כי $m + k \le j + k < i$, כלומר $i - m > k$ ולכן $L_{im} = 0$.
(ii) עבור $m > j$, הנחנו כי L היא משולשית תחתונה ולכן יתקיים $L_{jm} = 0$.
לכן לכל המקרים נקבל כי $\left(LL^T\right)_{ij} = (K)_{ij} = 0$, כלומר K היא $k-band$, כנדרש.

נפתח:

$$\begin{aligned}Tr(SK) - log(|K|) &= Tr(SLL^T) - log(|LL^T|)\\ &= Tr(L^TSL) - 2log(|L|)\end{aligned}$$

לכן הבעיה המקורית שקולה לבעיה:

$$\begin{aligned}&\min_{L\in\mathbb{R}^{n\times n}} Tr(L^TSL)-2log(|L|)\\ &s.t \qquad L_{ij}=0\ \forall(i,j)\in E\\ &\qquad\quad L_{ii}>0\end{aligned}$$

כאשר נגדיר: $G=\langle V,E\rangle$, עם $V=[n^2]$ ו־ $E=\{(i,j):|i-j|>k\ \vee\ j>i\}$.
האילוץ $(i,j)\in E$ הוא בדיוק האילוץ של L להיות מטריצה משולשית תחתונה ו־$k-band$.
אזי נסמן פשוט ב־$\mathcal{L}$ את קבוצת המטריצות שמקיימות את אילוצי הבעיה, והבעיה כעת שקולה לבעיה הבאה:

$$\min_{L\in\mathcal{L}} Tr(L^TSL)-2log(|L|)$$

נפתח עוד ונקבל:

$$\begin{aligned}Tr(L^TSL)-2log(|L|)&=\sum_{i=1}^n l_i^TSl_i-2\sum_{i=1}^n log([l_i]_i)\\ &=2\sum_{i=1}^n\left(\frac{1}{2}l_i^TSl_i-log([l_i]_i)\right)\end{aligned}$$

כאשר l_i היא העמודה ה־i של המטריצה L, ו־$[l_i]_i$ היא הקואורדינטה ה־i בוקטור l_i, כלומר $[l_i]_i=L_{ii}$.
סה"כ קיבלנו את הבעיה:

$$\min_{L\in\mathcal{L}} 2\sum_{i=1}^n\left(\frac{1}{2}l_i^TSl_i-log([l_i]_i)\right)$$

נבצע מינימיזציה לפי l_i, תחת האילוצים:

$$[l_i]_i>0\ ;\ \forall 1\leq j<i:[l_i]_j=0\ \forall i\ ;\ \forall r\leq n\ s.t\ |i-r|>k:[l_i]_r=0$$

ונחלק את הבעיה ל־n תת־בעיות שתלויות בכל עמודה של L ונקבל את הבעיה:

$$\min_{l_i\in\mathbb{R}^n} l_i^TSl_i-2log([l_i]_i)$$

לכל $i \in [n]$.
הגרדיאנט לכל $i \in [n]$:

$$\nabla f(l_i) = 2Sl_i - \frac{2}{[l_i]_i}e_i = 2\left(Sl_i - \frac{1}{[l_i]_i}e_i\right)$$

ההסיאן לכל $i \in [n]$:

$$\nabla^2 f(l_i) = 2\left(S + \frac{1}{[l_i]_i^2}E_{ii}\right)$$

כאשר E_{ij} היא מטריצת שבקואורדינטה ה־i, j יש 1 ובשאר אפסים.
אכן ניתן לראות כי הפונקציה קמורה ־ ההסיאן חיובי מאחר ו־$S \succeq 0$, ומאחר ו־$\frac{1}{[l_i]_i^2} > 0$ נקבל כי $S + \frac{1}{[l_i]_i^2}E_{ii} \succeq 0$ לכל $i \in [n]$.

2 האלגוריתם

- האלגוריתם בו בחרתי להשתמש הוא $Newton - Raphson$.
- על מנת לשפר את הביצועים של האלגוריתם, הגדרתי וקטור l'_i, מטריצה S_i ו־E'_{ii} בעל מימדים קטנים יותר, שמתעלם מהאפסים ב־l_i. ניתן לעשות זאת מאחר ואני יודע בדיוק אילו קואורדינטות מתאפסות ב־l_i. נעשה זאת כך:
עבור $i \leq n-k-1$ נגדיר $l'_i \in \mathbb{R}^{k+1}$ כך ש: $[l'_i]_j = [l_i]_{j+i-1}$ לכל $j \in [k+1]$.
אחרת, אם $i > n-k-1$, נגדיר את l'_i להיות ב־$\mathbb{R}^{k+1-r}$ לכל $0 \leq r \leq k$ כך ש: $[l'_i]_j = [l_i]_{j+i-1}$ לכל $j \in [k+1-r]$.
לכל i, כך שמימד הוקטור l'_i הוא p, נגדיר את $S_i \in \mathbb{S}^p_+$ כך שלכל i נתקדם בכיוון האלכסון של S וכך תהיה מוגדרת S_i. נעשה את אותו הדבר עבור E_{ii}.
בכך, חסכתי כפל מיותר של וקטור מרופד באפסים עם מטריצות.
בסוף, נרפד ל־L את הוקטורים שקיבלנו באמצעות האלגוריתם, תוך הוספת אפסים במקומות הנכונים, ונחזיר את LL^T.