אנליזה נומרית 1

© ארזים

23 באפריל 2017

1 פירוק LU

ראינו שהמצב בו $a_{k,k}^{(k)}$ קטן עשוי להיות בעייתי. ראינו דוגמאות, שבאחת מהן קיבלנו פתרון נכון ובשנייה לא. הזכרנו שתי גישות לפתרון הבעיה, שכעת נפרט אותן.

1. שיטת Full Pivoting - בשלב k מוצאים $i_0, m_0 \geq k$ עבורם

$$\left|a_{i_0,m_0}^{(k)}\right| = \max_{i,m\geq k}\left|a_{i,m}^{(k)}\right|$$

מחליפים את השורה k עם השורה i_0 ואת העמודה k עם העמודה m_0, וממשיכים. העלות היא מציאת מקסימום של $(n-k)^2$ מספרים, ובסך הכל על פני כל האיטרציות, העלות היא $O\left(n^3\right)$, שזה כמו כל האלגוריתם - יקר.

2. שיטת Partial Pivoting - מחליפים רק שורות. בשלב k מוצאים i_0 עבורו

$$\left|a_{i_0,k}^{(k)}\right| = \max_{i\geq k}\left|a_{i,k}^{(k)}\right|$$

ואז מחליפים את השורה k עם השורה i_0. העלות היא מציאת מקסימום של $n-k$ מספרים בכל איטרציה, ובסך הכל $O\left(n^2\right)$.

ננסה את השיטה השנייה על הדוגמה הבעייתית:

$$\begin{aligned}10^{-6}x_1 + x_2 &= 1\\ x_1 + x_2 &= 2\end{aligned}$$

נחליף שורות:

$$\begin{aligned}x_1 + x_2 &= 2\\ 10^{-6}x_1 + x_2 &= 1\end{aligned}$$

כעת נקבל

$$\begin{aligned}x_1 + x_2 &= 2\\ x_2 &= 1\end{aligned}$$

ומכאן נקבל גם $x_1 = 1$.

הערה 1.1 כאשר פותרים מערכת משוואות עם Full Pivoting, החלפת עמודות $l \leftrightarrow k$ צריכה להיות מלווה בהחלפת משתנים $x_l \leftrightarrow x_k$.

הערה 1.2 אם משתמשים בשיטת Partial Pivoting אז המטריצות L,U מקיימות $LU = PA$, כאשר P היא מטריצת פרמוטציה המתארת את החלפת השורות.

מטריצת פרמוטציה היא מטריצה שמתקבלת מהפעלת החלפות שורה על מטריצת היחידה.

סיכום של האלגוריתם: רוצים לפתור את $Ax = b$. מפרקים $LU = PA$ ואז

$$
\begin{aligned}
Ax &= b \\
PAx &= Pb \\
LUx &= Pb
\end{aligned}
$$

וחישוב Pb לוקח n פעולות.

יש מקרים מיוחדים, בהם פירוק LU לוקח פחות מאשר $\frac{2}{3}n^3$:

1. מטריצה סימטרית מוגדרת חיובית. במקרה זה יש מטריצה משולשית תחתונה L המקיימת $A = LL^T$, וזה נקרא פירוק קולסקי. אפשר לבצע אותו מהר פי שתיים מאשר פירוק LU רגיל.
2. מטריצה תלת אלכסונית - יש ערכים שאינם 0 רק באלכסון הראשי ובשני האלכסונים שצמודים לו (מעל ומתחת). במקרה זה פירוק LU דורש $O(n)$ פעולות.

2 נורמות

הגדרה 2.1 נורמה על $\mathbb{R}^n$ היא פונקציה $||\cdot|| : \mathbb{R}^n \to [0,\infty)$ המקיימת:

1. לכל $x \in \mathbb{R}^n$ מתקיים $||x|| \geq 0$.
2. $||x|| = 0$ אם ורק אם $x = 0$.
3. $||\alpha x|| = |\alpha| \, ||x||$.
4. $||x+y|| \leq ||x|| + ||y||$.

נורמות שימושיות:

1. נורמת l_∞:

$$||x||_\infty = \max_{1\leq i\leq n} |x_i|$$

2. נורמת l_2:

$$||x||_2 = \left(\sum_{i=1}^{n} x_i^2\right)^{\frac{1}{2}} = \sqrt{x^T x}$$

3. נורמת l_1:

$$||x||_1 = \sum_{i=1}^{n} |x_i|$$

4. נורמת l_p:

$$||x||_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}}$$

2.1 שקילות נורמות

משפט 2.2 יהיו $||\cdot||$, $||\cdot||'$ שתי נורמות כלשהן על $\mathbb{R}^n$. אז קיימים שני קבועים m, M עבורם לכל $x \in \mathbb{R}^n$ מתקיים

$$m\,||x||' \leq ||x|| \leq M\,||x||'$$

הוכחה: מספיק להוכיח שכל נורמה $||\cdot||$ שקולה לנורמה $||\cdot||_2$. נכתוב את הווקטור x בתור

$$x = \sum_j x_j e_j$$

כאשר e_j הם ווקטורי הבסיס הסטנדרטי. אזי

$$||x|| = \left|\left|\sum_j x_j e_j\right|\right| \leq \sum_j |x_j|\,||e_j|| \leq ||x||_2 \underbrace{\sum_j ||e_j||}_{M} = M\,||x||_2$$

כדי להראות את החלק השני של השקילות, נראה תחילה כי $||\cdot||$ היא רציפה, כלומר אם

$$\lim_{n\to\infty} \left|\left|x^{(n)} - x\right|\right|_2 = 0$$

אזי גם

$$\lim_{n\to\infty} \left|\left|x^{(n)}\right|\right| = ||x||$$

הסיבה לזה היא:

$$\left|\left|\left|x^{(n)}\right|\right| - ||x||\right| \leq \left|\left|x^{(n)} - x\right|\right| \leq M \left|\left|x^{(n)} - x\right|\right|_2 \to 0$$

נתבונן בקבוצה $S = \{y \in \mathbb{R}^n \mid ||y||_2 = 1\}$. זו קבוצה סגורה וחסומה, ולכן ממשפט ויירשטראס, הפונקציה הרציפה $||\cdot||$ מקבלת עליה מינימום. כלומר, קיימת נקודה y_0 כך שלכל y עם $||y||_2 = 1$, מתקיים

$$||y|| \leq ||y_0||$$

נסמן $m = ||y_0|| > 0$, כי $y_0 \neq 0$. ניקח כעת $x \in \mathbb{R}^n$ ואז

$$||x|| = ||x||_2 \cdot \left|\left|\frac{x}{||x||_2}\right|\right| \geq m\, ||x||_2$$

וסיימנו. ■

נעיר שלצרכים פרקטיים יש הבדל בין הנורמות, למשל אם רוצים למצוא נקודה שממזערת מרחק מאוסף נקודות נתון - בנורמת 2 נקבל ממוצע, ובנורמת 1 חציון.

2.2 נורמות של מטריצות

אלה הן נורמות שמוגדרות על מטריצות בגודל $n \times n$.

משפט 2.3 אם $||\cdot||$ היא נורמה כלשהי על $\mathbb{R}^n$, אזי

$$f(A) = \max_{||x||=1} ||Ax||$$

היא נורמה על מטריצות. נורמה כזו נקראת נורמה מושרית.

מתקיים

$$||A|| := f(A) = \max_{||x||=1} ||Ax|| = \max_{x \neq 0} \frac{||Ax||}{||x||} = \max_{x \neq 0} \left|\left| A \frac{x}{||x||} \right|\right|$$

למה 2.4 לכל נורמה מושרית מתקיים

$$||Ax|| \leq ||A||\, ||x||$$

הוכחה:

$$\frac{||Ax||}{||x||} \leq \max_{x \neq 0} \frac{||Ax||}{||x||} = ||A||$$

■

משפט 2.5 תהי A מטריצה בגודל $n \times n$. אזי

$$||A||_\infty = \max_{1 \le i \le n} \sum_{j=1}^{n} |a_{i,j}|$$

הוכחה:

$$||Ax||_\infty = \max_i |(Ax)_i| = \max_i \left| \sum_j a_{ij} x_j \right| \le \max_i \sum_j |x_j| \, |a_{ij}| \le$$
$$\le \max_j |x_j| \max_i \sum_j |a_{ij}| = ||x||_\infty \max_i \sum_j |a_{ij}|$$

לכן נקבל

$$||A||_\infty = \max_{x \ne 0} \frac{||Ax||_\infty}{||x||_\infty} \le \max_i \sum_j |a_{ij}|$$

בכיוון השני, תהי i_0 השורה שעליה המקסימום מתקבל. נגדיר

$$\tilde{x}_j = \begin{cases} 1 & a_{i_0,j} > 0 \\ -1 & a_{i_0,j} < 0 \end{cases}$$

בבירור, $||\tilde{x}||_\infty = 1$. כמו כן,

$$||A\tilde{x}||_\infty = \max_j |(A\tilde{x})_i| \ge \left|(A\tilde{x})_{i_0}\right| = \left| \sum_{j=1}^{n} a_{i_0,j} \tilde{x}_j \right| = \sum_{j=1}^{n} |a_{i_0,j}|$$

אזי

$$||A||_\infty = \max_{||x||=1} ||Ax|| \ge ||A\tilde{x}|| \ge \max_i \sum_j |a_{ij}|$$

■

הגדרה 2.6 נורמת פרובניוס של מטריצה A מוגדרת על ידי

$$||A||_F = \sqrt{\sum_{i,j=1}^{\infty} |a_{i,j}|^2}$$

הערה 2.7 נורמת פרובניוס אינה מושרית.

הוכחה: לכל נורמה מושרית מתקיים

$$||I|| = \max_{x\neq 0} \frac{||Ix||}{||x||} = 1$$

אבל

$$||I||_F = \sqrt{n}$$

■

הגדרה 2.8 הרדיוס הספקטרלי של מטריצה מוגדר על ידי $\rho\left(A\right) = \max_i \left|\lambda_i\right|$, כלומר הערך המוחלט המקסימי של ערך עצמי של A (λ_i הם הערכים העצמיים של A).

משפט 2.9

$$||A||_2 = \sqrt{\rho\left(A^T A\right)}$$

הוכחה: בשיעור הבא. ■

מסקנה 2.10 אם $A = A^T$ אזי

$$||A||_2 = \rho\left(A\right)$$

הוכחה:

$$||A||_2^2 = \rho\left(A^T A\right) = \rho\left(A^2\right) = \rho\left(A\right)^2$$

■

למה 2.11 לכל נורמה מושרית $||A||$ מתקיים

$$\rho\left(A\right) \leq ||A||$$

הוכחה: נניח כי $\rho\left(A\right) = \left|\lambda_n\right|$, כאשר $Av_n = \lambda_n v_n$. אזי

$$||A|| = \max_{x\neq 0} \frac{||Ax||}{||x||} \geq \frac{||Av_n||}{||v_n||} = \frac{\left|\lambda_n\right| ||v_n||}{||v_n||} = \rho\left(A\right)$$

■