

אנליזה נומרית 1

© ארזים

7 במאי 2017

1 אלגברה לינארית

1.1 ניתוח שגיאות

אנחנו רוצים לפתור את המערכת $Ax = b$, אבל במקום זה, בגלל שעובדים עם מחשב, פותרים את $A\tilde{x} = \tilde{b}$. נסמן

$$r = \|\tilde{b} - b\|$$

נשים לב שמתקיים

$$\tilde{x} - x = A^{-1}\tilde{b} - A^{-1}b = A^{-1}r$$

ועל כן

$$\|\tilde{x} - x\| \leq \|A^{-1}\| \|r\|$$

ניזכר בהגדרה של Condition number של מטריצה:

$$\text{Cond}(A) = \|A\| \|A^{-1}\|$$

ואז נקבל

$$\frac{\|\tilde{x} - x\|}{\|x\|} \leq \text{Cond}(A) \frac{\|r\|}{\|b\|}$$

את כל זה ראינו בשיעור שעבר.

1.1.1 הפרעות במטריצה A

רוצים לפתור את $Ax = b$, אבל פותרים את

$$(A + E)\tilde{x} = b = Ax$$

נפשט ונקבל

$$\begin{aligned} A(\tilde{x} - x) + E\tilde{x} &= 0 \\ \tilde{x} - x &= -A^{-1}E\tilde{x} \\ \|\tilde{x} - x\| &\leq \|A^{-1}\| \|E\| \|\tilde{x}\| \\ \frac{\|\tilde{x} - x\|}{\|\tilde{x}\|} &\leq \|A^{-1}\| \|E\| = \text{Cond}(A) \frac{\|E\|}{\|A\|} \end{aligned}$$

זו כמעט השגיאה היחסית (יש $\tilde{x}$ במכנה במקום x).

משפט 1.1 (שלא נוכיח) נניח כי $Ax = b$ וכן $(A + E)\tilde{x} = b + r$, כאשר $\det(A) \neq 0$ וכן

$$\|E\| \leq \frac{1}{\|A\|^{-1}}$$

אזי

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \frac{\text{Cond}(A)}{1 - \text{Cond}(A) \frac{\|E\|}{\|A\|}} \left(\frac{\|r\|}{\|b\|} + \frac{\|E\|}{\|A\|} \right)$$

1.2 פירוק LU

פותרם מערכת משוואות על ידי $LUx = b$. החסם שראינו הוא

$$\frac{\|x - \tilde{x}\|}{\|x\|} \leq \text{Cond}(A) \frac{\|r\|}{\|b\|}$$

נמצא חסם:

$$\begin{aligned} \text{Cond}(A) &= \|A\| \|A^{-1}\| = \|LU\| \|(LU)^{-1}\| \leq \\ &\leq \|L\| \|U\| \|U^{-1}\| \|L^{-1}\| = \text{Cond}(U) \text{Cond}(L) \end{aligned}$$

לכאורה החסם לא ידידותי לנו - נראה כי שימוש בפירוק LU יכול להגדיל את רגישות המערכת.

"הדוגמה הבאה היא של פון ניומן הגדול. אני מניח שכולכם נתקלתם בו בהקשר
זו או אחר" - המרצה
"לא" - הסטודנטים

דוגמא שאפשר להכליל לכל מימד של מטריצה. ניקח

$$A = \begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ -1 & 1 & 0 & 0 & 1 \\ -1 & -1 & 1 & 0 & 1 \\ -1 & -1 & -1 & 1 & 1 \\ -1 & -1 & -1 & -1 & 1 \end{pmatrix}, A^{-1} = \begin{pmatrix} \frac{1}{2} & -\frac{1}{4} & -\frac{1}{8} & -\frac{1}{16} & -\frac{1}{16} \\ 0 & \frac{1}{2} & -\frac{1}{4} & -\frac{1}{8} & -\frac{1}{8} \\ 0 & 0 & \frac{1}{2} & -\frac{1}{4} & -\frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{2} & -\frac{1}{2} \\ \frac{1}{2} & \frac{1}{4} & \frac{1}{8} & \frac{1}{16} & \frac{1}{16} \end{pmatrix}$$

שמוגדרת על ידי

$$A = \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 \\ -1 & -1 & -1 & 1 & 0 \\ -1 & -1 & -1 & -1 & 1 \end{pmatrix}}_L \underbrace{\begin{pmatrix} 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 0 & 2 \\ 0 & 0 & 1 & 0 & 4 \\ 0 & 0 & 0 & 1 & 8 \\ 0 & 0 & 0 & 0 & 16 \end{pmatrix}}_U$$

ואז

$$L^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 2 & 1 & 1 & 0 & 0 \\ 4 & 2 & 1 & 1 & 0 \\ 8 & 4 & 2 & 1 & 1 \end{pmatrix}, U^{-1} = \begin{pmatrix} 1 & 0 & 0 & 0 & -\frac{1}{16} \\ 0 & 1 & 0 & 0 & -\frac{1}{8} \\ 0 & 0 & 1 & 0 & -\frac{1}{4} \\ 0 & 0 & 0 & 1 & -\frac{1}{2} \\ 0 & 0 & 0 & 0 & \frac{1}{16} \end{pmatrix}$$

כעת, מתקיים

$$\|A\|_{\infty} = n = 5, \|A^{-1}\|_{\infty} = 1$$

ולכן

$$\text{Cond}(A, \infty) = n = 5$$

לעומת זאת

$$\|L\|_{\infty} = n = 5, \|L^{-1}\|_{\infty} = 2^{n-1} = 16$$

$$\|U\|_{\infty} = 2^{n-1} = 16, \|U^{-1}\|_{\infty} = \frac{3}{2}$$

כלומר

$$\text{Cond}(L, \infty) = n2^{n-1}, \text{Cond}(U, \infty) = \frac{3}{2}2^{n-1}$$

1.2.1 שגיאות עיגול

בכל איטרציה של אלגוריתם מבצעים שגיאות עיגול בעצם מתקיים $LU = A + E$ והשאלה היא כמה קטן $\|E\|$?

משפט 1.2 (ללא הוכחה) תהי A מטריצה בגודל $n \times n$ עם $\det(A) \neq 0$ ונניח שמחשבים LU על ידי אלמיניציית גאוס (עם pivoting) אז

$$\|E\|_{\infty} \leq n^2 \max_{1 \leq i, j, k \leq n} |a_{i,j}^{(k)}| \varepsilon_M$$

נסמן

$$p \cdot \|A\|^\infty = \max_{1 \leq i,j,k \leq n} |a_{i,j}^{(k)}|$$

ואז מתקיים

$$\frac{\|E\|_\infty}{\|A\|_\infty} \leq n^2 p \varepsilon_M$$

ואז

$$\frac{\|x - \tilde{x}\|_\infty}{\|\tilde{x}\|_\infty} \leq \text{Cond}(A) \frac{\|E\|_\infty}{\|A\|_\infty} \leq \text{Cond}(A) n^2 p \varepsilon_M$$

מסקנה 1.3 אם $\text{Cond}(A)$ "קטן" אז פירוק LU הוא תהליך יציב (קרוב לאמת).

2 פתרון מערכת לינאריות

נתבונן בפונקציה

$$f(x) = \frac{1}{2} x^T A x - b^T x$$

נניח כי A סימטרית ומוגדרת חיובית. נמצא מינימום של הפונקציה. נחשב ∇_f :

$$\begin{aligned} f(x) &= \frac{1}{2} \sum_{i,j} a_{i,j} x_i x_j - \sum_i b_i x_i \\ \frac{\partial f}{\partial x_k} &= \frac{1}{2} \left(\sum_{j \neq k} a_{k,j} x_j + \sum_{i \neq k} a_{i,k} x_i + 2a_{k,k} x_k \right) - b_k = \\ &= \frac{1}{2} \left(\sum_j a_{k,j} x_j + \sum_i a_{i,k} x_i \right) - b_k = \\ &= \frac{1}{2} \left(2 \sum_j a_{k,j} x_j \right) - b_k = \sum_j a_{k,j} x_j - b_k = (Ax)_k - b_k = (Ax - b)_k \end{aligned}$$

לכן

$$\nabla_f(x) = (Ax - b)$$

יש נקודת קיצון של f בנקודה $Ax = b$, והקיצון הזה הוא מינימום (תרגיל).

החלפנו בעיה של פתרון מערכת משוואות בבעיה של מציאת מינימום. אז איך נמצא מינימום?

נתחיל מנקודה כלשהי x_0 , ונחליק בכיוון בו f קטנה הכי מהר (שזה $-\nabla_f(x_0)$ $(b - Ax_0)$). נגיע לנקודה x_1 (מיד נחשב את גודל הצעד), ונחזור על התהליך. נסמן $r_i = b - Ax_i$, כאשר r_i מודד כמה אנחנו רחוקים מאגף ימין הנכון. מתקיים

$$r_i = b - Ax_i = -\nabla_f(x_i)$$

r_i הוא הכיוון בו f קטנה הכי מהר בנקודה x_i . x_1 היא מהצורה $x_1 = x_0 + \alpha r_0$. הפונקציה $x_1(\alpha) = x_0 + \alpha r_0$ מגדירה ישר. נמצא α שיביא את f למינימום לאורך הישר $x_1(\alpha)$. נדרוש

$$\nabla_f(x_1) r_0 = \nabla_f(x_1) \frac{\partial f}{\partial \alpha}(\alpha) = \frac{\partial (f \circ x_1)}{\partial \alpha}(\alpha) = 0$$

נבחר x_1 עבור $\nabla_f(x_1)$ ניצב לווקטור r_0 . ניזכרת כי $\nabla_f(x_1) = -r_1$ ולכן נדרוש

$$r_1^T r_0 = 0$$

כעת נפשט את התנאי:

$$r_1^T r_0 = 0$$

$$(b - Ax_1)^T r_0 = 0$$

$$(b - A(x_0 + \alpha r_0))^T r_0 = 0$$

$$(b - Ax_0)^T r_0 - \alpha (Ar_0)^T r_0 = 0$$

$$r_0^T r_0 - \alpha r_0^T Ar_0 = 0$$

$$\alpha = \frac{r_0^T r_0}{r_0^T Ar_0}$$

האלגוריתם (Steepest Descent) מתחילים מנקודה x_0 כלשהי (למשל 0). כעת חוזר עד (נחליט על תנאי התחלה בעתיד):

$$r_i = b - Ax_i$$

$$\alpha_i = \frac{r_i^T r_i}{r_i^T Ar_i}$$

$$x_{i+1} = x_i + \alpha_i r_i$$

באופן שקול אפשר לכתוב

$$\alpha_i = \frac{r_i^T r_i}{r_i^T Ar_i}$$

$$x_{i+1} = x_i + \alpha_i r_i$$

$$r_{i+1} = r_i - \alpha_i Ar_i$$

למה זה טוב יותר? רוב הסיבוכיות קשורה להכפלה של A בווקטור. כאן אנו צריכים לכפול את A רק בווקטור r_i . בשיטה הקודמת כפלנו גם Ax_i .
 בנינו סדרה של פתרונות מקורבים $x_1, x_2, \dots$ כך שמתקיים $x_n \rightarrow A^{-1}b$. אנחנו מתכנסים כיוון שבכל שלב אנו מתקדמים לכיוון שבו הפונקציה יותר נמוכה.
 כמה צעדים נדרשים כדי לקבל דיוק מסויים? היינו רוצים לחסום את $\|x_i - x\|$.

למה 2.1 תהי A מטריצה סימטרית מוגדרת חיובית. אזי $\|x\|_A = \sqrt{x^T A x}$ מגדירה נורמה.

■

הוכחה: נשארת כתרגיל.

נשתמש בזה בשיעור הבא כדי לחסום את מספר הצעדים.