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

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

נשים לב שמתקיים

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

ועל כן

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

ניזכר בהגדרה של Condition number של מטריצה:

$$\mathrm{Cond}\left(A\right) = \left|\left|A\right|\right| \left|\left|A^{-1}\right|\right|$$

ואז נקבל

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

את כל זה ראינו בשיעור שעבר.

1.1.1 הפרעות במטריצה A

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

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

נפשט ונקבל

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

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

משפט 1.1 (שלא נוכיח) נניח כי $Ax=b$ וכן $\left(A+E\right)\tilde{x}=b+r$, כאשר $\det\left(A\right)\neq0$ וכן

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

אזי

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

1.2 פירוק LU

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

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

נמצא חסם:

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

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

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

ולכן

$$\mathrm{Cond}\left(A,\infty\right)=n=5$$

לעומת זאת

$$\begin{aligned}||L||_{\infty}&=n=5,\left|\left|L^{-1}\right|\right|_{\infty}=2^{n-1}=16\\||U||_{\infty}&=2^{n-1}=16,\left|\left|u^{-1}\right|\right|_{\infty}=\frac{3}{2}\end{aligned}$$

כלומר

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

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 $\det\left(A\right)\neq0$ ונניח שמחשבים LU על ידיד אלימיניציית גאוס (עם pivoting) אז

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

נסמן

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

ואז מתקיים

$$\frac{||E||_{\infty}}{||A||_{\infty}}\le n^{2}p\varepsilon_{M}$$

ואז

$$\frac{||x-\tilde{x}||_{\infty}}{||\tilde{x}||_{\infty}}\le \mathrm{Cond}\left(A\right)\frac{||E||_{\infty}}{||A||_{\infty}}\le \mathrm{Cond}\left(A\right)n^{2}p\varepsilon_{M}$$

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

2 פתרון מערכת לינאריות

נתבונן בפונקציה

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

נניח כי A סימטרית ומוגדרת חיובית. נמצא מינימום של הפונקציה. נחשב ∇_{f}:

$$\begin{aligned}
f\left(x\right) &= \frac{1}{2}\sum_{i,j}a_{i,j}x_{i}x_{j}-\sum_{i}b_{i}x_{i}\\
\frac{\partial f}{\partial x_{k}} &= \frac{1}{2}\left(\sum_{j\neq k}a_{k,j}x_{j}+\sum_{i\neq k}a_{i,k}x_{i}+2a_{k,k}x_{k}\right)-b_{k}=\\
&= \frac{1}{2}\left(\sum_{j}a_{k,j}x_{j}+\sum_{i}a_{i,k}x_{i}\right)-b_{k}=\\
&= \frac{1}{2}\left(2\sum_{j}a_{k,j}x_{j}\right)-b_{k}=\sum_{j}a_{k,j}x_{j}-b_{k}=\left(Ax\right)_{k}-b_{k}=\left(Ax-b\right)_{k}
\end{aligned}$$

לכן

$$\nabla_{f}\left(x\right)=\left(Ax-b\right)$$

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

החלפנו בעיה של פתרון מערכת משוואות בבעייה 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:

$$\begin{aligned}
r_1^T r_0 &= 0 \\
(b - Ax_1)^T r_0 &= 0 \\
(b - A(x_0 + \alpha r_0))^T r_0 &= 0 \\
(b - Ax_0)^T r_0 - \alpha (Ar_0)^T r_0 &= 0 \\
r_0^T r_0 - \alpha r_0^T A r_0 &= A \\
\alpha &= \frac{r_0^T r_0}{r_0^T A r_0}
\end{aligned}$$

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

$$\begin{aligned}
r_i &= b - Ax_i \\
\alpha_i &= \frac{r_i^T r_i}{r_i^T A r_i} \\
x_{i+1} &= x_i + \alpha_i r_i
\end{aligned}$$

באופן שקול אפשר לכתוב

$$\begin{aligned}
\alpha_i &= \frac{r_i^T r_i}{r_i^T A r_i} \\
x_{i+1} &= x_i + \alpha_i r_i \\
r_{i+1} &= r_i - \alpha_i A r_i
\end{aligned}$$

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 $x_1, x_2, \ldots$ כך שמתקיים $x_n \to A^{-1}b$. 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.