3.5.09

אלגברה לינארית - שיעור 7

בשבועות האחרונים עסקנו (ונמשיך לעסוק עד הסוף) בדרכים שונות לפתור מערכת משוואות לינארית מהצורה $Ax=b$ בצורה יעילה. ובמיקוד כיוונו לשפר סדר אדל כמה רבה כל האלגוריתמים שלנו. כמובן שכל זה בלוותי רק כאשר מתמקדים על מטריצות דלולות מאוד. נתרכז את השיעור על קבלת מושג בסיסי לגבי מה חשוב לקבוע לפתרון מטריצות דלולות, ומה ניתן בכלל.

נגזרת שנייה - חד ממדי

נניח שיש לנו פונקציה u, המוגדרת בקטע $[0,1]$. נגדיר:

$$u(0)=g_L \quad , \quad u(1)=g_R$$

ונתבונן f המקיימת

$$f(x)=u_{xx}=\frac{d^2u}{dx^2}$$

נגדיר על הקטע $[0,1]$ רשת, כלומר נבחר $h=\frac{1}{N}$ (עבור N גדול) ונגדיר $x_j=j\cdot h$. אז אנחנו יודעים כי מתקיים:

$$u_{xx}(x_j)=\frac{1}{h^2}\left[u(x_{j-1})-2u(x_j)+u(x_{j+1})\right]+O(h^2)$$

לכן נוכל לחשב (בקירוב) נגזרת שנייה על נקודה ע"פ נוסחה זו.

נחבר מטריצה שבעזרתה נמצא את האופרטור של הנגזרת השנייה:

$$\frac{1}{h^2}\underbrace{\begin{pmatrix} -2 & 1 & 0 & 0 & 0 & 0 \\ 1 & -2 & 1 & 0 & 0 & 0 \\ 0 & 1 & -2 & 1 & 0 & 0 \\ 0 & 0 & \ddots & \ddots & \ddots & 0 \\ 0 & 0 & 0 & 1 & -2 & 1 \\ 0 & 0 & 0 & 0 & 1 & -2 \end{pmatrix}}_{(N-1)\times(N-1)}\begin{pmatrix} u_1 \\ \vdots \\ u_j \\ \vdots \\ u_{N-1} \end{pmatrix}+\frac{1}{h^2}\begin{pmatrix} g_L \\ 0 \\ \vdots \\ 0 \\ \vdots \\ g_R \end{pmatrix}^{*}=\begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ \vdots \\ f_{N-1} \end{pmatrix}$$

* המטרה של הוקטור האחרון הזה לתקן את השפה הקלות של המטריצה. מי שלא מבין למה זה הגיון יכול להתעלם ממנו לחלוטין שהוא מבין את המשוואה הפלאית הזו לא מעניין חשוב.

במקרה זה נראה להכפיל מטריצה בוקטור באופן מדויק. אם, מצד שני, נתונים לנו f ואנחנו רוצים למצוא את u (האינטגרל הפוך שלה) נצטרך לפתור מערכת משוואות לינארית.

נגזרת שנייה - דו מימדי

נגדיר הפעם:

$$\nabla^2 u(x,y) = u_{xx} + u_{yy} := f(x,y)$$

$$u_{jk} := u(x_j, y_k)$$

ומקבלים, בדומה למקרה החד מימדי:

$$\nabla^2 u_{jk} = \frac{1}{h^2}\left[u_{j+1,k} + u_{j-1,k} + u_{j,k+1} + u_{j,k-1} - 4u_{j,k}\right]$$

ובשביל את המטריצה נקבל מטריצה בגודל $N^2 \times N^2$ עם 5 אלכסונים בכל שורה, כלומר מטריצה דלילה.

כאן, אנו מגיעים לשאלה שאמרנו מהי המוטיבציה, ברור לכולנו שפירוק LU ופירוק QR זה לא מספיק, ואנחנו משתמקים לשמור דברים נוספת לפתור מערכת משוואות לינאריות. נעשה זאת.

פירוק Schur

בדומה לפירוקים הקודמים, נמצא דרך נוחה יותר לייצג את A.

בצורה זו:

$$A = Q \cdot T \cdot Q^*$$

אונטרית, משולש עליון, $= Q^{-1}$

נזכר כי $A^* = \overline{A^T}$, ומטריצה אונטרית היא מטריצה המקיימת $Q^* = Q^{-1}$.

נרצה להוכיח כי בשביל קיום פירוק כזה לכל A. נעשה זאת באינדוקציה על הגודל של A.

בסיס: אם A מגודל 1×1 אז

$$\underset{A}{(a)} = \underset{Q}{(1)}\,\underset{T}{(a)}\,\underset{Q^*}{(1)}$$

צעד האינדוקציה: נניח כי לכל מטריצה בגודל N-1×N-1 קיים פירוק Schur, כי ל-A בגודל N×N קיים פירוק כזה.

נמתר ע"ע של A נקרא לו λ, ונניח כי ה-ו"ע שלו (המנורמל) הוא v. נבנה מטריצה אונטרית U שהעמודה הראשונה שלה היא v (שאר העמודות נוכל לבנות ע"י גרם-שמידט).

אז מתקיים:

$$U^*AU = \begin{pmatrix} \lambda & B \\ 0 & \\ 0 & C \\ 0 & \end{pmatrix}$$

המטריצה המתאימה היא מטריצה עליונית, כאשר λ אלא יהיה, C מטריצה בגודל $N-1\times N-1$ ו- B הוא וקטור שורה בגודל $1\times N-1$.

מכיוון ש- C בגודל $N-1\times N-1$, מהנחת האינדוקציה יש ל- C פירוק Schur, כלומר קיימים V אונטרית ו- T משולשית עליונית כך ש- $C=VTV^*$.

נגדיר כאן: $Q = U\begin{pmatrix} 1 & 0 \\ 0 & V \end{pmatrix}$,

ואז, מהכפלת מטריצות קל לראות כי:

$$Q^*AQ = \begin{pmatrix} \lambda & VB \\ 0 & T \end{pmatrix}$$ (זו מטריצה עליונית)

לכן:

$$A = Q\begin{pmatrix} \lambda & VB \\ 0 & T \end{pmatrix}Q^*$$

זה פירוק Schur של A.

רדיוס ספקטרלי

הגדרה: $\rho(A)=\max_{i=1,\dots,N}|\lambda_i|$ כאשר λ_i הם הערכים העצמיים של A.

נשים לב כי $\rho(A)$ אינם נורמה. אחד התנאים להיות פונקציה נורמה הוא ש- $\|A\|=0 \iff A=0$ (במובן שבלי ימין הכוונה הוא לעמודות השמאל), אך נשים לב כי $\rho\begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}=0$.

מצד שני, כן קיים קשר כלשהו בין הרדיוס הספקטרלי לנורמות על A:

משפט: לכל נורמה $\|\cdot\|$ מתקיים: $\forall A,\ \rho(A)\le\|A\|$

הוכחה: נניח כי $\rho(A)=\lambda$ ו- v הוא הוקטור העצמי המנורמל של λ. אזי:

$$\rho(A)=|\lambda|=|\lambda|\|v\|=\|\lambda v\|=\|Av\|\le\max_{\|v\|=1}\|Av\|=\|A\|$$

משפט: לכל סדר מטריצה m $(A\in\mathbb{C}^{m\times m})$ ולכל $\varepsilon>0$ קיימת נורמה $\|\cdot\|$ כך ש- $\rho(A)\le\|A\|\le\rho(A)+\varepsilon$

הוכחה:

מפירוק Schur קיימת P אוניטרית ו-B משולשת עליונה כך ש:

$$PAP^{-1}=B$$

נגדיר את Λ המורה האלכסון של B, ואת U המורה השאר, כך:

$$\underbrace{\begin{pmatrix} b_{11} & b_{12} & \cdots & b_{1n} \\ 0 & b_{22} & b_{23} & \cdots b_{2n} \\ 0 & 0 & b_{33} & \cdots \\ 0 & 0 & 0 & \ddots \\ 0 & 0 & 0 & b_{nn} \end{pmatrix}}_{B} = \underbrace{\begin{pmatrix} b_{11} & 0 & 0 & 0 \\ 0 & b_{22} & 0 & 0 \\ 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & b_{nn} \end{pmatrix}}_{\Lambda} + \underbrace{\begin{pmatrix} 0 & b_{12} & b_{13} & \cdots b_{1n} \\ 0 & 0 & b_{23} & \cdots b_{2n} \\ 0 & 0 & 0 & \vdots \\ 0 & 0 & 0 & b_{n-1,n} \\ 0 & 0 & 0 & 0 \end{pmatrix}}_{U}$$

בנוסף, נגדיר את המטריצה D כך:

$$D=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & \frac{1}{\delta} & 0 & 0 & 0 \\ 0 & 0 & \frac{1}{\delta^2} & 0 & 0 \\ 0 & 0 & 0 & \ddots & 0 \\ 0 & 0 & 0 & 0 & \frac{1}{\delta^n} \end{pmatrix}$$

עבור δ קטן מספיק.

נגדיר את C להיות:

$$C:=DBD^{-1}=D\Lambda D^{-1}+\underline{DUD^{-1}}$$

נגדיר את E להיות הביטוי הבא:

נשים לב שאיברי E בעלי שווי:

$$E_{ij}=\begin{cases} 0 & i\le j \\ u_{ij}\delta^{j-1} & \text{אחרת} \end{cases}$$

ולכן E_{ij} קטן כרצוננו (בהתאם ל קטנות כרצוננו) לכל i,j.

עשינו נישום מאוד, נתאמץ מכל הה סדרות האקראיות האלה, ונתחיל לעשות סדר בדברים. אנחנו יודעים כי מתקיים:

$$A=P^{-1}BP$$

ומן:

$$B=D^{-1}CD$$

לכן:

$$A=P^{-1}D^{-1}CDP=(DP)^{-1}\cdot C\cdot DP$$

נזכר מה האלגברה הכבדים שלנו לעזרה, שנאמר ל הגדיר נורמה בעזרת מטריצה ומרבלה קנונית מי שלא נזכר, אבל רשם נקרא בוודאי תיכ)

מכיוון שמטריצה DP הפיכה, נוכל להגדיר את הנורמה הווקטורית הבאה:

$$\|x\|^2 = \langle x, P^*D^*DPx\rangle = \langle DPx, DPx\rangle$$

השוויון הזה מתקיים כי DP הוא מקיים לגמרי שזה לגדרה שלה.

כמו כן, נגדיר את הנורמה המושרית של מטריצות בצורה הרגילה:

$$\|A\| = \max_{\|v\|=1} \|Av\|$$

עכשיו סימנו עד ההגדרות נראה מה אנחנו יכולים להסיק:

$$\|Av\|^2 \underset{\text{ס"ב הגדרת הנורמה}}{=} \langle DPAv, DPAv\rangle \underset{\text{ס"ב המשוואה האחרונה כאשר הקודם}}{=} \langle CDPv, CDPv\rangle$$

נסמן (שזה מאוד שימושי להגדיר) $z = DPv$ ונקבל:

$$\|Av\|^2 = \langle Cz, Cz\rangle = \langle z, C^*Cz\rangle$$

נראה כי גם C^*C נקבל:

$$C^*C = (\Lambda^* + E^*)(\Lambda + E) = \Lambda^*\Lambda + M(\delta)$$

מטריצה שכל איבריה $\leq \delta$

לכן: $\langle Cz, Cz\rangle \leq \max_j(|\lambda_j|^2 z^*z) + \langle z, M(\delta)z\rangle \leq [\rho(A)^2 + o(\delta)]z^*z$

נשים לב שמ כי $1 = \|v\|^2 = \langle DPv, DPv\rangle = \langle z, z\rangle = z^*z = 1$

ולכן נ נסיק כי $\|A\| \leq [\rho^2(A) + o(\delta)]^{\frac{1}{2}} = \rho(A) + o(\delta)$.

בהתאמה δ קטן מספיק, $o(\delta) < \varepsilon$ וקיבלנו את הנורמה המבוקשת.

<u>הגדרה</u>: הסדרה $\{A^{(k)}\} \in \mathbb{C}^{n\times n}$ מתכנסת למטריצה A אם $\lim_{k\to\infty} \|A^{(k)} - A\| = 0$.

משפט:

I. $\lim_{k\to\infty} A^k = 0 \iff \rho(A) < 1$

II. הטור $\sum_{k=0}^{\infty} A^k$ מתכנס אמ"מ $\rho(A)<1$ ואז $\sum_{k=0}^{\infty} A^k = (I-A)^{-1}$

III. אם $\rho(A)<1$ אז $I-A$ הפיכה ומתקיים $\frac{1}{1+\|A\|} \le \|(I-A)^{-1}\| \le \frac{1}{1-\|A\|}$.

הוכחה:

I. $\rho(A)<1$ לכן קיימת נורמה $\|\cdot\|$:

$$\|A\| < \rho(A)-\varepsilon < 1$$

ואז מתקיים:

$$\|A^k\| = \|A\|^k \xrightarrow{k\to\infty} 0$$

ומכן נובע:

$$A^k \xrightarrow{k\to\infty} 0$$

מטריצת האפס כמובן

נוכיח את הכיוון השני, נניח כי $A^k\to 0$ ו λ ע"ע של A, x וקטור:

$$A^k x = \lambda^k x$$

בכיוון שבנתון ש- $A^k\to 0$ נובע ש:

$$\lambda^k \to 0$$
$$\Downarrow$$
$$\lambda < 1$$

לכן הוכחנו שכל הערכים העצמיים קטנים מ-1, לכן $\rho(A)<1$.

II. נניח כי $\rho(A)<1$, אז לפי הסעיף הראשון, $A^k\to 0$, ואז נסתכל על

$$(I-A)\underbrace{(I+A+A^2+\dots+A^k)}_{\sum_{j=0}^{k} A^j} = (I-A^{k+1})$$

ואז, לפי הנתון

$$\lim_{k\to\infty} (I-A)(I+A+A^2+\dots+A^k) = I$$

ולכן:

$$\sum_{j=0}^{\infty} A^k = (I-A)^{-1}$$

הכיוון השני קל. אם הטור מתכנס ברור כי $A^k\to 0$, ולכן לפי הסעיף הראשון $\rho(A)<1$.

III. נניח כי $\rho(A)<1$. אז:

$$1=\|I\|=\|(I-A)(I-A)^{-1}\|\le\|I-A\|\cdot\|(I-A)^{-1}\|\le(1+\|A\|)\|(I-A)^{-1}\|$$

$$\Rightarrow \frac{1}{1+\|A\|}\le\|(1-A)^{-1}\|$$

מכאן נראה את החסם מהכיוון השני:

$$I=I-A+A \qquad /\cdot(I-A)^{-1}$$

$$(I-A)^{-1}=I+A(I-A)^{-1} \qquad /\ \text{נורמה}$$

$$\|(I-A)^{-1}\|\le 1+\|A\|\|(I-A)^{-1}\| \qquad /-\|A\|\cdot\|(I-A)^{-1}\|$$

$$\|(I-A^{-1}\|(1-\|A\|)\le 1$$

$$\|(I-A^{-1})\|\le\frac{1}{1-\|A\|}$$

פתרון איטרטיבי למערכת משוואות לינאריות (LAS)

לאחר שהוכרנו לעומק הקשר של הרדיוס הספקטרלי, נחזור לאינטואיציה שלנו בלאחר כמה שיעורי דרכים לפתור את המשוואה $Ax=b$. והפעם, נתקדם מאמצעות הגישה האיטרטיבית.

הכוונה סדרת וקטורים המוגדרת ע"י $x^{(k+1)}=Bx^{(k)}+f$. B נקרא ה-Iteration Matrix של הסדרה.

<u>הגדרה</u>: בהתלך $x^{(k+1)}=Bx^{(k)}+f$ נקרא <u>קונסיסטנטי</u> עם $Ax=b$ אם $f=(I-B)\underbrace{A^{-1}b}_{x}$

נגדיר את השגיאה של הסדרה כמו $e^{(k)}=x^{(k)}-x$

<u>משפט</u>: יהי התלך $x^{(k+1)}=Bx^{(k)}+f$ קונסיסטנטי עם $Ax=b$, אז הסדרה $\{x^{(k)}\}$ מתכנס לפתרון של $Ax=b$ לכל $x^{(0)}$ בתנאי אמ"מ $\rho(B)<1$.

הוכחה:

אנחנו יודעים כי:

$$x^{(k+1)} = Bx^{(k)} + f$$

וגם:

$$x = Bx + f$$

לכן, תישאר מהטעות כי:

$$e^{(k+1)} = Be^{(k)}$$

לכן, כמו שהוכחנו בשיעור קודם,

$$\rho(B) < 1 \iff \forall e^{(0)}.\ \lim_{k\to\infty} B^k e^{(0)} = 0$$

(כי $B^k \to 0$).

במידה שאם $\rho(B) \geq 1$ אז קיים ערך עצמי $\lambda \geq 1$ ואז אם ניקח וקטור עצמי שלו, לעולם לא נקבל את $e^{(k)}$.

מעלה, אז הוכחנו שאם סדרה קונסטרוקטיבית עם משוואה $Ax=B$, ו-B^k מתכנסת לאפס (כך שקול ל- $\rho(B)<1$ לפי שלא צריך), אז הסדרה מתכנסת לפתרון. עכשיו נלמד כיצד ניצור ~~משוואה~~ מטריצה משוואת לינארית סדרה כזו שמתכנסת לנו לפתרון. נכיר את שיטת ~~יעקבי~~, יעקובי, על אל שיש סוף.

שיטת יעקובי:

בהינתן המשוואה $Ax=b$, נגדיר את N בתור האלכסון של A, ואת P בתור השאר. כלומר,

$$A = N + P$$

כאשר N אלכסונית (והאלכסון של P הוא אפס). כעת, נמלא אלינו אריתמטיקה פשוטה מספר מהחוקים אלגבריים:

$$Ax = b$$

$$(N+P)x = b$$

$$Nx = -Px + b$$

$$x = \underbrace{-(N^{-1})P}_{B}x + \underbrace{N^{-1}b}_{f}$$

נגדיר את התהליך האיטרטיבי שלנו בתור:

$$x^{(k+1)} = \underbrace{-N^{-1}P}_{B}x + \underbrace{N^{-1}b}_{f}$$

כך יכולנו איטרציה "בהשראת" המשוואה המקורית שלנו. ברור כי פתרון של המשוואה יהיה נקודת שבת של האיטרציה, כי משם נגיע $x^{(k+1)}=x^{(k)}$ ונקבל את המשוואה שממנה יכולנו את האיטרציה.

אם שיטת יעקובי נראית לכם יפה ואלגנטית, הנה הנוסחא שלמדנו בשיעור:

$$x_k^{(n+1)} = \frac{1}{a_{kk}}\left(b_k - \sum_{\substack{j=1 \\ j\neq k}}^{n} a_{kj}\, x_j^{(n)}\right)$$