Predavanja 4

Linearni sistemi

Četrti sklop izročkov

Fakulteta za računalništvo in informatiko
Univerza v Ljubljani

25. oktober 2021

LU razcep s kompletnim pivotiranjem

Pri kompletnem pivotiranju pred eliminacijo v j-tem stolpcu poiščemo element z največjo absolutno vrednostjo v podmatriki $A(j:n, j:n)$ in nato izvedemo ustrezni menjavi vrstic in stolpcev.

Dodatno delo pri LU razcepu s kompletnim pivotiranjem je $\mathcal{O}(n^3)$ primerjanj in menjav. Torej je skupna cena dražja od LU razcepa z delnim pivotiranjem in se v praksi redko uporablja.

Stabilnost LU razcepa matrike

Naj bo $A = \left[a_{ij}\right]_{i,j}$ $n \times n$ matrika. Z $|A|$ označimo matriko $\left[|a_{ij}|\right]_{i,j}$. Brez dokaza navedimo naslednjo trditev o stabilnosti LU razcepa.

Trditev

Naj bo A obrnljiva matrika, pri kateri se izvede LU razcep brez pivotiranja. Za izračunani matriki $\widehat{L}$, $\widehat{U}$ velja $A = \widehat{L}\widehat{U} + E$, kjer je

$$|E| \leqslant nu|\widehat{L}||\widehat{U}|.$$

Lahko se zgodi, da je $|\widehat{L}||\widehat{U}|$ zelo veliko v primerjavi z A, zato računanje LU razcepa v splošnem ni obratno stabilna metoda.

- Če želimo omejiti $|\widehat{L}|$, moramo uporabiti pivotiranje. Po konstrukciji bo veljalo $|\ell_{ij}| \leqslant 1$.
- Oceniti pa moramo še $|U|$. Vpeljemo pivotno rast

$$g := \frac{\max_{i,j} |u_{ij}|}{\max_{i,j} |a_{ij}|}.$$

Sledi

$$\max_{ij} |E_{ij}| \leqslant n^2 u g \max_{ij} |a_{ij}|.$$

Če je pivotna rast g majhna, lahko pričakujemo majhno napako.

Trditev

Pri delnem pivotiranju je pivotna rast omejena z 2^{n-1}.

Velja namreč $|\ell_{ij}| \leqslant 1$, a_{ij} pa na vsakem od največ $n-1$ korakov izračunamo kot $a_{ij} = a_{ij} - \ell_{ik} a_{kj}$. Torej se absolutna vrednost največjega elementa v matriki kvečjemu podvoji.

Žal pa za vsak n obstajajo matrike s pivotno rastjo 2^{n-1}, tako da LU razcep z delnim pivotiranjem teoretično ni obratno stabilen.

Primer

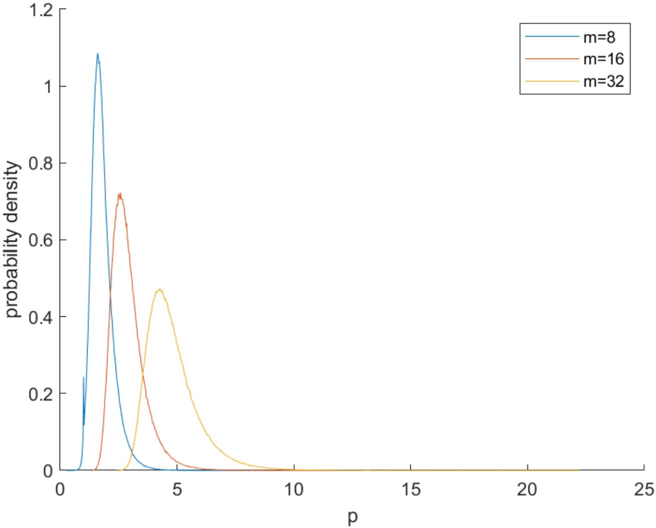
Matrika

$$A_n = \begin{pmatrix} 1 & 0 & \cdots & 0 & 1 \\ -1 & 1 & \ddots & \vdots & 1 \\ \vdots & \ddots & \ddots & 0 & \vdots \\ \vdots & & \ddots & \ddots & \vdots \\ -1 & \cdots & \cdots & -1 & 1 \end{pmatrix}$$

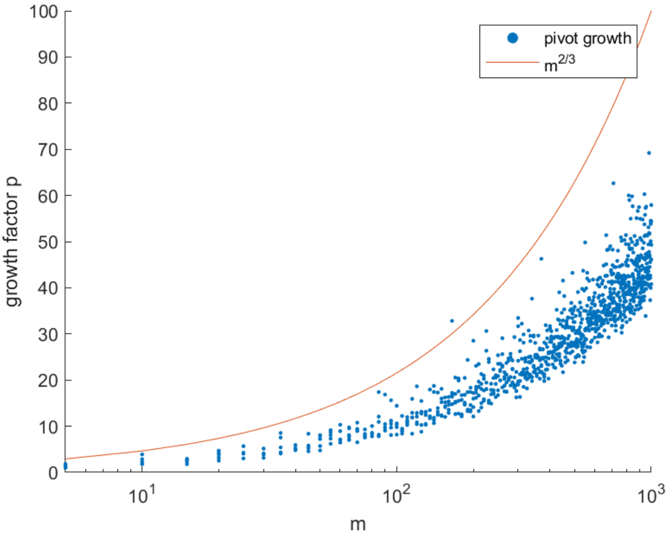
ima pivotno rast 2^{n-1}.

Statistično pa velja, da je pričakovana vrednost pivotne rasti $\mathcal{O}(n^{2/3})$, tako da LU razcep z delnim pivotiranjem v praksi je obratno stabilen.

Verjetnostne porazdelitve slučajne spremenljivke ρ, generirane z milijon naključnimi matrikami velikosti $m \times m$ (tj. vsak vhod naključen element iz enakomerne zvezne porazdelitve na intervalu $[-1, 1]$):



Pivotna rast 200 naključnih matrik velikosti $m \times m$ (tj. vsak vhod naključen element iz enakomerne zvezne porazdelitve na intervalu $[-1, 1]$):



Občutljivost sistema $Ax = b$

Vprašanje. Denimo, da smo sistem $Ax = b$ z neko metodo (npr. LU razcepom) rešili numerično in dobili približek $\hat{x}$. Kako dober je ta približek?

Pišimo $\hat{x} = x + \Delta x$. V resnici velja

$$A(x + \Delta x) = b + \Delta b.$$

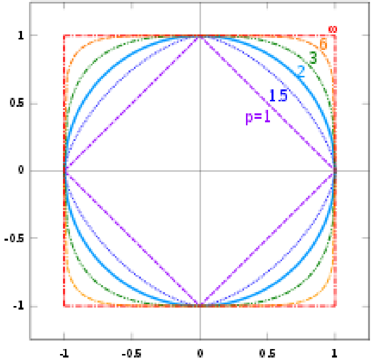
Če vključimo še numerično napako v A, velja

$$(A + \Delta A)(x + \Delta x) = b + \Delta b. \tag{1}$$

Radi bi ocenili velikostni razred Δx v primerjavi z x. V ta namen potrebujemo nekaj osnov vektorskih in matričnih norm.

Vektorska norma je preslikava $\|\cdot\| : \mathbb{C}^n \to \mathbb{R}$, ki zadošča:

1. Pozitivna definitnost: $\|x\| \geqslant 0$ za vsak $x \in \mathbb{C}^n$ in $\|x\| = 0 \Leftrightarrow x = 0$.
2. Homogenost: $\|\alpha x\| = |\alpha| \, \|x\|$ za vsaka $\alpha \in \mathbb{C}$ in $x \in \mathbb{C}^n$
3. Trikotniška neenakost: $\|x + y\| \leqslant \|x\| + \|y\|$ za vsaka $x, y \in \mathbb{C}^n$.



Matrična norma je preslikava $\|\cdot\| : \mathbb{C}^{n\times n} \to \mathbb{R}$, ki zadošča:

1. Pozitivna definitnost: $\|A\| \geqslant 0$ za vsak $A \in \mathbb{C}^{n\times n}$ in $\|A\| = 0 \Leftrightarrow A = 0$.
2. Homogenost: $\|\alpha A\| = |\alpha|\,\|A\|$ za vsaka $\alpha \in \mathbb{C}$ in $A \in \mathbb{C}^{n\times n}$.
3. Trikotniška neenakost: $\|A + B\| \leqslant \|A\| + \|B\|$ za vsaka $A, B \in \mathbb{C}^{n\times n}$.
4. Submultiplikativnost: $\|AB\| \leqslant \|A\|\,\|B\|$ za vsaka $A, B \in \mathbb{C}^{n\times n}$.

Trditev

Naj bo $\|\cdot\|_*$ vektorska norma na $\mathbb{C}^n$. Potem predpis

$$\|A\|_* := \max_{\|x\|=1} \|Ax\|_* \tag{2}$$

določa matrično normo na $\mathbb{C}^{n\times n}$. (2) pa lahko ekvivalentno definiramo tudi kot

$$\|A\|_* := \max_{\|x\|\neq 0} \frac{\|Ax\|}{\|x\|}.$$

Primer

Pišimo $x = (x_1, x_2, \dots, x_n) \in \mathbb{C}^n$ *in*

$$A = [a_{ij}]_{i,j=1,\dots,n} = \begin{bmatrix} \mathbf{a}_1 & \mathbf{a}_2 & \dots & \mathbf{a}_n \end{bmatrix} \in \mathbb{C}^{n\times n},$$

kjer so $\mathbf{a}_1, \dots, \mathbf{a}_n$ *stolpci matrike* A*. Nekaj vektorskih in matričnih norm:*

1. $\|x\|_1 = \sum_{i=1}^{n} |x_i|$ *(1–norma),* $\qquad \|A\|_1 = \max_{j=1,\dots,n} \sum_{i=1}^{n} |a_{ij}|.$

   Dokaz enakosti za $\|A\|_1$*.*

$$\begin{aligned} \|Ax\|_1 &= \|\mathbf{a}_1 x_1 + \dots + \mathbf{a}_n x_n\|_1 \leqslant \|\mathbf{a}_1\|_1 |x_1| + \dots + \|\mathbf{a}_n\|_1 |x_n| \\ &\leqslant \max_{j=1,\dots,n} \|\mathbf{a}_j\|_1 (|x_1| + \dots + |x_n|) = \max_{j=1,\dots,n} \|\mathbf{a}_j\|_1 \|x\|_1 \\ &= \max_{j=1,\dots,n} \|\mathbf{a}_j\|_1 . \end{aligned}$$

   Ker je $Ae_j = \mathbf{a}_j$ *(*e_j *je standardni vektor z 1 na j–tem mestu in 0 drugje), je enakost dosežena.*

2. $\|x\|_2 = \sqrt{\sum_{i=1}^{n} |x_i|^2} \quad$ (2−norma),

$$\|A\|_2 = \sqrt{\max_{j=1,\ldots,n} \lambda_j(A^T A)} \quad \text{(spektralna norma)}$$

Tu $\lambda_j(X)$ označuje j–to lastno vrednost matrike X.

3. $\|x\|_\infty = \max_{i=1,\ldots,n} |x_i| \quad$ (∞−norma), $\quad \|A\|_\infty = \max_{i=1,\ldots,n} \sum_{j=1}^{n} |a_{ij}|$.

4. $\|A\|_F = \sqrt{\sum_{i,j=1}^{n} |a_{ij}|^2} \quad$ (Frobeniusova matrična norma).

Pozor: $N(A) := \max_{i,j=1,\ldots,n} |a_{ij}|$ ni matrična norma. Protiprimer za submultiplikativnost $N(AB) \leqslant N(A)N(B)$ sta matriki $A = B = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$, saj je

$$N(AB) = N\left(\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix}\right) = 2 > N(A)N(B) = 1.$$

Zakaj imeti več matričnih norm?

Nekatere norme je bistveno zahtevneje izračunati od ostalih. Zahtevno je npr. določanje spektralne norme $\|\cdot\|_2$, saj je računanje lastnih vrednosti zahtevna naloga. Poceni pa je izračunati 1–normo, ∞–normo in F–normo. Iz različnih ocen, kot so

$$
\begin{aligned}
\frac{1}{\sqrt{n}}\|A\|_F \leqslant & \quad \|A\|_2 & \leqslant & \quad \|A\|_F, \\
\frac{1}{\sqrt{n}}\|A\|_1 \leqslant & \quad \|A\|_2 & \leqslant & \quad \sqrt{n}\|A\|_1, \\
\frac{1}{\sqrt{n}}\|A\|_\infty \leqslant & \quad \|A\|_2 & \leqslant & \quad \sqrt{n}\|A\|_\infty, \\
N(A) \leqslant & \quad \|A\|_2 & \leqslant & \quad n \cdot N(A),
\end{aligned}
$$

pa lahko dobro ocenimo $\|A\|_2$.