

1. Obratne napake in občutljivosti lastnih vrednosti

Bor Plestenjak

NLA

18. marec 2015

Nesimetrični problem lastnih vrednosti

Dana je $A \in \mathbb{R}^{n \times n}$. Če je $Ax = \lambda x$, kjer $x \in \mathbb{C}^n$, $x \neq 0$ in $\lambda \in \mathbb{C}$, potem je λ **lastna vrednost**, x pa **(desni) lastni vektor**.

Vektor $y \neq 0$, pri katerem je $y^H A = \lambda y^H$, je **levi lastni vektor**.

Če je $\det(S) \neq 0$, sta A in $S^{-1}AS$ **podobni matriki** in imata enake lastne vrednosti.

Izrek (Schurova forma)

Za vsako matriko A obstajata unitarna matrika Q in zgornja trikotna matrika T , da je $Q^H A Q = T$.

Izrek (realna Schurova forma)

Za vsako realno matriko A obstajata ortogonalna matrika Q in kvazi zgornja trikotna matrika T , da je $Q^T A Q = T$.

Matrika

$$A = \begin{bmatrix} -149 & -50 & -154 \\ 537 & 180 & 546 \\ -27 & -9 & -25 \end{bmatrix}$$

ima lastne vrednosti $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$. V Matlabu `l=eig(A)` vrne

$$\begin{aligned} \hat{\lambda}_1 &= 1.0000000000010722 \\ \hat{\lambda}_2 &= 1.999999999991790 \\ \hat{\lambda}_3 &= 2.999999999997399. \end{aligned}$$

Zgled 2

Matrika Godunova

$$A = \begin{bmatrix} 289 & 2064 & 336 & 128 & 80 & 32 & 16 \\ 1152 & 30 & 1312 & 512 & 288 & 128 & 32 \\ -29 & -2000 & 756 & 384 & 1008 & 224 & 48 \\ 512 & 128 & 640 & 0 & 640 & 512 & 128 \\ 1053 & 2256 & -504 & -384 & -756 & 800 & 208 \\ -287 & -16 & 1712 & -128 & 1968 & -30 & 2032 \\ -2176 & -287 & -1565 & -512 & -541 & -1152 & -289 \end{bmatrix}$$

ima eksaktne lastne vrednosti $-4, -2, -1, 0, 1, 2, 4$.

Matlab vrne $\text{eig}(A) =$

$$\begin{array}{lcl} 5.0968 & + & 1.8932i \\ 5.0968 & - & 1.8932i \\ 1.2157 & + & 4.3784i \\ 1.2157 & - & 4.3784i \\ -5.6738 & & \\ -3.4756 & + & 3.4463i \\ -3.4756 & - & 3.4463i \end{array}$$

Lastne vrednosti zmotene matrike

Za koliko se premaknejo lastne vrednosti, če matriko A malo zmotimo?

Izrek (Bauer-Fike)

Naj se da matriko A diagonalizirati kot $A = X\Lambda X^{-1}$, kjer je $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$. Potem vse lastne vrednosti matrike $A + \varepsilon E$ ležijo v uniji n krogov

$$B_i := \{z \in \mathbb{C} : |z - \lambda_i| \leq \varepsilon \kappa(X) \|E\|\}, \quad i = 1, \dots, n,$$

kjer za normo lahko vzamemo katerokoli izmed norm $\|\cdot\|_1$, $\|\cdot\|_2$ ali $\|\cdot\|_\infty$.

Opomba: Če unija krogov razpade na povezane komponente, potem vsaka komponenta vsebuje toliko lastnih vrednosti, kot je v njej krogov.

Posledica (Weylov izrek)

Če je $A = A^T$ in $E = E^T$, potem za vsako lastno vrednost $\hat{\lambda}$ matrike $A + \varepsilon E$ obstaja lastna vrednost λ matrike A , da je

$$|\hat{\lambda} - \lambda| \leq \varepsilon \|E\|.$$

Gerschgorinov izrek

Pri mnogo algoritmihi poskusimo matriko spraviti v diagonalno obliko. Končamo, ko so izvendiagonalni elementi dovolj majhni in diagonalne elemente vzamemo kot pribliške za lastne vrednosti. Kako dobri pribliški so?

Izrek (Gerschgorin)

Vse lastne vrednosti matrike A ležijo v uniji *Gerschgorinovih krogov*

$$G_i = \left\{ z \in \mathbb{C} : |z - a_{ii}| \leq \sum_{\substack{k=1 \\ k \neq i}}^n |a_{ik}| \right\}, \quad i = 1, \dots, n.$$

Opomba: Če unija krogov razpade na povezane komponente, potem vsaka komponenta vsebuje toliko lastnih vrednosti, kot je v njej krogov.

Izboljšava v primeru izoliranih krogov

$$A = \begin{bmatrix} 1 & 0.01 & 0.02 \\ 0.03 & 2 & 0.02 \\ -0.01 & -0.02 & 3 \end{bmatrix} \text{ ima radije Gerschgorinovih krogov } 0.03, 0.05, 0.03.$$

Še boljše ocene lahko dobimo iz DAD^{-1} , kjer je $D = \text{diag}(1, \dots, 1, \textcolor{red}{c}, 1, \dots, 1)$ s faktorjem c na i -tem mestu. Faktor $\textcolor{red}{c}$ izberemo minimalen možen, da je i -ti krog še ločen od ostalih, npr.

$$DAD^{-1} = \begin{bmatrix} 1 & 0.01\textcolor{red}{c} & 0.02\textcolor{red}{c} \\ 0.03/\textcolor{red}{c} & 2 & 0.02 \\ -0.01/\textcolor{red}{c} & -0.02 & 3 \end{bmatrix}.$$

Če vzamemo $D = \text{diag}(0.04, 1, 1)$, ima DAD^{-1} radije 0.0008, 0.75, 0.25, torej

$$|\lambda - 1| \leq 0.0008.$$

Točne lastne vrednosti A so:

$$\lambda_1 = 1 - 2.07 \cdot 10^{-4}, \quad \lambda_2 = 2 + 7.14 \cdot 10^{-4}, \quad \lambda_3 = 3 - 5.07 \cdot 10^{-4}.$$

Občutljivost posamezne lastne vrednosti

Izrek

Naj bo λ_i enostavna lastna vrednost matrike A , x_i ustrežni desni, y_i levi lastni vektor in $\|x_i\|_2 = \|y_i\|_2 = 1$. Če je $\lambda_i + \delta\lambda_i$ ustrežna lastna vrednost zmotene matrike $A + \delta A$, potem velja

$$\lambda_i + \delta\lambda_i = \lambda_i + \frac{y_i^H \delta A x_i}{y_i^H x_i} + \mathcal{O}(\|\delta A\|^2).$$

Z izrazom ustrežna lastna vrednost je mišljeno, da si izberemo tisto lastno vrednost zmotene matrike, za katero velja $\lim_{\|\delta A\| \rightarrow 0} (\delta\lambda_i) = 0$.

Definicija

Če definiramo

$$s_i := \frac{y_i^H x_i}{\|x_i\|_2 \|y_i\|_2},$$

je s_i^{-1} občutljivost enostavne lastne vrednosti λ_i . Če je λ_i večkratna lastna vrednost, je njena občutljivost enaka ∞ .

Predpostavimo, da se da matriko A diagonalizirati.

Izrek

Naj bo lastna vrednost λ_i matrike A enostavna in x_i ustrezeni desni, y_i levi lastni vektor, $\|x_i\|_2 = \|y_i\|_2 = 1$ za $i = 1, \dots, n$. Če je $\lambda_i + \delta\lambda_i$ ustrezna lastna vrednost zmotene matrike $A + \delta A$ in je $x_i + \delta x_i$ pripadajoči lastni vektor, potem velja

$$x_i + \delta x_i = x_i + \sum_{\substack{k=1 \\ k \neq i}}^n \frac{y_k^H \delta A x_i}{(\lambda_i - \lambda_k) s_k} x_k + \mathcal{O}(\|\delta A\|^2).$$

Lema

Če je λ_i enostavna lastna vrednost, potem je $s_i \neq 0$.

Opomba: V primeru večkratne lastne vrednosti lahko vektorja x_i in y_i določimo tako, da bo $s_i = 0$, ni pa to nujno res za poljubne lastne vektorje večkratne lastne vrednosti.

Matrika ne more imeti natanko ene zelo občutljive lastne vrednosti

Lema

Naj bo $A = X\Lambda X^{-1}$, kjer je $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_n)$, $X = [x_1 \cdots x_n]$ matrika normiranih desnih in $Y = [y_1 \cdots y_n]$ matrika normiranih levih lastnih vektorjev. Potem velja

$$X^{-1} = \begin{bmatrix} \frac{1}{s_1} y_1^H \\ \vdots \\ \frac{1}{s_n} y_n^H \end{bmatrix}.$$

Izrek

Matrika ne more imeti natanko ene zelo občutljive lastne vrednosti.

Opomba: Velika občutljivost lastne vrednosti pomeni, da je blizu večkratne lastne vrednosti, to pa seveda pomeni, da obstaja vsaj še ena taka lastna vrednost.

Občutljivost in oddaljenost do najbližjega singularnega problema

Pri občutljivosti matrike za reševanje linearnega sistema smo videli, da je recipročna vrednost občutljivosti enaka oddaljenosti od najbližjega singularnega sistema. Podobno tudi sedaj velja, da je občutljivost enostavne lastne vrednosti povezana z oddaljenostjo od najbližje matrike z večkratno lastno vrednostjo.

Izrek

Naj bo λ_i enostavna lastna vrednost matrike A z normiranimi levim in desnim lastnim vektorjem y_i in x_i in naj bo $s_i = y_i^ x_i$. Potem obstaja matrika $A + \delta A$ z večkratno lastno vrednostjo λ_i in*

$$\frac{\|\delta A\|_2}{\|A\|_2} \leq \frac{|s_i|}{\sqrt{1 - |s_i|^2}}.$$

Obratna napaka

Naj bo $(\hat{\lambda}, \hat{x})$ približek za lastni par matrike A . Ali lahko ocenimo napako?

Definicija

Po normi relativna obratna napaka približka lastnega para $(\hat{\lambda}, \hat{x})$ je

$$\eta(\hat{\lambda}, \hat{x}) = \min \left\{ \epsilon > 0 : (A + \delta A)\hat{x} = \hat{\lambda}\hat{x}, \|\delta A\| \leq \epsilon \|A\| \right\}.$$

Lema

Velja

- 1) $\eta(\hat{\lambda}, \hat{x}) = \frac{\|A\hat{x} - \hat{\lambda}\hat{x}\|}{\|A\|\|\hat{x}\|},$
- 2) $\eta(\hat{\lambda}) = \min_{\hat{x} \neq 0} \eta(\hat{\lambda}, \hat{x}) = \frac{\sigma_{\min}(A - \hat{\lambda}I)}{\|A\|},$
- 3) $\eta(\hat{x}) = \min_{\hat{\lambda}} \eta(\hat{\lambda}, \hat{x}) = \frac{\|A\hat{x} - \rho(A, \hat{x})\hat{x}\|}{\|A\|\|\hat{x}\|}.$

Naj bo $(\hat{\lambda}, \hat{x})$ približek za lastni par matrike A in

$$r = A\hat{x} - \hat{\lambda}\hat{x}.$$

Če poznamo tudi približek za levi lastni vektor $\hat{y}$, lahko ocenimo

$$|\hat{\lambda} - \lambda| \leq \frac{\|r\|}{|s|},$$

kjer je

$$s := \frac{\hat{y}^H \hat{x}}{\|\hat{x}\| \|\hat{y}\|}.$$

$$A = \begin{bmatrix} -149 & -50 & -154 \\ 537 & 180 & 546 \\ -27 & -9 & -25 \end{bmatrix}$$

ima lastne vrednosti $\lambda_1 = 1$, $\lambda_2 = 2$, $\lambda_3 = 3$. V Matlabu `l=eig(A)` vrne

$$\begin{aligned} \hat{\lambda}_1 &= 1.000000000010722 \\ \hat{\lambda}_2 &= 1.999999999991790 \\ \hat{\lambda}_3 &= 2.999999999997399. \end{aligned}$$

Če izračunamo še leve in desne lastne vektorje in ostanke, dobimo ocene:

$\ r_i\ $	$ s_i ^{-1}$	$\ r_i\ / s_i $	$ \hat{\lambda}_i - \lambda_i $
$3.8859e - 014$	$6.0364e + 002$	$2.3457e - 011$	$1.0722e - 011$
$2.9063e - 014$	$3.9524e + 002$	$1.1487e - 011$	$8.2105e - 012$
$9.3727e - 014$	$2.1929e + 002$	$2.0554e - 011$	$2.6010e - 012$

Zgled 2

Matrika Godunova

$$A = \begin{bmatrix} 289 & 2064 & 336 & 128 & 80 & 32 & 16 \\ 1152 & 30 & 1312 & 512 & 288 & 128 & 32 \\ -29 & -2000 & 756 & 384 & 1008 & 224 & 48 \\ 512 & 128 & 640 & 0 & 640 & 512 & 128 \\ 1053 & 2256 & -504 & -384 & -756 & 800 & 208 \\ -287 & -16 & 1712 & -128 & 1968 & -30 & 2032 \\ -2176 & -287 & -1565 & -512 & -541 & -1152 & -289 \end{bmatrix}$$

ima eksaktne lastne vrednosti $-4, -2, -1, 0, 1, 2, 4$.

Matlab vrne $\text{eig}(A) =$

$$\begin{array}{lcl} 5.0968 & + & 1.8932i \\ 5.0968 & - & 1.8932i \\ 1.2157 & + & 4.3784i \\ 1.2157 & - & 4.3784i \\ -5.6738 & & \\ -3.4756 & + & 3.4463i \\ -3.4756 & - & 3.4463i \end{array}$$

Zgled 3

Matrika

$$A = \begin{bmatrix} -9 & 11 & -21 & 63 & -252 \\ 70 & -69 & 141 & -421 & 1684 \\ -575 & 575 & -1149 & 3451 & -13801 \\ 3891 & -3891 & 7782 & -23345 & 93365 \\ 1024 & -1024 & 2048 & -6144 & 24572 \end{bmatrix}$$

ima petkratno lastno vrednost 0, saj je $A^5 = 0$.

Matlab vrne $\text{eig}(A) =$

$$\begin{aligned} & -4.0520e - 002 \\ & -1.1779e - 002 + 3.8286e - 002i \\ & -1.1779e - 002 - 3.8286e - 002i \\ & 3.2040e - 002 + 2.2812e - 002i \\ & 3.2040e - 002 - 2.2812e - 002i \end{aligned}$$