

Analisi della perturbazione dello spettro di una matrice tridiagonale di Toeplitz e Applicazioni

S. Noschese

Due Giorni di Algebra Lineare Numerica
16 - 17 Febbraio 2012, Genova.

Matrici Tridiagonali di Toeplitz. Motivazioni

- Le matrici tridiagonali di Toeplitz compaiono in varie applicazioni, tra le quali la soluzione di ODE e PDE.
- Analisi del condizionamento e condizionamento strutturato dello spettro. Studio dell' ε -pseudospettro. Derivazione di espressioni esplicite. In collaborazione con L. Pasquini e L. Reichel.
- Applicazione alla costruzione dell' ε -pseudospettro strutturato di una matrice tridiagonale di Toeplitz. In collaborazione con P. Buttà e N. Guglielmi.

Matrici Tridiagonali di Toeplitz. Notazioni

$$T = (n; \sigma, \delta, \tau) = \begin{bmatrix} \delta & \tau & & & & & O \\ \sigma & \delta & \tau & & & & \\ & \sigma & \cdot & \cdot & & & \\ & & \cdot & \cdot & \cdot & & \\ & & & \cdot & \cdot & \cdot & \\ & & & & \cdot & \cdot & \tau \\ O & & & & & \sigma & \delta \end{bmatrix} \in \mathbb{C}^{n \times n}$$

$$\alpha = \arg \sigma, \quad \beta = \arg \tau, \quad \gamma = \arg \delta.$$

Autovalori ed autovettori di T

- Lo spettro di $T = (n; \sigma, \delta, \tau)$ è noto:

$$\lambda_h(T) = \delta + 2\sqrt{|\sigma\tau|} e^{i(\alpha+\beta)/2} \cos \frac{h\pi}{n+1}, \quad h = 1 : n.$$

- Se $\sigma\tau \neq 0$, le componenti dell'autovettore destro $x_h = [x_{h,1}, x_{h,2}, \dots, x_{h,n}]^T$ associato all'autovalore $\lambda_h(T)$ sono date da

$$x_{h,k} = (\sigma/\tau)^{k/2} \sin \frac{hk\pi}{n+1}, \quad k = 1 : n, \quad h = 1 : n.$$

Il corrispondente autovettore sinistro $y_h = [y_{h,1}, y_{h,2}, \dots, y_{h,n}]^T$ ha componenti

$$y_{h,k} = (\bar{\tau}/\bar{\sigma})^{k/2} \sin \frac{hk\pi}{n+1}, \quad k = 1 : n, \quad h = 1 : n.$$

Sensibilità dello spettro. Misure di non-normalità

- La matrice $T = (n; \sigma, \delta, \tau)$ è normale se e solo se

$$|\sigma| = |\tau|.$$

- La *distanza dalla normalità* nella norma di Frobenius della matrice T è data da

$$d_F(T, \mathcal{N}) = \min_{T_N \in \mathcal{N}} \|T - T_N\|_F,$$

dove \mathcal{N} è la varietà algebrica delle matrici normali in $\mathbb{C}^{n \times n}$.

- La *distanza strutturata dalla normalità* nella norma di Frobenius della matrice T è data da

$$d_F(T, \mathcal{N}_{\mathcal{T}}) = \min_{T_N \in \mathcal{N}_{\mathcal{T}}} \|T - T_N\|_F,$$

dove \mathcal{T} è il sottospazio di $\mathbb{C}^{n \times n}$ formato dalle matrici tridiagonali di Toeplitz e $\mathcal{N}_{\mathcal{T}}$ è $\mathcal{N} \cap \mathcal{T}$.

Sensibilità dello spettro. Matrice T^*

La formula per la distanza strutturata dalla normalità nella norma di Frobenius di matrici bandate di Toeplitz è data in [S.N.-Reichel, BIT(2009)].

Sia $T = (n; \sigma, \delta, \tau)$ una matrice in \mathcal{T} . Esiste una unica matrice $T^* = (n; \sigma^*, \delta^*, \tau^*) \in \mathcal{N}_{\mathcal{T}}$ che minimizza $\|T_{\mathcal{N}} - T\|_F$ su $\mathcal{N}_{\mathcal{T}}$:

$$\sigma^* = \frac{|\sigma| + |\tau|}{2} e^{i\alpha}, \quad \delta^* = \delta, \quad \tau^* = \frac{|\sigma| + |\tau|}{2} e^{i\beta}.$$

La distanza strutturata dalla normalità di T dipende dalla differenza tra i valori assoluti degli elementi sotto- e sopra- diagonali.

Sensibilità dello spettro. Distanza strutturata dalla normalità

$$d_F(T, \mathcal{N}_{\mathcal{T}}) = \sqrt{\frac{n-1}{2}} (\max\{|\sigma|, |\tau|\} - \min\{|\sigma|, |\tau|\}).$$

$$0 \leq \frac{d_F(T, \mathcal{N}_{\mathcal{T}})}{\|T\|_F} \leq \frac{\sqrt{2}}{2},$$

$$\frac{d_F(T, \mathcal{N}_{\mathcal{T}})}{\|T\|_F} = \frac{\sqrt{2}}{2} \text{ sse } [\delta, \sigma\tau] = [0, 0] \text{ e } \frac{d_F(T, \mathcal{N}_{\mathcal{T}})}{\|T\|_F} = 0 \text{ sse } |\sigma| = |\tau|.$$

Sensibilità dello spettro. L' ε -pseudospettro

Per un dato $\varepsilon > 0$, l' ε -pseudospettro di $T = (n; \sigma, \delta, \tau) \in \mathbb{C}^{n \times n}$ è

$$\Lambda_\varepsilon(T) = \{z : \|(zI - T)^{-1}\| \geq \varepsilon^{-1}\}.$$

Gli ε -pseudospettri di matrici di Toeplitz bandate sono stati analizzati in dettaglio in [Reichel-Trefethen, LAA(1992)], [Trefethen-Embree, 2005], e [Böttcher-Grudsky, 2005].

Quando $\varepsilon \rightarrow 0$ e $n \rightarrow \infty$, $\Lambda_\varepsilon(T)$ approssima lo spettro dell'operatore di Toeplitz $T_\infty = (\infty; \sigma, \delta, \tau)$. Il simbolo di T è $g(z) = \sigma z + \delta + \tau z^{-1}$, e la frontiera di $\Lambda(T_\infty)$ è $g(S) = \{g(z) : z \in \mathbb{C}, |z| = 1\}$, dove S è il cerchio unitario.

Per ogni $n \geq 1$ finito, $\lambda(T)$ vive nell'intervallo tra i fuochi dell'ellisse $g(S)$

$$\mathcal{S}_{\text{foci}} = \left\{ \delta + t e^{i(\alpha+\beta)/2}, t \in \mathbb{R}, |t| \leq 2\sqrt{|\sigma\tau|} \right\},$$

e $\lambda(T^*)$, dove T^* è la matrice più vicina a T in \mathcal{N}_T , nell'asse maggiore di $g(S)$ [S.N.-Pasquini-Reichel, NLAA(2011)]

$$\mathcal{S}_{\text{major_axis}} = \left\{ \delta + t e^{i(\alpha+\beta)/2}, t \in \mathbb{R}, |t| \leq |\sigma| + |\tau| \right\}.$$

$$T = \left(50; \frac{4 + 3i}{10}, 16 - 3i, -5 \right).$$

Usiamo le formule (o l'algoritmo QR implementato dalla funzione **eig** del MATLAB) per il calcolo degli autovalori.

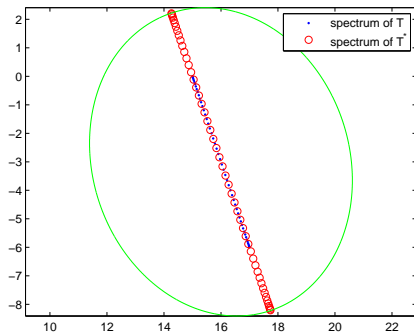


Figure: Spettri della matrice T (in blu) e della matrice tridiagonale di Toeplitz normale più vicina T^* (in rosso) e immagine del cerchio unitario sotto il simbolo di T (in verde).

Sensibilità dello spettro. Condizionamento tradizionale degli autovalori

$$\text{Individuale } \kappa(\lambda) = \frac{\|x\|_2 \|y\|_2}{|y^H x|} \implies \text{Globale } \kappa_F(\lambda) = \sum_{h=1}^n \kappa(\lambda_h)$$

Sia $r = \min\{|\sigma|, |\tau|\} / \max\{|\sigma|, |\tau|\}$. In [S.N.-Pasquini-Reichel, NLAA(2011)] si mostra che

$$\kappa(\lambda_h(T)) = \frac{(1 - r^{n+1})(1 + r)(1 - \cos \frac{2h\pi}{n+1})}{r^{(n-1)/2}(n+1)(1-r)(1+r^2 - 2r \cos \frac{2h\pi}{n+1})}. \quad (*)$$

Per il numero di condizionamento globale si ha

$$\frac{K_{n,r}}{2} \leq \kappa_F(\lambda) \leq 2K_{n,r}, \text{ dove } K_{n,r} = \frac{1}{r^{(n-1)/2}} \frac{1 - r^{n+1}}{1 - r} (1 + r) \frac{n}{n+1}.$$

Sensibilità dello spettro. Condizionamento tradizionale degli autovalori

- La sensibilità degli autovalori è indipendente dall'elemento diagonale δ e dagli argomenti α e β degli elementi non-diagonali;
- gli autovalori nel centro dello spettro hanno il peggior condizionamento;
- se $0 < r < 1$, il numero di condizionamento $\kappa(\lambda_h(T))$ cresce esponenzialmente con n ;
- se T è normale, i.e. $|\sigma| = |\tau|$ (e $r = 1$),

$$\kappa(\lambda_h(T)) = \frac{\|x_h\|_2 \|y_h\|_2}{|y_h^H x_h|} = 1;$$

- si ha $\kappa(\lambda_h(T)) \rightarrow 1$ quando $r \rightarrow 1$;
- si ha $\kappa(\lambda_h(T)) \rightarrow \infty$ quando $r \rightarrow 0$. E, in tal caso, si ha la stima

$$\kappa(\lambda_h(T)) \approx \frac{1 - \cos \frac{2h\pi}{n+1}}{n+1} \left(\frac{1}{r}\right)^{\frac{n-1}{2}} \quad h = 1 : n.$$

Sensibilità dello spettro. Condizionamento tradizionale degli autovalori

Sia $a = \sigma/\tau$. Gli autovalori sono molto malcondizionati quando o $|a|$ o il suo reciproco, è vicino a zero.

Si assuma che gli autovettori destro e sinistro associati ad un autovalore λ hanno norma unitaria. Ipotizziamo che $|a|$, o $1/|a|$, è vicino a zero. Vale che

- quanto più è piccolo $|a| < 1$, tanto più è grande la prima componente di x e l'ultima componente di y ;
- quanto più è grande $|a| > 1$, tanto più è grande l'ultima componente di x e la prima componente di y .

Si consideri la *perturbazione di Wilkinson* $W = yx^H$. La perturbazione (di norma unitaria) W di T induce la più grande perturbazione in λ . Nell'ipotesi che $|a|$ o $1/|a|$ è vicino a zero, gli elementi di massimo modulo in W sono prossimi all'angolo in basso a sinistra se $|a| < 1$ o in alto a destra se $|a| > 1$.

$$T = \left(50; -5, 16 - 3i, \frac{4 + 3i}{10} \right).$$

Usiamo l'algoritmo QR. In alto a destra errori di arrotondamento dell'algoritmo *non* strutturato.

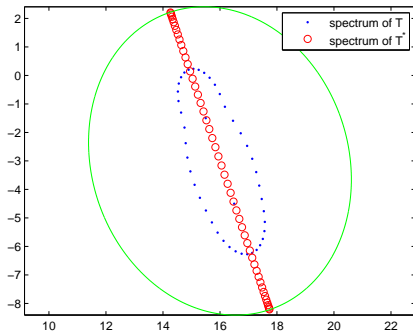


Figure: Spettri della matrice T (in blu) e della matrice tridiagonale di Toeplitz normale più vicina T^* (in rosso) e immagine del cerchio unitario sotto g (in verde).

Sensibilità dello spettro. Condizionamento strutturato degli autovalori

Si ammettano solo perturbazioni strutturate. Per il numero di condizionamento strutturato di λ vale [S.N.-Pasquini, JCAM(2006)]

$$\kappa_{\mathcal{T}}(\lambda) = \kappa(\lambda) \|W|_{\mathcal{T}}\|_F,$$

dove $W|_{\mathcal{T}}$ è la proiezione di W nel sottospazio \mathcal{T} delle matrici tridiagonali di Toeplitz. Inoltre tale matrice - normalizzata - induce la più grande perturbazione strutturata (di norma unitaria) in λ .

Nell'ipotesi che $|a|$ o $1/|a|$ è vicino a zero, gli elementi di massimo modulo in W sono prossimi all'angolo in basso a sinistra se $|a| < 1$ o in alto a destra se $|a| > 1$. Quindi gli elementi di massimo modulo non sono in $W|_{\mathcal{T}}$. Ne segue che un grande $\kappa(\lambda)$ non implica necessariamente che $\kappa_{\mathcal{T}}(\lambda)$ sia grande.

Un autovalore λ può essere *molto* più sensibile a una perturbazione qualsiasi di T che non a una perturbazione strutturata.

Confronto tra condizionamento tradizionale e strutturato

λ	$\kappa(\lambda(T))$	$\kappa_{\mathcal{T}}(\lambda(T))$
λ_1	$7.0463 \cdot 10^4$	$8.7215 \cdot 10^{-1}$
λ_2	$2.5759 \cdot 10^5$	$8.2610 \cdot 10^{-1}$
λ_3	$5.0517 \cdot 10^5$	$7.5194 \cdot 10^{-1}$
λ_4	$7.5633 \cdot 10^5$	$6.5374 \cdot 10^{-1}$
λ_5	$9.7209 \cdot 10^5$	$5.3790 \cdot 10^{-1}$
λ_6	$1.1325 \cdot 10^6$	$4.1511 \cdot 10^{-1}$
λ_7	$1.2300 \cdot 10^6$	$3.0680 \cdot 10^{-1}$
λ_8	$1.2626 \cdot 10^6$	$2.5820 \cdot 10^{-1}$
λ_9	$1.2300 \cdot 10^6$	$3.0680 \cdot 10^{-1}$
λ_{10}	$1.1325 \cdot 10^6$	$4.1511 \cdot 10^{-1}$
λ_{11}	$9.7209 \cdot 10^5$	$5.3790 \cdot 10^{-1}$
λ_{12}	$7.5633 \cdot 10^5$	$6.5374 \cdot 10^{-1}$
λ_{13}	$5.0517 \cdot 10^5$	$7.5194 \cdot 10^{-1}$
λ_{14}	$2.5759 \cdot 10^5$	$8.2610 \cdot 10^{-1}$
λ_{15}	$7.0463 \cdot 10^4$	$8.7215 \cdot 10^{-1}$

Table: Numeri di condizionamento tradizionale e strutturato degli autovalori di $T = (15; -i, 11 - 2i, 6 + 8i)$.

Confronto tra condizionamento tradizionale e strutturato

Gli autovalori nel centro dello spettro hanno il peggior condizionamento rispetto a perturbazioni *non strutturate*. In maggior dettaglio, nella formula (*) il fattore che dipende da h assume il massimo per $h = \lfloor n/2 \rfloor$

$$\frac{1}{2} \leq \frac{1 - \cos \frac{2h\pi}{n+1}}{1 + r^2 - 2r \cos \frac{2h\pi}{n+1}} \leq 2.$$

D'altra parte, più gli autovalori di T sono lontani da δ , maggiore risulta la loro sensibilità a perturbazioni *strutturate*. In effetti, lo Jacobiano di $f = \lambda(T)$ rispetto a σ e τ è

$$J_f(\sigma, \tau) = \begin{bmatrix} \sqrt{\frac{\tau}{\sigma}} \cos \frac{\pi}{n+1} & \sqrt{\frac{\sigma}{\tau}} \cos \frac{\pi}{n+1} \\ \sqrt{\frac{\tau}{\sigma}} \cos \frac{2\pi}{n+1} & \sqrt{\frac{\sigma}{\tau}} \cos \frac{2\pi}{n+1} \\ \cdot & \cdot \\ \cdot & \cdot \\ \sqrt{\frac{\tau}{\sigma}} \cos \frac{n\pi}{n+1} & \sqrt{\frac{\sigma}{\tau}} \cos \frac{n\pi}{n+1} \end{bmatrix}.$$

Quindi la sensibilità strutturata di $\lambda_h(T)$ aumenta con la sua distanza da δ .

Applicazione all'approssimazione dell'ascissa pseudospettrale strutturata

In [Guglielmi-Overton, SIMAX(2011)] un algoritmo trova l'ascissa pseudospettrale di una matrice A .

Un nuovo algoritmo è stato costruito per approssimare l'ascissa dell' ε -pseudospettro strutturato di T [Buttà-Guglielmi-S.N.(2012)].

Algoritmo. Sia λ_0 l'autovalore più a destra di T . Sia

$$T_1 = T + \varepsilon \frac{W^0|_{\mathcal{T}}}{\|W^0|_{\mathcal{T}}\|_F}.$$

Per $k = 1, 2, \dots$, sia λ_k l'autovalore più a destra di T_k . Sia

$$T_{k+1} = T + \varepsilon \frac{W^k|_{\mathcal{T}}}{\|W^k|_{\mathcal{T}}\|_F}. \diamond$$

Si ha convergenza locale dell'algoritmo verso lo pseudoautovalore strutturato più a destra di T (*unico* punto stazionario dell'algoritmo). Naturali generalizzazioni permettono di approssimare il raggio pseudospettrale strutturato e di disegnare sezioni significative del bordo dello pseudospettro strutturato. Estensione ad altre pattern di simmetria (Toeplitz bandate e non, Hankel complesse, ...).

$$T = \left(15; \frac{-1+i}{10}, \frac{-3+4i}{10}, 2+i\right).$$

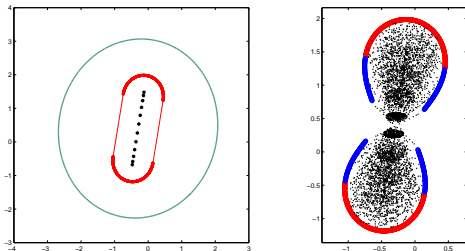


Figure: Figura a sinistra: frontiera dello pseudospettro non strutturato (ovale grigio) e sezione della frontiera dell' ε -pseudospettro strutturato (in rosso), con $\varepsilon = 0.5$, di T . La linea sottile ne delinea l'involucro convesso. Figura a destra: i punti neri sono gli spettri di 1000 matrici perturbate ottenute aggiungendo a T perturbazioni tridiagonali random di Toeplitz di norma di Frobenius ε . In blu ulteriore sezione dello pseudospettro strutturato.