

Un approccio root-finding ai problemi polinomiali agli autovalori

Vanni Noferini (Università di Pisa)
In collaborazione con D.A. Bini, L. Gemignani
Genova - DGALN 2012

17 Febbraio 2012

Problemi polinomiali agli autovalori (PEP)

Setting: Problemi polinomiali agli autovalori per polinomi **regolari**

$$P(\lambda) = \sum_{i=0}^k P_i \lambda^i, \quad P_i \in \mathbb{C}^{n \times n}, \quad \det(P(\lambda)) \neq 0$$

Si cerca λ_0 tale che $P(\lambda_0)x_0 = 0$ per qualche vettore $0 \neq x_0 \in \mathbb{C}^n$.

λ_0 è un autovalore se e solo se è una radice di
 $p(\lambda) := \det(P(\lambda)) = 0$.

Metodi numerici per PEP

L'approccio usuale in letteratura:

- Si cerca una linearizzazione per il polinomio. $L(x)$ è una linearizzazione di $P(x)$ se esistono due polinomi unimodulari $E(x), F(x)$ tali che $E(x)L(x)F(x) = \begin{pmatrix} P(x) & 0 \\ 0 & I \end{pmatrix}$
- Si applica a $L(x)$ un metodo matriciale iterativo (per es. QZ)
- Complessità del QZ: $O(k^3 n^3)$ operazioni

Metodi numerici per PEP

Oggi presenteremo un approccio alternativo basato sull'applicazione di metodi root-finding all'equazione caratteristica del polinomio di matrici. In particolare il **metodo di Ehrlich-Aberth** ha una storia di implementazioni robuste nel caso scalare: vogliamo estenderne l'uso al problema matriciale.

Non è desiderabile dal punto di vista dell'analisi numerica calcolare esplicitamente $p(\lambda)$ per evitare possibili criticità nel calcolo dei coefficienti del determinante (overflow, underflow, instabilità...).

Il metodo di Ehrlich-Aberth

Il metodo di Ehrlich-Aberth approssima simultaneamente tutte le radici di un polinomio scalare di grado N :

$$z_j^{(k+1)} = z_j^{(k)} - \frac{p(z_j^{(k)})/p'(z_j^{(k)})}{1 - \frac{p(z_j^{(k)})}{p'(z_j^{(k)})} \sum_{\ell=1, \ell \neq j}^N \frac{1}{z_j^{(k)} - z_\ell^{(k)}}}, \quad 1 \leq j \leq N.$$

Bini e Fiorentino (2000): implementazione robusta ed efficiente.

Applicazione ai PEPs

È preferibile stimare la correzione di Newton tramite metodi indiretti e operanti sui coefficienti matriciali. Possiamo usare la formula di Jacobi:

$$\frac{d \det P(\lambda)}{d\lambda} = \det(P(\lambda)) \cdot \operatorname{tr}(P^{-1}(\lambda) \frac{dP(\lambda)}{d\lambda})$$

Applicazione ai PEPs

È dunque cruciale lo sviluppo di un algoritmo che calcoli efficientemente la traccia di $P^{-1}(\lambda) \frac{dP(\lambda)}{d\lambda}$.

Diverse opzioni:

- Usare una linearizzazione companion ed eseguire una fattorizzazione LQ sfruttando la struttura a blocchi, infine calcolare la traccia. Complessità $O(n^3 + kn^2)$
- Calcolare una linearizzazione in forma Hessenberg-triangolare e sfruttarne la struttura semiseparabile per calcolare la traccia. Complessità $O(n^2k^2)$
- Non linearizzare, ma valutare con il metodo di Horner $P(x)$ e $P'(x)$. Complessità: $O(n^3 + kn^2)$

Linearizzare?

Principale vantaggio nell'uso della linearizzazione: maggiore flessibilità verso il caso di grandi valori di n .

Principale vantaggio nell'evitare la linearizzazione: si evita ogni possibile rischio di peggioramento del condizionamento dell'autovalore.

Complessità

La complessità totale dell'algoritmo è di ordine $O(ts)$ dove s è il costo relativo a una singola valutazione delle traccia e t è il numero di valutazioni necessari per la convergenza.

Da esperimenti numerici: Quando $k^2 > n$ tramite uso di opportune strategie per la scelta delle approssimazioni iniziali si riesce ad ottenere una dipendenza lineare di t rispetto al numero totale di approssimazioni da calcolare nk . La dipendenza è leggermente superlineare quando $n \gg k^2$. Nel caso strutturato si ha linearità anche in questa regione.

- **Se $k^2/n \gg 1$:** bene! Dipendenza solo quadratica in k (quarta potenza in n)
- **Se $n/k^2 \gg 1$:** Dipendenza (super)cubica da nk ($\approx kn^{3.1}$)

Esperimenti numerici: efficienza

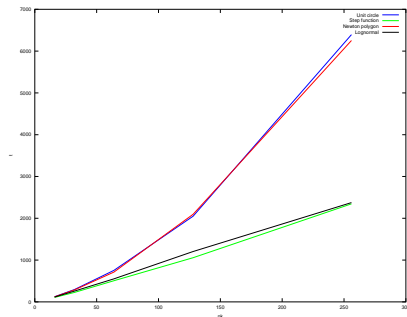
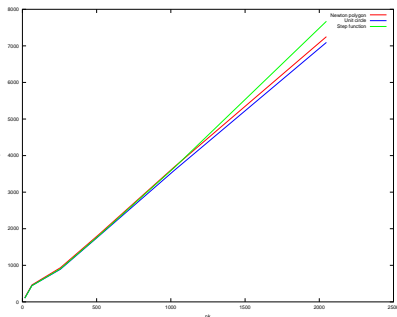


Fig. 1. Dipendenza di t da nk per $n = 2$ (sx.) e per $k = 2$ (dx.)

Esperimenti numerici: accuratezza

Problema hospital in NLEVP:

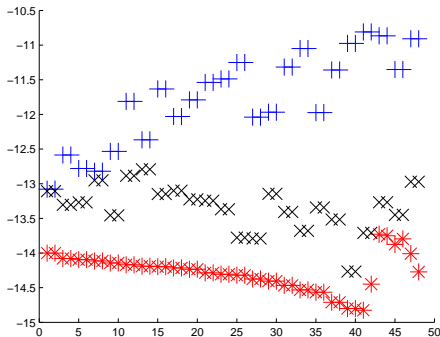


Fig. 2. Errori assoluti in avanti: *polyeig* (blu), *quadeig* (nero), *EAI* (rosso)

Autovettori

Gli autovettori possono essere ottenuti associando la EAI con altri metodi (es. spazio nullo con SVD, iterazione inversa). Non aumenta il costo computazionale.

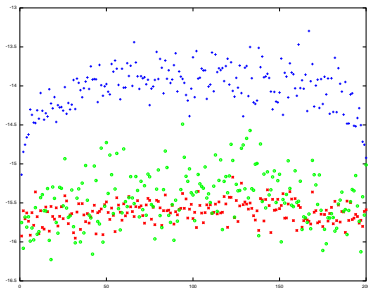


Fig. 3. Residui autovettori: *polyeig* (blu), *EAI+II* (verde), *EAI+SVD* (rosso)

PEPs strutturati...

Alcuni problemi polinomiali agli autovalori di interesse per ingegneri, fisici, matematici, etc. presentano peculiari strutture nei coefficienti di $P(x)$ che a loro volta inducono una struttura nello spettro degli autovalori.

È di interesse teorico e di utilità applicativa lo sviluppo di metodi numerici che siano sufficientemente flessibili da poter catturare esattamente le simmetrie nello spettro.

...e spettri strutturati

Supponiamo che $f(x)$ sia tale che $f(f(x)) = x \forall x$ e che, a causa di qualche simmetria nota del problema, gli autovalori si presentino in coppie $\{\lambda_i, f(\lambda_i)\}$.

Possiamo fattorizzare $p(\lambda) = \prod_i p_i(\lambda)$, dove abbiamo definito $p_i(\lambda) = \lambda^2 - (\lambda_i + f(\lambda_i))\lambda + \lambda_i f(\lambda_i)$.

Esempi di funzioni che coincidono con la propria inversa:

- Razionali del tipo $f(x) = \frac{ax+b}{cx-a}$ purché $a^2 + bc \neq 0$.
- Qualsiasi $h = g \circ f \circ g^{-1}$ con g invertibile ed f self-inverse (su \mathbb{C} , solo le razionali mostrate sopra)

EA e spettri strutturati

Sono possibili diverse strategie per costruire varianti di E-A capaci di catturare lo spettro:

- Nel caso speciale $f(x) = x$, ogni autovalore ha molteplicità pari, $p(\lambda) = q(\lambda)^2$, e si può applicare il metodo a $q(\lambda)$.
- Si può approssimare solo metà degli autovalori forzando ad ogni iterazione l'altra metà $z_j^{(k)} = f(z_{j-nk/2}^{(k)})$ per $j = nk/2 + 1, \dots, nk$.
- Supponiamo che esista una trasformazione $\mu = \mu(\lambda)$ tale che $p_i(\lambda(\mu))$ è lineare in μ . Si può allora applicare il metodo a $p(\lambda(\mu))$.
- Si può esplicitamente costruire un nuovo polinomio $M(\mu)$ il cui determinante è $p(\lambda(\mu))^2$, riconducendosi al primo caso.

Applicazioni

Alcuni esempi pratici di strutture:

- $f(x) = x$
Polinomi skew-Hamiltonian; polinomi antisimmetrici con n pari
- $f(x) = -x$
Polinomi even/odd; polinomi Hamiltonian/skew-Hamiltonian
- $f(x) = (\pm)\frac{1}{x}$
Polinomi palindromici/antipalindromici; matrici simplettiche

Polinomi T-palindromici

Struttura nei coefficienti: $P_i = P_{k-i}^T$; quindi $P(\lambda^{-1}) = \lambda^{-k} P^T(\lambda)$



Struttura spettrale: λ_0 autovalore $\leftrightarrow \lambda_0^{-1}$ autovalore

Polinomi T-palindromici

Sia $P(\lambda)$ un polinomio T-pal. di grade pari.

Usiamo la *trasformazione di Dickson* $\mu = \lambda + \lambda^{-1}$

Trucco: usare i polinomi di Dickson

$$\begin{cases} \phi_0(\mu) = 2 \\ \phi_1(\mu) = \mu \\ \forall j \geq 1, \mu \phi_j(\mu) = \phi_{j+1}(\mu) + \phi_{j-1}(\mu) \end{cases}$$

(Sostanzialmente sono polinomi trigonometrici, i.e.,

$$\lambda = e^{i\alpha} \Rightarrow \phi_j = 2 \cos(j\alpha)$$

Polinomi T-palindromici

Usiamo la *trasformazione di Dickson* $\mu = \lambda + \lambda^{-1}$

Un po' di algebra porta a un polinomio $M(\mu)$ il cui determinante è $[p(\lambda(\mu))]^2$. Usare la base di Dickson aiuta qualora si passi per la linearizzazione.

Se $\mu = \lambda + \lambda^{-1}$ and $\omega = \lambda - \lambda^{-1}$ allora esistono polinomi B e C per cui $P(\lambda) = B(\mu) + \omega C(\mu)$. Definiamo dunque

$$M(\mu) := \begin{pmatrix} B(\mu) & \omega^2 C(\mu) \\ C(\mu) & B(\mu) \end{pmatrix}.$$

Polinomi T-palindromici

Osservazioni:

- M è in effetti un polinomio in μ in quanto $\omega^2 = \mu^2 - 4$
- $\det M(\mu) = [p(\lambda(\mu))]^2$
- I coefficienti di M espressi nella base di Dickson sono ottenibili in $O(kn^2)$ operazioni.
- Risultati teorici consentono di ottenere anche autovettori e vettori di Jordan.

Un teorema sul cambio di variabile

Theorem

Siano $m, p \in \mathbb{N}_0$ e $n(y), d(y) \in \mathbb{F}[y]$, g il grade di $P(x)$,
 $G = \max(\deg n(y), \deg d(y))$. e gG il grade di
 $Q(y) = \Phi_g(P(x)) := [d(y)]^g P\left(\frac{n(y)}{d(y)}\right) \in (\mathbb{F}[y])^{m \times p}$. Allora $x_0 \in \mathbb{F}^*$ è
autovalore di $P(x) \in (\mathbb{F}[x])^{m \times p}$ associato ai divisori elementari
 $(x - x_0)^{\ell_1}, \dots, (x - x_0)^{\ell_j}$ sss $\forall y_0 \in T_{x_0}$ y_0 è autovalore di $Q(y)$
associato ai divisori elementari $(y - y_0)^{m_0 \ell_1}, \dots, (y - y_0)^{m_0 \ell_j}$, dove
 m_0 è la molteplicità di y_0 . Inoltre:

- gli indici minimi destri di $P(x)$ sono β_1, \dots, β_s sss gli indici minimi destri di $Q(y)$ sono $G\beta_1, \dots, G\beta_s$;
- gli indici minimi sinistri di $P(x)$ sono $\gamma_1, \dots, \gamma_r$ sss gli indici minimi sinistri di $Q(y)$ sono $G\gamma_1, \dots, G\gamma_r$.

Polinomi even/odd

Struttura nei coefficienti: alternativamente matrici simmetriche e antisimmetriche. Dunque $P^T(x) = \pm P(-x)$.

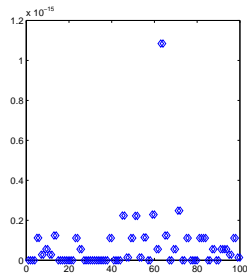
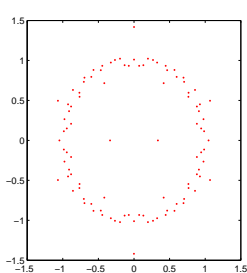


Spettro strutturato: λ autovalore $\leftrightarrow -\lambda$ autovalore

Possiamo usare la trasformazione quadratica $\mu = \lambda^2$

Esperimenti numerici: accuratezza

Polinomio odd con coefficienti generati casualmente, $k = 50$, $n = 2$.
Usiamo $\mu = \lambda^2$, applichiamo EA a $q(\mu) := p(\lambda(\mu))$, cosicché
 $\lambda_{i,nk+1-i} = \pm \sqrt{\mu_i}$.
Approssimazioni calcolate ed errori assoluti (rispetto ai valori esatti
calcolati con Mathematica e MPSolve [Bini Fiorentino 2000]):



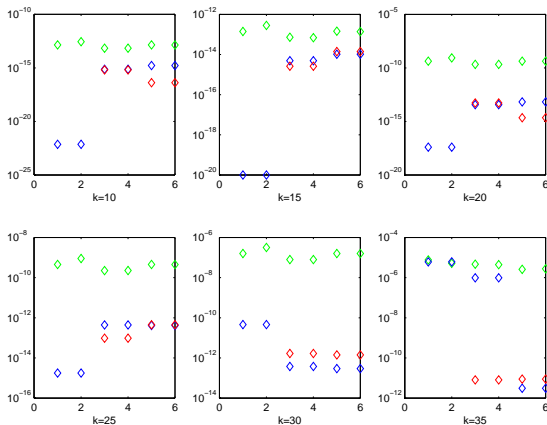
Esperimenti numerici: accuratezza

Un test più probante [Kressner-Schroeder-Watkins 2009]:

$$A = \begin{pmatrix} 0 & 1 - \phi & 0 \\ \phi - 1 & 0 & i \\ 0 & -i & 0 \end{pmatrix}, M = \begin{pmatrix} A & 0 \\ 0 & 2A \end{pmatrix}$$

Gli autovalori di $\lambda I_6 + M$ sono noti: $\{0, 0, \pm w, \pm 2w\}$, $w = \sqrt{2\phi - \phi^2}$
Il numero di condizionamento non strutturato cresce quando $\phi \rightarrow 0$, pertanto il QR viene superato in accuratezza da metodi materiali strutturati come il QR palindromo [KSW 09]. Come si comportano le varianti strutturate di EA?

Errori assoluti per diversi valori di k e $\phi = 2^{-k}$. QR (verde), EA (blu), EA con deflazione dell'autovalore teoricamente noto 0 (rosso)



Autovalori eccezionali

Alcune delle varianti strutturate della EAI presentano una perdita di accuratezza per gli autovalori eccezionali (punti fissi di $f(x)$) o quasi-eccezionali.

Desideriamo introdurre un raffinamento strutturato in modo da ovviare al problema senza perdere l'esatta riproduzione della struttura spettrale.

Polinomi scalari intrecciati

Un polinomio scalare $p(x)$ viene detto $f(x)$ -intrecciato di grade pari se le sue radici si presentano in coppie $\{\lambda, f(\lambda)\}$ con $f(x)$ funzione autoinversa.

Divisione strutturata di polinomi intrecciati

Theorem

Sia $s(z) = \prod_{j=1}^m (z - \xi_j)(z - f(\xi_j))$, $\xi_j \neq \xi_k$ for $j \neq k$, un polinomio $f(x)$ -intrecciato di grade $n = 2m$. Per ogni $f(x)$ -egt $a(z)$ di grade n e $\xi \in \mathbb{C}$ t.c. $\xi \neq \xi_j$, $1 \leq j \leq m$, esistono unicamente determinati un $f(x)$ -egt $q(z)$ di grade $n - 2$ e uno scalare $r \in \mathbb{C}$ t. c.

$$a(z) = q(z)(z - \xi)(z - f(\xi)) + rs(z). \quad (1)$$

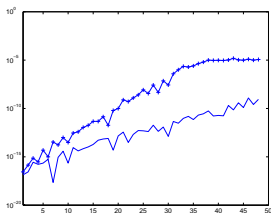
Un raffinamento tramite metodo di Newton

Dal teorema precedente si ottengono metodi di Newton per il raffinamento strutturato di alcune radici. Ad esempio per i palindromi:

$$\xi_{\ell+1} = \xi_{\ell} - \frac{a(\xi_{\ell})/a'(\xi_{\ell})}{1 - \frac{a(\xi_{\ell})}{a'(\xi_{\ell})} \left(\sum_{j=1}^k \frac{2\xi_{\ell} + \alpha_j}{\xi_{\ell}^2 + \alpha_j \xi_{\ell} + 1} + \sum_{j=1}^{m-k} \frac{2\xi_{\ell} + \theta_j}{\xi_{\ell}^2 + \theta_j \xi_{\ell} + 1} \right)}, \quad \ell \geq 0, \quad (2)$$

Un affinamento tramite metodo di Newton

Risultati del affinamento di autovalori quasi eccezionali per un polinomio T-palindromo, al variare di un parametro k . La distanza degli autovalori da 1 è proporzionale a 2^{-k} .



La figura mostra gli errori in avanti dopo la SEAI (crocette) e dopo il affinamento (continua).

Risultati comparabili con i metodi matriciali strutturati.

Conclusioni

- È possibile applicare il metodo di Ehrlich-Aberth ai PEPs, con vantaggi computazionali nel caso di grado alto.
- Varianti di EA consentono di catturare esattamente strutture spettrali derivanti da considerazioni teoriche.
- In generale, si ottengono errori in avanti minori del QZ. Nelle varianti strutturate, talvolta (ma non sempre) c'è perdita di accuratezza per gli autovalori eccezionali; si può ovviare con un raffinamento strutturato.
- **Obiettivi di ricerca:** migliore strategia per la scelta dei punti iniziali particolarmente nel caso di grandi valori di n ; approfondimento dell'analisi comparativa delle varie opzioni (linearizzazione/Horner, diverse varianti strutturate); comparazione EA/altri root-finders (es. Durand-Kerner, varianti di EA, etc.).