Iterazioni razionali per funzioni di matrici

Bruno Iannazzo

26 febbraio 2007

Outline

Motivazioni

- Risoluzione di equazioni matriciali non lineari

Motivazioni

- Risoluzione di equazioni matriciali non lineari
- Calcolo numerico di funzioni di matrici.
  Se $M^{-1}AM = \mathrm{diag}(J_1, \ldots, J_r)$

$$f(A) \doteq M\mathrm{diag}(f(J_1), \ldots, f(J_r))M^{-1}$$

dove

$$f(J) = \begin{bmatrix} f(\lambda) & f'(\lambda) & \ldots & \frac{f^{[d-1]}}{(d-1)!} \\ & f(\lambda) & \ddots & \ddots \\ & & \ddots & \ddots \\ & & & f(\lambda) \end{bmatrix}.$$

Iterazioni razionali scalari e matriciali

Un'iterazione razionale scalare è del tipo:

$$x_{k+1} = \varphi(x_k) = \frac{p(x_k)}{q(x_k)}.$$

Qual è la sua analoga matriciale?

Iterazioni razionali scalari e matriciali

Un'iterazione razionale scalare è del tipo:

$$x_{k+1} = \varphi(x_k) = \frac{p(x_k)}{q(x_k)}.$$

Qual è la sua analoga matriciale?

- variabile scalare $\longrightarrow$ matrice
- moltiplicazione $\longrightarrow$ prodotto righe per colonne
- divisione $\longrightarrow$ inversione

Cosa succede alle costanti scalari?

Due possibili generalizzazioni:

- coefficienti $\longrightarrow$ multipli dell'identità, $X_{k+1} = \varphi(X_k)$;

Due possibili generalizzazioni:

- coefficienti $\longrightarrow$ multipli dell'identità,
  $X_{k+1} = \varphi(X_k)$;
- coefficienti $\longrightarrow$ matrici,
  $X_{k+1} = \psi(X_k, A_0, \dots, A_n)$.

Due possibili generalizzazioni:

- coefficienti $\longrightarrow$ multipli dell'identità,
  $X_{k+1} = \varphi(X_k)$;
- coefficienti $\longrightarrow$ matrici,
  $X_{k+1} = \psi(X_k, A_0, \ldots, A_n)$.

Chiamiamo **iterazione razionale** la prima e **iterazione razionale dipendente da n parametri** la seconda.

Classe studiata

Si studia una classe intermedia tra le due e cioè le **iterazioni razionali dipendenti da un parametro**:

$$X_{k+1} = \psi(X_k, A),$$

dove $\psi(z, t)$ è una funzione razionale delle due variabili e $X_0 = p(A)$.

Classe studiata

Si studia una classe intermedia tra le due e cioè le **iterazioni razionali dipendenti da un parametro**:

$$X_{k+1} = \psi(X_k, A),$$

dove $\psi(z, t)$ è una funzione razionale delle due variabili e $X_0 = p(A)$.

- Questa classe comprende le iterazioni razionali matriciali;

Classe studiata

Si studia una classe intermedia tra le due e cioè le **iterazioni razionali dipendenti da un parametro**:

$$X_{k+1} = \psi(X_k, A),$$

dove $\psi(z, t)$ è una funzione razionale delle due variabili e $X_0 = p(A)$.

- Questa classe comprende le iterazioni razionali matriciali;
- Per ogni k, $X_k = s_k(A)$, con $s_k(z)$ funzione razionale;

Classe studiata

Si studia una classe intermedia tra le due e cioè le **iterazioni razionali dipendenti da un parametro**:

$$X_{k+1} = \psi(X_k, A),$$

dove $\psi(z, t)$ è una funzione razionale delle due variabili e $X_0 = p(A)$.

- Questa classe comprende le iterazioni razionali matriciali;
- Per ogni k, $X_k = s_k(A)$, con $s_k(z)$ funzione razionale;
- Tutti gli elementi commutano;

Classe studiata

Si studia una classe intermedia tra le due e cioè le **iterazioni razionali dipendenti da un parametro**:

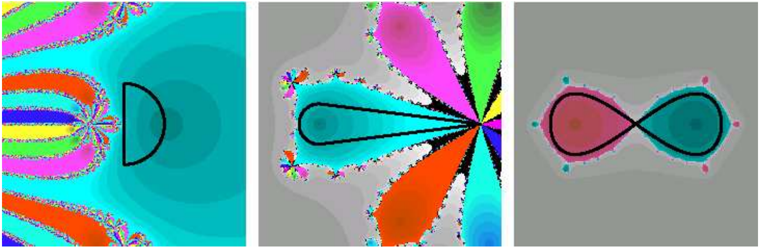
$$X_{k+1} = \psi(X_k, A),$$

dove $\psi(z, t)$ è una funzione razionale delle due variabili e $X_0 = p(A)$.

- Questa classe comprende le iterazioni razionali matriciali;
- Per ogni k, $X_k = s_k(A)$, con $s_k(z)$ funzione razionale;
- Tutti gli elementi commutano;
- Tutte le iterate si trasformano contemporaneamente in blocchi di Jordan tramite la stessa similitudine.

Problemi: convergenza scalare

Problemi insiti nell'iterazione razionale scalare.



[I., SIMAX '06; Guo, Higham, SIMAX '06; Kenney, Laub, SIMAX '91]

Trovare regioni di convergenza può essere *logorante*.

Problemi: convergenza matriciale

Se

$$MAM^{-1} = \mathrm{diag}(J_1, \ldots, J_r),$$

allora

$$Ms_k(A)M^{-1} = \mathrm{diag}(s_k(J_1), \ldots s_k(J_r)).$$

Problemi: convergenza matriciale

Se

$$MAM^{-1} = \mathrm{diag}(J_1, \ldots, J_r),$$

allora

$$Ms_k(A)M^{-1} = \mathrm{diag}(s_k(J_1), \ldots s_k(J_r)).$$

Ma la convergenza sugli autovalori **non implica** la convergenza sui blocchi, per $x_{k+1} = x_k^2$,

$$X_0 = \begin{bmatrix} 1 & 1 \\ 0 & 1 \end{bmatrix}, \qquad X_k = \begin{bmatrix} 1 & k+1 \\ 0 & 1 \end{bmatrix}.$$

Problemi: convergenza matriciale

Teorema (Higham, 2006)

Se

$$\left|\frac{\partial\psi}{\partial z}(z,t)|_{(f(\lambda),\lambda)}\right| = 0,$$

allora l'iterazione matriciale converge.

- Non comprende iterazioni a convergenza lineare

Problemi: convergenza matriciale

Teorema (Higham, 2006)

Se

$$\left| \frac{\partial \psi}{\partial z}(z,t)|_{(f(\lambda),\lambda)} \right| = 0,$$

allora l'iterazione matriciale converge.

- Non comprende iterazioni a convergenza lineare
- Non viene detto qual è limite matriciale

Un teorema di convergenza

Teorema (I., 2007)

Se

$$\left|\frac{\partial\psi}{\partial z}(z,t)|_{(f(\lambda),\lambda)}\right| < 1$$

allora la successione scalare converge uniformemente a f e la successione matriciale converge a $f(A)$.

Un teorema di convergenza

Teorema (I., 2007)

Se
$$\left|\frac{\partial\psi}{\partial z}(z,t)|_{(f(\lambda),\lambda)}\right| < 1$$
allora la successione scalare converge uniformemente a f e la successione matriciale converge a $f(A)$.

- Convergenza degli autovalori a punti fissi stabili implica convergenza matriciale;

Un teorema di convergenza

Teorema (I., 2007)

Se

$$\left|\frac{\partial \psi}{\partial z}(z,t)|_{(f(\lambda),\lambda)}\right| < 1$$

allora la successione scalare converge uniformemente a f e la successione matriciale converge a $f(A)$.

- Convergenza degli autovalori a punti fissi stabili implica convergenza matriciale;
- Similitudine tra le due convergenze, $\|s_k(J) - f(J)\|_\infty \leqslant \gamma \|s_k(z) - f(z)\|_K$;

Un teorema di convergenza

Teorema (I., 2007)

Se

$$\left|\frac{\partial\psi}{\partial z}(z,t)|_{(f(\lambda),\lambda)}\right| < 1$$

allora la successione scalare converge uniformemente a f e la successione matriciale converge a $f(A)$.

- Convergenza degli autovalori a punti fissi stabili implica convergenza matriciale;
- Similitudine tra le due convergenze, $\|s_k(J) - f(J)\|_\infty \leqslant \gamma \|s_k(z) - f(z)\|_K$;
- Nessuno strumento di dinamica olomorfa nella dimostrazione.

Corollario

Corollario

Se per tutti gli autovalori l'iterazione converge a punti fissi stabili, allora **il limite di una iterazione razionale matriciale è diagonalizzabile**.

Corollario

Corollario

Se per tutti gli autovalori l'iterazione converge a punti fissi stabili, allora **il limite di una iterazione razionale matriciale è diagonalizzabile**.

- Se una funzione di matrice è limite di iterazioni razionali allora è **sempre** diagonalizzabile (es. funzione matrice segno, matrice settore);

Corollario

Corollario

Se per tutti gli autovalori l'iterazione converge a punti fissi stabili, allora **il limite di una iterazione razionale matriciale è diagonalizzabile**.

- Se una funzione di matrice è limite di iterazioni razionali allora è **sempre** diagonalizzabile (es. funzione matrice segno, matrice settore);
- Se una funzione di matrice **non è sempre** diagonalizzabile allora non esiste alcuna iterazione razionale che converga ad essa (es. radici di matrici).

Problemi: stabilità

Come cambia la convergenza in presenza di perturbazioni?

Problemi: stabilità

Come cambia la convergenza in presenza di perturbazioni?

In generale: convergenza locale $\Longrightarrow$ stabilità.

Problemi: stabilità

Come cambia la convergenza in presenza di perturbazioni?

In generale: convergenza locale $\Longrightarrow$ stabilità.

Ma: convergenza locale scalare $\not\Longrightarrow$ convergenza locale matriciale!

Problemi: stabilità

Come cambia la convergenza in presenza di perturbazioni?

In generale: convergenza locale $\Longrightarrow$ stabilità.

Ma: convergenza locale scalare $\not\Longrightarrow$ convergenza locale matriciale!

Le soluzioni spesso non sono isolate.

Differenziale o derivata di Fréchet

Al posto della convergenza locale si può richiedere la condizione.

Definizione (Cheng, Higham, Kenney, Laub, 2001)

Un'iterazione $X_{k+1} = \psi(X_k)$ è **stabile nell'intorno di un punto fisso** X se il differenziale $d\psi$ nel punto X ha potenze limitate.

Se X_k è *vicino* a X e l'errore è E_k,

$$X_{k+1} - X = d\psi_X[E_k] + o(\|E_k\|).$$

Differenziale o derivata di Fréchet

Al posto della convergenza locale si può richiedere la condizione.

Definizione (Cheng, Higham, Kenney, Laub, 2001)

Un'iterazione $X_{k+1} = \psi(X_k)$ è **stabile nell'intorno di un punto fisso** X se il differenziale $d\psi$ nel punto X ha potenze limitate.

Se X_k è *vicino* a X e l'errore è E_k,

$$X_{k+1} - X = d\psi_X[E_k] + o(\|E_k\|).$$

Altri possibili approcci: **usare una struttura** in cui ci sia convergenza locale o **lavorare solo a livello matriciale**.

Esempio: radice P-esima di una matrice

Per risolvere l'equazione $X^p = A$, il metodo

$$X_{k+1} = \frac{(p-1)X_k + AX^{1-p}}{p}, \quad X_0 = I,$$

non è stabile, ma la variante

$$\begin{cases} X_{k+1} = X_k \left(\frac{(p-1)I+M_k}{p}\right), & X_0 = I, \\ M_{k+1} = \left(\frac{(p-1)I+M_k}{p}\right)^{-p} M_k, & M_0 = A. \end{cases}$$

è stabile [I., SIMAX, 2006]

Gruppi di automorfismi

Su $\mathbb{K} = \mathbb{C}$ o $\mathbb{R}$, una matrice M non singolare fornisce un prodotto scalare

$$\langle x, y \rangle_{\mathbf{M}} = \begin{cases} x^T \mathbf{M} y & \text{per forme bilineari;} \\ x^H \mathbf{M} y & \text{per forme sesquilineari.} \end{cases}$$

Il gruppo di automorfismi associato è

$$\mathbb{G} = \{A \in \mathbb{K}^{n \times n} : \langle Ax, Ay \rangle_M = \langle x, y \rangle_M, \forall x, y \in \mathbb{K}^n\}.$$

Esempi: matrici ortogonali, unitarie, simplettiche, perplettiche. . .

Un teorema

Teorema (Higham, Mackey, Mackey, Tisseur, 05)

Sia $\mathbb{G}$ un gruppo di automorfismi associato a un prodotto scalare come sopra. Una funzione f preserva la struttura se e solo se può essere espressa come

$$\pm z^k \frac{p(z)}{revp(z)}$$

con p reale, k intero positivo.

Un'iterazione $x_{k+1} = f(x_k)$ con f come sopra preserva la struttura di $\mathbb{G}$.

Principal Padé iterations

Dalla tabella di Padé di

$$h(\xi) = (1 - x^2)^{-1/2}$$

si generano iterazioni razionali di ordine r per matrice segno e radice quadrata.

Teorema (Kenney, Laub, 1992, Higham, Mackey, Mackey, Tisseur, 2006)

Le iterazioni di ordine dispari costruite a partire dalla diagonale della tabella di Padé convergono per matrici che non hanno autovalori immaginari puri e preservano la struttura di $\mathbb{G}$.

Iterazioni basate sulla tabella di Padè

$$r = 2 \qquad \frac{2x}{1+x^2}$$

$$r = 3 \qquad \frac{x(3+x^2)}{1+3x^2}$$

$$r = 4 \qquad \frac{4x(1+x^2)}{1+6x^2+x^4}$$

$$r = 5 \qquad \frac{x(5+10x^2+x^4)}{1+10x^2+5x^4}$$

La famiglia di König

La famiglia di König's per gli zeri di f

$$K_{f,\sigma} = \mathrm{Id} + (\sigma - 1)\frac{(1/f)^{[\sigma-2]}}{(1/f)^{[\sigma-1]}},$$

dove $(1/f)^{[k]}$ è la k-esima derivata di $1/f$.
La convergenza alle radici semplici di f è di ordine σ.

- $K_{f,2}$ è il metodo di Newton;
- $K_{f,3}$ è il metodo di Halley.

La famiglia di König è dovuta a Schröder (1870)

La famiglia di König per $x^2 - 1$

$$r = 2 \qquad \frac{1+x^2}{2x}$$

$$r = 3 \qquad \frac{x(3+x^2)}{1+3x^2}$$

$$r = 4 \qquad \frac{1+6x^2+x^4}{4x(1+x^2)}$$

$$r = 5 \qquad \frac{x(5+10x^2+x^4)}{1+10x^2+5x^4}$$

Se scopri un nuovo metodo iterativo
per la soluzione di equazioni non lineari,
probabilmente è dovuto a Schröder
— **Householder** [?]

Legami con Padè e strutture

Teorema (I., 2007)

Il metodo di König per il polinomio $x^2 - 1$ coincide con il metodo basato sulla tabella di Padè o con il suo reciproco.

Teorema (I., 2007)

Se $\sigma \equiv 3 \pmod{p}$, il metodo di König di ordine σ per il polinomio $x^p - 1$ ha la forma

$$z \frac{p(z)}{\mathit{rev}\, p(z)},$$

cioè preserva la struttura di gruppo.

Il metodo di Halley

Il metodo con $\sigma = 3$ (metodo di Halley) preserva la struttura per ogni p.

Teorema (I., 2007)

L'iterazione scalare data dal metodo di Halley converge per ogni valore iniziale con parte reale positiva.

A partire dal metodo di Halley si trova un'iterazione stabile e che preserva la struttura di gruppo per calcolare radici di matrici.

- C. Kenney, A.J. Laub, *Rational iterative methods for the matrix sign function*, SIMAX, 1991.
- N. Higham, S. Mackey, N. Mackey, F. Tisseur, *Functions preserving matrix groups and iterations for the matrix square root*, SIMAX, 2005.
- B. Iannazzo, *On the Newton method for the matrix pth root*, SIMAX, 2006.
- N. Higham and C.-H. Guo, *A Schur-Newton method for the matrix pth root and its inverse*, SIMAX, 2006.
- N. Higham, *Function of a matrix: theory and computation*, libro in preparazione.