

# Calcolare medie geometriche di matrici

Bruno **Iannazzo**, Università di **Perugia**  
con Dario **Bini**, Beatrice **Meini** e Federico **Poloni**

GALN 09, 16–17 febbraio 2009

## Medie di matrici

Media aritmetica:

$$\frac{A + B}{2}.$$

Media armonica:

$$2(A^{-1} + B^{-1})^{-1}.$$

Media geometrica di matrici hermitiane definite positive (positive):

$$(AB)^{1/2},$$

**non è sempre definita, inoltre  $(AB)^{1/2} \neq (BA)^{1/2}$ .**

## Media geometrica di due matrici positive

Abbastanza presto si è convenuto sulla seguente definizione

$$A\#B := A(A^{-1}B)^{1/2},$$

(leggi “*A sharp B*”), dove la radice è intesa nel senso delle funzioni di matrici [Pusz-Woronowicz, Trapp, Ando, anni '70].

1. Perché questo oggetto non cambia se **si scambiano**  $A$  e  $B$ ?

$$A(A^{-1}B)^{1/2} = A(A^{-1}B)(A^{-1}B)^{-1/2} = B(B^{-1}A)^{1/2}$$

2. Perché questo oggetto è **positivo**?

# Una dimostrazione

## Teorema

Se  $f : \mathbb{R}_+ \rightarrow \mathbb{R}_+$  e  $A$  e  $B$  sono positive allora  $Af(A^{-1}B)$  è positiva.

## Dimostrazione.

$$\begin{aligned} Af(A^{-1}B) &= A^{1/2}A^{1/2}f(A^{-1/2}A^{-1/2}BA^{-1/2}A^{1/2}) \\ &= A^{1/2}f(A^{-1/2}BA^{-1/2})A^{1/2} \end{aligned}$$

Si è usata la proprietà  $f(M^{-1}AM) = M^{-1}f(A)M$ . □

Anche  $A \exp(A^{-1}B)$  è positiva, mentre  $A \log(A^{-1}B)$  è hermitiana.

## Proprietà della media geometrica

- ▶ Unica soluzione positiva di  $XA^{-1}X = B$  (eq. di Riccati);
- ▶ Massima  $X$  tale che  $\begin{bmatrix} A & X \\ X & B \end{bmatrix} \geq 0$ ;
- ▶ Parentela con funzioni di matrici, di cui ha parlato Beatrice Meini nel precedente talk;
- ▶ Punto medio di una geodetica.

## Applicazioni numeriche della media tra due matrici

Come nel caso delle funzioni di matrici i problemi sono:

- ▶ Calcolare  $A\#B$  con  $A$  e  $B$  dense e di dimensioni moderate (reti elettriche, MRI, base per calcolo di medie tra più matrici) [Trapp, Arsigny et al., Moakher];
- ▶ Calcolare  $(A\#B)v$  con  $A$  e  $B$  sparse e di grandi dimensioni e  $v$  vettore (elementi finiti).

In generale può essere interessante calcolare  $A(A^{-1}B)^t v$ , per  $t \in [-1, 1]$  [Arioli-Loghin].

## Interpretazione geometrica

Dotiamo  $\mathbb{P}_n$  (matrici positive) di struttura di varietà Riemanniana<sup>1</sup> con prodotto scalare in  $T_A$ ,

$$g(M_1, M_2) = \langle A^{-1}M_1, A^{-1}M_2 \rangle.$$

Il punto finale della geodetica che parte da  $A$  e tangente a  $V$  è  $\exp_A(V) = A \exp(A^{-1}V)$ .

Esiste una sola geodetica che unisce  $A$  e  $B$  di parametrizzazione

$$\gamma(t) = \exp_A(t \log_A(B)) = A(A^{-1}B)^t.$$

**La media geometrica tra due matrici è il punto medio della geodetica che le unisce.**

---

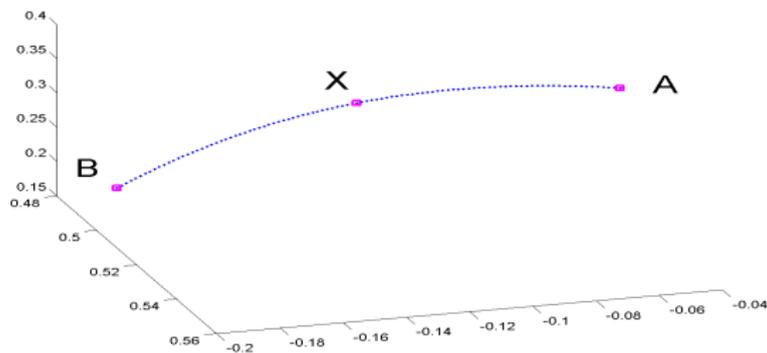
<sup>1</sup>con curvatura (sezionale) costante negativa

## Citazione

*E così riempiamo di significato geometrico qualcosa che di geometrico aveva solo il nome.*

Elena Addis

## Breve dimostrazione



Se  $X$  è il punto medio della geodetica da  $A$  a  $B$ , allora

$$B = X(X^{-1}A)^{-1} = XA^{-1}X.$$

## Algoritmo 1. Scaled averaging technique

**Calcolando ripetutamente la media aritmetica e quella armonica si converge a quella geometrica.**

Si ottiene l'iterazione [Anderson, Morley e Trapp]:

$$\begin{cases} A_0 = A, & B_0 = B, \\ A_{k+1} = (A_k + B_k)/2, \\ B_{k+1} = 2(A_k : B_k), \end{cases}$$

dove  $A_{k+1}$  e  $B_{k+1}$  sono, rispettivamente, la media aritmetica e armonica di  $A_k$  e  $B_k$  ( $(A_k : B_k) := A_k(A_k + B_k)^{-1}B_k$ ).

## Algoritmo 1. Scaled averaging technique, varianti

La stessa iterazione è stata ottenuta anche per altra via

1. Matrice segno [Higham et al.]

$$\text{sign} \left( \begin{bmatrix} 0 & A \\ B^{-1} & 0 \end{bmatrix} \right) = \begin{bmatrix} 0 & A\#B \\ (A\#B)^{-1} & 0 \end{bmatrix}.$$

2. Riduzione ciclica palindromica [Iannazzo-Meini]

$$\begin{cases} P_0 = \frac{1}{4}(B - A), & Q_0 = \frac{1}{2}(B + A), \\ P_{k+1} = -P_k Q_k^{-1} P_k, \\ Q_{k+1} = Q_k - 2P_k Q_k^{-1} P_k, \end{cases}$$

$$\text{e } \lim Q_k = A(A^{-1}B)^{1/2}.$$

## Algoritmo 1. Scaled averaging technique, implementazione

L'iterazione si può implementare in modo efficiente.

### Algoritmo (Scaled averaging iteration)

$$\begin{cases} X_0 = A^{-1}, & Y_0 = B, \\ \gamma_k = |\det(X_k Y_k)|^{-1/(2n)} \\ X_{k+1} = (\gamma_k X_k + (\gamma_k Y_k)^{-1})/2, \\ Y_{k+1} = (\gamma_k Y_k + (\gamma_k X_k)^{-1})/2, \end{cases} \quad k = 1, 2, \dots$$

Il costo è di circa  $4n^3$  per passo e la convergenza avviene in pochi passi.

## Algoritmo 2. Decomposizione polare

### Algoritmo (Decomposizione polare)

*Date  $A$  e  $B$  positive.*

- 1. Calcola la fattorizzazione di Cholesky di  $A^{-1} = L_{A^{-1}}^* L_{A^{-1}}$ ;*
- 2. Calcola il fattore unitario polare  $U$  di  $L_B L_{A^{-1}}^*$ ;*
- 3.  $A \# B = L_{A^{-1}}^{-1} U^* L_B$ .*

Il costo totale è  $64/3n^3$ .

È una generalizzazione dell'iterazione di Higham per la radice quadrata di una matrice positiva.

## Algoritmo 2. Decomposizione polare

**Con grossa sorpresa** se il fattore polare è calcolato come limite di  $Z_{k+1} = \frac{1}{2}(Z_k + Z_k^{-*})$ , si ha:

### Teorema

*Per la sequenza  $A_k$  generata dall'averaging technique e  $Z_k$  come sopra vale  $A_k = L_{A^{-1}}^{-1} Z_k^* L_B$*

Conseguenza: l'algoritmo di Higham per la radice quadrata di una matrice positiva può essere visto come una variante del metodo di Newton!

## Algorithm 3. Quadratura gaussiana

Un'utile identità integrale

$$\int_0^1 \frac{t^{\alpha-1}(1-t)^{\beta-1}}{(ta^{-1} + (1-t)b^{-1})^{\alpha+\beta}} dt = \frac{\Gamma(\alpha)\Gamma(\beta)}{\Gamma(\alpha+\beta)} a^\alpha b^\beta.$$

permette di calcolare  $A^\alpha$ , la matrice segno, il fattore polare e la media di due matrici per varie scelte di  $\alpha, \beta, a, b$ .

Questa formula si ottiene anche trasformando al formula di Cauchy

$$f(A) = \frac{1}{2\pi i} \oint f(z)(zI - A)^{-1} dz$$

già usata nell'algoritmo “triple Nick” [Hale, Higham, Trefethen].

## Algorithm 3. Quadratura gaussiana

Il precedente integrale con  $\alpha = \beta = 1/2$  e  $a = A$ ,  $b = B$ , ci dà

$$A\#B = \frac{2}{\pi} \int_{-1}^1 \frac{((1+t)B^{-1} + (1-t)A^{-1})^{-1}}{\sqrt{1-t^2}} dt, \quad (1)$$

### Algoritmo

*Date A e B positive. Calcola l'integrale (1) usando (qualche schema adattivo per) la quadratura di Gauss-Chebyshev.*

$$T_{n+1} = \frac{2}{\pi} S_{n+1} = \frac{2}{n+1} \sum_{k=0}^n ((1+x_k)B^{-1} + (1-x_k)A^{-1})^{-1}$$

*ottenendo  $A\#B = \lim_{n \rightarrow \infty} T_{n+1}$ .*

**Adatto a calcolare  $(A\#B)v$ .**

## Algoritmo 3. Quadratura gaussiana

Con ancora più grossa sorpresa

### Teorema

Se  $T_k$  è data dalla formula di quadratura e  $B_k$  è la sequenza generata dall'averaging technique allora

$$T_{2^k} = B_k.$$

## Algoritmo 4. Formula dei trapezi

$A\#B$  è il termine centrale dell'inverso del polinomio di Laurent

$$\mathcal{H}(z) = \frac{1}{4}(A - B)z + \frac{1}{2}(A + B) + \frac{1}{4}(A - B)z^{-1},$$

che è dato dall'integrale

$$\frac{1}{2\pi} \int_0^{2\pi} \mathcal{H}(e^{it}) dt,$$

che può essere calcolato con la formula dei trapezi ( $\mathcal{H}$  è regolare e periodica).

## Algoritmo 4. Formula dei trapezi

La formula integrale di Cauchy tramite sostituzioni fornisce

$$A\#B = \frac{2}{\pi} \int_0^{\pi/2} (\cos^2(2t)A^{-1} + \sin^2(2t)B^{-1})^{-1} dt$$

La formula dei trapezi vista sopra fornisce

$$A\#B = \frac{2}{\pi} \int_0^{\pi/2} (\sin^2(t)A^{-1} + \cos^2(t)B^{-1})^{-1} dt$$

**Nonostante siano espressioni molto simili questa volta gli algoritmi risultanti sono differenti!**

## Medie di più matrici

Cos'è la media di 3 matrici?

La risposta sembra molto difficile: infatti nonostante l'insieme dato dalle tre geodetiche

$$S_0 = \{A\#_t B, A\#_t C, B\#_t C, t \in [0, 1]\}$$

sia una curva, l'insieme

$$S_1 = \bigcup_{M, N \in S_0} M\#_t N$$

non è una superficie, ma qualcosa di **“più grasso”**.

# Medie di più matrici

Due possibili approcci:

- ▶ Assiomatico/modellistico: si danno alcune proprietà che la media deve verificare;
- ▶ Geometrico: si cerca il centro di massa (che è unico secondo Cartan).

Problema numerico: **scegliere una media adatta al modello e calcolarla in modo efficiente.**

## Medie di più matrici

Alcune proprietà desiderate, se  $G(A, B, C)$  è la media:

- ▶ Coerenza con gli scalari: se  $A, B, C$  commutano,  $G(A, B, C) = (ABC)^{1/3}$ ;
- ▶ Invarianza per permutazioni;
- ▶ Continuità;
- ▶ Invarianza per congruenza:  
 $G(S^*AS, S^*BS, S^*CS) = S^*G(A, B, C)S$
- ▶ Omogeneità:  $G(\alpha A, B, C) = \alpha^{1/3}G(A, B, C)$ .

Esistono infinite medie che verificano questi e altri assiomi detti ALM (Ando-Li-Mathias).

## Medie di più matrici

Procedimenti iterativi: a partire da  $A_0 = A$ ,  $B_0 = B$ ,  $C_0 = C$ , si costruisce una sequenza che converge a un limite che verifica gli assiomi.

- ▶ Ando-Li-Mathias:  $A_{k+1} = B_k \# C_k$ ,  
Convergenza lineare.
- ▶ Bini-Meini-Poloni:  $A_{k+1} = A_k \#_{2/3}(B_k \# C_k)$ .  
Convergenza cubica.

Ne parlerà Federico Poloni nel prossimo talk.

# Problemi

Due interessanti problemi:

1. Problema matematico: **descrivere l'insieme delle matrici che verificano gli assiomi ALM**;
2. Problema modellistico: le medie che verificano gli assiomi ALM sono “vicine”; **dare una palla che le racchiude tutte**. Se  $G$  verifica gli assiomi e  $A$  e  $H$  sono le medie armonica e geometrica

$$\|G - (A + H)/2\| \leq \|(A - H)/2\|_2,$$

questa stima è più precisa dell'errore sui dati sperimentali nei problemi di elasticità.

## Media di Cartan

È l'**unica** matrice  $X$  che minimizza la quantità

$$\delta^2(X, A) + \delta^2(X, B) + \delta^2(X, C),$$

ovvero risolve l'equazione matriciale

$$\log(X^{-1}A) + \log(X^{-1}B) + \log(X^{-1}C) = 0,$$

cioè è l'**unico** punto  $X$  tale che sia nulla la media aritmetica dei vettori tangenti a tutte le geodetiche che uniscono  $X$  ad  $A$ ,  $B$  e  $C$ .

# Media di Cartan: work in progress

## Problemi

- ▶ Dimostrare che verifica tutti gli assiomi ALM o viceversa trovare un controesempio;
- ▶ Calcolarla numericamente.  
Non esistono algoritmi generali per il calcolo.