Una media di matrici efficiente che soddisfa le proprietà di Ando–Li–Mathias

D. A. Bini[1] B. Iannazzo[2] B. Meini[1] Federico Poloni[3]

[1]Dipartimento di matematica, Università di Pisa
[2]Università di Perugia
[3]Scuola Normale Superiore, Pisa

Perugia, 16 Febbraio 2009

Il problema fisico

Esperimenti sull'elasticità [Hearmon, 1952; Moakher, 2006]:
Diverse misure sperimentali del tensore di rigidità (stiffness tensor) e del suo inverso (compliance tensor)

Problema

Come farne una media?

Richiesta: fare la media tra gli inversi (*compliance*) dovrebbe essere la stessa cosa che fare la media tra i dati originali (*stiffness*) e poi invertire:

$$M(A, B, C, \dots)^{-1} = M(A^{-1}, B^{-1}, C^{-1}, \dots)$$

Media geometrica

Tra le medie classiche, l'unica che fa questo lavoro è la media geometrica

Definizione

Media geometrica di k reali positivi:

$$GM(a_1, \dots, a_k) = (a_1 \dots a_k)^{1/k}$$

Esempio

Una scatola ha dimensioni $10 \times 15 \times 20$ cm; quanto dev'essere il lato di una scatola cubica che ha lo stesso volume?

Esempio

Su una somma investita, ottengo in tre anni consecutivi un tasso di interesse (composto) dell'1%, 1.5%, 2%. Qual è l'interesse medio?

Il problema matematico

Parallelamente, studiato come problema di algebra lineare [Ando–Li–Mathias, 2003; Bhatia, 2005 e altri]

Problema

Come generalizzare a matrici SPD il concetto di media geometrica di numeri reali positivi?

Approccio "da matematico": cominciamo a definire un po' di assiomi che dovrebbe soddisfare. . .

Cosa ci aspettiamo da una media (geometrica)?

[Ando–Li–Mathias, 2003]: dieci proprietà che ogni "buona" media geometrica dovrebbe soddisfare

- compatibilità con gli scalari: se A, B, C commutano, $GM(A, B, C) = (ABC)^{1/3}$
- simmetria: $GM(A, B, C) = GM(B, A, C) = \dots$
- monotonia: $A < A' \Rightarrow GM(A, B, C) < GM(A', B, C)$
- invarianza per congruenza: $GM(S^*AS, S^*BS, S^*CS) = S^*GM(A, B, C)S$
- invarianza per inversione: $GM(A^{-1}, B^{-1}, C^{-1}) = GM(A, B, C)^{-1}$

...e altro (concavità, continuità...)

Osservazione

Questi assiomi non determinano univocamente una funzione $GM(\cdot)$!

Media di due matrici

Problema

Come definire la media di *due* matrici?

- $GM(A,B) = (AB)^{1/2}$, oppure $GM(A,B) = A^{1/2}B^{1/2}$ (diverse!) non funzionano: non simmetriche
- $GM(A,B) = \exp\left(\frac{\log A + \log B}{2}\right)$ neppure: non monotona

...sorprendentemente, funziona $GM(A,B) = A\left(A^{-1}B\right)^{1/2}$

Osservazione (Ando–Li–Mathias)

Compatibilità con gli scalari e invarianza per congruenza determinano univocamente la *GM* di due matrici

Geometria Riemanniana su matrici

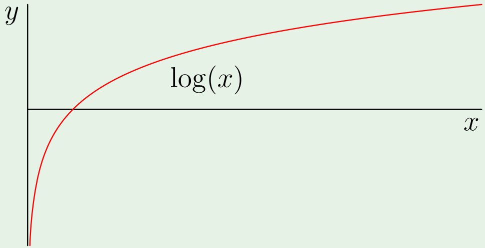
Sulle matrici SPD, geometria Riemanniana determinata dalla forma metrica

$$g_A(X,Y) = \langle A^{-1}X, A^{-1}Y \rangle_{\text{Frobenius}} \qquad X, Y \in \text{Tangent space}_A$$

È una metrica che "si curva" sempre di più quando M si avvicina alla singolarità

Esempio

In dimensione 1, è la consueta scala logaritmica sui numeri positivi



Medie e geometria

Teorema

La geodetica che congiunge X e Y è

$$\gamma(t) = A \#_t B := A \left(A^{-1}B\right)^t, \quad t \in [0,1]$$

Media geometrica tra due matrici = punto medio della geodetica

Problema

Come definire la media di tre o più matrici?

Non c'è una soluzione semplice

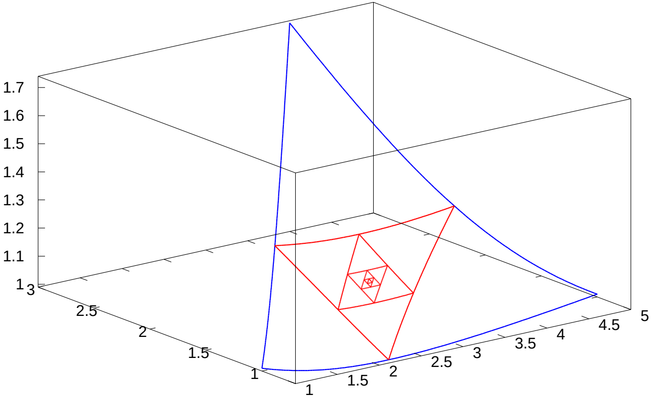
Le proprietà ALM non determinano univocamente una media

Proposta [Ando–Li–Mathias, 2003]: simmetrizzazione

$$
\begin{aligned}
A_1 &= GM(B, C) & \qquad A_2 &= GM(B_1, C_1) & \\
B_1 &= GM(C, A) & \qquad B_2 &= GM(C_1, A_1) & \qquad \dots \\
C_1 &= GM(A, B) & \qquad C_2 &= GM(A_1, B_1) &
\end{aligned}
$$

Converge a una matrice, la definiamo media geometrica (GM_{ALM})

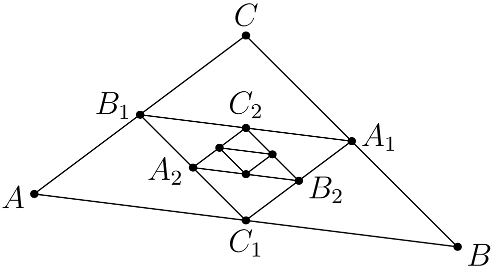
Costruzione ALM



Cosa c'è sotto?

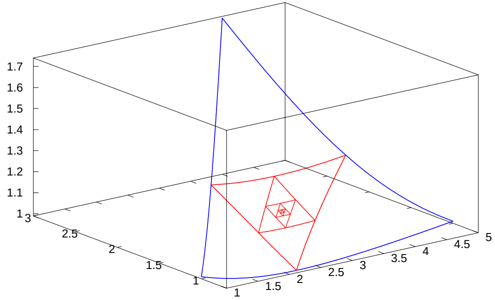
Teorema

La stessa costruzione fatta nel piano converge al baricentro del triangolo ABC.



Proprietà della media ALM

- soddisfa le dieci proprietà
- generalizzabile a $k \geq 4$ matrici: $A_1 = GM(B, C, \dots)$
- calcolabile direttamente dalla definizione
- problema: calcolo laborioso (soprattutto se k cresce)

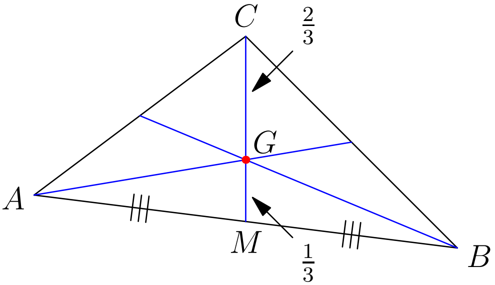


Proprietà del baricentro

Mediane del triangolo ABC: segmento che congiunge A al punto medio di BC, e ciclici

Teorema

Le tre mediane di un triangolo si incontrano (a 2/3 della loro lunghezza) nel baricentro

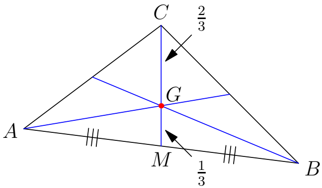


Proprietà del baricentro

Teorema

Le tre mediane di un triangolo si incontrano (a 2/3 della loro lunghezza) nel baricentro

Dimostrazione: sorprendentemente semplice con geometria analitica. Sulla mediana CM, costruiamo un "candidato G" che stia alla lunghezza giusta della mediana...



$$x_M = \frac{1}{2}(x_A + x_B)$$

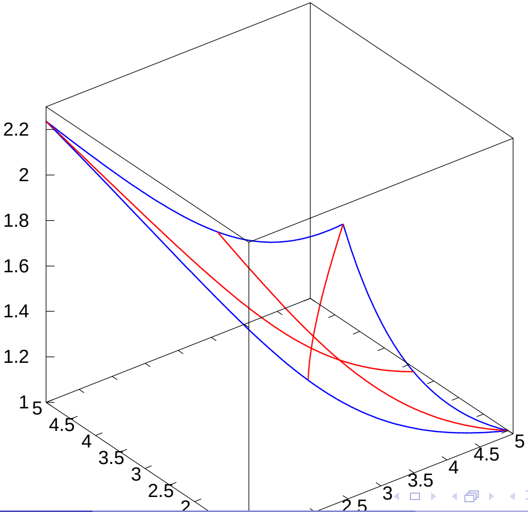
$$x_G = \frac{2}{3}x_M + \frac{1}{3}x_C = \frac{1}{3}(x_A + x_B + x_C)$$

Simmetrica in A, B, C: se rifaccio la costruzione su un altro lato, finisco nello stesso punto. □

Una nuova media

Vorrei fare la stessa costruzione nelle matrici SPD...

Problema: le mediane non si incontrano!



Soluzione

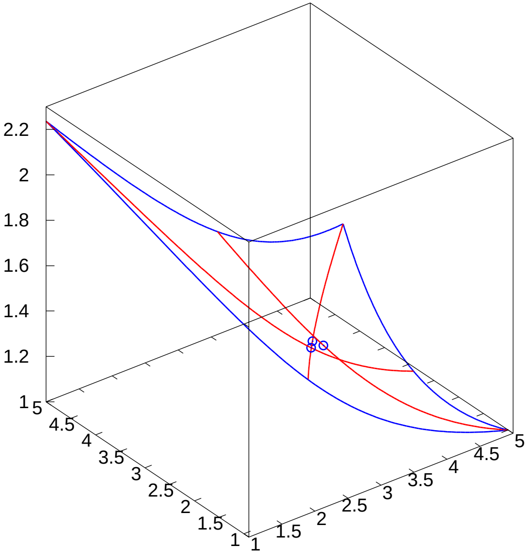
Definiamo A_1, B_1, C_1 = i punti a $2/3$ delle mediane di ABC
(diversi tra loro in generale)

A_2, B_2, C_2 definiti nello stesso modo a partire da $A_1 B_1 C_1$ e così via

Teorema

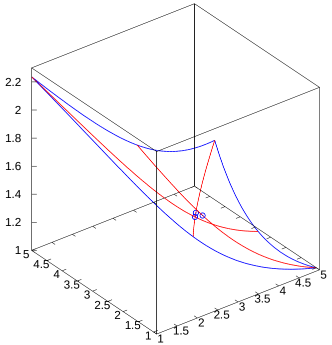
A_i, B_i, C_i convergono a uno stesso punto. Lo definiamo media di A, B, C

Una nuova media



Proprietà della nuova media

- diversa in generale dalla GM_{ALM}
- soddisfa le dieci proprietà
- generalizzabile a $k \geq 4$ matrici
- calcolabile direttamente dalla definizione
- convergenza di ordine 3: molto più rapida di GM_{ALM}



Qualche numero

5
1.92542947898189
2.90969918536362
2.35774114351751
2.61639158463414
2.48316587472793
2.54876054375880
2.51571460655576
2.53217471946628
2.52392903948587
2.52804796243998
2.52598752310721
2.52701749813482
2.52650244948321
2.52675995852183
2.52663120018107
2.52669557839604
2.52666338904971
2.52667948366316
2.52667143634151

5
2.59890269690271
2.53027293208879
2.53025171828977
2.53025171828977

Esempio

$$A = \begin{bmatrix} 5 & 2 \\ 2 & 1 \end{bmatrix} \quad B = \begin{bmatrix} 4 & 3 \\ 3 & 3 \end{bmatrix} \quad C = \begin{bmatrix} 1 & 0 \\ 0 & 5 \end{bmatrix}$$

Visualizzato A_{11} al passo i
A sinistra: media ALM
A destra: nuova media

Misure di elasticità: con dati di [Hearmon, '52] (fino a 6 matrici 6×6 con alcuni blocchi nulli): velocità fino a $100\times$

Vantaggi

- numero di iterazioni molto ridotto
- quando k (numero di matrici da mediare) cresce di poco ($k = 4, 5, 6, 7$), i vantaggi si accumulano: numero di iterazioni ridotto *ad ogni passo*.
- alti valori di k ancora non fattibili: resta dipendenza come $k!$.
- c'è un'idea di B. Iannazzo per eliminare questo problema, ancora work in progress!

Altri spunti

Nuova media : $A \to A \#_{\frac{2}{3}} \left(B \#_{\frac{1}{2}} C \right)$

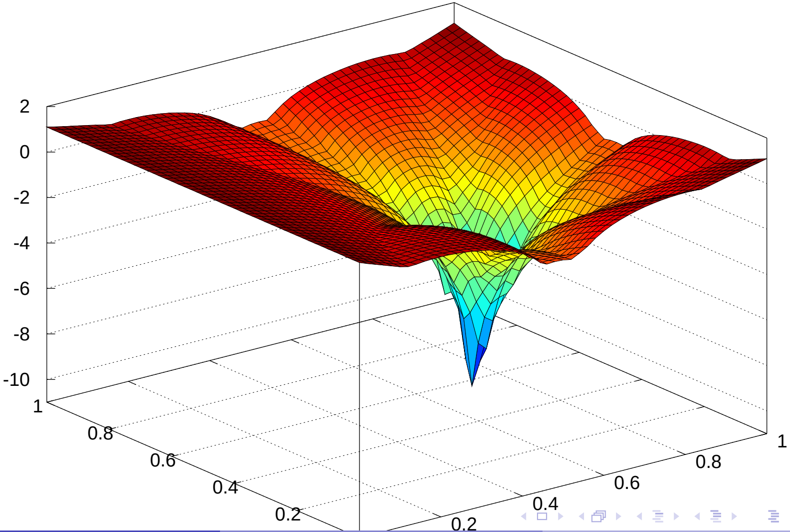
Media ALM : $A \to A \#_{1} \left(B \#_{\frac{1}{2}} C \right)$

Generalizzo con due parametri s, t:

$$A \to A \#_{s} (B \#_{t} C)$$

Chi sono gli s e t migliori?

Altri spunti



Altri spunti

Anche la media di Cartan ha un'interpretazione in geometria euclidea:

Teorema

Il baricentro è il punto che minimizza la somma delle tre distanze dai vertici:

$$G = \operatorname{argmin}_P(AP^2 + BP^2 + CP^2)$$

Di nuovo semplice con geometria analitica:

$$\begin{aligned} AP^2 + BP^2 + CP^2 &= \\ &= (x_P - x_A)^2 + (x_P - x_B)^2 + (x_P - x_C)^2 + \dots \\ &= 3x_P^2 - 2x_P(x_A + x_B + x_C) + (x_A^2 + x_B^2 + x_C^2) + \dots \end{aligned}$$

parabola, ha minimo in $x_P = \frac{x_A + x_B + x_C}{3}$

Altri spunti

La geometria Euclidea finora è stata una miniera di idee

Un'altra proprietà che può tornare utile. . .

Teorema

Il baricentro è il punto che sta a $\frac{2}{3}$ del segmento che congiunge ortocentro e circocentro.

Chi sono ortocentro (incrocio delle altezze) e circocentro (incrocio degli assi) nel setting delle matrici?

Grazie dell'attenzione!