There is no free mean

Constructing cheaper matrix geometric means (or the impossibility thereof)

Federico Poloni[1,2]

[1]Technische Universität Berlin (A. von Humboldt postdoctoral fellow) now
[2]Scuola Normale Superiore, Pisa until end 2010

17$^{\text{th}}$ ILAS Conference
Braunschweig, 24.08.2011

What is a matrix geometric mean?

A map (Hermitian (S)PD matrices)^n $\to$ Hermitian (S)PD matrices satisfying 10 properties: [Ando, Li, Mathias '04]

Consistency with scalars if A, B, C commute, $G(A, B, C) = A^{1/3}B^{1/3}C^{1/3}$

Permutation invariance $G(A, B, C) = G(A, C, B) = G(B, A, C) = \dots$

Joint homogeneity $G(\alpha A, B, C) = \alpha^{1/3} G(A, B, C)$

Monotonicity $A \le A' \Rightarrow G(A, B, C) \le G(A', B, C)$

Congruence invariance $PG(A, B, C)P^* = G(PAP^*, PBP^*, PCP^*)$

+ others self-duality, concavity…

$n = 2$

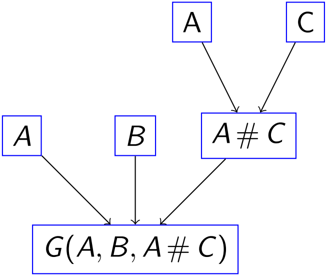
The mean of two matrices is $A \#_t B := A(A^{-1}B)^t$ for $t = 1/2$

$n \ge 3$

Many answers…

Composing matrix geometric means

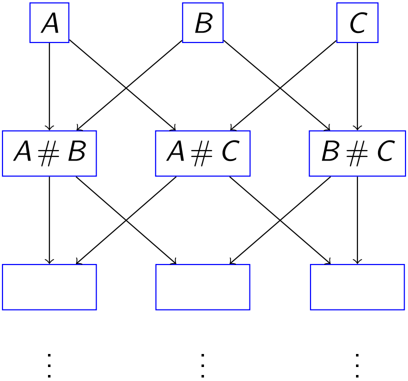
A common way to generate new functions that "behave like matrix geometric means" is composing means of less variables. . .



(G=any mean of 3 matrices)

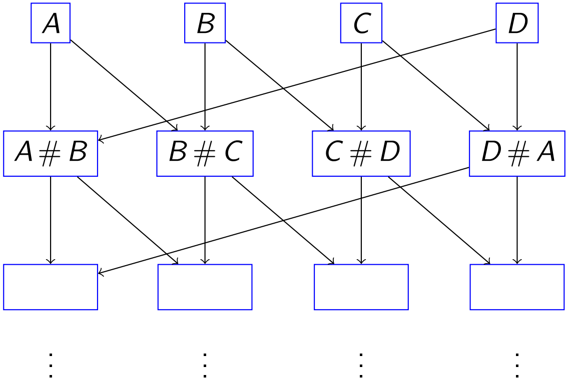
This includes many studied means:

Ando–Li–Mathias mean [Ando, Li, Mathias '04]



Repeat until convergence

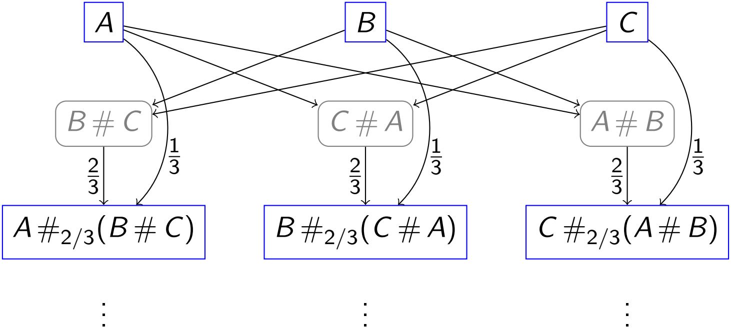
Pálfia's "graph-based" mean [Pálfia, '05 & '11]



Repeat until convergence

Does not respect permutation invariance: $G(A, B, C, D) \neq G(B, A, C, D)$ in general!

Cubically convergent mean [Nakamura, '09] [Bini, Meini, P. '10]



Repeat until convergence

Cost of existing means

ALM and BMP satisfy all ten "geometric mean" properties, but are expensive to compute:

- need a limit process
- every "iteration" requires n means of $n-1$ variables
  $\Rightarrow$ computing a mean of n matrices requires $O(n! \cdot \text{steps})$ operations

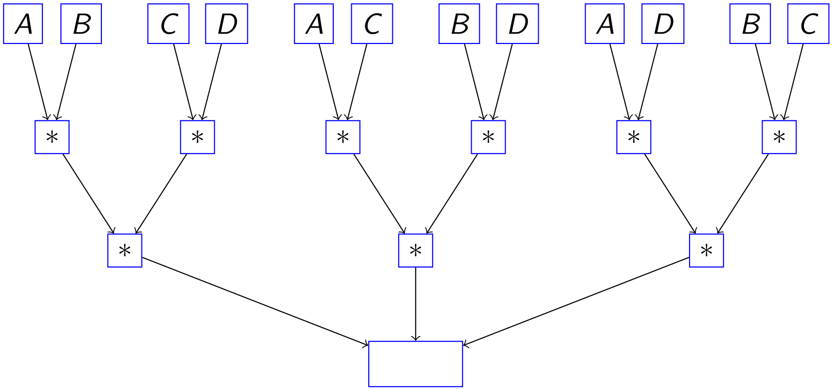
The mean proposed by Pálfia is cheaper to compute, but does not satisfy permutation invariance

Question

Are there means that are cheaper than $O(n! \cdot \text{steps})$, and satisfy all axioms?

Maybe we just need to look more carefully. . .

For instance. . .



Tournament mean: consider all 3 essentially different "tennis tournament" arrangements of 4 matrices, and average them [P., '10]

Satisfies ALM properties, significantly cheaper than existing 4-means

Composing existing means

What is easy and what is hard in assembling existing means?

- Most properties carry over easily to function composition (concavity, congruence invariance...)
- Some need to be adapted, since weights $(\frac{1}{n}, \frac{1}{n}, \dots, \frac{1}{n})$ are "hardcoded" in the ALM properties
  e.g., if $G(A, B, C) := (A \# B) \# C$, then

$$G(\alpha A, \beta B, \gamma C) = \alpha^{1/4} \beta^{1/4} \gamma^{1/2} G(A, B, C)$$

- Important exception: permutation invariance doesn't hold in general!

It is not easy to construct permutation-invariant means by composition

Quasi-means

Definition

A quasi-mean is a function satisfying a modified set of ALM properties:

- Allow for different "weights" in some axioms, e.g.

$$G(\alpha A, \beta B, \gamma C) = \alpha^{w_1} \beta^{w_2} \gamma^{w_3} G(A, B, C)$$

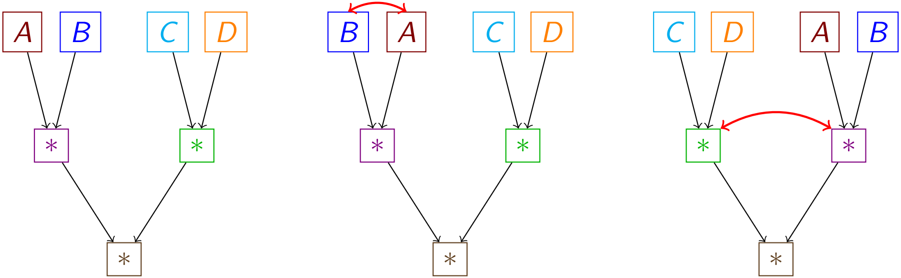
- Do not require permutation invariance

Easy to prove that:

- Composition and limit of quasi-means are quasi-means
- Quasi mean + permutation invariance = geometric mean (all 10 properties would hold)

Invariance groups

Some permutation invariance properties follow from those of the underlying means:



Definition

The invariance group of a quasi-mean is the group of permutations of its arguments that leave it unchanged

Invariance groups

Definition

The invariance group $\text{inv}(G)$ of a quasi-mean is the group of permutations of its arguments that leave it unchanged

E.g., inv(previous slide) $=$ dihedral group $\mathfrak{D}_4$ (symmetries of a square)

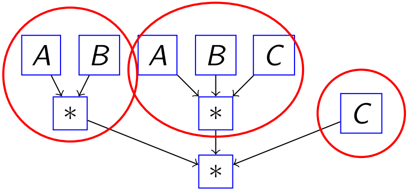
- it must be a group (closed by composition...)

What can we deduce on inv() of "composite means" based on inv() of their components?

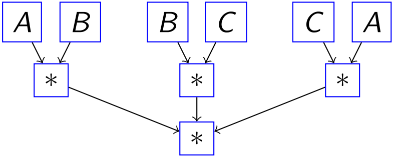
First, some reductions...

Simplifying the problem

- Only two levels



- Only compositions where the means in the "upper level" are the same, up to permutation of their arguments — they have larger inv()



- may assume it's *all* such permutations (add dummy arguments)

Some group theory

Theorem [P., '10]

Let $H = \mathrm{inv}(\text{upper means})$, $G = \mathrm{inv}(\text{lower mean})$, $\rho_H(\sigma) =$ action of σ on the left cosets $\{H\tau\}$, K the largest group s.t. $\rho_H(K) \subseteq G$. Then, $K \subseteq \mathrm{inv}(\text{composition})$

Describes all the symmetry properties deriving from the "base means"

$$+$$

Known fact

For $n \geq 5$, the group of all even permutations $\mathfrak{A}_n$ (of size $\frac{n!}{2}$) is simple

$$=$$

$$\dots$$

There ain't no free mean

Theorem

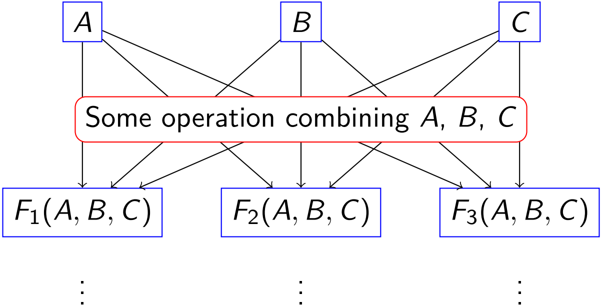
If we can prove using the theorem that $\mathfrak{A}_n \subseteq \text{inv}(\text{composition})$, then either $\mathfrak{A}_n \subseteq \text{inv}(\text{upper means})$ or $\mathfrak{A}_n \subseteq \text{inv}(\text{lower mean})$

i.e., we cannot hope to find a geometric mean by composing means of less matrices and playing with their symmetry properties

A (finite) "composition-based" matrix geometric mean isn't impossible, but would require new techniques

Ok, but all "named means" are based on a limit process; what about them?

Considering limit processes



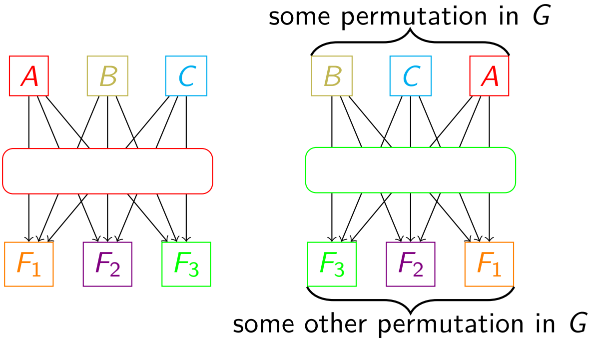
Repeat until convergence

Again, we wish to abstract all "symmetry properties" proofs in this setting

Group preservation

Definition

Let G be a permutation subgroup. The limit process preserves G if permuting the inputs according to $\sigma \in G$, we obtain the same outputs permuted according to $\tau \in G$



"No free mean" for limits

Theorem [P., '10]

If a quasi-mean limit process preserves G (and its component converge to a common limit), then $G \subseteq \text{inv}(\text{limit mean})$

Extends all the existing proofs

$$+$$

Known fact

The only "large" subgroups of the permutation group $\mathfrak{S}_n$ are $\mathfrak{A}_n$ and $\mathfrak{S}_{n-1}$

$$=$$

$$\dots$$

“No free mean” for limits

Theorem

If a limit process preserves $G = \mathfrak{S}_n$, then either

- $\operatorname{inv}(F_i) \subseteq \mathfrak{A}_n$ for some i (one of the components was already “almost” a n-mean)
- $\operatorname{inv}(F_j) \subseteq \mathfrak{S}_{n-1}$ for all j (n means of $n-1$ matrices, à la ALM/BMP)

Again, the existing techniques cannot be improved

Nothing better than $O(n! \cdot \text{steps})$

Conclusions

- The existing techniques cannot yield anything essentially cheaper than the existing means, for $n \geq 5$
- For $n = 4$, our "tournament mean" is cheaper than ALM/BMP
- After the work in [Lawson, Lim '10 arXiv], another nail in the coffin of ALM-like, composition-based means
- Need new techniques, or new means, to break the $O(n! \cdot \text{steps})$ barrier
- Interesting question: is there a quasi-mean F with $\text{inv}(F) = \mathfrak{A}_n$?

F. Poloni, *Constructing matrix geometric means.* Electr. J. Linear Algebra 20 (2010)

Conclusions

- The existing techniques cannot yield anything essentially cheaper than the existing means, for $n \geq 5$
- For $n = 4$, our "tournament mean" is cheaper than ALM/BMP
- After the work in [Lawson, Lim '10 arXiv], another nail in the coffin of ALM-like, composition-based means
- Need new techniques, or new means, to break the $O(n! \cdot \text{steps})$ barrier
- Interesting question: is there a quasi-mean F with $\text{inv}(F) = \mathfrak{A}_n$?

Thanks for your attention!

F. Poloni, *Constructing matrix geometric means.*
Electr. J. Linear Algebra 20 (2010)