Matrix Power Means

Jose A. Franco

University of North Florida
Department of Mathematics & Statistics

NOVA - March, 2017
Power Means of Numbers

- The term "power means" is usually used to describe the family of means parametrized by a parameter $t$ and defined by:

$$Q_t(a, b) = \left(\frac{a^t + b^t}{2}\right)^{1/t}$$

for positive real numbers $a$ and $b$.

- More generally, if $\mathbb{A} = \{a_i : 1 \leq i \leq n\}$ is a set of $n$ positive real numbers and $\{\omega_i : 1 \leq i \leq n\}$ is such that $\sum_i \omega_i = 1$, the weighted power means can be defined by:

$$Q_t(\omega; \mathbb{A}) := \left(\sum_i \omega_i a_i^t\right)^{1/t}.$$
Extension of Power Means to Matrices

- For the rest of the talk, matrices will be assumed to be positive definite unless otherwise specified.
- Power means of numbers can be extended to the set of positive definite matrices in different ways. Naïvely, one would extend the power means as:

\[ Q_t(\omega; A, B) = \left( \omega A^t + (1 - \omega) B^t \right)^{1/t}. \]

This extension is well-defined for any \( t \in \mathbb{R} \). Some care must be taken when approaching zero, where the limit is considered.

- A less naïve extension is the Kubo-Ando extension of this norm, for \( t \in [-1, 1] \)

\[ P_t(\omega; A, B) = A^{1/2} \left( \omega I + (1 - \omega)(A^{-1/2} B A^{-1/2})^t \right)^{1/t} A^{1/2}. \]
Which extension is best?

- Each of the extensions have their own benefits.
- For example, each Kubo-Ando mean has an operator monotone function that generates it. As such, if two means $M$ and $N$ are generated by two functions $f : [0, \infty) \to \mathbb{R}$ and $g : [0, \infty) \to \mathbb{R}$, respectively, such that $f \geq g$ point-wise, then $M(A, B) \geq N(A, B)$ for all positive definite matrices $A$ and $B$.
- The power means $P_t$ have been used to study the Karcher means by Lim and Palfia.
- The power means $Q_t$ were used to study the log Euclidean mean by Bhagwat and Subramanian.
On the first part of the talk we will explore some properties of \( P_t(A, B) := P_t(1/2; A, B) \) and their relations with the Arithmetic-Geometric-Harmonic means.

Then, we will introduce several Arithmetic-Geometric-Harmonic interpolations and compare them to the power means.
Monotonicity of $P_t(\omega; A, B)$

Notice that

$$P_t(\omega; A, B) = A^{1/2} f_t(A^{-1/2} BA^{-1/2}) A^{1/2},$$

where $f_t(x) = (\omega + (1 - \omega)x^t)^{1/t}$ for $\omega \in (0, 1)$. It is easy to show that $f_t$ is a matrix monotone function on $[0, \infty)$ for $t \in [-1, 1]$. Therefore, if $B \geq D > 0$ we have that

$$A^{-1/2} BA^{-1/2} \geq A^{-1/2} DA^{-1/2}$$

and so $f_t(A^{-1/2} BA^{-1/2}) \geq f_t(A^{-1/2} DA^{-1/2})$. Consequently, for $D \geq B$

$$P_t(\omega; A, B) \geq P_t(\omega; A, D).$$

More is true, if $A \geq C$ and $B \geq D,$

$$P_t(\omega; A, B) \geq P_t(\omega; C, D).$$
A larger interval of $t$

While the function $f_t$ is not matrix monotone outside of $[-1, 1]$, it can still be shown to be an increasing function of $t$ on $\mathbb{R}$. Hence, the following inequality holds for $-\infty < t \leq s < \infty$

$$P_t(\omega; A, B) \leq P_s(\omega; A, B).$$

In particular, since $P_{-1}(\omega; A, B) = (\omega A^{-1} + (1 - \omega)B^{-1})^{-1}$ corresponds with the weighted harmonic means and $P_1(\omega; A, B) = \omega A + (1 - \omega)B$ corresponds with the weighted arithmetic means, the power means provide an interpolation between these two means.
Connection to the Geometric Mean

Not so obviously, at \( t = 0 \) we consider the limit. Lim and Palfia showed that:

\[
\lim_{t \to 0} P_t(\omega; A, B) = A^{1/2} (A^{-1/2} BA^{-1/2})^{1-\omega} A^{1/2} =: A^{\#}_{1-\omega} B,
\]

which is the Kubo-Ando extension of the geometric mean for positive definite matrices.

Therefore, for \( \omega = 1/2 \) the power means also provide a Geometric-Arithmetic mean interpolation. For the rest of the talk we will consider the weight to be \( \omega = 1/2 \) unless otherwise specified.
The most straightforward of these Geometric-Arithmetic mean interpolations are the Heron means:

\[ H_t(A, B) = t \frac{A + B}{2} + (1 - t)A\#B, \]

where \( A\#B = A^{1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2} \).

Since these interpolations are also of Kubo-Ando type, to compare them to the power-means it suffices to compare their generating functions.
For $t \in [0, 1/2]$ and positive matrices $A, B$:

$$H_t(A, B) \geq P_t(A, B).$$

The reversed inequality holds for $t \in [1/2, 0]$. Equality is obtained only for $t \in \{0, 1/2, 1\}$.

The proof is based on a clever rewriting of the numerical inequality

$$\left(\frac{1 + x^t}{2}\right)^{1/t} \leq t \frac{1 + x}{2} + (1 - t)x^{1/2}$$
Figure: Graph of $\lambda_3(H_t(A, B))$ and $\lambda_3(P_t(A, B))$ on $t \in [0, 1]$ for two $3 \times 3$ positive definite matrices $A$ and $B$. 
A not so popular interpolation

Similar to the Heron means, we can define the “straight-line" interpolation of the Harmonic-Geometric means

\[ F_t(A, B) = t \frac{A! B}{2} + (1 - t) A\#B, \]

and study the behavior of this interpolation against the behavior of the power means. This result is easy to show, as the power means for numbers are convex:

**Theorem**

*For* \( t \in [0, 1] \) *and positive matrices* \( A, B \):

\[ F_t(A, B) \geq P_{-t}(A, B). \]
Figure: Graph of $\lambda_3(F_{-t}(A, B))$ and $\lambda_3(P_t(A, B))$ on $t \in [-1, 0]$ for two $3 \times 3$ positive definite matrices $A$ and $B$. 
An even less popular interpolation

Now, we define the linear interpolation between the harmonic and arithmetic mean with the parameter $t \in [-1, 1]$:

$$K_t(A, B) := \frac{t + 1}{2} \left( \frac{A + B}{2} \right) + \frac{1 - t}{2} A!B.$$

And, we compare this to the other interpolations.
Theorem

Let $A$ and $B$ be positive definite matrices. Then:

1. For $t \in [-1, 0]$,
   \[ K_t(A, B) \geq F_{-t}(A, B). \]

2. For $t \in [0, 1/2]$,
   \[ K_t(A, B) \geq H_t(A, B). \]

3. For $t \in [1/2, 1]$,
   \[ K_t(A, B) \geq P_t(A, B). \]
Figure: Graph of $\lambda_2(K_t(A, B))$ and $\lambda_2(P_t(A, B))$ on $t \in [-1, 1]$, $\lambda_2(F_{-t}(A, B))$ on $[-1, 0]$, and $\lambda_2(H_t(A, B))$ on $t \in [0, 1]$ for two $3 \times 3$ positive definite matrices $A$ and $B$. 
For Item 3, the result follows from the inequality for $x > 0$ and $t \in [1/2, 1]$:

$$\frac{t + 1}{2} \left( \frac{1 + x}{2} \right) + (1 - t) \left( \frac{x}{x + 1} \right) \geq \left( \frac{1 + x^t}{2} \right)^{1/t}.$$

It has been shown that the function on right-hand-side is concave for $t \in [1/2, 1]$. Thus, it is bounded above by its tangent line at $t = 1$,

$$y = \frac{1}{2} \left( x \ln x - (x + 1) \ln \frac{x + 1}{2} \right) (t - 1) + \frac{x + 1}{2}$$

on the interval $[1/2, 1]$. 
Hence, it suffices to show that the following the inequality holds for $x > 0$ and $t \in [1/2, 1]$,

\[
\frac{t + 1}{2} \left( \frac{1 + x}{2} \right) + (1 - t) \left( \frac{x}{x + 1} \right) \\
\geq \frac{1}{2} \left( x \ln x - (x + 1) \ln \frac{x + 1}{2} \right) (t - 1) + \frac{x + 1}{2}.
\]

After several simplifications and arguments using the derivative, this reduces to showing

\[(x - 1)^2 \geq 0,
\]

which is obviously true.
With a somewhat unusual notation, in this section we will denote the Kubo-Ando extension of the Heinz means for matrices by $G_t(A,B)$

$$G_t(A, B) = \frac{1}{2}(A\#_t B + A\#_{1-t} B).$$

**Theorem**

*For positive definite matrices $A$ and $B$ and for $t \in [1/2, 3/4]$*

$$G_t(A, B) \leq P_{2t-1}(A, B) \leq H_{2t-1}(A, B)$$

*and for $t \in [3/4, 1]$,*

$$G_t(A, B) \leq H_{2t-1}(A, B) \leq P_{2t-1}(A, B)$$
Figure: Graph of $\lambda_2(G_t(A, B))$, $\lambda_2(H_{2t-1}(A, B))$, and $\lambda_2(P_{2t-1}(A, B))$ on $t \in [1/2, 1]$ for two $3 \times 3$ positive definite matrices $A$ and $B$. 
One last proof

I will only show,

\[ G_t(A, B) \leq H_{2t-1}(A, B), \]

for \( t \in [1/2, 1] \). This reduces to show the following inequality,

\[
\frac{2t - 1}{2} (1 + x) + (2 - 2t)x^{1/2} \geq \frac{x^t + x^{1-t}}{2}
\]  \hspace{1cm} (1)

By dividing by \( x^{1/2} \) and substituting \( x \) by \( e^y \), (1) becomes

\[
(2t - 1) \cosh \left( \frac{y}{2} \right) + (2 - 2t) \geq \cosh \left( \frac{y}{2} (2t - 1) \right).
\]
One last proof

By the concavity of the function $x \mapsto x^{2t-1}$ on this interval,

$$\cosh \left( \frac{y}{2} \right)^{2t-1} = \left( \frac{e^{y/2} + e^{-y/2}}{2} \right)^{2t-1} \geq \left( \frac{e^{(2t-1)y/2} + e^{-(2t-1)y/2}}{2} \right) = \cosh \left( \frac{y}{2} (2t - 1) \right).$$

So, the desired inequality follows if we show:

$$(2t - 1) \cosh \left( \frac{y}{2} \right) + (2 - 2t) \geq \cosh \left( \frac{y}{2} \right)^{2t-1}.$$
Equivalently,

$$za + (1 - z) \geq a^z$$

or

$$z(a - 1) + 1 \geq a^z$$

for any positive real $a$ and $0 \leq z \leq 1$. However, this is just Bernoulli’s inequality,

$$(x + 1)^r \leq 1 + rx$$

for $x = a + 1$ and $z = r$. This completes the proof.
Thank you! Questions?