The Abdus Salam International Centre for Theoretical Physics

United Nations Educational, Scientific and Cultural Organization

IAEA International Atomic Energy Agency

2044-7

Summer School and Advanced Workshop on Trends and Developments in Linear Algebra

22 June - 10 July, 2009

Extending Means to Higher Orders: A General Framework (joint work with Yongdo Lim)

Jimmie Lawson

Louisiana State University Baton Rouge
LA 70803
USA

Extending Means to Higher Orders: A General Framework

Jimmie Lawson

Joint Work with Yongdo Lim

Department of Mathematics

Louisiana State University

Baton Rouge, LA 70803, USA

Introduction

Formally a mean of order n, or n-mean for short, on a set X is a function

$$\mu : X^n \to X \text{ satisfying } \ \forall x \in X, \mu(x, x, \ldots, x) = x.$$

It is frequently assumed in the definition of the mean that a mean is invariant under any permutation of variables; we call these symmetric means.

The mean $\mu : X^n \to X$ is a topological mean if X is Hausdorff and μ is continuous.

Typically a mean represents some type of averaging operator.

A Bit of History

The subject of means dates back into antiquity. The Greeks, motivated by their interest in proportions, defined up to eleven different means, the arithmetic, geometric, harmonic, and golden being the best known.

In the twentieth century interest emerged in the theory of topological symmetric means on topological spaces. This work was pioneered by G. Aumann (1934), who showed among other things that no sphere admits such a mean. The problem of characterizing those spaces, particularly metric continua, that admit the structure of a topological mean has attracted considerable attention up to the present day. See J. Charatonic's web overview of both older and recent work.

Origins of Operator Means

Operator means are of more recent vintage, but have a substantial literature that has grown out of foundational papers, particularly that of Kubo and Ando (1980).

A key result is that the continuous, invariant (under congruence transformations), monotone 2-means on the positive operators on Hilbert space are given by

$$\mu(A, B) = A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2},$$

where $f : \mathbb{R}^+ \to \mathbb{R}^+$ is a continuous, nondecreasing, operator monotone function with $f(1) = 1$.

The theory has found a variety of applications, including the establishment of important inequalities, some of which find application in quantum mechanical calculations.

The Challenge

The various successful applications of operator means has motivated a search for extending them to means of higher orders. However, finding a general method of extending operator means of two variables to means of a higher number of variables has proved elusive, and the question of how to do this has remained an open problem.

Included in the problem are the subproblems of defining precisely what one means by such an extension, what properties it should have, and to what extent it is unique.

Recent Progress

Similar inductive approaches to extending means to higher orders have been carried out by Horwitz (2002) for the case of means on the positive real numbers and by Ando, Li, and Mathias (2004) for the case of the geometric mean on the positive (semi)definite Hermitian matrices. This approach has also been adopted and generalized beyond the case of the matrix geometric mean by Petz and Temesi (2005), although in the general setting they only obtain existence of the higher order means for ordered tuples.

Bhatia and Holbrook (2006) have proposed an alternative generalization for the geometric mean via a geometric approach linked to work of E. Cartan.

A General Approach Via Metric Spaces

Our purpose is to develop a method of extending means to higher orders that appears to offer a viable general approach. We show that the basic approach of Horwitz, of Ando, Li and Mathias, and of Petz and Temesi can be generalized to means on metric spaces and develop the theory of extensions in this context.

The resulting theory is attractive for the generality of its results and for the resulting uniqueness and preservation properties of the extensions.

The Barycentric Operator

Given a set X and a k-mean $\mu : X^k \to X$, the barycentric operator $\beta = \beta_\mu : X^{k+1} \to X^{k+1}$ is defined by

$$\beta(\mathbf{x}) := \big(\mu(\pi_{\neq 1}\mathbf{x}), \ldots, \mu(\pi_{\neq k+1}\mathbf{x})\big),$$

where $\mathbf{x} = (x_1, \ldots, x_{k+1}) \in X^{k+1}$ and
$\pi_{\neq j}\mathbf{x} := (x_1, \ldots, x_{j-1}, x_{j+1}, \ldots, x_{k+1}) \in X^k$.
For a topological k-mean, we say that the barycentric map β is power convergent if for each $\mathbf{x} \in X^{k+1}$, we have $\lim_n \beta^n(\mathbf{x}) = (x^*, \ldots, x^*)$ for some $x^* \in X$.

An Intuitive Example

As a motivating geometric example for the terminology consider the 3-mean in $\mathbb{R}^3$ that assigns to any three points the centroid of the triangle for which they are the vertices, i.e., the point where the three medians meet. If we take now the four vertices of a 3-simplex or tetrahedron in $\mathbb{R}^3$, the barycentric operator applied to the 4-tuple consisting of the four vertices replaces each vertex with the centroid (barycenter) of the face opposite it, the face with vertices the remaining three vertices. Thus one may envision the tetrahedron with vertices the four centroids of the four faces as the result. Repeating this process, one obtains a shrinking family of tetrahedra whose intersection is the barycenter of the original tetrahedron, represented by the 4-tuple with all entries equal to that point.

Extending One's Means

We define our first notion of an extension in terms of the barycentric operator.

A mean $\nu : X^{k+1} \to X$ is a β-invariant extension of $\mu : X^k \to X$ if $\nu \circ \beta_\mu = \nu$, i.e., $\forall \mathbf{x} = (x_1, \ldots, x_{k+1}) \in X^{k+1}$

$$\nu(\mathbf{x}) = \nu\big(\mu(\pi_{\neq 1}\mathbf{x}), \ldots, \mu(\pi_{\neq k+1}\mathbf{x})\big). \tag{1}$$

The notion of a β-invariant extension was introduced by Horwitz, who called it type I invariance.

An Existence and Uniqueness Result

Proposition. Assume that $\mu : X^k \to X$ is a topological k-mean and that the corresponding barycentric operator β is power convergent. Define $\tilde{\mu} : X^{k+1} \to X$ by $\tilde{\mu}(\mathbf{x}) = x^*$ where $\lim_n \beta^n(\mathbf{x}) = (x^*, \ldots, x^*)$.

(i) $\tilde{\mu} : X^{k+1} \to X$ is a $(k+1)$-mean on X that is a β-invariant extension of μ.

(ii) Any continuous mean on X^{k+1} that is a β-invariant extension of μ must equal $\tilde{\mu}$.

(iii) If μ is symmetric, so is $\tilde{\mu}$.

Powerful Extensions

We seek a notion of mean extension that allows one both to deduce readily that a large number of properties transfer from a mean to its extension and also is applicable to a wide variety of means. The preceding proposition provides the ingredients for this definition.

A $(k+1)$-mean ν is a β-extension of a topological k-mean μ (or β-extends μ) if for each $\mathbf{x} \in X^{k+1}$, $\lim_n \beta^n(\mathbf{x}) = (\nu(\mathbf{x}), \ldots, \nu(\mathbf{x}))$. In this case we say that β power converges to ν, written $\beta_\mu^n \to \nu$.

We note (using the previous proposition) that a β-extension ν is a β-invariant extension, and hence any continuous β-invariant extension of μ must equal ν.

Tying Some Things Together

A. Horwitz and later D. Petz and R. Temesi consider means on the positive reals and show that any continuous symmetric 2-mean that is strict ($\min(a,b) < \mu(a,b) < \max(a,b)$ for $a \neq b$) and order-preserving in each variable has a power convergent barycentric map, and hence has a unique β-extension to a 3-mean. Petz and Temesi point out that the argument for power convergence extends to higher order variables, and thus one can inductively define β-extensions for all $n > 2$.

For the arithmetic, geometric, and harmonic means the extensions yield the usual corresponding means of n-variables. To check this, one has only to note that they are continuous and are β-invariant extensions, then apply the earlier uniqueness result.

Preserving Convexity

Let μ be a topological mean on a metric space X. In a direct fashion one may define (in terms of μ) convex sets, local convexity, uniform local convexity, closed ball convexity, etc. and directly show that if μ admits a β-extension ν, then ν also inherits the corresponding properties. One also concludes that ν is also a topological mean, provided that (X, μ) is locally convex.

Two Key Notions

(1) A k-mean μ on a metric space X is called nonexpansive if it satisfies for all $\mathbf{x} = (x_1, \ldots, x_k), \mathbf{y} = (y_1, \ldots, y_k) \in X^k$,

$$d(\mu(\mathbf{x}), \mu(\mathbf{y})) \leq \max\{d(x_j, y_j) : 1 \leq j \leq k\},$$

i.e., μ has Lipschitz constant 1 wrt the sup metric on X^k.

(2) For $0 < \rho < 1$, we say that μ is coordinatewise ρ-contractive if for any $\mathbf{x}, \mathbf{y} \in X^k$ that differ only in one coordinate, say $x_j \neq y_j$,

$$d(\mu(\mathbf{x}), \mu(\mathbf{y})) \leq \rho\, d(x_j, y_j).$$

The End That Determines Means

We come now to our principal tool for extending means.

Theorem. Let X be a complete metric space equipped with a nonexpansive, coordinatewise ρ-contractive $(0 < \rho < 1)$ k-mean $\mu : X^k \to X$, $k \geq 2$. Then there exists uniquely a family of continuous means $\mu_n : X^n \to X$, one for every $n > k$, such that each is a β-extension of the previous one. Furthermore, each μ_n is nonexpansive and coordinatewise ρ-contractive.

Metric Convex Means

A general metric space may have none, one, or many midpoints between two given points. (Recall that m is a midpoint of a and b if $d(m,a) = d(m,b) = (1/2)d(a,b)$.) We consider the setting where possibly many midpoints may exist, but there is a distinguished midpoint that appears in a "convex" manner.

A symmetric mean $\mu : X \times X \to X$, written $\mu(x,y) = x\#y$, on a complete metric space X is called a convex mean if it satisfies the *basic convexity condition*

$$d(x\#z, y\#z) \leq \frac{1}{2}d(x,y) \text{ for all } x, y, z \in X.$$

Note. For a convex mean, $x\#y$ is a midpoint for all x, y.

Extension of Convex Means

Proposition. A convex mean inductively β-extends to a symmetric, nonexpansive, coordinatewise $(1/2)$-contractive n-mean for every $n > 2$.

Example. Let X be a Banach space (or a closed convex subset thereof) and define the symmetric 2-mean $\mu(x, y) = (1/2)(x + y)$. This is the midpoint with respect to the norm metric, and is easily seen to be a convex mean. Setting $\mu_k(x_1, \ldots, x_k) = (1/k)\sum_{i=1}^{k} x_i$, one verifies directly that μ_{k+1} is the β-invariant extension of μ_k, so μ inductively β-extends to the standard arithmetic mean μ_n for all n.

Hadamard Spaces

A metric space X is said to satisfy the semiparallelogram law if for any two points $x_1, x_2, \in X$, there exists $z \in X$ that satisfies for all $x \in X$:

$$d(x_1, x_2)^2 + 4d(x, z)^2 \leq 2d(x, x_1)^2 + 2d(x, x_2)^2.$$

It follows readily that z is the unique midpoint between x_1 and x_2. A Hadamard space (occasionally called a Bruhat-Tits space) is a complete metric space that satisfies the semiparallelogram law.

CAT(0)-Spaces

Using a metric notion for an upper bound of curvature (geodesic triangles in the metric space satisfy certain inequalities when compared with euclidean test triangles), one calls a metric space a CAT(κ)-space if it is a geodesic space (each pair of points can be connected by a metric geodesic) satisfying the curvature bound condition for the real number κ. The CAT(0)-spaces are the non-positively curved spaces. A metric space has an alternative characterization as a Hadamard space: it is a simply connected, complete, geodesic CAT(0)-space.

Hadamard Means

Remark. Defining a 2-mean $\mu(x, y)$ on a Hadamard space (X, d) as the unique midpoint between x and y defines a convex mean in the sense of the preceding section. Hence this mean may be β-extended to an n-mean for every $n > 2$. This result provides a new and interesting operation of “barycenter" for any finite subset of a Hadamard spaces.

A wide variety of Hadamard spaces and constructions for new Hadamard spaces from old appear in the literature. Some examples include Hadamard manifolds (simply connected complete Riemannian manifolds with nonpositive sectional curvature), particularly simply connected symmetric spaces of noncompact type, finite-dimensional hyperbolic geometries over the reals, complexes, and quaternions, symmetric cones, Tits buildings, and various examples obtained by coning and gluing.

Positive Definite Matrices

Of particular interest to us is the example of the manifold X of positive definite (real or complex) matrices. Endowed with the usual Riemannian metric, the trace metric, X is a Hadamard manifold and the midpoint mean operation is precisely the geometric mean of the two positive definite matrices. Using the fact that the length metric satisfies the semiparallelogram law, hence is a convex metric with the midpoint operation being a convex mean, we obtain the following alternative derivation of the principal result of Ando, Li, and Mathias:

Corollary The midpoint operation for the trace length metric, which is precisely the geometric mean, defines a convex 2-mean, which β-extends to an n-mean for each $n > 2$.

Iterated Means

A standard construction technique for means is iteration, the arithmetic-geometric or Gauss mean being the best known example. We derive a result for showing that certain iterated means are coordinatewise ρ-contractive and nonexpansive, hence admit β-extensions of all orders.

Definition. Let λ, ν be 2-means on a complete metric space X. Starting with $\lambda_1 = \lambda$ and $\nu_1 = \nu$, we inductively obtain sequences of means $\{\lambda_n\}$ and $\{\nu_n\}$:

$$\lambda_{n+1}(x,y) = \lambda(\lambda_n(x,y), \nu_n(x,y)),\ \nu_{n+1}(x,y) = \nu(\lambda_n(x,y), \nu_n(x,y)).$$

If there exists a 2-mean μ such that for all $x, y \in X$, $\lim_n \lambda_n(x,y) = \mu(x,y) = \lim_n \nu_n(x,y)$, then μ is called the *iterated composition* of λ and ν and denoted $\mu = \lambda * \nu$.

Extending Iterated Means

The following result enables the extension of a significant class of iterated means.

Proposition. Suppose in a complete metric space X that λ is a convex mean and ν is coordinatewise ρ'-contractive, $0 < \rho' < 1$, and nonexpansive. Then the iterated composition $\lambda * \nu$ exists, is coordinatewise ρ-contractive, $\rho = \max\{1/2, \rho'\}$ and nonexpansive, and hence β-extends to all orders greater than two.

Positive Definite Operators

Consider the space Ω of positive definite operators on a Hilbert space with the Thompson metric

$$d(A, B) = \log(\max\{M(A/B), M(B/A)\}),$$

where $M(A/B) = \inf\{t \geq 0 : A \leq tB\}$.

The Thompson metric is a complete metric on Ω and the operator geometric mean

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

is a convex mean wrt the Thompson metric, hence (uniquely) β-extends to all higher orders.

Basic Properties

We order the open cone Ω of positive definite operators by the Loewner order:

$$A \le B \Leftrightarrow B - A \text{ is positive semidefinite.}$$

By definition the order interval $[A, B] := \{C \in \Omega : A \le C \le B\}$.

A 2-mean μ on Ω is monotonic if $A_1 \le B_1$ and $A_2 \le B_2$ imply $\mu(A_1, B_1) \le \mu(A_2, B_2)$. It is invariant or satisfies the transformer equality with respect to a given metric d if $d(PAP^*, PBP^*) = d(A, B)$ for $A, B \in \Omega$ and P invertible.

Proposition. Monotonicity and invariance are preserved by β-extensions of means.

Extending Iterated Operator Means

The following theorem is the main one for forming and extending iterated operator means.

Theorem. Let Ω denote the set of positive operators on a Hilbert space, let $\lambda(A, B) = A\#B$ denote the geometric mean of A, B, and let ν be a continuous, invariant, monotonic 2-mean on Ω that is also nonexpansive and coordinatewise ρ_n-contractive for $0 < \rho_n < 1$ on the order interval $[(1/n)I, nI]$ for each n with respect to the Thompson metric. Then the iterated composition $\mu = \lambda * \nu$ exists, is a coordinatewise ρ_n-contractive, nonexpansive mean when restricted to $[(1/n)I, nI]$ for each n, and hence inductively β-converges to a β-extension for each $n > 2$.

Gauss and Logarithmic Means

The following lemma is important for forming and extending iterated means.

Lemma. The arithmetic and harmonic operator means are continuous, invariant, monotonic 2-means on Ω that are also nonexpansive and coordinatewise ρ_n-contractive for $0 < \rho_n < 1$ on the order interval $[(1/n)I, nI]$ for each n with respect to the Thompson metric.

The Gauss mean and the logarithmic mean may be formed as appropriate iterated means of the geometric and arithmetic means, and hence by our theorem on extensions of iterated means may be inductively β-extended for each $n > 2$.

References

1. T. Ando, C.-K. Li, and R. Mathias, Geometric means, Linear Algebra and Appl. 385(2004), 305–334.
2. G. Aumann, Aufbau von Mittelwerten mehrerer Argumente I, Math. Ann. **109** (1934), 235-253.
3. R. Bhatia and J. Holbrook, Riemannian geometry and matrix geometric means, Linear Algebra Appl. **413** (2006), 594-618.
4. J. Charatonik, Means on arc-like continua, http://web.mst.edu/∼continua/4_Bubu.pdf.
5. A. Horwitz, Invariant means, J. Math. Anal Appl. **270** (2002), 499-518.
6. F. Kudo and T. Ando, Means of positive linear operators, Math. Ann. **246** (1980), 205-224.
7. D. Petz and R. Temesi, Means of positive numbers and matrices, SIAM J. Matrix Anal. Appl. 27 (2005), 712-720.