The Karcher Barycentric Map on Positive Cones of Operators

Jimmie Lawson
Lousiana State University

Joint work with Yongdo Lim

July, 2016

There has been significant work in recent years on multivariable matrix means, and more generally operator means. In this report we consider the extension of one of these, the Karcher mean, to integrable measures on the open cone $\mathbb{P}$ of positive operators on a Hilbert space. For ease of presentation we restrict our considerations to separable metric spaces and separable Hilbert spaces.

Borel Measures

A *Borel probability measure* on a metric space (X, d) is a countably additive non-negative measure μ on the Borel algebra $\mathcal{B}(X)$, the smallest Σ-algebra containing the open sets, such that $\mu(X) = 1$. We denote the set of all probability measures on $(X, \mathcal{B}(X))$ by $\mathcal{P}(X)$.

Let $\mathcal{P}_0(X)$ be the set of all uniform finitely supported probability measures, i.e., all $\mu \in \mathcal{P}(X)$ of the form $\mu = \frac{1}{n}\sum_{j=1}^{n} \delta_{x_j}$ for some $n \in \mathbb{N}$, where δ_x is the point measure of mass 1 at x.

The *support of* $\mu \in \mathcal{P}(X)$ is given by

$$\mathrm{supp}(\mu) = \{x \in X : \; x \in U \text{ open } \Rightarrow \mu(U) > 0\}.$$

Integrable Measures

For $p \in [1, \infty)$ let $\mathcal{P}^p(X) \subseteq \mathcal{P}(X)$ be the set of probability measures with *finite p-moment*: for some (and hence all) $x \in X$,

$$\int_X d^p(x, y) d\mu(y) < \infty.$$

For $p = \infty$, $\mathcal{P}^\infty(X)$ denotes the set of probability measures with bounded support.

The set $\mathcal{P}^1(X)$ is the largest of these sets, and is also called the set of *integrable* measures. It will be our focus in this talk.

Couplings

For metric spaces X and Y, a continuous $f : X \to Y$ induces a *push-forward* map $f_* : \mathcal{P}(X) \to \mathcal{P}(Y)$ defined by $f_*(\mu)(B) = \mu(f^{-1}(B))$ for $\mu \in \mathcal{P}(X)$ and $B \in \mathcal{B}(Y)$.

We say that $\omega \in \mathcal{P}(X \times X)$ is a *coupling* for $\mu, \nu \in \mathcal{P}(X)$ and that μ, ν are *marginals* for ω if for all $B \in \mathcal{B}(X)$

$$\omega(B \times X) = \mu(B) \qquad \text{and} \qquad \omega(X \times B) = \nu(B).$$

Equivalently μ and ν are the push-forwards of ω under the projection maps π_1 and π_2 resp. We note that one such coupling is the product measure $\mu \times \nu$, and that for any coupling ω it must be the case that $\mathrm{supp}(\omega) \subseteq \mathrm{supp}(\mu) \times \mathrm{supp}(\nu)$. We denote the set of all couplings by $\Pi(\mu, \nu)$.

The Wasserstein Metric

The Wasserstein distance d^W (alternatively Kantorovich-Rubinstein distance) on $\mathcal{P}^1(X)$ is defined by

$$d^W(\mu_1, \mu_2) := \inf_{\pi \in \Pi(\mu_1, \mu_2)} \int_{X \times X} d(x, y) d\pi(x, y).$$

It is known that d^W is a complete metric on $\mathcal{P}^1(X)$ whenever X is a complete metric space and $\mathcal{P}_0(X)$ is d^W-dense in $\mathcal{P}^1(X)$.

Barycenters

Given $G : \mathcal{P}_0(X) \to X$ satisfying $G(\delta_x) = x$ for all $x \in X$, we define $G_n : X^n \to X$ for $n \geq 2$ by $G_n(x_1, \ldots, x_n) = G(\frac{1}{n}\sum_{k=1}^{n} \delta_{x_k})$.
(1) If for each $n \geq 2$ and $(x_1, \ldots, x_n), (y_1, \ldots, y_n) \in X^n$,

$$d(G_n(x_1, \ldots, x_n), G_n(y_1, \ldots, y_n)) \leq \frac{1}{n} \sum_{k=1}^{n} d(x_k, y_k),$$

then G is *contractive*, i.e., has Lipschitz constant≤ 1, from $(\mathcal{P}_0(X), d^W)$ to (X, d).
(2) If (X, d) is complete and G is contractive, then G uniquely extends to a contractive map $\beta_G : \mathcal{P}^1(X) \to X$, called the *barycentric map* associated with G.

Example: Banach Spaces

For X a Banach space, we define the arithmetic mean $\mathbf{A}$ by

$$\mathbf{A}_n(x_1, \ldots, x_n) = \mathbf{A}\left(\frac{1}{n}\sum_{i=1}^{n}\delta_{x_i}\right) = \frac{1}{n}\sum_{i=1}^{n} x_i.$$

It is easily checked that the arithmetic mean $\mathbf{A}$ is contractive, so uniquely extends to a contractive barycentric map $\beta_A : \mathcal{P}^1(X) \to X$.

Hadamard Spaces

Hadamard spaces are complete metric spaces satisfying for all x, y, z the semiparallelogram law

$$d^2(x,y) \leq \frac{1}{2}d^2(x,z) + \frac{1}{2}d^2(y,z) - \frac{1}{4}d^2(x\#y,z),$$

where $x\#y$ is the midpoint, necessarily unique, between x and y. In the literature these are often referred to as (global) CAT(0)-spaces or spaces of nonpositive curvature (NPC).

In a Hadamard space there is a unique metric geodesic $t \mapsto a\#_t b$, $0 \leq t \leq 1$, connecting a and b, where $a\#_t b$ is uniquely determined by

$$d(a,b) = (1-t)d(a, a\#_t b) + td(a\#_t b, b).$$

The Least Squares Mean and Cartan Barycentric Map

(1) In a Hadamard metric space X the *least squares mean* Λ is given by $\Lambda_n(a_1, \ldots, a_n) = x^*$, where x^* is the unique minimizer of the function $x \mapsto \sum_{i=1}^n d^2(x, a_i)$:

$$\Lambda_n(a_1, \ldots, a_n) = \arg\min_{x \in X} \sum_{i=1}^{n} d(x, a_i)^2.$$

(2) The *Cartan barycentric map* on a Hadamard space (X, d) carries $\mu \in \mathcal{P}^1(X)$ to unique minimizer (independent of y)

$$\beta_\Lambda(\mu) = \arg\min_{z \in X} \int_X [d^2(z, x) - d^2(y, x)] d\mu(x).$$

The least squares mean is contractive and the Cartan barycentric map is its contractive extension.

The open cone, and hence manifold, $\mathbb{P}_n$ of $n \times n$ positive definite matrices admits a natural Riemannian metric, called the trace metric, making it a simply connected Riemannian manifold of negative curvature. All such manifolds are Hadamard spaces with respect the the length metric (the length of the unique metric geodesic between two points), which in this case is given by the sum of the eigenvalues equal the trace of $A^{-1/2}BA^{-1/2}$:

$$d(A,B) = \sum_{i=1}^{n} \lambda_i(A^{-1/2}BA^{-1/2}) = \mathrm{tr}(A^{-1/2}BA^{-1/2}),$$

The metric geodesic between A and B in $\mathbb{P}_n$ is given by $t \mapsto A\#_t B := A^{1/2}(A^{-1/2}BA^{-1/2})^t A^{1/2}$, called the *t-weighted geometric mean*.

Thus the least squares mean Λ on $\mathbb{P}_n$ induces a Cartan barycentric map $\beta_\Lambda : \mathcal{P}^1(\mathbb{P}_n) \to \mathbb{P}_n$.

Karcher Equation

Consider for $a_1, \ldots, a_n \in X$, a metric space, the sum of distances squared function $\sigma : X \to \mathbb{R}$ given by $\sigma(x) = \sum_{i=1}^{n} d(x, a_i)^2$. For simply connected Riemannian manifolds of non-positive curvature, Karcher showed that the least squared mean $\Lambda_n(a_1, \ldots, a_n)$ could be characterized by the vanishing of the gradient of σ. For $\mathbb{P}_n$ this yields that $X = \Lambda_n(A_1, \ldots, A_n)$ is the unique solution of the *Karcher equation*

$$\sum_{i=1}^{n} \log(X^{-1/2} A_i X^{-1/2}) = 0.$$

Karcher Barycentric Map

For the open cone $\mathbb{P}$ of positive operators on an infinite-dimensional Hilbert space, the Hadamard space and least squares machinery disappears. However, somewhat suprisingly, one can show that the Karcher equation still has a unique solution [L.,Lim], giving rise to a *Karcher mean* even in the infinite-dimensional setting. This mean turns out to have many of the nice properties of the least squares mean. In particular, it is contractive for the Thompson metric on $\mathbb{P}$, so gives rise to a *Karcher barycentric map* $\beta : \mathcal{P}^1(\mathbb{P}) \to \mathbb{P}$.

The Lim-Palfia Cyclic Approximation Theorem

Let $(a_1, \ldots, a_n) \in X$, a Hadamard metric space. We recycle over and over through these points by setting $a_k = a_j$ if $k > n$, $1 \leq j \leq n$, and k is congruent to j modulo n. We inductively define a sequence y_k by $y_1 = a_1$, $y_2 = a_1 \# a_2$, the midpoint of a_1 and a_2, and in general $y_k = y_{k-1} \#_{1/k} x_k$, that is, we move from y_{k-1} toward x_k a distance of $\frac{1}{k} d(y_{k-1}, x_k)$.

Theorem (Lim, Palfia)

The sequence $\{y_n\}$ converges to the least squares mean $\Lambda(a_1, \ldots, a_n)$.

Another Cyclic Approximation Theorem

What happens if we repeat this process for $(A_1, \dots, A_n) \in \mathbb{P}$, the open cone of positive operators on an infinite-dimesional Hilbert space?

Theorem (L.,Lim)

The sequence $\{Y_n\}$ converges in the strong operator topology to the Karcher mean $\Lambda(A_1, \dots, A_n)$.

A subset $K \subseteq \mathbb{P}$ is *convex* (short for geodesically convex) if given $A, B \in K$, the geodesic segment $\{A \#_t B : 0 \leq t \leq 1\}$ is contained in K.

Corollary

For $\mu \in \mathcal{P}^1(\mathbb{P})$ the Karcher barycenter $\beta(\mu)$ belongs to the strong closure of the convex hull of $\mathrm{supp}(\mu)$, the support of μ.

Proposition (Jensen's Inequality)

Let $F : \mathbb{P} \to \mathbb{R}$ be lower semicontinuous with respect to the strong topology and convex, i.e., $F(A \#_t B) \leq (1-t)F(A) + tF(B)$ for all $A, B \in \mathbb{P}$. Then for any measure $\mu \in \mathcal{P}^1(\mathbb{P})$, $F(\beta_\Lambda(\mu)) \leq \int_{\mathbb{P}} F(X)\, d\mu(X)$.

An Open Problem and Future Work

K.-T. Sturm has derived a version of the Strong Law of Large Numbers (SLLN) for random variables into a Hadamard space. An interesting question is whether a similar SLLN holds for the space $\mathbb{P}(H)$ of positive operators on a Hilbert space. The $\mathbb{P}$-version of the Lim-Palfia theorem may be viewed as a deterministic version of this result for measures with finite support. More generally, the existence of the Karcher barycenter map for $\mathbb{P}$ suggests the possibility of a theory of $\mathbb{P}$-valued random variables.