The α-z-Bures Wasserstein divergence and quantum α-z-fidelity

Trung Hoa Dinh
Troy University

SEAM40 - Gainesville
March 19, 2024

Introduction

Let $A_1, A_2, \cdots, A_m$ be positive definite matrices. The problem of finding the averaging of A_i is important.

- The least squares problem (LSP):

$$\min \lim_{X>0} \sum_{i=1}^{m} \delta^2(X, A_i),$$

  where $\delta(A, B)$ is some distance function.

- Algebraic approach:

$$X = \sum \omega_i f(X, A_i),$$

  or

$$F(X, A_1, A_2, \cdots, A_n) = 0$$

The solution to these problems (if they exist) is called the Karcher mean of A_i.

Introduction

Let A, B be positive definite matrices.

- Moakher (2014), and Bhatia and Holbrook (2015):

$$\min_{X>0}(\delta_2^2(X, A_1) + \delta_2^2(X, A_2)), \tag{1}$$

  where $\delta_2(A, B) = ||\log(A^{-1}B)||_2$ is the Riemannian distance between A and B.

- Puzs and Woronowicz (1975): The algebraic Riccati equation:

$$XA^{-1}X = B.$$

- The solution of both problems is the geometric mean

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

Introduction

Let $A_1, A_2, \cdots, A_m$ be positive definite matrices.

- The Karcher equation:

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

- Lim-Palfia (JFA, 2012):

$$X = \sum_{i=1}^{m} \frac{1}{n} X \sharp_t A_i.$$

- An equation that occurs in many applied fields:

$$X^p = A + \sum_{i}^{m} M_i^* f(X) M_i,$$

Introduction

- Introduce new distance functions, consider if they are applicable or not (for example, Data Processing Inequality in quantum information theory).
- Study the LSP with respect to new distance functions.
- Study properties of the solution of LSP.
- Compare solutions of the LSPs in different distance functions.
- Find algorithms to approximate the solution of the LSP.
- etc.

Presentation Outline

α-z-Bures Wasserstein divergence and α-z-weighted right mean
- α-z-Bures Wasserstein divergence and the least squares problem

Quantum α-z-fidelity and α-z-fidelity between unitary orbits
- Variational formulas for quantum α-z-fidelity
- α-z-fidelity between unitary orbits

Quantum divergences

A map $\Phi : \mathcal{D}_n \times \mathcal{D}_n \to [0, +\infty)$ ($\mathcal{D}_n$ is the set of positive definite matrices) is a quantum divergence if it satisfies the following properties:

(A) $\Phi(A, B) \geq 0$ and $\Phi(A, B) = 0$ if and only if $A = B$.

(B) The first derivative with respect to the second variable vanishes on the diagonal:

$$\frac{\partial \Phi}{\partial B}(A, B)\big|_{B=A}(C) = 0 \quad \text{for all Hermitian matrices } C.$$

(C) The second derivative is non-negative on the diagonal:

$$\frac{\partial^2 \Phi}{\partial B^2}(A, B)\big|_{B=A}(C, C) \geq 0 \quad \text{for all Hermitian matrices } C.$$

Quantum Hellinger divergences

- With the geometric mean (Bhatia, Gaubert, Jain, 2019):

$$d_3(A, B) = \mathrm{Tr}\,\frac{A+B}{2} - \mathrm{Tr}\,(A\sharp B)\,.$$

- With a general Kubo-Ando mean (Pitrik, Virosztek, 2020):

$$d_4(A, B) = \mathrm{Tr}\,((1-c)A + cB - A\sigma B),$$

  where σ is a Kubo-Ando mean, f_σ is the representing function of σ, and $c = f'_\sigma(1) \in (0, 1)$ is the weight of σ.
- Kubo-Ando's theory:

$$A\sigma B = A^{1/2} f_\sigma(A^{-1/2} B A^{-1/2}) A^{1/2}.$$

Quantum divergences

- Bures-Wasserstein distance in the theory of optimal transport:

$$d_b^2(A,B) = \mathrm{Tr}(A+B)/2 - \mathrm{Tr}\left(\left(A^{1/2}BA^{1/2}\right)^{1/2}\right).$$

- The Log-Determinant distance or S-distance in machine learning and and quantum information:

$$d_S^2(A,B) = \log\det\frac{A+B}{2} - \frac{1}{2}\log\det(AB).$$

- The Hellinger metric or Bhattacharya metric in quantum information:

$$d_h^2(A,B) = \mathrm{Tr}(A+B)/2 - \mathrm{Tr}\left(A^{1/2}B^{1/2}\right).$$

Quantum fidelity and Relative entropies

Let $P_{\alpha,z}(A,B) = (B^{\frac{1-\alpha}{2z}} A^{\frac{\alpha}{z}} B^{\frac{1-\alpha}{2z}})^z$.

- **Sandwiched quasi-relative entropy** (Lieb, Ruskai, 1973): For $A, B \geq 0$, $t \in [0,1]$,

$$F_t(A,B) = \operatorname{Tr}(A^{\frac{1-t}{2t}} B A^{\frac{1-t}{2t}})^t = P_{t,t}(B,A)$$

  which appears in the definition of the weighted Bures-Wasserstein distance as

$$d_{b,t}(A,B) = (\operatorname{Tr}((1-t)A + tB) - \operatorname{Tr}(F_t(A,B))^{1/2}.$$

- **α-z-Renyi relative entropy** (Audenaert, Datta, 2015): For positive definite matrices A, B, $\alpha, z \in \mathbb{R}$,

$$D_{\alpha,z}(A||B) = \frac{1}{\alpha - 1} \log \operatorname{Tr}(P_{\alpha,z}(A,B)).$$

α-z-Bures Wasserstein divergence

- The weighted Bures-Wasserstein distance:

$$d_{b,t}(A,B) = \left(\mathrm{Tr}\left((1-t)A + tB\right) - \mathrm{Tr}\left(F_t(A,B)\right)\right)^{1/2}.$$

- Mention that $(B^{\frac{1-\alpha}{2z}} A^{\frac{\alpha}{z}} B^{\frac{1-\alpha}{2z}})^z$ and $(A^{\frac{1-t}{2t}} B A^{\frac{1-t}{2t}})^t$ are matrix generalizations of the geometric mean $\sqrt{ab}$ and the weighted geometric mean $a^{1-t}b^t$ of positive numbers a and b, respectively.
- **α-z-Bures Wasserstein divergence** (D., Le, Vo, Vuong, LAA, 2021): For $0 \leq \alpha \leq z \leq 1$, the following is a quantum divergence:

$$\Phi(A,B) = \mathrm{Tr}\left((1-\alpha)A + \alpha B\right) - \mathrm{Tr}\left(Q_{\alpha,z}(A,B)\right),$$

where $Q_{\alpha,z}(A,B) = P_{\alpha,z}(B,A)$ is the matrix function in the α-z-Renyi relative entropy.

Data Processing Inequality

- **The data processing inequality** with respect to a quantum divergence Ψ means that for any completely positive trace preserving map $\mathcal{E}$ and for any positive semidefinite matrices A and B,

$$\Psi(\mathcal{E}(A), \mathcal{E}(B)) \leq \Psi(A, B).$$

- A map $\Psi(A, B)$ is **jointly convex**, **unitarily invariant and invariant under tensor product**, then Ψ is monotone with respect to all completely positive trace-preserving maps.
- The α-z-Bures Wasserstein divergence $\Phi(A, B)$ is jointly convex since the trace function $\mathrm{Tr}\left[\left(X^{\frac{q}{2}} Y^p X^{\frac{q}{2}}\right)^s\right]$ is jointly concave for $0 \leq p, q \leq 1$ and $0 \leq s \leq 1/(p+q)$. It is unitarily invariant and invariant under tensor product, hence, satisfies the Data Processing Inequality.

The least squares problem

Theorem (D., Le, Vo, Vuong, LAA, 2021). *For* $0 \le \alpha \le z \le 1$, *the function*

$$F(X) = \sum_{i=1}^{m} \omega_i \Phi(A_i, X)$$

attains minimum at X_0, *where* X_0 *is the unique positive definite solution of the following matrix equation*

$$\sum_{j=1}^{m} w_j (X^{\frac{\alpha}{2z}} A_j^{\frac{1-\alpha}{z}} X^{\frac{\alpha}{2z}})^z = \sum_{j=1}^{m} w_j Q_{1-\alpha,z}(X, A_i) = X.$$

- The solution X_0 is called the **α-z-weighted right mean**, denoted by $\mathcal{R}_{\alpha,z}(\omega, \mathbb{A})$.
- Some properties of this mean were recently studied by Jeong and Kim (2023).

Variational formulas for quantum α-z-fidelity

- Audenaert and Datta (2015) introduced the quantity so-called α-z-fidelity

$$f_{\alpha,z}(\rho,\sigma) := \operatorname{Tr}\left(\rho^{\alpha/2z}\sigma^{(1-\alpha)/z}\rho^{\alpha/2z}\right)^z = \operatorname{Tr}\left(\sigma^{(1-\alpha)/2z}\rho^{\alpha/z}\sigma^{(1-\alpha)/2z}\right)^z. \quad (2)$$

Bhatia and coauthors (2018) established some variational formulas for $f_{\alpha,\alpha}(\rho,\sigma)$ via extreme values of the following matrix functions.

Let ρ and σ be positive definite matrices and let $0 < \alpha < 1$. Then

- $f_{\alpha,\alpha}(\rho,\sigma) = \min_{X>0} \operatorname{Tr}[(1-\alpha)\left(\sigma^{\frac{\alpha-1}{2\alpha}}X\sigma^{\frac{\alpha-1}{2\alpha}}\right)^{\frac{\alpha}{\alpha-1}} + \alpha X\rho]$;
- $f_{\alpha,\alpha}(\rho,\sigma) = \min_{X>0} \operatorname{Tr}[(\sigma^{\frac{\alpha-1}{2\alpha}}X\sigma^{\frac{\alpha-1}{2\alpha}})^{\frac{\alpha}{\alpha-1}}]^{1-\alpha}[\operatorname{Tr}(X\rho)]^{\alpha}$;
- $f_{\alpha,\alpha}(\rho,\sigma) = \min_{X>0} \operatorname{Tr}[\alpha\sigma^{\frac{1-\alpha}{\alpha}}X + (1-\alpha)(\rho^{-\frac{1}{2}}X\rho^{-\frac{1}{2}})^{\frac{\alpha}{\alpha-1}}]$;
- $f_{\alpha,\alpha}(\rho,\sigma) = \min_{X>0}[\operatorname{Tr}\sigma^{\frac{1-\alpha}{\alpha}}X]^{\alpha}[\operatorname{Tr}((\rho^{-\frac{1}{2}}X\rho^{-\frac{1}{2}})^{\frac{\alpha}{\alpha-1}}]^{1-\alpha}$.

Variational formulas for quantum α-z-fidelity

- Chehade (2020) used the classical matrix inequalities to prove that for $\alpha > 1$ and $z > 1$,

$$f_{\alpha,z}(\rho,\sigma) = \max_{X>0} P(X),$$

where

$$P(X) = z\,\mathrm{Tr}\left(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}X\right) - (z-1)\,\mathrm{Tr}\left(\left(\sigma^{\frac{z-1}{2z}}X\sigma^{\frac{z-1}{2z}}\right)^{\frac{z}{z-1}}\right).$$

- D., Le, Vuong (Int. J. Quantum Info, 2023): $f_{\alpha,z}(\rho,\sigma)$ is also the minimum of $P(X)$ when $0 < \alpha < z < 1$. In addition, $f_{\alpha,z}(\rho,\sigma) = \min_{X>0} Q(X)$, where

$$Q(X) = \left(\mathrm{Tr}(\sigma^{\frac{z-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{z-\alpha}{2z}}X)\right)^{z}.\left(\mathrm{Tr}(\sigma^{\frac{z-1}{2z}}X\sigma^{\frac{z-1}{2x}})^{\frac{z}{z-1}}\right)^{1-z}.$$

The α-z-fidelity between unitary orbits

- Let $U(\mathbb{H})$ be the set of $n \times n$ unitary matrices, and $\mathbb{D}_n$ the set of density matrices. For $\rho \in \mathbb{D}_n$, its unitary orbit is defined as

$$U_\rho = \{U\rho U^* : U \in U(\mathbb{H})\}.$$

- The study of "distance" between two quantum states under general local unitary dynamics was initiated by Zhang et. al. (2014). In particular, they found the maximum and minimum fidelity and relative entropy between two unitary orbits.
- Bhatia and Congedo (2019) investigated the optimization of several functions between unitary orbits including the Bures-Wasserstein distance, the Kullback-Leibler divergence, the Bhattacharyya divergence and the log-determinant divergence.
- Yan et. al. (2020) studied this problem for the quantum α-fidelity function.

The α-z-fidelity between unitary orbits

- Let ρ and $\sigma \in \mathbb{D}_n$, the α-z-fidelity $f_{\alpha,z}(\rho,\sigma) = \operatorname{Tr}\left(\sigma^{\frac{1-\alpha}{2z}}\rho^{\frac{\alpha}{z}}\sigma^{\frac{1-\alpha}{2z}}\right)^z$ between the unitary orbits U_ρ and U_σ satisfies

$$\max_{U \in U(\mathbb{H})} f_{\alpha,z}(\rho, U\sigma U^*) = \sum_{i=1}^n \lambda_i^\downarrow(\rho)^\alpha \lambda_i^\downarrow(\sigma)^{1-\alpha},$$

and

$$\min_{U \in U(\mathbb{H})} f_{\alpha,z}(\rho, U\sigma U^*) = \sum_{i=1}^n \lambda_i^\downarrow(\rho)^\alpha \lambda_i^\uparrow(\sigma)^{1-\alpha},$$

where $\lambda(\rho) = (\lambda_1, \ldots, \lambda_n)$ are the eigenvalues of ρ and $\lambda^\downarrow(\rho)$ (resp. $\lambda^\uparrow(\rho)$) is a rearrangement of $\lambda(\rho)$ in decreasing order (resp. increasing order).

The α-z-fidelity between unitary orbits

- D., Le, Vuong (Int. J. Quantum Info, 2023): For $0 \le \alpha \le z \le 1$,

$$\{f_{\alpha,z}(\rho, U\sigma U^*) : U \in U(\mathbb{H})\} = \Big[\sum_{i=1}^{n} \lambda_i^{\downarrow}(\rho)^{\alpha} \lambda_i^{\uparrow}(\sigma)^{1-\alpha}, \sum_{i=1}^{n} \lambda_i^{\downarrow}(\rho)^{\alpha} \lambda_i^{\downarrow}(\sigma)^{1-\alpha}\Big].$$

Hope you slept comfortably