

General Notes

Jayadev Naram

Contents

1	Inequalities for the Wasserstein Mean of Positive Definite Matrices	2
----------	--	----------

1 Inequalities for the Wasserstein Mean of Positive Definite Matrices

Definition 1.1. Let \mathbb{P} be the space of $n \times n$ complex positive definite matrices. The **Bures-Wasserstein distance** on \mathbb{P} is the metric defined as

$$d(A, B) = \left[\text{tr}(A + B) - 2\text{tr}(A^{1/2}BA^{1/2})^{1/2} \right]^{1/2}. \quad (1)$$

Definition 1.2. Let A_1, \dots, A_m be given positive definite matrices and let $w = (w_1, \dots, w_m)$ be a vector of weights, i.e., $w_j \geq 0$ and $\sum_{j=1}^m w_j = 1$. Then the **(weighted) Wasserstein mean**, or the **Wasserstein barycentre** of A_1, \dots, A_m is defined as

$$\Omega(w; A_1, \dots, A_m) = \arg \min_{X \in \mathbb{P}} \sum_{j=1}^m w_j d^2(X, A_j). \quad (2)$$

Remark 1.3. The function defined on \mathbb{P} by the sum on the right hand side of (2) has a unique minimizer.

Remark 1.4. The Wasserstein barycentre Ω defined in (2) is the unique positive definite solution of the equation

$$X = \sum_{j=1}^m w_j (X^{1/2} A_j X^{1/2})^{1/2}. \quad (3)$$

Consider the special case $m = 2$, writing $(A_1, A_2) = (A, B)$ and $(w_1, w_2) = (1 - t, t)$, where $0 \leq t \leq 1$. We rewrite the above equation as

$$\begin{aligned} X &= (1 - t)(X^{1/2} A X^{1/2})^{1/2} + t(X^{1/2} B X^{1/2})^{1/2} \\ \implies X^{1/4} X^{1/2} X^{1/4} &= X^{1/4} ((1 - t)A^{1/2} + tB^{1/2}) X^{1/4} \\ \implies X^{1/2} &= (1 - t)A^{1/2} + tB^{1/2} \\ \implies X &= ((1 - t)A^{1/2} + tB^{1/2})^2. \end{aligned}$$

Thus, we obtained an explicit formula for Ω . Denoting this by $A \diamond_t B$ we have

$$A \diamond_t B = (1 - t)^2 A + t^2 B + t(1 - t)[(AB)^{1/2} + (BA)^{1/2}]. \quad (4)$$

The metric d in (1) is the distance corresponding to an underlying metric on \mathbb{P} , and (4) is an equation for the geodesic segment between two points A and B in the manifold \mathbb{P} . The special choice $t = 1/2$ gives the midpoint of this geodesic. This is denoted by

$$A \diamond B = \frac{1}{4}[A + B + (AB)^{1/2} + (BA)^{1/2}], \quad (5)$$

and can be thought of as the **Wasserstein mean** of A and B .

Definition 1.5. The **Cartan metric**

$$\delta(A, B) = \|\log A^{-1/2} B A^{-1/2}\|_2, \quad (6)$$

where $\|\cdot\|_2$ is the Forbenius norm on matrices. The **weighted Cartan mean** (or the **weighted geometric mean**) of A_1, \dots, A_m is defined as

$$G(w; A_1, \dots, A_m) = \arg \min_{X \in \mathbb{P}} \sum_{j=1}^m w_j \delta^2(X, A_j). \quad (7)$$

Remark 1.6. The (unique) solution of this minimization problem is also the positive definite solution of the equation

$$\sum_{j=1}^m w_j \log(X^{-1/2} A_j X^{-1/2}) = 0. \quad (8)$$

Analogous to (4) the equation of the geodesic segment joining A and B with respect to the metric δ is

$$A \#_t B = A^{1/2} (A^{-1/2} B A^{-1/2})^t A^{1/2}, \quad 0 \leq t \leq 1. \quad (9)$$

This is also the t -weighted geometric mean of A and B . When $t = 1/2$, this reduces to

$$A \# B = A^{1/2} (A^{-1/2} B A^{-1/2})^{1/2} A^{1/2}, \quad (10)$$

and is called the geometric mean of A and B .