Sliced-Wasserstein on Symmetric Positive Definite Matrices for M/EEG Signals

Artyom Matveev

Moscow Institute of Physics and Technology

matveev.as@phystech.edu

October 17, 2023

Background[1]

- Space of all symmetric matrices: $S(n) = \{\mathbf{S} \in M(n), \mathbf{S}^T = \mathbf{S}\}$
- Space of all SPD matrices: $P(n) = \{\mathbf{P} \in S(n), \mathbf{u}^T\mathbf{P}\mathbf{u} > 0, \forall \mathbf{u} \in \mathbb{R}^n\}$, $P(n) \equiv \mathcal{M}(n)$, $\mathcal{M}(n)$ is a differentiable Riemannian manifold
- Spatial Covariance Matrix: $\Sigma = \mathbb{E}\{(\mathbf{x_t} - \mathbb{E}\{\mathbf{x_t}\})(\mathbf{x_t} - \mathbb{E}\{\mathbf{x_t}\})^T\}$, $\mathbf{x_t} \in \mathbb{R}^n$
- Matrix of trials: $\mathbf{X_i} = [\mathbf{x}_{\mathbf{t}+T_i}, \dots, \mathbf{x}_{\mathbf{t}+T_i+T_s-1}]$, $\mathbf{X_i} \in \mathbb{R}^{n \times T_s}$
- Sample Covariance Matrix: $\mathbf{P_i} = \frac{1}{T_s - 1}\mathbf{X_i}\mathbf{X_i}^T$, $\mathbf{P_i} \in \mathbb{R}^{n \times n}$

Riemannian Geodesic distance:

$$\delta_R(\mathbf{P_1}, \mathbf{P_2}) = \|\log(\mathbf{P_1}^{-1}\mathbf{P_2})\|_F = \left[\sum_{i=1}^{n} \log^2 \lambda_i\right]^{1/2}$$

Riemannian mean:

$$\mathfrak{G}(\mathbf{P}_1, \dots, \mathbf{P}_N) = \underset{\mathbf{P} \in P(n)}{\arg\min} \sum_{i=1}^{N} \delta_R^2(\mathbf{P}, \mathbf{P}_i)$$

[1] **Barachant, A.**, Bonnet, S., et al. Multiclass brain-computer interface classification by Riemannian geometry. 2012

Euclidean Wasserstein distance[2]

For $\mu, \nu \in \mathcal{P}_p(\mathbb{R}^d)$ two measures with finite moments of order $p \geq 1$, the Wasserstein distance is defined as

$$W_p^p(\mu, \nu) = \inf_{\gamma \in \Pi(\mu, \nu)} \int \|x - y\|_2^p \,\mathrm{d}\gamma(x, y), \tag{1}$$

where $\Pi(\mu, \nu) = \{\gamma \in \mathcal{P}(\mathbb{R}^d \times \mathbb{R}^d), \pi^1_{\#}\gamma = \mu, \pi^2_{\#}\gamma = \nu\}$ denotes the set of couplings between μ and ν, $\pi^1 \colon (x, y) \mapsto x$ and $\pi^2 \colon (x, y) \mapsto y$ the projections on the first and second coordinate and $\#$ is the push-forward operator, defined as a mapping on all borelian $A \subset \mathbb{R}^d$, such that $T_{\#\mu}(A) = \mu(T^{-1}(A))$.
The computation complexity of (1) is $\mathcal{O}(n^3 \log n)$.

[2] **Bonet, C.**, Malézieux, B., et al. Sliced-Wasserstein on Symmetric Positive Definite Matrices for M/EEG Signals. 2023

Euclidean Sliced-Wasserstein distance

The average of the Wasserstein (SW) distance between one dimensional projections of the measures in all directions, i.e. for $\mu, \nu \in \mathcal{P}_p(\mathbb{R})$,

$$SW_p^p(\mu, \nu) = \int_{S^{d-1}} W_p^p(t_\#^\theta \mu, t_\#^\theta \nu)\, \mathrm{d}\lambda(\theta), \tag{2}$$

where λ is the uniform distribution on the sphere $S^{d-1} = \{\theta \in \mathbb{R}, \|\theta\|_2 = 1\}$ and t^θ is the coordinate of the projection on the line span(θ), i.e. $t^\theta(x) = \langle x, \theta \rangle$ for $x \in \mathbb{R}^d$, $\theta \in S^{d-1}$. Computation complexity here is $\mathcal{O}(L \cdot n \cdot (d + \log n))$ with L projections using Monte-Carlo method.

Definition

Let λ_S be the uniform distribution on $\{A \in S_d(\mathbb{R}), \|A\|_F = 1\}$. Let $p \geq 1$ and $\mu, \nu \in \mathcal{P}_p(S_d^{++}(\mathbb{R}))$, then the SPDSW discrepancy is defined as

$$SPDSW_p^p(\mu, \nu) = \int_{S_d(\mathbb{R})} W_p^p(t_\#^A \mu, t_\#^A \nu)\, \mathrm{d}\lambda_S(A) \tag{3}$$