Classication of SPD matrices using a Riemannian-based kernel

Anna Kuzina

Skoltech, HSE

anna.kuzina@skoltech.ru

June 8, 2018

Motivation

Where do SPD matrices come from?

- fMRI (functional magnetic resonance imaging) enables us to build a graph, representing connections of different parts of the brain

- Structural connectome can be seen as a symmetric adjacency matrix

- We want to take into account geometry of the data

From SI to SPD

1. Laplacian of the SI adjacency matrix — SPSD matrix

$$L_S = D_S - S$$

2. Add regularization

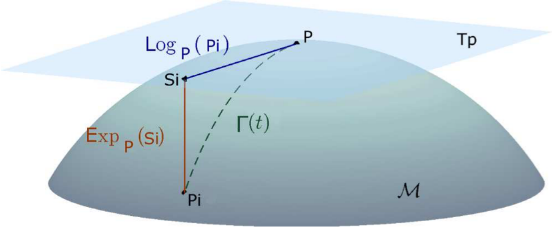
$$P = L_S + \lambda I, \; \lambda > 0$$

Basic notations

- $\mathcal{M}(n)$ — space of $n \times n$ real matrices
- $\{S \in \mathcal{M}(n) : S^T = S\}$ — space of $n \times n$ symmetric matrices
- $\{P \in S(n) : u^T P u > 0 \; \forall u \in \mathbb{R}^n\}$ — space of $n \times n$ SPD matrices
- $\exp(S) = C \, \mathrm{diag}(\exp(\lambda_1), \ldots, \exp(\lambda_n)) \, C^T \in P(n)$
- $\log(P) = C \, \mathrm{diag}(\log(\lambda_1), \ldots, \log(\lambda_n)) \, C^T \in S(n)$
- $P^{\frac{1}{2}}$ — symmetric matrix A, s.t. $AA = P$

Riemannian manifold

- Smooth manifold
- Tangent finite-dimensional Euclidean space at every point with inner product g_P
- g_P, Riemannian metric, varies smoothly from point to point



Affine-invariant metrics

Scalar product on the Tangent space at identity generates metrics on the whole Lie group.

Dot product at Identity

$$\langle S_1, S_2 \rangle_I = \mathrm{Tr}(S_1 S_2^T) = \mathrm{Tr}(S_1 S_2), \quad S_1, S_1 \in T_I \mathcal{M}$$

Action of Linear group on a Lie group

$$A \odot P = APA^T, \ A \in GL(n)$$

Invariance under congruent transformations

$$\forall g \in T_P \quad \langle S_1, S_2 \rangle_P = \langle g \odot S_1, g \odot S_2 \rangle_{g \odot P}$$

Affine-invariant metric

$$\langle S_1, S_2 \rangle_P = \mathrm{Tr}(S_1 P^{-1} S_2 P^{-1}) = \mathrm{Tr}(P^{-1} S_1 P^{-1} S_2)$$

Riemannian metrics

Induced Norm

$$\|S\|_P^2 = \langle S, S\rangle_P = \mathrm{Tr}((SP^{-1})^2)$$

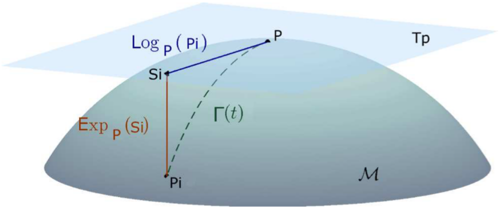
$P = I$:

$$\|S\|_I^2 = \mathrm{Tr}(S^2) = \mathrm{Tr}(S^T S) = \|S\|_F^2 = \|\mathrm{vect}(S)\|_2^2$$

$$\mathrm{vect}(S) = \left[S_{1,1}, \sqrt{2}S_{1,2}, S_{2,2}, \sqrt{2}S_{1,3}, \sqrt{2}S_{2,3}, S_{3,3}, \ldots, S_{n,n}\right]^T$$

Riemannian Geodesic distance

The geodesic — shortest path between two points
Riemannian distance — length of the geodesic



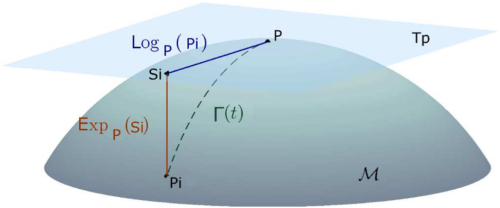
Riemannian Geodesic distance

The geodesic — shortest path between two points
Riemannian distance — length of the geodesic

$$\Gamma(t) : [0,1] \to P(n)$$

Norm of the tangent vector — instantaneous speed of the geodesic

$$L(\Gamma(t)) = \int_0^1 \|\dot{\Gamma(t)}\|_{\Gamma(t)} dt$$



Geodesic

For the invariant metric, geodesic is generated by the action of one-parameter subgroup

One-parameter subgroup

$$x = \exp(tS)$$

Geodesic going through I with tangent vector S

$$\Gamma_{I,S}(t) = \exp(tS)^{\frac{1}{2}} I \exp(tS)^{\frac{1}{2}} = \exp(tS)$$

Geodesic

Geodesic, starting from identity

$$\Gamma_{I,S}(t) = \exp(tS)$$

Using invariance

$$A \odot \Gamma_{P,S}(t) = \Gamma_{A\odot P, A\odot S}(t), \quad A = P^{-\frac{1}{2}}$$

$$\Gamma_{P,S}(t) = P^{\frac{1}{2}}\Gamma_{I, A^{-\frac{1}{2}}SA^{-\frac{1}{2}}}(t)P^{\frac{1}{2}} = P^{\frac{1}{2}}\exp(tP^{-\frac{1}{2}}SP^{-\frac{1}{2}})P^{\frac{1}{2}}$$

Exponential map

Exponential mapping (from tangent space to the manifold)

$$\exp_P(S_i) = P_i = P^{\frac{1}{2}} \exp(P^{-\frac{1}{2}} S_i P^{-\frac{1}{2}}) P^{\frac{1}{2}}$$

Logarithmic mapping (from the manifold on the tangent space)

$$\log_P(P_i) = S_i = P^{\frac{1}{2}} \log(P^{-\frac{1}{2}} P_i P^{-\frac{1}{2}}) P^{\frac{1}{2}}$$

Riemannian distance

$$L(\Gamma(t)) = \int_0^1 \|\dot{\Gamma(t)}\|_{\Gamma(t)} dt = \int_0^1 \sqrt{\langle \dot{\Gamma(t)}, \dot{\Gamma(t)} \rangle_{\Gamma(t)}} dt =$$
$$= \int_0^1 \sqrt{Tr\left((\dot{\Gamma(t)}\Gamma(t)^{-1})^2\right)} dt$$

Again we start from the identity.
Having $S = C\text{diag}(\lambda_i)C^T$, where C is a unitary matrix:

$$\dot{\Gamma}_{I,S}(t) = C\text{diag}(\lambda_i \exp(t\lambda_i))C^T$$
$$\Gamma_{I,S}(t)^{-1} = C\text{diag}(\exp(-t\lambda_i))C^T$$
$$\dot{\Gamma}_{I,S}(t)\Gamma_{I,S}(t)^{-1} = C\text{diag}(\lambda_i \exp(t\lambda_i))\text{diag}(\exp(-t\lambda_i))C^T =$$
$$= C\text{diag}(\lambda_i)C^T = S$$

Riemannian distance

$$L(\Gamma(t)) = \int_0^1 \sqrt{Tr\left((\dot{\Gamma(t)}\Gamma(t)^{-1})^2\right)}dt$$

Distance to the Identity

$$\sigma_R(I, P) = L(\Gamma_{I,S}(t)) = \int_0^1 \sqrt{Tr\left(S^2\right)}dt = \|S\|_F = \|\log(P)\|_F$$

Affine invariance

$$\sigma_R(P_1, P_2) = \sigma_R(A \odot P_1, A \odot P_1), \quad A = P_1^{-\frac{1}{2}}$$

$$\sigma_R(P_1, P_2) = \sigma_R(I, P_1^{-\frac{1}{2}} P_2 P_1^{-\frac{1}{2}}) = \|\log(P_1^{-1}P_2)\|_F = \left(\sum \log^2 \lambda_i\right)^{\frac{1}{2}}$$

Classification in the Riemannian manifold

- Minimum distance to Riemanian Mean

Geometric mean

$$P^* = \arg\min_P \sum_{i=1}^{m} \sigma_R^2(P, P_i)$$

- kNN with Riemannian distances

Kernel Approach

Mapping function

$$\phi(P_i) = \log_{P^*}(P_i)$$

Riemannian base kernel

$$k_R(P_i, P_j) = \langle \phi(P_i), \phi(P_j) \rangle_{P^*} = \mathrm{Tr}(\log_{P^*}(P_i) P^{*-1} \log_{P^*}(P_j) P^{*-1}) =$$

$$= \mathrm{Tr}(\log(P^{*-\frac{1}{2}} P_i P^{*-\frac{1}{2}}) \log(P^{*-\frac{1}{2}} P_j P^{*-\frac{1}{2}})) = \mathrm{Tr}(\tilde{P}_i \tilde{P}_j) =$$

$$= \langle \tilde{P}_i, \tilde{P}_j \rangle_F = \mathrm{vect}(\tilde{P}_i)^T \mathrm{vect}(\tilde{P}_j) = \langle \mathrm{vect}(\tilde{P}_i), \mathrm{vect}(\tilde{P}_j) \rangle_2$$

Reference point

$$k_R(P_i, P_j) = \mathrm{Tr}(\log(P^{*-\frac{1}{2}} P_i P^{*-\frac{1}{2}}) \log(P^{*-\frac{1}{2}} P_j P^{*-\frac{1}{2}})) = \mathrm{Tr}(\tilde{P}_i \tilde{P}_j)$$

Reference point

$$k_R(P_i, P_j) = \mathrm{Tr}(\log(P^{*-\frac{1}{2}} P_i P^{*-\frac{1}{2}}) \log(P^{*-\frac{1}{2}} P_j P^{*-\frac{1}{2}})) = \mathrm{Tr}(\tilde{P}_i \tilde{P}_j)$$

log-Euclidean kernel

$$P^* = I$$

$$k_{LE}(P_i, P_j) = \mathrm{Tr}(\log(P_i) \log(P_j))$$

Reference point

$$k_R(P_i, P_j) = \mathrm{Tr}(\log(P^{*-\frac{1}{2}} P_i P^{*-\frac{1}{2}}) \log(P^{*-\frac{1}{2}} P_j P^{*-\frac{1}{2}})) = \mathrm{Tr}(\tilde{P}_i \tilde{P}_j)$$

log-Euclidean kernel

$$P^* = I$$

$$k_{LE}(P_i, P_j) = \mathrm{Tr}(\log(P_i) \log(P_j))$$

Geometric mean

$$P^* = \arg\min_P \sum_{i=1}^{m} \sigma_R^2(P, P_i)$$

Riemannian Network

BiMap Layer

$$X_k = W_k X_{k-1} W_k^T$$

W_k — row full-rank matrix

ReEig Layer

$$X_k = U_{k-1} \max(\varepsilon I, \Sigma_{k-1}) U_{k-1}^T$$

$X_{k-1} = U_{k-1} \Sigma_{k-1} U_{k-1}^T$

LogEig Layer

Projection on the tangent plane at Identity

$$X_k = U_{k-1} \log(\Sigma_{k-1}) U_{k-1}^T$$

References

M. Moakher (2005)
Symmetric Positive-Denite Matrices: From Geometry to Applications and Visualization

X. Pennec et al. (2006)
A Riemannian Framework For Tensor Computing

A. Barachant et al. (2010)
Riemannian Geometry Applied to BCI Classification

A. Barachant et al.(2012)
Multiclass Brain-Computer Interface Classication by Riemannian Geometry

Z. Huang, (2017)
A Riemannian Network for SPD Matrix Learning

A. Pimkin, M. Belyaev et al. (2017)
Classification of brain network structures using analysis of symmetric semidefinite matrices

The End