



HAL
open science

A Julia Package for manipulating Brain-Computer Interface Data in the Manifold of Positive Definite Matrices

Marco Congedo, Saloni Jain

► **To cite this version:**

Marco Congedo, Saloni Jain. A Julia Package for manipulating Brain-Computer Interface Data in the Manifold of Positive Definite Matrices. IEEE SMC 2019 International Conference on Systems, Man, and Cybernetics, Oct 2019, Bari, Italy. hal-02315153

HAL Id: hal-02315153

<https://hal.science/hal-02315153>

Submitted on 14 Oct 2019

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

A Julia Package for manipulating Brain-Computer Interface Data in the Manifold of Positive Definite Matrices

M. Congedo, *Member, IEEE*, S. Jain

Abstract— The Riemannian geometry of positive definite matrices yields state-of-the-art classification accuracy for brain-computer interface (BCI) data. The use of this framework is steadily increasing in the BCI community, sustained by its excellent classification accuracy and ability to operate transfer learning. Currently, open-source code libraries exist for the Matlab and Python programming language. Julia is a young open-source cross-platform language specifically conceived for scientific computing, which is rapidly gaining momentum in the data science community thanks to its efficiency and compatibility with the best available computing protocols. By means of this article we present and release a state-of-the-art open-source Julia package for the Riemannian geometry of positive definite matrices, named *PosDefManifold*. It supports nine metrics for the manifold of both real and complex positive definite matrices and includes all fundamental tools for manipulating data in them.

I. INTRODUCTION

Riemannian geometry is a branch of differential geometry that studies smooth manifolds, curved spaces with peculiar geometries. In these spaces notions of angles, shortest path between two points, distances, center of mass of several points, etc., allow to study analytic properties of mathematical operators from a geometrical perspectives. In the field of brain-computer interface (BCI) the manifold of positive definite matrices (PDMs) [1] has proved particularly useful, since multivariate EEG (electroencephalography) data in finite time windows can effectively be mapped as points onto this manifold through the estimation of some form of their covariance matrix [2-5]. This approach has led to the introduction of classifiers with remarkable characteristics as compared to the state-of-the-art. Among these characteristics we find the simplicity of use, the excellent classification accuracy, as demonstrated by the winning score obtained in five international BCI data classification competitions [6] and the ability to operate transfer learning [7-9]. For a formal introduction to the manifold of PDMs the reader is referred to [1], while for a review of its use in BCIs to [2, 6, 10].

Julia has been created at the Massachusetts Institute of Technology (MIT) to provide data scientist with an open-source and cross-platform language that is both high-level (e.g., like Matlab) and fast (e.g., like C and Fortran) [11]. Typically, Julia code runs at a speed within a factor of two relative to fully optimized C code, thus in general it is an order of magnitude faster as compared to Python or R. However, coding itself in Julia is as concise and pleasant as coding in

other high-level languages. Julia is an object-oriented, dynamic and compiled language. Its main characteristics are:

- Automatic generation of efficient and specialized code for different argument types.
- Multiple dispatch mechanism, providing the ability to define function behavior across many combinations of argument types.
- Efficient standard libraries for floating-point calculations, linear algebra, random number generation, and regular expression matching.
- Support for concurrent, parallel and distributed computing.
- Support for direct calling of C and Fortran libraries without wrappers or special APIs and easy calls to Python functions.
- User-defined types as fast and compact as built-ins ones.
- Elegant and extensible conversions and promotions for numeric and other types.
- Support for Unicode, including but not limited to UTF-8.
- Garbage-collection.
- An efficient built-in package manager.
- Lisp-like macros and other metaprogramming facilities.

Officially launched in 2012, v1.0 of Julia has been released in 2018, providing a stable version of the language. Currently, open-source code libraries for manipulating data in the Riemannian manifold of PDMs exist for Matlab [13] and Python [14]. The objective of this article is to present an open-source package for Julia, named *PosDefManifold*.

II. THE MANIFOLD OF POSITIVE DEFINITE MATRICES

A. Introduction

We are interested in operations on the metric space (\mathbf{P}_n, δ) of $n \times n$ positive definite matrices endowed with a distance or symmetric divergence δ , which hereafter with an abuse of terminology we will collectively name ‘metric’. A smooth manifold in differential geometry is a topological space that is locally similar to the Euclidean space and has a globally defined differential structure. The *tangent space* at point G is the vector space containing the tangent vectors to all curves on the manifold passing through G (Fig. 1). A smooth Riemannian manifold is equipped with an inner product on the tangent space, named a Riemannian metric, defined at each point and varying smoothly from point to point. Thus, a Riemannian metric turns the metric space (\mathbf{P}_n, δ) into a *Riemannian manifold*. This is the case, for example, of the

M. Congedo is with the GIPSA-lab, Grenoble Alpes University, CNRS, Grenoble-INP, Grenoble, FRANCE (phone: +33 4 76 82 62 52; e-mail: marco.congedo@gipsa-lab.fr).

S. Jain is with the Indian Institute of Technology, Kharagpur, INDIA (e-mail: tosalonijain@gmail.com) and has performed this research while visiting the GIPSA-lab.

Fisher metric, which has a central role in \mathbf{P}_n [7] and of the Wasserstein metric, fundamental in optimal transport theory.

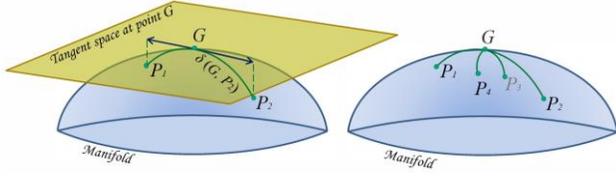


Figure 1. Schematic illustration of the Riemannian manifold of positive definite matrices. Left: geodesic (green curve) relying points P_1 and P_2 passing through its-mid-point (mean) G , tangent space at point G with tangent vectors to geodesic from G to P_1 and from G to P_2 (blue arrowed lines) and distance $\delta(G, P_2)$. Right: the *center of mass* (also named *mean*) G of points P_1, \dots, P_4 defined as the point minimizing the sum of the four squared distances $\delta^2(G, P_i)$, for $i=\{1, \dots, 4\}$.

B. Geodesic

The *geodesic* is the shortest path joining two points on the manifold, analogous to straight lines in the Euclidean space (Fig. 1). The points along the geodesic between P and Q with arc-length a are denoted $\gamma_a(P, Q)$. They can be understood as *weighted means* of P and Q . For example, the geodesic equation according to the Euclidean metric is $(1-a)P + aQ$, which is the way weighted means of two matrices are usually defined.

C. Distance

The *distance* $\delta(P, Q)$ is the length of the geodesic (at constant velocity) between points P and Q (Fig. 1).

D. Distance from I

The manifold \mathbf{P}_n endowed with the Fisher metric is symmetric around its origin, which is the identity matrix I_n . The distance between a point P and the origin, *i.e.*, $\delta(P, I)$, is analogous therein to the length of vectors in Euclidean space.

E. Mean

The mid-point along the geodesic relying P and Q is named the *mean* of P and Q . Using the Euclidean metric this is the usual arithmetic mean and using the inverse Euclidean metric this is the harmonic mean. Those are straightforward extensions of their scalar counterparts - see (5) and (9). Using the Fisher metric, the mid-point of the geodesic allows the proper generalization to matrices of the scalar *geometric mean* [1]. Other metrics allows other definitions of means, as we will see.

F. Fréchet Mean

Using Fréchet's variational approach we can extend to PDMs the concept of mean for a set of scalars [6]; as the midpoint G on the geodesic relying P_1 and P_2 is the minimizer of $\delta^2(P_1, G) + \delta^2(P_2, G)$ (Fig1: left), so the Fréchet mean $G(\mathbf{P}, w)$ of a set $\mathbf{P} = \{P_1, P_2, \dots, P_k\}$ with (optional) associated non-negative real weights summing up to 1, such as $w = \{w_1, w_2, \dots, w_k\}$: $\sum_i w_i = 1$ for $i=\{1, \dots, k\}$, is defined as the matrix G verifying (Fig1: right)

$$\operatorname{argmin}_G \sum_{i=1}^k w_i \delta^2(P_i, G). \quad (1)$$

Thus, every distance (or divergence) function induces a Fréchet mean. The ability to define means (or centers of mass)

for an arbitrary number of points in \mathbf{P}_n (typically those are observations) is exploited by all Riemannian classifiers.

G. Invariance Properties

The choice of the metric for \mathbf{P}_n engenders *invariance properties* (isometries) on the distance, which are in turn inherited by the mean. Let us here denote the set \mathbf{P} by $\{P_i\}$, where $i=\{1, \dots, k\}$ and by $G\{\mathbf{P}_i\}$ the mean of such set. Let us here also drop the weights for convenience; the fundamental invariance properties are summarized in Table 1.

TABLE I. INVARIANCE PROPERTIES INDUCED BY METRICS

	<i>Distance</i>	<i>Mean</i>
<i>Rotation</i>	$\delta(P, Q) = \delta(U^* P U, U^* Q U)$	$G\{U^* P_i U\} = U^* G\{P_i\} U$
<i>Affine</i>	$\delta(P, Q) = \delta(B^* P B, B^* Q B)$	$G\{B^* P_i B\} = B^* G\{P_i\} B$
<i>Inversion</i>	$\delta(P, Q) = \delta(P^{-1}, Q^{-1})$	$G\{P_i^{-1}\} = G^{-1}\{P_i\}$

for any unitary matrix U and non-singular matrix B

The affine invariance implies the rotation invariance and is also named *congruence invariance*. Invariance properties are instrumental to the performance of Riemannian classifiers [6].

III. IMPLEMENTED METRICS

A. Introduction

We are interested in distance or symmetric divergence functions, the difference between the two is that a symmetric divergence does not need to satisfy the triangle inequality. One can find several distances and divergences in the literature and they often turn out to be closely related to each other, see for example [14-16]. Nine of them are implemented in *PosDefManifold* and two of them are Riemannian metrics (the Fisher and Wasserstein metric). We give here the characterizations of their induced *distance* of a point P from the origin $\delta(P, I)$ and between two points P and Q $\delta(P, Q)$, the *geodesic* relying P to Q $\gamma_a(P, Q)$ and the *weighted Fréchet mean* $G(\mathbf{P}, w)$ of a set of $k > 2$ points $\{P_1, P_2, \dots, P_k\}$ with associated weights $\{w_1, w_2, \dots, w_k\}$ summing up to 1. For convenience, in the following let $b=1-a$.

B. Euclidean metric

$$\delta^2(P, I) \quad \|P - I\|^2 \quad (2)$$

$$\delta^2(P, Q) \quad \|P - Q\|^2 \quad (3)$$

$$\gamma_a(P, Q) \quad bP + aQ \quad (4)$$

$$G(\mathbf{P}, w) \quad \sum_{i=1}^k w_i P_i \quad (5)$$

This is the classical Euclidean distance leading to the usual arithmetic mean. It verifies only the rotation invariance; however, the mean also verifies the congruence invariance.

C. Inverse Euclidean metric

$$\delta^2(P, I) \quad \|P^{-1} - I\|^2 \quad (6)$$

$$\delta^2(P, Q) \quad \|P^{-1} - Q^{-1}\|^2 \quad (7)$$

$$\gamma_a(P, Q) \quad (bP^{-1} + aQ^{-1})^{-1} \quad (8)$$

$$G(\mathbf{P}, w) \quad \left(\sum_{i=1}^k w_i P_i^{-1} \right)^{-1} \quad (9)$$

This is the usual harmonic distance leading to the harmonic mean. It verifies only the rotation invariance; however, the mean also verifies the congruence invariance.

D. log Euclidean metric

$$\delta^2(P, I) \quad \|\log P\|^2 \quad (10)$$

$$\delta^2(P, Q) \quad \|\log P - \log Q\|^2 \quad (11)$$

$$\gamma_a(P, Q) \quad \exp(b \log P + a \log Q) \quad (12)$$

$$G(\mathbf{P}, w) \quad \exp\left(\sum_{i=1}^k w_i \log P_i\right) \quad (13)$$

This metric enjoys the rotation and inversion invariance. The log-Euclidean distance to I is the same as per the Fisher metric. The log Euclidean mean has the same determinant as the Fisher mean, and trace equal or superior to the trace of the Fisher mean. A minimum-trace log Euclidean mean approximating well the Fisher mean has been proposed in [19]. If matrices $\{P_1, P_2, \dots, P_k\}$ all pair-wise commute, then this metric coincides with the Fisher metric [17-18].

E. Fisher metric

$$\delta^2(P, I) \quad \|\log P\|^2 \quad (14)$$

$$\delta^2(P, Q) \quad \|\log(P^{\times} Q P^{\times})\|^2 \quad (15)$$

$$\gamma_a(P, Q) \quad P^{1/2} \left(P^{-1/2} Q P^{-1/2} \right)^a P^{1/2} \quad (16)$$

The Fisher metric, also known as *Rao, affine-invariant* and *natural* metric, among others, has a paramount importance for the \mathbf{P}_n manifold, standing out as the natural choice both from the perspective of differential geometry and information geometry. Endowed with the Fisher metric \mathbf{P}_n is a Riemannian manifold of non-positive curvature and is symmetric. This metric verifies all invariances listed in Table 1.

The geometric mean G of two matrices P and Q , that is, (16) with $a = 1/2$, is denoted shortly as $P\#Q$. Currently, this is an object of intense study because of its interesting mathematical properties [1]. Among others:

- it is the unique solution to Riccati equation $GQ^{-1}G=P$
- it is equal to $F^{-H}(D_1 D_2)^{1/2} F^{-1}$ for whatever joint diagonalizer F of P and Q , i.e., for whatever matrix F satisfying $F^H P F = D_1$ and $F^H Q F = D_2$, where D_1 and D_2 are positive definite diagonal matrices [19] and superscript H denotes complex conjugate transpose
- it enjoys all 10 properties of means postulated in the seminal work [23].

When P and Q commute, the geometric mean of two matrices reduces to $P^{1/2} Q^{1/2}$, which indeed in this case is the (equal-weight) log Euclidean mean $\exp(1/2 \log P + 1/2 \log Q)$ (13).

The Fréchet weighted mean does not have close-form solution in general. However it always exists. It is the unique PDM G satisfying [20, 21]

$$\sum_{i=1}^k w_i \log \left(G^{-1/2} P_i G^{-1/2} \right) = 0. \quad (17)$$

For estimating it, *PosDefManifold* uses a well-known gradient descent algorithm [2], but with a variable step-size. This mean

is known under many different names (*Pusz-Woronowicz, Cartan, Fréchet, Karcher, geometric, ...*).

The Fisher geodesic equation (16) is usually denoted $P\#_a Q$. Note that $I\#_a P = P^a$ and $P\#_a I = P^b$, where $b=1-a$. Also, it holds $P\#_a Q = Q\#_b P$ and $(P\#_a Q)^{-1} = P^{-1}\#_a Q^{-1}$. An interesting property of the Fisher metric is that using its geodesic equation we can extrapolate PDMs, always remaining in \mathbf{P}_n . That is to say, referring to (16), with:

- $0 < a < 1$ we move toward Q (attraction),
- $a > 1$ we move over and beyond Q (extrapolation),
- $a < 0$ we move back away from Q (repulsion).

The arithmetic, harmonic and geometric mean we have encountered are all members of the 1-parameter family of *power means* with parameter $p \in [-1, 1]$ introduced by [24] to generalize the concept of power means of scalars, which are also known as *Hölder means* or *generalized means*. The family of power means G with parameter p satisfies equation

$$G = \sum_{i=1}^k w_i \left(G \#_p P_i \right), \quad (18)$$

where $G\#_p P_i$ is the common notation for the Fisher geodesic equation (16). In particular:

- with $p = -1$ (18) yields the harmonic mean (9)
- with $p = +1$ (18) yields the arithmetic mean (5)
- at the limit of p evaluated at zero from both side (18) yields the geometric mean satisfying (17).

Thus, the family of power means (18) continuously interpolate between the arithmetic and harmonic mean passing through the geometric mean, as it is the case for scalars. All power means enjoy the congruence invariance, but only the geometric mean enjoys also the inversion invariance. When the matrices in a set all pairwise commute, it has been proved in [24] (see Property 1, p. 1502) that their power means are

$$\left(\sum_{i=1}^k w_i P_i^p \right)^{1/p}, \quad (19)$$

which are the straightforward extensions of scalar power means to matrices. As usual, such straightforward extensions work in commuting algebra, but not in general. In *PosDefManifold* we name *generalized means* expressions (19) and *power means* the solutions to (18). Both have a parameter $p \in [-1, 1]$. For the latter, the solution is implemented via the fixed-point MPM algorithm given in [22].

If the matrices in a set all pair-wise commute, the special case $p=1/2$ of (19) yields a mean proposed in a different context by [25] as a modified *Bhattacharyya mean*, a mean that we will encounter next under the name of *logdet-zero* mean.

F. logdet-zero metric

$$\delta^2(P, I) \quad \log \det \frac{1}{2} (P + I) - \frac{1}{2} \log \det (P) \quad (20)$$

$$\delta^2(P, Q) \quad \log \det \frac{1}{2} (P + Q) - \frac{1}{2} \log \det (PQ) \quad (21)$$

The logdet-zero divergence, also known as the square of the *Bhattacharyya divergence* [25], *Stein divergence* [26], *S-divergence* [16], or the log determinant *a-divergence* (with

$\alpha=0$) [14] is a Jensen-Bregman divergence enjoying all invariances listed in Table 1. Its square root has been shown to be a distance [16]. It behaves very similarly to the Fisher metric at short distances [15, 16, 25, 26] and the mean of two matrices it induces coincides with their geometric mean [25]. Thus, it is often used instead of the Fisher metric because it allows more efficient calculations.

The weighted mean does not have closed-form expression in general. It is the unique PDM G satisfying

$$\sum_{i=1}^k w_i \left(\frac{1}{2} P_i + \frac{1}{2} G \right)^{-1} = G^{-1}. \quad (22)$$

For estimating it *PosDefManifold* implements the fixed-point iterative algorithm proposed in [25]. The specific case $k=2$ yields solutions to the geodesic equation.

The logdet α -divergence family and associated means for α in $[-1, 1]$ [14] allows:

- the logdet-zero mean for $\alpha = 0$,
- the *left Kullback-Leibler* mean for $\alpha = -1$,
- the *right Kullback-Leibler* mean for $\alpha = 1$.

We do not consider the left and right Kullback-Leibler divergences because the related means are trivially the arithmetic and harmonic one, respectively [25]. As per the *symmetrized Kullback-Leibler divergence*, this is known as the *Jeffrey divergence*, which we consider next.

G. Jeffrey metric

$$\delta^2(P, I) = \frac{1}{2} \text{tr}(P + P^{-1}) - n \quad (23)$$

$$\delta^2(P, Q) = \frac{1}{2} \text{tr}(Q^{-1}P + P^{-1}Q) - n \quad (24)$$

This is a Jensen-Bregman divergence known as the symmetrized Kullback-Leibler divergence (see logdet-zero) [27]. This metric verifies all invariances listed in Table 1. As for the logdet-zero metric, the mean of two matrices it induces coincides with their geometric mean.

The *weighted mean* is the geometric mean of the weighted arithmetic and the weighted harmonic mean [25, 27]. We therefore use this definition for obtaining its *geodesics*. Note that when $k>2$, for scalars, but not for matrices in general, the geometric mean is the geometric mean of the arithmetic mean and harmonic mean.

H. Wasserstein metric

$$\delta^2(P, I) = \text{tr}(P + I) - 2\text{tr}(P^{1/2}) \quad (25)$$

$$\delta^2(P, Q) = \text{tr}(P + Q) - 2\text{tr}(P^{1/2}QP^{1/2})^{1/2} \quad (26)$$

$$\gamma_\alpha(P, Q) = b^2P + a^2Q + ab \left[(PQ)^{1/2} + (QP)^{1/2} \right] \quad (27)$$

This is an extension to matrices of the *Hellinger divergence* for vectors and is known as the *Bures distance* in quantum physics, where it is applied on density matrices (unit trace PDMs). It is also the metric naturally arising in the domain of *optimal transport*. It enjoys only the rotation invariance. Endowed with the Wasserstein metric the manifold \mathbf{P}_n has a Riemannian geometry of non-negative curvature. See for example [28].

The weighted mean does not have closed-form expression in general. It is the unique PDM G satisfying [29]

$$G = \sum_{i=1}^k w_i \left(G^{1/2} P_i G^{1/2} \right)^{1/2}. \quad (28)$$

For estimating it, *PosDefManifold* implements the fixed-point iterative algorithm given in [30]. If the matrices in a set all pair-wise commute, the Wasserstein mean coincides with the $p=1/2$ instance of both the generalized means (19) and the power means (18).

I. Cholesky Euclidean Metric

Let L_P be the lower triangular Cholesky factor of P , then

$$\delta^2(P, I) = \|L_P - I\|_F^2 \quad (29)$$

$$\delta^2(P, Q) = \|L_P - L_Q\|_F^2 \quad (30)$$

$$\gamma_\alpha(P, Q) = (bL_P + aL_Q) (bL_P^* + aL_Q^*) \quad (31)$$

$$G(P, w) = \left(\sum_{i=1}^k w_i L_{P_i} \right) \left(\sum_{i=1}^k w_i L_{P_i}^* \right) \quad (32)$$

This metric does not enjoy any invariance property, thus is rarely used (e.g., [31]).

L. von Neumann metric

$$\delta^2(P, I) = \frac{1}{2} \text{tr}(P \log P - \log P) \quad (33)$$

$$\delta^2(P, Q) = \frac{1}{2} \text{tr}(P \log P - P \log Q + Q \log Q - Q \log P) \quad (34)$$

This is a Jensen-Bregman divergence [32], again used in physics, which so far has been rarely used is signal processing and machine learning. It enjoys the rotation invariance. The geodesic and weighted mean for the von Neumann metric have not been defined to the best of the authors' knowledge.

IV. OTHER IMPLEMENTATIONS

Besides functions for computing geodesics, distances, (weighted) means and (weighted) Fréchet means for all above metrics, *PosDefManifold* implements several useful tools:

A. Tangent space operations

Since the tangent space is an Euclidean space, projecting onto it data from \mathbf{P}_n allows the application of classical machine learning methods such as SVM, logistic regression, deep learning, etc., which are usually formulated for the Euclidean geometry taking as input *feature vectors* (and not PDMs). It is actually using this procedure that five international machine learning competitions have been won using Riemannian geometry [6].

For any point P in \mathbf{P}_n , the *logarithmic map* projects P onto the tangent space at base point G , such as

$$\text{Log}_G(P) = S = G^{1/2} \log(G^{-1/2} P G^{-1/2}) G^{1/2}. \quad (35)$$

This results in a symmetric (or Hermitian) matrix S . In order to obtain features vectors for Euclidean machine learning algorithms, its lower (or upper) triangular part is vectorized assigning weight 1.0 to the diagonal elements and $\sqrt{2}$ to the non-diagonal elements [2].

The *exponential map* is the reverse operation, *i.e.*, projecting a point from the tangent space at base point G onto \mathbf{P}_n , as

$$\text{Exp}_G(S) = P = G^{\frac{1}{2}} \exp(G^{-\frac{1}{2}} S G^{-\frac{1}{2}}) G^{\frac{1}{2}}. \quad (36)$$

This results in a PDM. The reverse of the aforementioned vectorization operation is available as well in *PosDefManifold*.

Using *Parallel transport* a tangent vector S lying on the tangent space at base-point G , as per (35), can be transported to the tangent space at another base-point R . Parallel transport is instrumental for realizing transfer learning in a Riemannian framework [7-9]. It is defined by

$$G \parallel_R(S) = (R G^{-1})^{\frac{1}{2}} S (R G^{-1})^{\frac{1}{2}}. \quad (37)$$

The same transformation can be applied directly to a matrix in \mathbf{P}_n , *i.e.*, in (37) S is a PDM to be transported directly like if it was projected on the tangent space at base point G , parallel transported to the tangent space at base-point R and projected back onto the manifold [9].

B. Laplacians, Diffusion Maps and Spectral Embedding

Given a set of k PDMs $\{P_1, P_2, \dots, P_k\}$, consider all their pair-wise *squared inter-distances* $\delta^2(P_i, P_j)$ and build the symmetric matrix Δ using the Gaussian radial basis functions $\Delta_{ij} = \exp(-\varepsilon^{-1} \delta^2(P_i, P_j))$, where ε is the Gaussian scale parameter. *PosDefManifold* computes the symmetric *normalized Laplacian*, defined as

$$\Omega = D^{-\frac{1}{2}} \Delta D^{-\frac{1}{2}}, \quad (38)$$

where D is the diagonal matrix holding on the main diagonal the sum of the rows (or column) of Δ . The elements of matrix Ω holds a measure of proximity in \mathbf{P}_n between all points taken pair-wise. The eigenvectors of Ω associated with the largest eigenvalues (discarding the first one) allow to construct *Laplacian eigenmaps (diffusion maps)*, *i.e.*, to project the data points $\{P_1, P_2, \dots, P_k\}$ in lower Euclidean dimensions (typically two or three) for, among other purposes, classifying them, following their trajectories over time or other dimensions, etc. See [33] for references and some applications. This map from \mathbf{P}_n to a lower-dimensional Euclidean space is named *spectral embedding* and is instrumental to, among others, the field of *manifold learning*. *PosDefManifold* implements an efficient algorithm to estimate the eigenvectors of Ω associated with its largest eigenvalues, avoiding the computation of the eigenvalue-eigenvector decomposition of Ω , thus enabling the study of diffusion maps when k is large.

C. The Procrustes Problem

The well-known Procrustes problem in Euclidean space is formulated such as

$$\underset{U}{\operatorname{argmin}} \|Y - UX\|_F, \quad (39)$$

where X, Y are data matrices and U is the orthogonal/unitary matrix to be found so as to match as much as possible data in X and Y . If the data is represented by PDMs P and Q , the Procrustes problem in \mathbf{P}_n is naturally defined such as

$$\underset{U}{\operatorname{argmin}} \delta(P - U^* Q U) \quad (40)$$

for any desired metrics acting in \mathbf{P}_n . $U^* Q U$ is named in physics the *unitary orbit* of Q . In [34] it has been shown that using each of the aforementioned Fisher, Wasserstein and the whole family of α -divergences, including the logdet-zero and the left and right Kullback-Leibler divergence, the best approximant to P from the unitary orbit of Q commutes with P . Surprisingly, the solution is the same as the solution obtained using the Euclidean metric, which is known since a longtime [e.g., 35]. *PosDefManifold* computes the solution for both the argmin problem (40) and its argmax counterpart.

D. Miscellaneous

PosDefManifold implements functions to generate random PDMs and random orthogonal/unitary matrices, as well as a generic function to obtain any spectral functions of eigenvalues and low-level functions for normalizing matrices.

V. GENERAL CONSIDERATIONS

A. Machine Learning

Riemannian classification algorithms operating in \mathbf{P}_n [7] and on the tangent space (see section IV.A) are being implemented in a complementary Julia package named *PosDefManifoldML*.

B. Efficiency

Care has been injected in implementing state-of-the-art iterative algorithms and coding efficient functions. Besides leveraging on the speed of Julia's compiled code, all key functions in *PosDefManifold* are *multi-threaded*. On a standard personal computer this allow saving computational time, as compared to the aforementioned Matlab and Python packages, by a factor of two at least.

C. Code examples

PosDefManifold features a simple uniform API. Following our presentation here, functions work in the same way for all metrics. Here are some examples:

```
using PosDefManifold
metric=Fisher

P, Q = randP(20), randP(20) # random PMDs of size 20x20
d = distance(metric, P, Q) # δ(P, Q) (15)
R = geodesic(metric, P, Q, 0.1) # γ0.1(P, Q) (16)
G = mean(metric, P, Q) # G({P, Q}) = γ0.5(P, Q).
U = procrustes(P, Q) # section IV.C
S = logMap(metric, P, G) # tangent space proj. (35)
```

```
P = randP(20, 100) # random set P of 100 PDMs of size (20x20)
G = mean(metric, P) # G(P) satisfying (17)
λ, U, i, c = spEmb(metric, P, 2) # section IV.B
```

VI. CONCLUSION

Julia is a relatively new language whose popularity is steadily increasing thanks to its enormous potential for scientific computing. Research on the manifold of PDMs is currently very active and is finding more and more applications to real data problems, in particular, in the field of brain-computer interfaces. A state-of-the-art open-source package for manipulating data in the manifold of PDMs for the Julia language has been presented and therein released. This publication enables collaboration in the community to further expand it and for performing research using it. *PosDefManifold* features a rich and detailed documentation and a script for testing all implemented functions. Coupled with *PosDefManifoldML* the package is, as for today, the most comprehensive open-source package for manipulating data in the manifold of PDMs. Released under the liberal MIT license, it is available in the Julia package registry and for download, forking, cloning, etc. at <https://github.com/Marco-Congedo/PosDefManifold.jl>. The reference to package *PosDefManifoldML* can be found there as well. As a by-product, while presenting the package, in this article we have provided a succinct, yet comprehensive survey of useful concepts for manipulating data in \mathbf{P}_n .

ACKNOWLEDGMENT

The authors are grateful to Dr. Florent Bouchard for first bringing to their attention the potential of the Julia language, Raphaël Bacher for his invaluable help with building the documentation in github and Prof. Rajendra Bhatia for his useful comments on a preliminary version of this manuscript.

REFERENCES

- [1] R. Bhatia, *Positive Definite Matrices*, Princeton University Press, New Jersey, 2007.
- [2] A. Barachant, S. Bonnet, M. Congedo, C. Jutten, "Multi-class Brain Computer Interface Classification by Riemannian Geometry," *IEEE Trans. Biomed. Eng.*, vol. 59, no. 4, pp.920-928, 2012.
- [3] A. Barachant, S. Bonnet, M. Congedo, C. Jutten, "Classification of covariance matrices using a Riemannian-based kernel for BCI applications," *Neurocomputing*, vol. 112, pp. 172-178, 2013.
- [4] M. Congedo, "EEG Source Analysis," Habilitation à Diriger des Recherches, University Grenoble Alpes, 2013.
- [5] F. Lotte, L. Bougrain, A. Cichocki, M. Clerc, M. Congedo, A. Rakotomamonjy, et al., "A Review of Classification Algorithms for EEG-based Brain-Computer Interfaces: A 10-year Update," *J. Neural Eng.*, vol. 15, no. 5, 031005, 2018.
- [6] M. Congedo, A. Barachant, R. Bhatia, "Riemannian Geometry for EEG-based Brain-Computer Interfaces; a Primer and a Review," *Brain-Computer Interfaces*, vol. 4, no. 3, pp. 155-174, 2017.
- [7] P. Zanini, M. Congedo, C. Jutten, S. Said, Y. Berthoumieu, "Transfer Learning: a Riemannian geometry framework with applications to Brain-Computer Interfaces," *IEEE Trans. Biomed. Eng.*, vol. 65, no. 5, pp. 1107-1116, 2018.
- [8] P.L.C. Rodrigues, C. Jutten, M. Congedo, "Riemannian Procrustes Analysis : Transfer Learning for Brain-Computer Interfaces," *IEEE Trans. Biomed. Eng.*, in press, 2019.
- [9] O. Yair, M. Ben-Chen, and R. Talmon, "Parallel transport on the cone manifold of SPD matrices for domain adaptation," *IEEE Trans. Sig. Process.*, vol. 67, no. 7, pp. 1797-1811, 2019.
- [10] F. Yger, M. Berar, F. Lotte, "Riemannian approaches in brain-computer interfaces: a review," *IEEE Trans. Neural Syst. Rehabil. Eng.*, vol. 25, no. 10, pp. 1753-1762, 2017.
- [11] J. Bezanson, A. Edelman, S. Karpinski, V.B. Shah, "A fresh approach to numerical computing," *SIAM review*, vol. 59, no. 1, pp. 65-98, 2017.
- [12] A. Barachant, Covariance Toolbox, <https://github.com/alexandrebarachant/covariancetoolbox>.
- [13] A. Barachant, pyRiemann, <https://github.com/alexandrebarachant/pyRiemann>
- [14] Z. Chebbi, M. Moakher, "Means of Hermitian positive-definite matrices based on the log-determinant α -divergence function," *Linear. Algebra. Appl.*, vol. 436, no. 7, pp. 1872-1889, 2012.
- [15] A. Cichocki, S. Cruces, S-I- Amari, "Log-Determinant Divergences Revisited: Alpha-Beta and Gamma Log-Det Divergences," *Entropy*, vol. 17, no. 5, pp. 2988-3034, 2015.
- [16] S. Sra, "Positive definite matrices and the S-divergence," *Proc. Amer. Math. Soc.*, vol. 144 pp. 2787-2797, 2016.
- [17] V. Arsigny, P. Fillard, X. Pennec, N. Ayache, "Geometric means in a novel vector space structure on symmetric positive-definite matrices," *SIAM J. Matrix Anal. Appl.*, vol. 29, no. 1, pp. 328-347, 2007.
- [18] R. Bhatia, S. Gaubert, T. Jain, "Matrix versions of the Hellinger distance," *arXiv:1901.01378*, 2019.
- [19] M. Congedo, B. Afsari, A. Barachant, M Moakher, "Approximate Joint Diagonalization and Geometric Mean of Symmetric Positive Definite Matrices," *PLoS ONE*, vol. 10, no. 4, e0121423, 2015.
- [20] R. Bhatia, J. Holbrook, "Riemannian geometry and matrix geometric means," *Linear. Algebra. Appl.*, vol. 413, no. 2-3, pp. 594-618, 2006.
- [21] M. Moakher, "A Differential Geometric Approach to the Geometric Mean of Symmetric Positive-Definite Matrices," *SIAM J. Matrix Anal. Appl.*, vol. 26, no. (3), pp. 735-747, 2005.
- [22] M. Congedo, A. Barachant, E. Kharati Koopaie, "Fixed Point Algorithms for Estimating Power Means of Positive Definite Matrices," *IEEE Trans. Signal Process.*, vol. 65, no. 9, pp. 2211-2220, 2017.
- [23] T. Ando, C.-K. Li, R. Mathias, "Geometric means," *Linear Algebra. Appl.*, vol. 385, no. 1, pp. 305-334, 2004.
- [24] Y. Lim, M. Pálfia, "Matrix power means and the Karcher mean", *J. Funct. Anal.*, vol. 262, no. 4, pp. 1498-1514, 2012.
- [25] M. Moakher, "Divergence measures and means of symmetric positive-definite matrices," in D.H Lailaw and A. Vilanova (Eds) "New Developments in the Visualization and Processing of Tensor Fields", Springer, Berlin, 2012.
- [26] M.T. Harandi, R. Hartley, B. Lovell, C. Sanderson, "Sparse coding on symmetric positive definite manifolds using bregman divergences," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 27, no. 6, pp. 1294-1306, 2016.
- [27] M. Faraki, M. Harandi, F. Porikli, "More About VLAD: A Leap from Euclidean to Riemannian Manifolds," in *Proc. IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*, Boston, 2015.
- [28] R. Bhatia, T. Jain, Y. Lim, "Inequalities for the Wasserstein mean of positive definite matrices," *Linear. Algebra. Appl.*, in press, 2019.
- [29] M. Agueh G. Carlier, "Barycenters in the Wasserstein space," *SIAM J. Mat. Anal. Appl.*, vol. 43, pp. 904-924, 2011.
- [30] P. C. Álvarez-Esteban, E. del Barrio, J.A. Cuesta-Albertos, C. Matrán, "A fixed-point approach to barycenters in Wasserstein space," *J. Math. Anal. Appl.*, vol. 441, no. 2, 744-762, 2016.
- [31] X. Dai, S. Khamis, Y. Zhang, L.S. Davis, "Parameterizing region covariance: an efficient way to apply sparse codes on second order statistics," *arXiv, 1602.02822*, 2016.
- [32] J. Taghia, M. Bänkestad, F. Lindsten, T.B. Schön, "Constructing the Matrix Multilayer Perceptron and its Application to the VAE," *arXiv, 1902.01182v1*, 2019.
- [33] P. L. C. Rodrigues, M. Congedo, C Jutten, "Multivariate Time-Series Analysis Via Manifold Learning," in *Proc. of the the IEEE Statistical Signal Processing Workshop (SSP 2018)*, Fribourg-en-Brisgau, Germany, Jun 2018.
- [34] R. Bhatia, M. Congedo, "Procrustes problems in manifolds of positive definite matrices," *Linear Algebra. Appl.*, vol. 563, pp. 440-445, 2019.
- [35] S. Uemeyama, "An Eigendecomposition Approach to Weighted Graph Matching Problems," *IEEE Trans. Pattern. Anal. Mach. Intell.*, vol. 10, no. 5, pp. 695-703, 1988.