



HAL
open science

Geodesics and Curvature of the Quotient-Affine Metrics on Full-Rank Correlation Matrices

Yann Thanwerdas, Xavier Pennec

► **To cite this version:**

Yann Thanwerdas, Xavier Pennec. Geodesics and Curvature of the Quotient-Affine Metrics on Full-Rank Correlation Matrices. GSI 2021 - 5th conference on Geometric Science of Information, Jul 2021, Paris, France. pp.93-102, 10.1007/978-3-030-80209-7_11 . hal-03157992v3

HAL Id: hal-03157992

<https://hal.science/hal-03157992v3>

Submitted on 11 May 2021

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.

Geodesics and Curvature of the Quotient-Affine Metrics on Full-Rank Correlation Matrices

Yann Thanwerdas and Xavier Pennec

Université Côte d'Azur, Inria, Epione, France
yann.thanwerdas@inria.fr, xavier.pennec@inria.fr

Abstract. Correlation matrices are used in many domains of neurosciences such as fMRI, EEG, MEG. However, statistical analyses often rely on embeddings into a Euclidean space or into Symmetric Positive Definite matrices which do not provide intrinsic tools. The quotient-affine metric was recently introduced as the quotient of the affine-invariant metric on SPD matrices by the action of diagonal matrices. In this work, we provide most of the fundamental Riemannian operations of the quotient-affine metric: the expression of the metric itself, the geodesics with initial tangent vector, the Levi-Civita connection and the curvature.

Keywords: Riemannian geometry · Quotient manifold · Quotient-affine metric · SPD matrices · Correlation matrices.

1 Introduction

Correlation matrices are used in many domains with time series data such as functional brain connectivity in functional MRI, electroencephalography (EEG) or magnetoencephalography (MEG) signals. Full-rank correlation matrices form a strict sub-manifold of the cone of Symmetric Positive Definite (SPD) matrices sometimes called the (open) ellipsope [1]. However, very few geometric tools were defined for intrinsic computations with correlation matrices. For example, [2] rely on a surjection from a product of n spheres of dimension $n - 1$ onto the space of correlation matrices in order to sample valid correlation matrices for financial applications: a point in the former space can be represented by an $n \times n$ matrix A with normed rows and therefore encodes a correlation matrix AA^T . Since low-rank matrices have null measure, one gets a full-rank correlation matrix almost surely. More recently, the open ellipsope was endowed with the Hilbert projective geometry [3] which relies on its convexity.

Since there exist efficient tools on SPD matrices (affine-invariant/Fisher-Rao metric, log-Euclidean metric...), correlation matrices are often treated as SPD matrices [4]. Nevertheless, these extrinsic tools do not respect the geometry of correlation matrices. Moreover, most of these tools on SPD matrices are invariant under orthogonal transformations which is not compatible with correlation matrices. The ellipsope is not even stable by the orthogonal action. Hence using the tools designed for SPD matrices may not be relevant for correlation matrices. One could restrict a Riemannian metric from the cone of SPD matrices to the open ellipsope to define intrinsic tools by embedding. To the best of our knowledge, this hasn't been studied in depth. However, a key property of correlation

matrices in applications with respect to covariance matrices is that the scale is renormalized independently for each axis. This physically corresponds to an invariance under a group action, namely the action of diagonal matrices on SPD matrices. The previously cited structures do not rely on this physical reality. This is why we are more interested in the recently introduced quotient-affine metric which corresponds to the quotient of the affine-invariant metric on SPD matrices by the action of diagonal matrices [5, 6].

In this work, we investigate the geometry of this quotient-affine metric and we contribute additional tools to compute on this manifold. Based on the formalization of the quotient manifold with vertical and horizontal distributions, we compute in closed form some fundamental Riemannian operations of the quotient-affine metrics, notably the exponential map and the curvature. In Section 2, we recall how quotient-affine metrics are introduced, as well as the basics on quotient manifolds. In Section 3, we provide the following fundamental quotient and Riemannian operations of quotient-affine metrics: the vertical and horizontal projections, the metric, the exponential map, the Levi-Civita connection and the sectional curvature. This opens the way to many practical algorithms on the open ellipsope for different applications. Considering SPD matrices as the Cartesian product of positive diagonal matrices and full-rank correlation matrices, it also allows to define new Riemannian metrics which preserve the quotient-affine geometry on correlation matrices. Thus in Section 4, we illustrate the quotient-affine metric in dimension 2 by coupling it with the diagonal power-Euclidean metrics $g_D^{E(p)}(\Delta, \Delta) = \text{tr}(D^{2(p-1)}\Delta^2)$ for $p \in \{-1, 0, 1, 2\}$, where D is positive diagonal and Δ diagonal, and then by comparing it with the affine-invariant and the log-Euclidean metrics on SPD matrices.

2 Quotient-affine metrics

2.1 The quotient manifold of full-rank correlation matrices

The group of positive diagonal matrices $\text{Diag}^+(n)$ acts on the manifold of SPD matrices $\text{Sym}^+(n)$ via the congruence action $(\Sigma, D) \in \text{Sym}^+(n) \times \text{Diag}^+(n) \mapsto D\Sigma D \in \text{Sym}^+(n)$. The manifold of full-rank correlation matrices $\text{Cor}^+(n)$ can be seen as the quotient manifold $\text{Sym}^+(n)/\text{Diag}^+(n)$ via the invariant submersion π which computes the correlation matrix from a covariance matrix, $\pi : \Sigma \in \text{Sym}^+(n) \mapsto \text{Diag}(\Sigma)^{-1/2} \Sigma \text{Diag}(\Sigma)^{-1/2} \in \text{Cor}^+(n)$, where $\text{Diag}(\Sigma) = \text{diag}(\Sigma_{11}, \dots, \Sigma_{nn})$. Hence, any Riemannian metric G on $\text{Sym}^+(n)$ which is invariant under $\text{Diag}^+(n)$ induces a quotient metric g on $\text{Cor}^+(n)$. The steps to define it are the following.

1. Vertical distribution. $\mathcal{V}_\Sigma = \ker d_\Sigma \pi$ for all $\Sigma \in \text{Sym}^+(n)$.
2. Horizontal distribution. $\mathcal{H}_\Sigma := \mathcal{V}_\Sigma^\perp$ for all $\Sigma \in \text{Sym}^+(n)$, where the orthogonality \perp refers to the inner product G_Σ .
3. Horizontal lift. The linear map $d_\Sigma \pi$ restricted to the horizontal space \mathcal{H}_Σ is a linear isomorphism onto the tangent space of full-rank correlation matrices $(d_\Sigma \pi)|_{\mathcal{H}_\Sigma} : \mathcal{H}_\Sigma \xrightarrow{\sim} T_{\pi(\Sigma)} \text{Cor}^+(n)$. The horizontal lift $\#$ is its inverse:

$$\# : X \in T_{\pi(\Sigma)} \text{Cor}^+(n) \xrightarrow{\sim} X^\# \in \mathcal{H}_\Sigma. \quad (1)$$

4. Quotient metric. It is defined by pullback through the horizontal lift:

$$\forall C \in \text{Cor}^+(n), \forall X \in T_C \text{Cor}^+(n), g_C(X, X) = G_\Sigma(X^\#, X^\#), \quad (2)$$

where $\Sigma \in \pi^{-1}(C)$ and the definition does not depend on the chosen Σ .

So the only missing ingredient is a Riemannian metric on SPD matrices which is invariant under the congruence action of positive diagonal matrices. In [5, 6], the authors chose to use the famous affine-invariant/Fisher-Rao metric.

2.2 The affine-invariant metrics and the quotient-affine metrics

The affine-invariant metric is the Riemannian metric defined on SPD matrices by $G_\Sigma(V, V) = \text{tr}(\Sigma^{-1}V\Sigma^{-1}V)$ for all $\Sigma \in \text{Sym}^+(n)$ and $V \in T_\Sigma \text{Sym}^+(n)$ [7–9]. It is invariant under the congruence action of the whole real general linear group $\text{GL}(n)$ which contains $\text{Diag}^+(n)$ as a subgroup. It provides a Riemannian symmetric structure to the manifold of SPD matrices, hence it is geodesically complete and the geodesics are given by the group action of one-parameter subgroups. We recall the exponential map, the Levi-Civita connection and the sectional curvature below for all $\Sigma \in \text{Sym}^+(n)$ and vector fields $V, W \in \Gamma(T\text{Sym}^+(n))$ where we also denote $V \equiv V_\Sigma$ and $W \equiv W_\Sigma \in T_\Sigma \text{Sym}^+(n)$:

$$\text{Exp}_\Sigma^G(V) = \Sigma^{1/2} \exp(\Sigma^{-1/2}V\Sigma^{-1/2})\Sigma^{1/2}, \quad (3)$$

$$(\nabla_V^G W)|_\Sigma = d_\Sigma W(V) - \frac{1}{2}(V\Sigma^{-1}W + W\Sigma^{-1}V), \quad (4)$$

$$\kappa_\Sigma^G(V, W) = \frac{1}{4} \frac{\text{tr}((\Sigma^{-1}V\Sigma^{-1}W - \Sigma^{-1}W\Sigma^{-1}V)^2)}{G(V, V)G(W, W) - G(V, W)^2} \leq 0. \quad (5)$$

The metrics that are invariant under the congruence action of the general linear group $\text{GL}(n)$ actually form a two-parameter family of metrics indexed by $\alpha > 0$ and $\beta > -\alpha/n$ [10]: $G_\Sigma^{\alpha, \beta}(V, V) = \alpha \text{tr}(\Sigma^{-1}V\Sigma^{-1}V) + \beta \text{tr}(\Sigma^{-1}V)^2$. We call them all affine-invariant metrics. In particular, these metrics are invariant under the congruence action of diagonal matrices so they are good candidates to define Riemannian metrics on full-rank correlation matrices by quotient. In [5, 6], the authors rely on the ‘‘classical’’ affine-invariant metric ($\alpha = 1, \beta = 0$). We generalize their definition below.

Definition 1 (Quotient-affine metrics on full-rank correlation matrices). *The quotient-affine metric of parameters $\alpha > 0$ and $\beta > -\alpha/n$ is the quotient metric $g^{\alpha, \beta}$ on $\text{Cor}^+(n)$ induced by the affine-invariant metric $G^{\alpha, \beta}$ via the submersion $\pi : \Sigma \in \text{Sym}^+(n) \mapsto \text{Diag}(\Sigma)^{-1/2} \Sigma \text{Diag}(\Sigma)^{-1/2} \in \text{Cor}^+(n)$.*

3 Fundamental Riemannian operations

In this section, we detail the quotient geometry of quotient-affine metrics $g^{\alpha, \beta}$. We give the vertical and horizontal distributions and projections in Section 3.1. We contribute the formulae of the metric itself in Section 3.2, the exponential map in Section 3.3, and finally the Levi-Civita connection and the sectional curvature in Section 3.4. To the best of our knowledge, all these formulae are new. They are proved in Appendix A.

3.1 Vertical and horizontal distributions and projections

- Let \bullet be the Hadamard product on matrices defined by $[X \bullet Y]_{ij} = X_{ij}Y_{ij}$.
- Let $A : \Sigma \in \text{Sym}^+(n) \mapsto A(\Sigma) = \Sigma \bullet \Sigma^{-1} \in \text{Sym}^+(n)$. This smooth map is invariant under the action of positive diagonal matrices. The Schur product theorem ensures that $A(\Sigma) \in \text{Sym}^+(n)$. A fortiori, $I_n + A(\Sigma) \in \text{Sym}^+(n)$.
- Let $\psi : \mu \in \mathbb{R}^n \mapsto (\mu \mathbf{1}^\top + \mathbf{1} \mu^\top) \in \text{Sym}(n)$. This is an injective linear map.
- Let $\mathcal{S}_\Sigma(V)$ the unique solution of the Sylvester equation $\Sigma \mathcal{S}_\Sigma(V) + \mathcal{S}_\Sigma(V) \Sigma = V$ for $\Sigma \in \text{Sym}^+(n)$ and $V \in \text{Sym}(n)$.
- Let $\text{Hol}^S(n)$ be the vector space of symmetric matrices with vanishing diagonal (symmetric hollow matrices). Each tangent space of the manifold of full-rank correlation matrices can be seen as a copy of this vector space.

Theorem 1 (Vertical and horizontal distributions and projections).

The vertical distribution is given by $\mathcal{V}_\Sigma = \Sigma \bullet \psi(\mathbb{R}^n)$ and the horizontal distribution is given by $\mathcal{H}_\Sigma = \mathcal{S}_{\Sigma^{-1}}(\text{Hol}^S(n))$. The vertical projection is:

$$\text{ver} : V \in T_\Sigma \text{Sym}^+(n) \mapsto \Sigma \bullet \psi((I_n + A(\Sigma))^{-1} \text{Diag}(\Sigma^{-1}V) \mathbf{1}) \in \mathcal{V}_\Sigma. \quad (6)$$

Then, the horizontal projection is simply $\text{hor}(V) = V - \text{ver}(V)$.

3.2 Horizontal lift and metric

Theorem 2 (Horizontal lift). Let $\Sigma \in \text{Sym}^+(n)$ and $C = \pi(\Sigma) \in \text{Cor}^+(n)$. The horizontal lift at Σ of $X \in T_C \text{Cor}^+(n)$ is $X^\# = \text{hor}(\Delta_\Sigma X \Delta_\Sigma)$ with $\Delta_\Sigma = \text{Diag}(\Sigma)^{1/2}$. In particular, the horizontal lift at $C \in \text{Sym}^+(n)$ is $X^\# = \text{hor}(X)$.

Theorem 3 (Expression of quotient-affine metrics). For all $C \in \text{Cor}^+(n)$ and $X \in T_C \text{Cor}^+(n)$, $g_C^{\alpha, \beta}(X, X) = \alpha g_C^{\text{QA}}(X, X)$ (independent from β) where:

$$g_C^{\text{QA}}(X, X) = \text{tr}((C^{-1}X)^2) - 2\mathbf{1}^\top \text{Diag}(C^{-1}X)(I_n + A(C))^{-1} \text{Diag}(C^{-1}X) \mathbf{1}. \quad (7)$$

3.3 Geodesics

The geodesics of a quotient metric are the projections of the horizontal geodesics of the original metric. This allows us to obtain the exponential map of the quotient-affine metrics.

Theorem 4 (Geodesics of quotient-affine metrics). The geodesic from $C \in \text{Cor}^+(n)$ with initial tangent vector $X \in T_C \text{Cor}^+(n)$ is:

$$\forall t \in \mathbb{R}, \gamma_{(C, X)}^{\text{QA}}(t) = \text{Exp}_C^{\text{QA}}(tX) = \pi(C^{1/2} \exp(t C^{-1/2} \text{hor}(X) C^{-1/2}) C^{1/2}). \quad (8)$$

In particular, the quotient-affine metric is geodesically complete.

The Riemannian logarithm between C_1 and $C_2 \in \text{Cor}^+(n)$ is much more complicated to compute. It amounts to find $\Sigma \in \text{Sym}^+(n)$ in the fiber above C_2 such that $\text{Log}_{C_1}^G(\Sigma)$ is horizontal. Then we have $\text{Log}_{C_1}^{\text{QA}}(C_2) = d_{C_1} \pi(\text{Log}_{C_1}^G(\Sigma))$. This means finding Σ that minimizes the affine-invariant distance in the fiber:

$$D = \arg \min_{D \in \text{Diag}^+(n)} d(C_1, DC_2D),$$

from which we get $\Sigma = DC_2D$. This is the method used in the original paper [5, 6]. Note that the uniqueness of the minimizer has not been proved yet.

3.4 Levi-Civita connection and sectional curvature

In this section, we give the Levi-Civita connection and the curvature. The computations are based on the fundamental equations of submersions [11]. We denote $\text{sym}(M) = \frac{1}{2}(M + M^\top)$ the symmetric part of a matrix.

Theorem 5 (Levi-Civita connection and sectional curvature of quotient-affine metrics). *The Levi-Civita connection of quotient-affine metrics is:*

$$\begin{aligned} (\nabla_X^{\text{QA}} Y)|_C &= d_C Y(X) + \text{sym}[\text{Diag}(X^\#)Y^\# + \text{Diag}(Y^\#)X^\# + \text{Diag}(X^\#C^{-1}Y^\#)C \\ &\quad - X^\#C^{-1}Y^\# - \frac{1}{2}\text{Diag}(X^\#)C\text{Diag}(Y^\#) - \frac{3}{2}\text{Diag}(X^\#)\text{Diag}(Y^\#)C]. \end{aligned} \quad (9)$$

The curvature of quotient-affine metrics is:

$$\kappa_C^{\text{QA}}(X, Y) = \kappa_C^G(X^\#, Y^\#) + \frac{3}{4} \frac{G_C(\text{ver}[X^\#, Y^\#], \text{ver}[X^\#, Y^\#])}{g_C(X, X)g_C(Y, Y) - g_C(X, Y)^2}, \quad (10)$$

$$= \frac{2 \text{tr}((C^{-1}X^\#C^{-1}Y^\# - C^{-1}Y^\#C^{-1}X^\#)^2) + 3 \mathbf{1}^\top D(I_n + A(C))^{-1} D \mathbf{1}}{8(g_C(X, X)g_C(Y, Y) - g_C(X, Y)^2)} \quad (11)$$

where $[V, W] = dW(V) - dV(W)$ is the Lie bracket on $\text{Sym}^+(n)$ and $D = D(X, Y) - D(Y, X)$ with $D(X, Y) = \text{Diag}(C^{-1}\text{Diag}(X^\#)Y^\# - C^{-1}Y^\#C^{-1}\text{Diag}(X^\#)C)$. There is a slight abuse of notation because $\text{ver}[X^\#, Y^\#]$ induces that $X^\#$ and $Y^\#$ are vector fields. Indeed here, they are horizontal vector fields extending the horizontal lifts at C .

The first term of the sectional curvature is negative, the second one is positive so we don't know in general the sign of the curvature of quotient-affine metrics.

The quotient-affine metrics not only provide a Riemannian framework on correlation matrices but also provide correlation-compatible statistical tools on SPD matrices if we consider that the space of SPD matrices is the Cartesian product of positive diagonal matrices and full-rank correlation matrices. We give a taste of such construction in the next section.

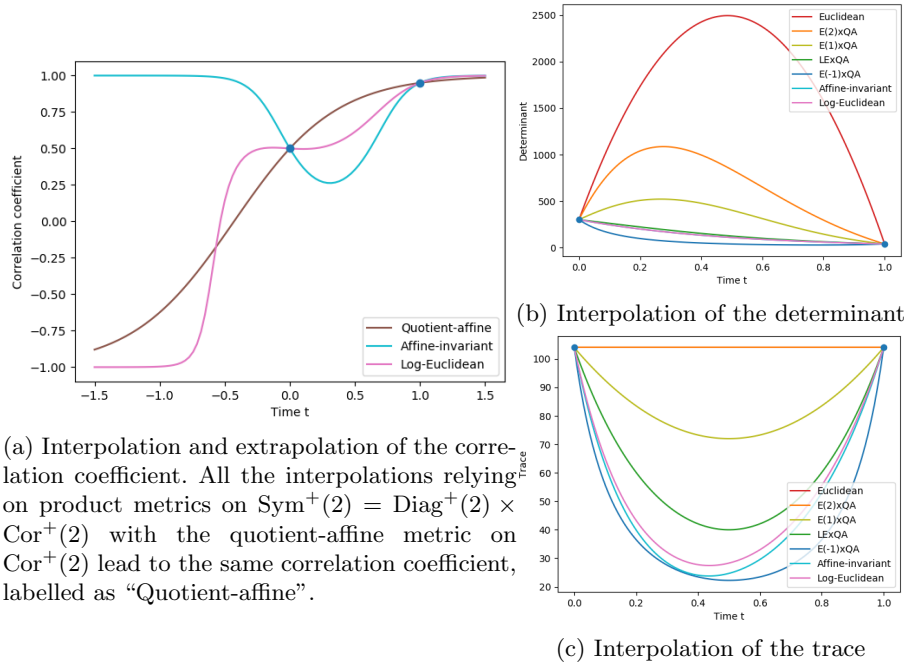
4 Illustration in dimension 2

In dimension 2, any correlation matrix $C \in \text{Cor}^+(2)$ writes $C = C(\rho) := \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$ with $\rho \in (-1, 1)$. In the following theorem, we give explicitly the logarithm, the distance and the interpolating geodesics of the quotient-affine metric.

Theorem 6 (Quotient-affine metrics in dimension 2). *Let $C_1 = C(\rho_1), C_2 = C(\rho_2) \in \text{Cor}^+(2)$ with $\rho_1, \rho_2 \in (-1, 1)$. We denote $f : \rho \in (-1, 1) \mapsto \frac{1+\rho}{1-\rho} \in (0, \infty)$ which is a smooth increasing map. We denote $\lambda = \lambda(\rho_1, \rho_2) = \frac{1}{2} \log \frac{f(\rho_2)}{f(\rho_1)}$ which has the same sign as $\rho_2 - \rho_1$. Then the quotient-affine operations are:*

1. (Logarithm) $\text{Log}_{C_1}^{\text{QA}}(C_2) = \lambda \begin{pmatrix} 0 & 1 - \rho_1^2 \\ 1 - \rho_1^2 & 0 \end{pmatrix}$, (distance) $d^{\text{QA}}(C_1, C_2) = \sqrt{2}|\lambda|$,
2. (Geodesics) $\gamma_{(C_1, C_2)}^{\text{QA}}(t) = C(\rho^{\text{QA}}(t))$ where $\rho^{\text{QA}}(t) = \frac{\rho_1 \cosh(\lambda t) + \sinh(\lambda t)}{\rho_1 \sinh(\lambda t) + \cosh(\lambda t)} \in (-1, 1)$ is monotonic (increasing if and only if $\rho_2 - \rho_1 > 0$).

Let $\Sigma_1, \Sigma_2 \in \text{Sym}^+(2)$ and C_1, C_2 their respective correlation matrices. We denote $\gamma^{\text{AI}}, \gamma^{\text{LE}}$ the geodesics between Σ_1 and Σ_2 for the affine-invariant and the log-Euclidean metrics respectively. We define $\rho^{\text{AI}}, \rho^{\text{LE}}$ such that the correlation matrices of $\gamma^{\text{AI}}(t), \gamma^{\text{LE}}(t)$ are $C(\rho^{\text{AI}}(t)), C(\rho^{\text{LE}}(t))$. Figure 1.(a) shows $\rho^{\text{AI}}, \rho^{\text{LE}}, \rho^{\text{QA}}$ with $\Sigma_1 = \begin{pmatrix} 4 & 1 \\ 1 & 100 \end{pmatrix}$ and $\Sigma_2 = \begin{pmatrix} 100 & 19 \\ 19 & 4 \end{pmatrix}$. When varying numerically Σ_1 and Σ_2 , it seems that ρ^{LE} and ρ^{AI} always have three inflection points. In contrast, ρ^{QA} always has one inflection point since $(\rho^{\text{QA}})'' = -2\lambda^2 \rho^{\text{QA}}(1 - (\rho^{\text{QA}})^2)$. Analogously, we compare the interpolations of the determinant (Fig. 1.(b)) and the trace (Fig. 1.(c)) using several Riemannian metrics: Euclidean (trace-monotonic); log-Euclidean and affine-invariant (determinant-monotonic); power-Euclidean \times quotient-affine (correlation-monotonic).



(a) Interpolation and extrapolation of the correlation coefficient. All the interpolations relying on product metrics on $\text{Sym}^+(2) = \text{Diag}^+(2) \times \text{Cor}^+(2)$ with the quotient-affine metric on $\text{Cor}^+(2)$ lead to the same correlation coefficient, labelled as “Quotient-affine”.

(b) Interpolation of the determinant

(c) Interpolation of the trace

Fig. 1: Extrapolation and interpolation between the SPD matrices Σ_1 and Σ_2 using various Riemannian metrics. The metrics $E(p) \times \text{QA}$ refer to the p -power-Euclidean metric on the diagonal part and the quotient-affine metric on the correlation part. When p tends to 0, $E(p)$ tends to the log-Euclidean metric $\text{LE} \equiv E(0)$.

On Figure 2, we compare the geodesics of these metrics. The determinant is the area of the ellipsis and the trace is the sum of the lengths of the axes. Thus the product metrics of the form power-Euclidean on the diagonal part and quotient-affine on the correlation part can be seen as performing a correlation-monotonic trade-off between the trace-monotonicity and the determinant-monotonicity.

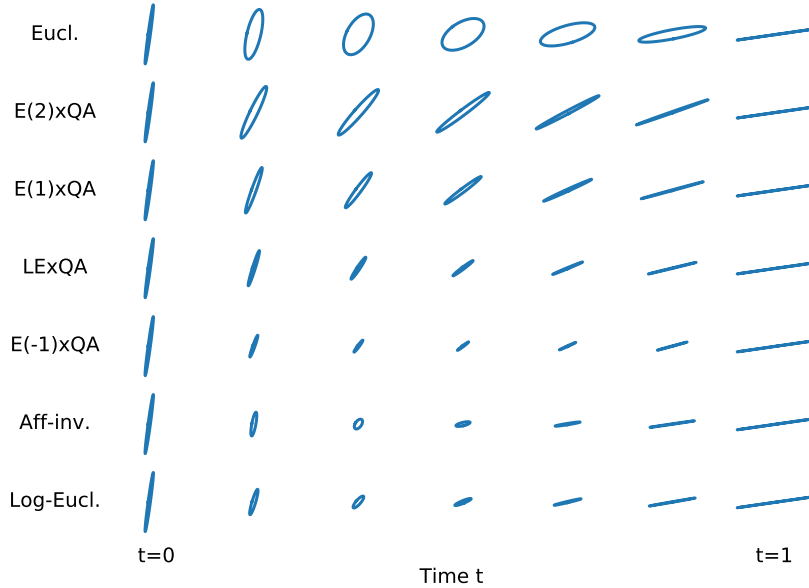


Fig. 2: Interpolations between SPD matrices Σ_1 and Σ_2 .

5 Conclusion

We investigated in this paper the very nice idea of quotienting the affine-invariant metrics on SPD matrices by the action of the positive diagonal group to obtain the principled quotient-affine metrics on full-rank correlation matrices. The quotient-affine metric with $\alpha = 1$ and $\beta = 0$ was first proposed in Paul David's thesis [6] and in the subsequent journal paper [5]. We contribute here exact formulae for the main Riemannian operations, including the exponential map, the connection and the sectional curvature. The exponential map is particularly interesting to rigorously project tangent computations to the space of full-rank correlation matrices. This opens the way to the implementation of a number of generic algorithms on Riemannian manifolds. However, we could not find a closed form expression for the logarithm which remains to be computed through an optimization procedure. Thus, computing distances with these metrics remains computationally expensive. In order to obtain more efficient methods, this leads us to look for other principled Riemannian metrics on correlation matrices for which the logarithm could be expressed in closed form.

Acknowledgements. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant G-Statistics agreement No 786854). This work has

been supported by the French government, through the UCAJEDI Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-15-IDEX-01 and through the 3IA Côte d’Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002. The authors would like to thank Frank Nielsen for insightful discussions on correlation matrices.

References

1. J. A. Tropp. Simplicial faces of the set of correlation matrices. *Discrete & Computational Geometry*, 60(2):512–529, 2018.
2. R. Rebonato and P. Jäckel. The most general methodology to create a valid correlation matrix for risk management and option pricing purposes. *Journal of Risk*, 2:17–27, 2000.
3. F. Nielsen and K. Sun. Clustering in hilbert’s projective geometry: The case studies of the probability simplex and the ellipsope of correlation matrices. In *Geometric Structures of Information*, pages 297–331. Springer, 2019.
4. G. Varoquaux, F. Baronnet, A. Kleinschmidt, P. Fillard, and B. Thirion. Detection of brain functional connectivity difference in post-stroke patients using group-level covariance modeling. *MICCAI*, 2010.
5. P. David and W. Gu. A Riemannian structure for correlation matrices. *Operators and Matrices*, 13(3):607–627, 2019.
6. P. David. *A Riemannian Quotient Structure for Correlation Matrices with Applications to Data Science*. PhD thesis, Institute of Mathematical Sciences, Claremont Graduate University, 2019.
7. C. Siegel. Symplectic Geometry. *American J. of Mathematics*, 65(1):1–86, 1943.
8. L. Skovgaard. A Riemannian geometry of the multivariate normal model. *Scandinavian Journal of Statistics*, 11(4):211–223, 1984.
9. S. Amari and H. Nagaoka. *Methods of Information Geometry*. American Mathematical Society, 2000.
10. X. Pennec. Statistical Computing on Manifolds: From Riemannian Geometry to Computational Anatomy. *Emerging Trends in Visual Comp.*, 5416:347–386, 2008.
11. B. O’Neill. The fundamental equations of a submersion. *Michigan Mathematical Journal*, 13(4):459–469, 1966.
12. S. Gallot, D. Hulin, and J. Lafontaine. *Riemannian Geometry*. Springer, 2004.

A Proofs

(Theorem 1) We denote $\Delta_\Sigma = \text{Diag}(\Sigma)^{1/2}$. Using Section 2.1, we have:

$$d_\Sigma \pi(V) = \Delta_\Sigma^{-1} \left[V - \frac{1}{2} (\Delta_\Sigma^{-2} \text{Diag}(V) \Sigma + \Sigma \text{Diag}(V) \Delta_\Sigma^{-2}) \right] \Delta_\Sigma^{-1}.$$

$$V \in \mathcal{V}_\Sigma \iff \forall i, j \in \{1, \dots, n\}, V_{ij} = \Sigma_{ij} \times \frac{1}{2} \left(\frac{V_{ii}}{\Sigma_{ii}} + \frac{V_{jj}}{\Sigma_{jj}} \right) \iff V \in \Sigma \bullet \psi(\mathbb{R}^n).$$

$$W \in \mathcal{H}_\Sigma \iff \forall V \in \mathcal{V}_\Sigma, \text{tr}(\Sigma^{-1} V \Sigma^{-1} W) = 0,$$

$$\iff \forall \mu \in \mathbb{R}^n, \sum_{ijkl} [\Sigma^{-1}]_{ij} \Sigma_{jk} (\mu_j + \mu_k) [\Sigma^{-1}]_{kl} W_{li} = 0,$$

$$\iff \forall D \in \text{Diag}(n), \text{tr}(\Sigma^{-1} DW + D \Sigma^{-1} W) = 0,$$

$$\iff (\Sigma^{-1} W + W \Sigma^{-1}) \in \text{Hol}^S(n) \iff W \in \mathcal{S}_{\Sigma^{-1}}(\text{Hol}^S(n)).$$

Thus, we have computed the horizontal space for the affine-invariant metric $\alpha = 1$ and $\beta = 0$. It is still valid for all $\alpha > 0$ and $\beta > -\alpha/n$ since the latter is included in the former (because $\text{tr}(\Sigma^{-1}W) = 0$) and they have the same dimension so they are equal. Now we compute the vertical projection. Let $V \in T_\Sigma \text{Sym}^+(n)$ and let $\mu \in \mathbb{R}^n$ and $W \in \mathcal{H}_\Sigma$ such that $V = \Sigma \bullet \psi(\mu) + W$. We are looking for μ . Since $\Sigma^{-1}V = \Sigma^{-1}(\Sigma \bullet \psi(\mu)) + \Sigma^{-1}W$, we have $[\Sigma^{-1}V]_{ii} = \sum_j [\Sigma^{-1}]_{ij} \Sigma_{ij} (\mu_i + \mu_j) = \mu_i + [A(\Sigma)\mu]_i$ so $\text{Diag}(\Sigma^{-1}V)\mathbf{1} = (I_n + A(\Sigma))\mu$, hence $\mu = (I_n + A(\Sigma))^{-1}\text{Diag}(\Sigma^{-1}V)\mathbf{1}$.

(Theorem 2) The horizontal lift is $X^\# = \text{hor}(\Delta_\Sigma X \Delta_\Sigma)$ since $d_\Sigma \pi(\Delta_\Sigma X \Delta_\Sigma) = \Delta_\Sigma^{-1} [V - \frac{1}{2}(\Delta_\Sigma^{-2}\text{Diag}(V)\Sigma + \Sigma\text{Diag}(V)\Delta_\Sigma^{-2})] \Delta_\Sigma^{-1} = X$, where $V = \Delta_\Sigma X \Delta_\Sigma$.

(Theorem 3) We use the definition of the quotient metric.

$$g_C^{\alpha,\beta}(X, X) = G_C^{\alpha,\beta}(\text{hor}(X), \text{hor}(X)) = G_C^{\alpha,\beta}(X, X) - G_C^{\alpha,\beta}(\text{ver}(X), \text{ver}(X)).$$

First, we observe that $\text{tr}(C^{-1}X) = \text{tr}(C^{-1}\text{ver}(X))$ since $\text{tr}(C^{-1}\text{hor}(X)) = 0$ so the terms in β vanish and $g_C^{\alpha,\beta}(X, X) = \alpha(\text{tr}(C^{-1}XC^{-1}X) - \text{tr}(C^{-1}\text{ver}(X)C^{-1}\text{ver}(X)))$ does not depend on β .

We denote $\text{ver}(X) = C \bullet \psi(\mu)$ with $\mu = (I_n + A(C))^{-1}\text{Diag}(C^{-1}X)\mathbf{1}$. Then:

$$\begin{aligned} \text{tr}(C^{-1}\text{ver}(X)C^{-1}\text{ver}(X)) &= \sum_{ijkl} [C^{-1}]_{ij} C_{jk} (\mu_j + \mu_k) [C^{-1}]_{kl} C_{li} (\mu_l + \mu_i), \\ &= 2 \sum_{ij} (\delta_{ij} + [C^{-1}]_{ij} C_{ij}) \mu_i \mu_j, \\ &= 2\mu^\top (I_n + A(C))\mu, \\ &= 2\mathbf{1}^\top \text{Diag}(C^{-1}X)(I_n + A(C))^{-1}\text{Diag}(C^{-1}X)\mathbf{1}. \end{aligned}$$

(Theorem 4) The geodesics are the the projections of horizontal geodesics [12] so they are given by $\text{Exp}_C^{\text{QA}}(X) = \pi(\text{Exp}_C^G(X^\#)) = \pi(C^{1/2} \exp(C^{-1/2}\text{hor}(X)C^{-1/2})C^{1/2})$.

(Theorem 5) According O'Neill's equations of submersions, the Levi-Civita connection of quotient-affine metrics is given by:

$$\begin{aligned} (\nabla_X^{\text{QA}} Y)|_C &= d_\Sigma \pi((\nabla_{X^\#}^G Y^\#)|_C), \\ &= d_C \pi(d_C Y^\#(X^\#) - \frac{1}{2}(X^\# C^{-1} Y^\# + Y^\# C^{-1} X^\#)), \\ &= d_C Y^\#(X^\#) - \frac{1}{2}(\text{Diag}(d_C Y^\#(X^\#))C + C\text{Diag}(d_C Y^\#(X^\#)) \\ &\quad - \frac{1}{2}(X^\# C^{-1} Y^\# + Y^\# C^{-1} X^\# - \text{Diag}(X^\# C^{-1} Y^\#)C - C\text{Diag}(X^\# C^{-1} Y^\#))). \end{aligned} \quad (12)$$

Since $Y_{\pi(\Sigma)} = \Delta_\Sigma^{-1} Y_\Sigma^\# \Delta_\Sigma^{-1} - \text{sym}(\Delta_\Sigma^{-3} \text{Diag}(Y_\Sigma^\#) \Sigma \Delta_\Sigma^{-1})$ where $\Delta_\Sigma = \text{Diag}(\Sigma)^{1/2}$, then:

$$\begin{aligned} d_{\pi(\Sigma)} Y(X) &= d_\Sigma(Y \circ \pi)(X^\#), \\ &= \Delta_\Sigma^{-1}(d_\Sigma Y^\#(X^\#))\Delta_\Sigma^{-1} - \text{sym}(\Delta_\Sigma^{-3} \text{Diag}(d_\Sigma Y^\#(X^\#))\Sigma\Delta_\Sigma^{-1}) \\ &\quad - \text{sym}[\Delta_\Sigma^{-3} \text{Diag}(X^\#)Y^\# \Delta_\Sigma^{-1} - \frac{3}{2}\Delta_\Sigma^{-5} \text{Diag}(X^\#)\text{Diag}(Y^\#)\Sigma\Delta_\Sigma^{-1} \\ &\quad + \Delta_\Sigma^{-3} \text{Diag}(Y^\#)X^\# \Delta_\Sigma^{-1} - \frac{1}{2}\Delta_\Sigma^{-3} \text{Diag}(Y^\#)\Sigma\text{Diag}(X^\#)\Delta_\Sigma^{-3}]. \end{aligned} \quad (13)$$

Since line (12) is line (13) at $\Sigma = C$, we finally get:

$$\begin{aligned} (\nabla_X^{\text{QA}} Y)|_C &= d_C Y(X) + \text{sym}[\text{Diag}(X^\#)Y^\# + \text{Diag}(Y^\#)X^\# + \text{Diag}(X^\# C^{-1} Y^\#)C \\ &\quad - X^\# C^{-1} Y^\# - \frac{1}{2}\text{Diag}(X^\#)C\text{Diag}(Y^\#) - \frac{3}{2}\text{Diag}(X^\#)\text{Diag}(Y^\#)C]. \end{aligned}$$

The curvature in formula (10) directly comes from the fundamental equations of submersions [11]. The curvature of the affine-invariant metric comes from [8]. Hence we only have to compute $G_C(\text{ver}[X^\#, Y^\#], \text{ver}[X^\#, Y^\#])$ where G is the affine-invariant metric, $[\cdot, \cdot]$ is the Lie bracket on the manifold of SPD matrices (it is not the matrix commutator), $C \in \text{Cor}^+(n)$ and $X^\#, Y^\#$ are horizontal vector fields on $\text{SPD}(n)$ extending $X_C^\#$ and $Y_C^\#$ respectively, where $X, Y \in T_C \text{Cor}^+(n)$ are tangent vectors at C . For example, we can consider that X, Y are constant vector fields on $\text{Cor}^+(n)$ and simply define $X^\#, Y^\#$ as their horizontal lifts everywhere:

$$\begin{aligned} Y_\Sigma^\# &= \text{hor}(\Delta_\Sigma Y \Delta_\Sigma), \\ &= \Delta_\Sigma Y \Delta_\Sigma - \text{ver}(\Delta_\Sigma Y \Delta_\Sigma), \\ &= \Delta_\Sigma Y \Delta_\Sigma - 2\Sigma \bullet \psi((I_n + \Sigma \bullet \Sigma^{-1})^{-1} \text{Diag}(\Sigma^{-1} \Delta_\Sigma Y \Delta_\Sigma) \mathbf{1}), \\ &= \Delta_\Sigma Y \Delta_\Sigma - 2\Sigma \bullet \psi((I_n + \Sigma \bullet \Sigma^{-1})^{-1} \text{Diag}(\pi(\Sigma)^{-1} Y) \mathbf{1}), \end{aligned}$$

with $\Delta_\Sigma = \text{Diag}(\Sigma)^{1/2}$ and $\psi(\mu) = \mu \mathbf{1}^\top + \mathbf{1} \mu^\top$ for $\mu \in \mathbb{R}^n$. Now we can compute $[X^\#, Y^\#] = \partial_{X^\#} Y^\# - \partial_{Y^\#} X^\#$ and evaluate it at C . We have:

$$\begin{aligned} \partial_{X^\#} Y^\# &= \frac{1}{2} (\Delta_\Sigma^{-1} \text{Diag}(X^\#) Y \Delta_\Sigma + \Delta_\Sigma Y \text{Diag}(X^\#) \Delta_\Sigma^{-1}) \\ &\quad - X^\# \bullet \psi((I_n + \Sigma \bullet \Sigma^{-1})^{-1} \text{Diag}(\Sigma^{-1} Y) \mathbf{1}) + v_\Sigma(X, Y), \end{aligned}$$

where $v_\Sigma(X, Y) = \Sigma \bullet \psi[(I_n + \Sigma \bullet \Sigma^{-1})^{-1} (X^\# \bullet \Sigma^{-1} - \Sigma \bullet \Sigma^{-1} X^\# \Sigma^{-1}) (I_n + \Sigma \bullet \Sigma^{-1})^{-1} \text{Diag}(\Sigma^{-1} Y) \mathbf{1} + (I_n + \Sigma \bullet \Sigma^{-1})^{-1} \text{Diag}(\pi(\Sigma)^{-1} X \pi(\Sigma)^{-1} Y) \mathbf{1}] \in \mathcal{V}_\Sigma$.

We evaluate it at $C \in \text{Cor}^+(n)$ so that $\Delta_C = I_n$. We denote $B = (I_n + C \bullet C^{-1})^{-1}$ to simplify the notations. Note that the last term is symmetric in X and Y so we can define $v_C^0(X, Y) = C \bullet \psi[B(X^\# \bullet C^{-1} - C \bullet C^{-1} X^\# C^{-1}) B \text{Diag}(C^{-1} Y) \mathbf{1}]$. Then:

$$[X^\#, Y^\#] = \frac{1}{2} (\text{Diag}(X^\#) Y + Y \text{Diag}(X^\#) - \text{Diag}(Y^\#) X - X \text{Diag}(Y^\#)), \quad (14)$$

$$- (X^\# \bullet \psi(B \text{Diag}(C^{-1} Y) \mathbf{1}) - Y^\# \bullet \psi(B \text{Diag}(C^{-1} X) \mathbf{1})), \quad (15)$$

$$+ v_C^0(X, Y) - v_C^0(Y, X). \quad (16)$$

Fortunately, the computation of $\text{ver}[X^\#, Y^\#] = C \bullet \psi(B \text{Diag}(C^{-1} [X^\#, Y^\#]) \mathbf{1})$ brings some simplifications. Indeed, to simplify line (14), we plug $Y = Y^\# - \frac{1}{2} (\text{Diag}(Y^\#) C + C \text{Diag}(Y^\#))$. Firstly, we have $\text{Diag}(C^{-1} \text{Diag}(X^\#) Y) = \text{Diag}(C^{-1} \text{Diag}(X^\#) Y^\#) - \frac{1}{2} \text{Diag}(C^{-1} \text{Diag}(X^\#) \text{Diag}(Y^\#) C) - \frac{1}{2} \text{Diag}(C^{-1} \text{Diag}(X^\#) C \text{Diag}(Y^\#))$ whose second term is symmetric in X and Y . Secondly, the other term is $\text{Diag}(C^{-1} Y \text{Diag}(X^\#)) = \text{Diag}(C^{-1} Y^\# \text{Diag}(X^\#)) - \frac{1}{2} \text{Diag}(C^{-1} \text{Diag}(Y^\#) C \text{Diag}(X^\#)) - \frac{1}{2} \text{Diag}(Y^\#) \text{Diag}(X^\#)$ whose first term is null, whose third term is symmetric in X and Y and whose second term is the symmetric of the third term of the previous expression. Hence, the vertical projection of the line (14) reduces to the vertical projection of $\frac{1}{2} (\text{Diag}(X^\#) Y^\# - \text{Diag}(Y^\#) X^\#)$. To simplify line (15), we can show by computing coordinate by coordinate that $\text{Diag}(C^{-1} (X^\# \bullet \psi(B \text{Diag}(C^{-1} Y) \mathbf{1}))) \mathbf{1} = (C^{-1} \bullet X^\#) B \text{Diag}(C^{-1} Y) \mathbf{1}$. Hence, line (15) cancels the first term of v^0 . Finally, a nicer expression of the second term of v^0 can be obtained by noticing that $\text{Diag}(Y^\#) \mathbf{1} = -2B \text{Diag}(C^{-1} Y) \mathbf{1}$. After simple calculations, we get $(C \bullet C^{-1} X^\# C^{-1}) B \text{Diag}(C^{-1} Y) \mathbf{1} = -\frac{1}{2} \text{Diag}(C \text{Diag}(X^\#) C^{-1} Y^\# C^{-1}) \mathbf{1}$.

To summarize, we have $\text{ver}[X^\#, Y^\#] = C \bullet \psi(\frac{1}{2} B D \mathbf{1})$ with $D = D(X, Y) - D(Y, X) \in \text{Diag}(n)$ and $D(X, Y) = \text{Diag}(C^{-1} \text{Diag}(X^\#) Y^\# - C^{-1} Y^\# C^{-1} \text{Diag}(X^\#) C)$. Finally, since $G_C(C \bullet \psi(\mu), C \bullet \psi(\mu)) = 2\mu^\top B \mu$ for any vector $\mu \in \mathbb{R}^n$, we get $G_C(\text{ver}[X^\#, Y^\#], \text{ver}[X^\#, Y^\#]) = \frac{1}{2} \mathbf{1}^\top D B D \mathbf{1}$, as expected.