Curvature of the manifold of fixed-rank positive-semidefinite matrices endowed with the Bures–Wasserstein metric*

Estelle Massart[1], Julien M. Hendrickx[1,2], and P.-A. Absil[1]

[1] ICTEAM, UCLouvain, Louvain-la-Neuve, Belgium
[2] CISE resident scholar at Boston University
estelle.massart@uclouvain.be

Abstract

We consider the manifold of rank-p positive-semidefinite matrices of size n, seen as a quotient of the set of full-rank n-by-p matrices by the orthogonal group in dimension p. The resulting distance coincides with the Wasserstein distance between centered degenerate Gaussian distributions. We obtain expressions for the Riemannian curvature tensor and the sectional curvature of the manifold. We also provide tangent vectors spanning planes associated with the extreme values of the sectional curvature.

1 Introduction

Positive-semidefinite (PSD) matrices appear, e.g., as covariance matrices in statistics, kernels in machine learning, and variables in semidefinite optimization; see, e.g., [MA18] for pointers to the literature.

The set of PSD matrices of size $n \times n$ is a stratified space [Tak11, Thm. C], in which the strata are the manifolds

$$\mathbb{S}_+(p, n) = \{S \in \mathbb{R}^{n\times n} | S \succeq 0, \operatorname{rank}(S) = p\},$$

of PSD matrices of rank p, for $p = 0, \dots, n$. In many practical applications, the rank of all the datapoints can be truncated to a common value, so that algorithms can be restricted to handle datapoints lying on the same stratum (see [MA18] and references within). This is for example the case when the data points are low-rank approximations of large PSD matrices. Each stratum $\mathbb{S}_+(p, n)$, with $p \geq 1$, can be given a Riemannian structure.

Classical algorithms on Riemannian manifolds can thus be used for processing data on $\mathbb{S}_+(p, n)$. For example, optimization on $\mathbb{S}_+(p, n)$ has been used

* This work was supported by (i) the Fonds de la Recherche Scientifique – FNRS and the Fonds Wetenschappelijk Onderzoek – Vlaanderen under EOS Project no 30468160, (ii) "Communauté française de Belgique - Actions de Recherche Concertées" (contract ARC 14/19-060), (iii) the WBI-World Excellence Fellowship.

in [MBS11,MMS11,MHB+16] for distance learning, distance matrix completion, and role model extraction. The works [LB14,GMM+17,KDB+18,MGS+19] run interpolation algorithms on $\mathbb{S}_+(p,n)$ for generating protein conformation transitions, modeling wind field, video classification and parametric model order reduction.

In the full-rank case, i.e., when $p = n$, the manifold $\mathbb{S}_+(n,n)$ is classically identified to the reductive homogeneous space $\mathbb{S}_+(n,n) \simeq \mathrm{GL}_n/\mathcal{O}_n$, where GL_n is the general linear group. Therefore, there exists a GL_n-invariant metric on $\mathrm{GL}_n/\mathcal{O}_n$ which leads (up to a scaling factor) to the natural, affine-invariant metric, or Fisher-Rao metric on $\mathbb{S}_+(n,n)$, see [Smi05]. When $p \neq n$, the set $\mathbb{S}_+(p,n)$ can be identified to a homogeneous space (see [VAV13]), but this homogeneous space is shown to be nonreductive, and there is no metric invariant under the group action. There is thus no wide agreement on a preferred metric on $\mathbb{S}_+(p,n)$.

In this work, we consider the identification $\mathbb{S}_+(p,n) \simeq \mathbb{R}_*^{n\times p}/\mathcal{O}_p$, with $\mathbb{R}_*^{n\times p}$ the set of full-rank n-by-p matrices. The quotient manifold $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$ is endowed with the metric induced from the Euclidean metric in $\mathbb{R}_*^{n\times p}$. This geometry was already proposed in [JBAS10] (which contains, e.g., expressions for the Riemannian exponential and for the projector on the horizontal space) and more recently described in [MA18]. In this last paper, we obtained expressions for the Riemannian logarithm, the injectivity radius and the cut locus. We mention that several other geometries have been proposed on $\mathbb{S}_+(p,n)$: [VAV09] represents $\mathbb{S}_+(p,n)$ as an embedded submanifold of $\mathbb{R}^{n\times n}$, [BS09] identifies it to the quotient manifold $(\mathrm{St}(p,n) \times \mathbb{S}_+(p,p))/\mathcal{O}_p$, and, as already mentioned, [VAV13] identifies $\mathbb{S}_+(p,n)$ to a homogeneous space endowed with a right-invariant metric.

Even though the metric resulting from the identification $\mathbb{S}_+(p,n) \simeq \mathbb{R}_*^{n\times p}/\mathcal{O}_p$ does not lead to a complete metric space, there are two main motivations to consider it. The first one is the low computation cost associated with the most common operations on the manifold. Indeed, the operations are directly performed on the representatives in $\mathbb{R}_*^{n\times p}$ of the matrices, which are smaller than the initial $n \times n$ matrices. As shown in [JBAS10,MA18], the Riemannian exponential and logarithm have a computational cost that evolves linearly with n. Among all the geometries proposed for $\mathbb{S}_+(p,n)$, this is to our knowledge the only one that leads to expressions for both the logarithm and the exponential maps that are cheap to evaluate.

The second motivation to consider this quotient geometry is its interpretation with respect to optimal transport theory. Indeed, there exists a bijection between the set of $n\times n$ PSD matrices and the set of (possibly degenerate) centered Gaussian distributions on $\mathbb{R}^n$. Let $C_1, C_2 \in \mathbb{S}_+(n,n)$, two nonsingular covariance matrices, and let $\mathrm{W}_2(\mu_1,\mu_2)$ be the 2-Wasserstein distance between the nondegenerate centered Gaussian distributions $\mu_1 := \mathcal{N}(0,C_1)$ and $\mu_2 := \mathcal{N}(0,C_2)$. It is well-known that $\mathrm{W}_2(\mu_1,\mu_2)$ coincides with the Riemannian distance between C_1 and C_2, for the metric inherited from the quotient representation $\mathbb{S}_+(n,n) \simeq \mathrm{GL}(n)/\mathcal{O}_n$ (see, e.g., [Tak11, BJL18]). When $C_1, C_2 \in \mathbb{S}_+(p,n)$, for $p < n$, the same conclusion holds: $\mathrm{W}_2(\mu_1,\mu_2)$ is equal to the Riemannian distance between the low-rank covariance matrices C_1 and C_2, for the metric induced by the quotient $\mathbb{S}_+(p,n) \simeq \mathbb{R}_*^{n\times p}/\mathcal{O}_p$ [Gel90, Cor. 2.5]. Specifically, the distance is given by (see [MA18, §2.10]):

$$\mathrm{d}(C_1,C_2) = \left[\mathrm{tr}\,(C_1) + \mathrm{tr}\,(C_2) - 2\mathrm{tr}\left(\left(C_1^{1/2} C_2 C_1^{1/2}\right)^{1/2}\right)\right]^{1/2}.$$

The Wasserstein metric is also known as the Bures metric in quantum theory (see [BJL18] and references therein).

Geometric properties of the manifold $\mathbb{S}_+(n,n) \simeq \mathrm{GL}(n)/\mathcal{O}_n$ have been widely studied, see, e.g., [Tak11, BJL18, MMP18]. In particular, its sectional curvature has been computed in [Tak11]. The contribution of this paper is to compute the Riemannian curvature tensor and the sectional curvature of the manifold $\mathbb{S}_+(p,n) \simeq \mathbb{R}_*^{n\times p}/\mathcal{O}_p$. We also provide tangent vectors spanning tangent planes associated with the maximal and minimal sectional curvatures. Bounds on the curvature of the manifold appear, e.g., in some optimization algorithms and associated convergence results on manifolds [ATV13, Bon13], and in guarantees for the continuity of the result of some curve fitting algorithms [AGSW16]. The Riemannian curvature tensor is, e.g., used in [SASK12] for curve fitting on manifolds. We show that the sectional curvature is non-negative, and may become infinitely large when approaching the boundary of the manifold (specifically, if two singular values go simultaneously to zero). A consequence is that some of the above-mentioned results involving bounds on the sectional curvature (in optimization or curve fitting) do not directly apply on this manifold. Our conclusions agree with the work [Dit95], which computes the curvature of the manifold of density matrices ($n\times n$ positive-definite complex matrices of unit trace), endowed with the Bures metric, and observes a similar unboundedness of the sectional curvature as the rank of the matrix goes to $n-2$.

The structure of this paper is as follows. Section 2 presents a brief summary of the geometry of $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$. In Section 3, we derive expressions for the Riemannian curvature tensor and the sectional curvature. Finally, we compute in Section 4 the extreme values of the sectional curvature.

2 Geometry of the manifold $\mathbb{S}_+(p,n) \simeq \mathbb{R}_*^{n\times p}/\mathcal{O}_p$

This quotient geometry, described in [MA18], relies on the characterization $\mathbb{S}_+(p,n) = \{YY^\top | Y \in \mathbb{R}_*^{n\times p}\}$. The quotient representation comes from the fact that the set of points $Y\mathcal{O}_p := \{YQ | Q \in \mathcal{O}_p\}$ is a fiber under the map $Y \mapsto YY^\top$. The tangent space $\mathcal{T}_Y \mathbb{R}_*^{n\times p} \simeq \mathbb{R}^{n\times p}$ is the direct sum of two orthogonal subspaces: the vertical space (the tangent space of the fiber $Y\mathcal{O}_p$), and the horizontal space (its orthogonal complement, with respect here to the Euclidean metric). The vertical space at Y is given by $\mathcal{V}_Y = \{Y\Omega | \Omega = -\Omega^\top \in \mathbb{R}^{p\times p}\}$, while the horizontal space is $\mathcal{H}_Y = \{\bar{\eta}_Y = Y(Y^\top Y)^{-1}S + Y_\perp K | S \in \mathbb{R}^{p\times p}, S = S^\top, K \in \mathbb{R}^{(n-p)\times p}\}$. Let $\pi : \mathbb{R}_*^{n\times p} \to \mathbb{R}_*^{n\times p}/\mathcal{O}_p$ be the quotient map, mapping points from $\mathbb{R}_*^{n\times p}$ to their fibers. For any $Y \in \mathbb{R}_*^{n\times p}$, any tangent vector $\xi_{\pi(Y)} \in \mathcal{T}_{\pi(Y)}\mathbb{R}_*^{n\times p}/\mathcal{O}_p$ is associated to a unique horizontal lift

$\bar{\xi}_Y \in \mathcal{H}_Y$, such that $\xi_{\pi(Y)} = D\pi(Y)[\bar{\xi}_Y]$. The metric in $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$ is defined as $g_{\pi(Y)}\left(\xi_{\pi(Y)}, \eta_{\pi(Y)}\right) := \operatorname{tr}\left(\bar{\xi}_Y^\top \bar{\eta}_Y\right)$, which turns the quotient map π into a Riemannian submersion. Finally, given two horizontal vector fields $\bar{\xi}$, $\bar{\eta}$, the projection on the vertical space of the bracket $[\bar{\xi}, \bar{\eta}]$ is:

$$\mathrm{P^v}_Y[\bar{\xi}, \bar{\eta}] = Y\mathbf{T}_{Y^\top Y}^{-1}\left(2\left(\bar{\eta}_Y^\top \bar{\xi}_Y - \bar{\xi}_Y^\top \bar{\eta}_Y\right)\right), \tag{1}$$

with $\mathbf{T}_{Y^\top Y}^{-1}(\Omega)$ the unique solution X to the Sylvester equation $Y^\top Y X + XY^\top Y = \Omega$, see [MA18, Prop. 2.37].

3 Curvature of the manifold $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$

In this section, we obtain expressions for the Riemannian curvature tensor and the sectional curvature of the manifold $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$. We rely on the fact that the operator $\pi : \mathbb{R}_*^{n\times p} \to \mathbb{R}_*^{n\times p}/\mathcal{O}_p$ is a Riemannian submersion.

Theorem 1. *Let ξ, η, α and β be vector fields on $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$, and let $\bar{\xi}, \bar{\eta}, \bar{\alpha}$ and $\bar{\beta}$ be their horizontal lifts. The Riemannian curvature tensor at $\pi(Y)$ satisfies:*

$$\begin{aligned}\langle R_{\mathbb{R}_*^{n\times p}/\mathcal{O}_p}(\xi_{\pi(Y)}, \eta_{\pi(Y)})\alpha_{\pi(Y)}, \beta_{\pi(Y)}\rangle &= \frac{1}{2}\langle \mathrm{P^v}_Y[\bar{\xi}, \bar{\eta}], \mathrm{P^v}_Y[\bar{\alpha}, \bar{\beta}]\rangle \\ &-\frac{1}{4}\left(\langle \mathrm{P^v}_Y[\bar{\eta}, \bar{\alpha}], \mathrm{P^v}_Y[\bar{\xi}, \bar{\beta}]\rangle - \langle \mathrm{P^v}_Y[\bar{\xi}, \bar{\alpha}], \mathrm{P^v}_Y[\bar{\eta}, \bar{\beta}]\rangle\right),\end{aligned}$$

with $\mathrm{P^v}_Y[\bar{\xi}, \bar{\eta}]$ given by (1).

Proof. According to [O'N66, Thm. 2], there holds:

$$\begin{aligned}&\langle R_{\mathbb{R}_*^{n\times p}/\mathcal{O}_p}(\xi_{\pi(Y)}, \eta_{\pi(Y)})\alpha_{\pi(Y)}, \beta_{\pi(Y)}\rangle = \langle R_{\mathbb{R}_*^{n\times p}}(\xi_Y, \eta_Y)\alpha_Y, \beta_Y\rangle \\ &+ \frac{1}{2}\langle \mathrm{P^v}_Y[\bar{\xi}, \bar{\eta}], \mathrm{P^v}_Y[\bar{\alpha}, \bar{\beta}]\rangle - \frac{1}{4}\langle \mathrm{P^v}_Y[\bar{\eta}, \bar{\alpha}], \mathrm{P^v}_Y[\bar{\xi}, \bar{\beta}]\rangle - \frac{1}{4}\langle \mathrm{P^v}_Y[\bar{\alpha}, \bar{\xi}], \mathrm{P^v}_Y[\bar{\eta}, \bar{\beta}]\rangle.\end{aligned}$$

Since $\mathbb{R}_*^{n\times p}$ is an open subset of $\mathbb{R}^{n\times p}$, its Riemannian curvature tensor is zero [O'N83, p.79], hence the first term of the previous expression vanishes. □

The sectional curvature is then obtained as a corollary, see [O'N66, Cor. 1, eq. 3]. In the case $n = p$, these results are already given in [Tak11].

Corollary 1. *Let $\xi_{\pi(Y)}, \eta_{\pi(Y)}$ be (independent) tangent vectors on $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$, with horizontal lifts $\bar{\xi}_Y$, $\bar{\eta}_Y$. The sectional curvature at $\pi(Y)$ in $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$ is*

$$K_{\mathbb{R}_*^{n\times p}/\mathcal{O}_p}(\xi_{\pi(Y)}, \eta_{\pi(Y)}) = \frac{3\left\|Y\mathbf{T}_{Y^\top Y}^{-1}\left(\bar{\eta}_Y^\top \bar{\xi}_Y - \bar{\xi}_Y^\top \bar{\eta}_Y\right)\right\|_{\mathrm{F}}^2}{\langle \bar{\xi}_Y, \bar{\xi}_Y\rangle\langle \bar{\eta}_Y, \bar{\eta}_Y\rangle - \langle \bar{\xi}_Y, \bar{\eta}_Y\rangle^2}. \tag{2}$$

The rest of the paper aims at computing the maximal and minimal sectional curvatures at an arbitrary $\pi(Y) \in \mathbb{R}_*^{n\times p}/\mathcal{O}_p$.

4 Extreme values of the sectional curvature

We first introduce two lemmas. The first one solves for X a Sylvester equation of the form $Y^\top Y X + X Y^\top Y = \Omega$, a step required to evaluate (2).

Lemma 1. *Let $Y \in \mathbb{R}_*^{n\times p}$, with $Y =: U\Sigma V^\top$ a singular value decomposition, with singular values $\sigma_1 \geq \cdots \geq \sigma_p > 0$, and let $\Omega \in \mathbb{R}^{p\times p}$. The solution X to the Sylvester equation $Y^\top Y X + X Y^\top Y = \Omega$ is*

$$X = V\tilde{X}V^\top, \text{ with } \tilde{X} \in \mathbb{R}^{p\times p}, \ \tilde{X}_{ij} := \frac{\tilde{\Omega}_{ij}}{(\sigma_i^2 + \sigma_j^2)}, \ \tilde{\Omega} := V^\top \Omega V. \tag{3}$$

Moreover, if the matrix Ω is skew-symmetric, then so are $\tilde{X}$ and X.

Proof. We sketch the proof, presented in [BR97, §10], for the reader's convenience. Since $Y^\top Y = V\Sigma^2 V^\top$, the Sylvester equation becomes: $V\Sigma^2 V^\top X + XV\Sigma^2 V^\top = \Omega$. Applying a similarity associated with V to both sides of the equation yields: $\Sigma^2 V^\top X V + V^\top X V \Sigma^2 = V^\top \Omega V$. Now, defining $\tilde{X} := V^\top X V$ and $\tilde{\Omega} := V^\top \Omega V$, the equation becomes: $\Sigma^2 \tilde{X} + \tilde{X}\Sigma^2 = \tilde{\Omega}$, which implies that $(\sigma_i^2 + \sigma_j^2)\tilde{X}_{ij} = \tilde{\Omega}_{ij}$. □

The second lemma provides an upper bound on the Frobenius norm of the skew part of the product of two matrices with unit norm. We will need this result when computing the maximal sectional curvature of $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$ at some point $\pi(Y) \in \mathbb{R}_*^{n\times p}/\mathcal{O}_p$.

Lemma 2. *Let $A, B \in \mathbb{R}^{n\times p}$, such that $\|A\|_\mathrm{F} = \|B\|_\mathrm{F} = 1$. Then,*

$$\left\|A^\top B - B^\top A\right\|_\mathrm{F}^2 \leq 2.$$

Proof. Let us consider the optimization problem:

$$\max_{\|A\|_\mathrm{F} = \|B\|_\mathrm{F} = 1} \left\|A^\top B - B^\top A\right\|_\mathrm{F}^2.$$

Observe that, by symmetry of the problem, the Lagrange multipliers associated with the constraints $\|A\|_\mathrm{F} = 1$ and $\|B\|_\mathrm{F} = 1$ are equal, and that the linear independence constraint qualification (LICQ) condition holds. Hence the KKT first-order necessary optimality conditions are:

$$\begin{cases} 2B(B^\top A - A^\top B) - \lambda A = 0 & \text{(4.a)} \\ -2A(B^\top A - A^\top B) - \lambda B = 0 & \text{(4.b)} \\ \|A\|_\mathrm{F} = \|B\|_\mathrm{F} = 1. & \text{(4.c)} \end{cases}$$

Premultiplying (4.a) by $A^\top$, (4.b) by $B^\top$, and taking the sum of the two yields:

$$\lambda(A^\top A + B^\top B) = 2(A^\top B - B^\top A)(B^\top A - A^\top B).$$

Taking the trace of both sides of the equation, we obtain:

$$\lambda = \operatorname{tr}\left((B^\top A - A^\top B)^\top (B^\top A - A^\top B)\right) = \left\|A^\top B - B^\top A\right\|_{\mathrm{F}}^2 .$$

We will show that $\lambda \leq 2$, which will conclude the proof. If $B = 0$, then the claim obviously holds, hence we assume from now on that $B \neq 0$. Let $B = U\Sigma V^\top$ be a compact singular value decomposition, where $U \in \mathbb{R}^{n\times r}$, $\Sigma \in \mathbb{R}^{r\times r}$ and $V \in \mathbb{R}^{p\times r}$, with r the rank of B and $U^\top U = V^\top V = I_r$. Equation (4.a) becomes:

$$2U\Sigma^2 U^\top A - 2U\Sigma V^\top A^\top U\Sigma V^\top = \lambda A.$$

Left- and right-multiplying this equation by respectively $U^\top$ and V yields:

$$2\Sigma^2 U^\top AV - 2\Sigma V^\top A^\top U\Sigma = \lambda U^\top AV.$$

Now, defining $\tilde{A} := \Sigma U^\top AV$, we get:

$$2\Sigma\tilde{A} - 2\Sigma\tilde{A}^\top = \lambda\Sigma^{-1}\tilde{A},$$

which can be written as:

$$2\Sigma^2\tilde{A} - 2\Sigma^2\tilde{A}^\top = \lambda\tilde{A}. \tag{5}$$

Assume first that $\tilde{A} \neq 0$. Then, if $r = 1$, $\lambda = 0$. If $r \geq 2$, the coefficients $\tilde{A}_{ij}$, $i, j = 1, \ldots, r$ of the matrix $\tilde{A}$ satisfy the equation:

$$\lambda(\tilde{A}_{ij} - \tilde{A}_{ji}) = 2(\sigma_i^2 + \sigma_j^2)(\tilde{A}_{ij} - \tilde{A}_{ji}).$$

If for some $i, j \in \{1, \ldots, r\}$, $\tilde{A}_{ij} \neq \tilde{A}_{ji}$ there holds $\lambda = 2(\sigma_i^2 + \sigma_j^2) \leq 2\left\|B\right\|_{\mathrm{F}}^2 = 2$. Otherwise (i.e., $\tilde{A} \neq 0$ is symmetric), $\lambda = 0$ by (5).

There remains to check the value of λ when $\tilde{A} = 0$. It can be readily checked that the matrix $[V, V_\perp]^\top B^\top A[V, V_\perp]$ is of the form:

$$[V, V_\perp]^\top B^\top A[V, V_\perp] = \begin{bmatrix} \tilde{A} & \Sigma U^\top AV_\perp \\ 0_{p-r\times r} & 0_{p-r\times p-r} \end{bmatrix}.$$

Since $\tilde{A} = 0$, the matrix is strictly upper triangular. There holds

$$\left\|B^\top A - A^\top B\right\|_{\mathrm{F}}^2 = \left\|[V, V_\perp]^\top (B^\top A - A^\top B)[V, V_\perp]\right\|_{\mathrm{F}}^2 = 2\left\|\Sigma U^\top AV_\perp\right\|_{\mathrm{F}}^2 \leq 2,$$

which concludes the proof. □

We are now able to compute the minimum and maximum values of the sectional curvature of $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$ at some point $\pi(Y)$. Observe that, since the sectional curvature is associated to a tangent plane, it does not depend on the choice of the vectors $\xi_{\pi(Y)}$, $\eta_{\pi(Y)}$ that span this tangent plane. As a result, we make the assumption in the rest of the document that the horizontal lifts $\bar{\xi}_Y$ and $\bar{\eta}_Y$ are orthonormal vectors, i.e., $\langle\bar{\xi}_Y, \bar{\xi}_Y\rangle = \langle\bar{\eta}_Y, \bar{\eta}_Y\rangle = 1$ and $\langle\bar{\xi}_Y, \bar{\eta}_Y\rangle = 0$. This makes the denominator of (2) equal to one.

Proposition 1. *The minimum of the sectional curvature at $\pi(Y)$ of the quotient manifold $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$ is always zero. If $p=1$, the sectional curvature is equal to zero.*

Proof. By (2), the sectional curvature associated with a pair of orthonormal tangent vectors $\xi_{\pi(Y)}, \eta_{\pi(Y)}$ is defined as:

$$K_{\mathbb{R}_*^{n\times p}/\mathcal{O}_p}(\xi_{\pi(Y)}, \eta_{\pi(Y)}) = 3\left\|Y\mathbf{T}^{-1}_{Y^\top Y}(\bar{\eta}_Y^\top\bar{\xi}_Y - \bar{\xi}_Y^\top\bar{\eta}_Y)\right\|_{\mathrm{F}}^2.$$

Using Lemma 1, with $Y = U\Sigma V^\top$ a singular value decomposition and $\tilde{\Omega} := V^\top(\bar{\eta}_Y^\top\bar{\xi}_Y - \bar{\xi}_Y^\top\bar{\eta}_Y)V$, there holds:

$$K_{\mathbb{R}_*^{n\times p}/\mathcal{O}_p}(\xi_{\pi(Y)}, \eta_{\pi(Y)}) = 3\left\|\left(U\Sigma V^\top\right)\left(V\tilde{X}V^\top\right)\right\|_{\mathrm{F}}^2, \quad \tilde{X}_{ij} = \frac{\tilde{\Omega}_{ij}}{(\sigma_i^2+\sigma_j^2)}.$$

Due to the unitarily invariance of the Frobenius norm, there holds:

$$K_{\mathbb{R}_*^{n\times p}/\mathcal{O}_p}(\xi_{\pi(Y)}, \eta_{\pi(Y)}) = 3\left\|\Sigma\tilde{X}\right\|_{\mathrm{F}}^2 = 3\sum_{i,j=1}^{p}\frac{\sigma_i^2\tilde{\Omega}_{ij}^2}{(\sigma_i^2+\sigma_j^2)^2}. \tag{6}$$

This is zero if and only if $\tilde{\Omega}$ is zero. If $p=1$, the sectional curvature is always zero since $\bar{\eta}_Y^\top\bar{\xi}_Y \in \mathbb{R}$. If $p\geq 2$, take for example $\bar{\xi}_Y = Y(Y^\top Y)^{-1}S$ with $S=S^\top$, and $\bar{\eta}_Y = Y\left\|Y\right\|_{\mathrm{F}}^{-1}$. Then, $\tilde{\Omega}=0$, and if the matrix S is chosen such that $\left\|\bar{\xi}_Y\right\|_{\mathrm{F}}=1$ and $\mathrm{Diag}\,(S)=0$, the two vectors $\bar{\xi}_Y$ and $\bar{\eta}_Y$ are orthonormal. □

The following result characterizes the maximum of the sectional curvature of $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$ at some point $\pi(Y)$.

Proposition 2. *Let $Y\in\mathbb{R}_*^{n\times p}$ and $Y=U\Sigma V^\top$ a singular value decomposition, with singular values $\sigma_1\geq\sigma_2\geq\cdots\geq\sigma_p>0$. If $p=1$, the sectional curvature is always zero. If $p\geq 2$, the maximum of the sectional curvature at $\pi(Y)$ of the quotient $\mathbb{R}_*^{n\times p}/\mathcal{O}_p$ is:*

$$K_{\mathbb{R}_*^{n\times p}/\mathcal{O}_p}(\xi^*_{\pi(Y)}, \eta^*_{\pi(Y)}) = \frac{3}{\sigma_{p-1}^2+\sigma_p^2}. \tag{7}$$

This value is reached for, e.g., $\xi^*_{\pi(Y)} = D\pi(Y)[\bar{\xi}^*_Y]$ and $\eta^*_{\pi(Y)} = D\pi(Y)[\bar{\eta}^*_Y]$, with $\bar{\xi}^*_Y = Y(Y^\top Y)^{-1}S_\xi$ and $\bar{\eta}^*_Y = Y(Y^\top Y)^{-1}S_\eta$, where

$$S_\xi := \frac{V(E_{p-1,p-1}-E_{p,p})V^\top}{\sqrt{\sigma_{p-1}^{-2}+\sigma_p^{-2}}} \qquad S_\eta := \frac{V(E_{p-1,p}+E_{p,p-1})V^\top}{\sqrt{\sigma_{p-1}^{-2}+\sigma_p^{-2}}},$$

with E_{ij} the matrix whose elements are zero excepted $E(i,j)=1$.

Proof. Similarly as in the proof of Proposition 1, let us write:

$$K_{\mathbb{R}_*^{n\times p}/\mathcal{O}_p}(\xi_{\pi(Y)},\eta_{\pi(Y)}) = 3\sum_{i,j=1}^{p}\frac{\sigma_i^2\tilde{\Omega}_{ij}^2}{(\sigma_i^2+\sigma_j^2)^2} = 3\sum_{i>j}\frac{\tilde{\Omega}_{ij}^2}{(\sigma_i^2+\sigma_j^2)}, \tag{8}$$

where the last inequality comes from the fact that $\tilde{\Omega} := V^\top(\bar{\eta}_Y^\top\bar{\xi}_Y - \bar{\xi}_Y^\top\bar{\eta}_Y)V$ is skew-symmetric. According to Lemma 2, the squared Frobenius norm of $\tilde{\Omega}$ is upper bounded by 2:

$$\left\|\tilde{\Omega}\right\|_{\mathrm{F}}^2 = \left\|V^\top(\bar{\eta}_Y^\top\bar{\xi}_Y - \bar{\xi}_Y^\top\bar{\eta}_Y)V\right\|_{\mathrm{F}}^2 = \left\|\bar{\eta}_Y^\top\bar{\xi}_Y - \bar{\xi}_Y^\top\bar{\eta}_Y\right\|_{\mathrm{F}}^2 \leq 2.$$

Therefore:

$$K_{\mathbb{R}_*^{n\times p}/\mathcal{O}_p}(\xi_{\pi(Y)},\eta_{\pi(Y)}) \leq \frac{3\sum_{i>j}\tilde{\Omega}_{ij}^2}{(\sigma_{p-1}^2+\sigma_p^2)} \leq \frac{3\|\tilde{\Omega}\|_{\mathrm{F}}^2}{2(\sigma_{p-1}^2+\sigma_p^2)} \leq \frac{3}{\sigma_{p-1}^2+\sigma_p^2}.$$

To finish the proof, we show that this bound is reached for the vectors $\bar{\xi}_Y^*$ and $\bar{\eta}_Y^*$ given in the proposition. It can be readily checked that $\bar{\xi}_Y^*$ and $\bar{\eta}_Y^*$ are orthogonal and have unit norm. There remains to compute $\tilde{\Omega}^*$:

$$\bar{\eta}_Y^{*\top}\bar{\xi}_Y^* = S_\eta(Y^\top Y)^{-1}S_\xi = \frac{V(E_{p-1,p}+E_{p,p-1})\Sigma^{-2}(E_{p-1,p-1}-E_{p,p})V^\top}{\sigma_{p-1}^{-2}+\sigma_p^{-2}},$$

which simply becomes

$$\bar{\eta}_Y^{*\top}\bar{\xi}_Y^* = \frac{V(\sigma_{p-1}^{-2}E_{p,p-1}-\sigma_p^{-2}E_{p-1,p})V^\top}{\sigma_{p-1}^{-2}+\sigma_p^{-2}}.$$

Therefore, $\tilde{\Omega}^*$ is:

$$\tilde{\Omega}^* = \frac{(\sigma_{p-1}^{-2}+\sigma_p^{-2})E_{p,p-1}-(\sigma_{p-1}^{-2}+\sigma_p^{-2})E_{p-1,p}}{\sigma_{p-1}^{-2}+\sigma_p^{-2}} = (E_{p,p-1}-E_{p-1,p}),$$

such that

$$K_{\mathbb{R}_*^{n\times p}/\mathcal{O}_p}(\xi_{\pi(Y)}^*,\eta_{\pi(Y)}^*) = \frac{3}{\sigma_{p-1}^2+\sigma_p^2}.$$ □

5 Conclusion

We have computed the curvature of the manifold $\mathbb{S}_+(p,n)$ endowed with the Bures–Wasserstein metric. We have provided expressions for the Riemannian curvature tensor and the sectional curvature of the manifold. We have shown that in the case $p = 1$ the sectional curvature is always zero. If $p \geq 2$, the minimum over the tangent planes of the sectional curvature is zero, while the maximum goes to infinity as the p^{th} and $p-1^{\mathrm{th}}$ eigenvalues of the PSD matrix go simultaneously to zero. Further works might aim at computing the curvature of $\mathbb{S}_+(p,n)$ endowed with the other metrics proposed in the literature (see [VAV09, BS09, VAV13]), which to our knowledge are still unknown.

References

AGSW16. P.-A. Absil, Pierre-Yves Gousenbourger, Paul Striewski, and Benedikt Wirth. Differentiable piecewise-Bézier surfaces on Riemannian manifolds. *SIAM Journal on Imaging Sciences*, 9(4):1788–1828, 2016. doi:10.1137/16M1057978.

ATV13. Bijan Afsari, Roberto Tron, and René Vidal. On the convergence of gradient descent for finding the Riemannian center of mass. *SIAM Journal on Control and Optimization*, 51(3):2230–2260, 2013. doi:10.1137/12086282X.

BJL18. Rajendra Bhatia, Tanvi Jain, and Yongdo Lim. On the Bures–Wasserstein distance between positive definite matrices. *Expositiones Mathematicae*, 2018. doi:10.1016/j.exmath.2018.01.002.

Bon13. Silvère Bonnabel. Stochastic gradient descent on Riemannian manifolds. *IEEE Transactions Automatic Control*, 58(9):2217–2229, 2013. doi:10.1109/TAC.2013.2254619.

BR97. Rajendra Bhatia and Peter Rosenthal. How and why to solve the operator equation AX- XB= Y. *Bulletin of the London Mathematical Society*, 29(1):1–21, 1997.

BS09. Silvère Bonnabel and Rodolphe Sepulchre. Riemannian metric and geometric mean for positive semidefinite matrices of fixed rank. *SIAM Journal on Matrix Analysis and Applications*, 31(3):1055–1070, 2009. doi:10.1137/080731347.

Dit95. Jochen Dittmann. On the Riemannian metric on the space of density matrices. *Reports on Mathematical Physics*, 36(2-3):309–315, 1995. doi:10.1016/0034-4877(96)83627-5.

Gel90. Matthias Gelbrich. On a formula for the L^2 Wasserstein metric between measures on Euclidean and Hilbert spaces. *Mathematische Nachrichten*, 147(1):185–203, 1990. doi:10.1002/mana.19901470121.

GMM+17. P.-Y. Gousenbourger, E. Massart, A. Musolas, P.-A. Absil, L. Jacques, J. M. Hendrickx, and Y. Marzouk. Piecewise-Bézier C^1 smoothing on manifolds with application to wind field estimation. *Proceedings of the* 25^{th} *European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning (ESANN)*, pages 305–310, 2017.

JBAS10. M. Journée, F. Bach, P.-A. Absil, and R. Sepulchre. Low-rank optimization on the cone of positive semidefinite matrices. *SIAM Journal on Optimization*, 20(5):2327–2351, 2010. doi:10.1137/080731359.

KDB+18. Anis Kacem, Mohamed Daoudi, Boulbaba Ben Amor, Stefano Berretti, and Juan Carlos Alvarez-Paiva. A Novel Geometric Framework on Gram Matrix Trajectories for Human Behavior Understanding. *IEEE Transactions on Pattern Analysis and Machine Intelligence (T-PAMI)*, 2018. doi:10.1109/tpami.2018.2872564.

LB14. X.-B. Li and F. J. Burkowski. Conformational transitions and principal geodesic analysis on the positive semidefinite matrix manifold. In Wang J. Basu M., Pan Y., editor, *Bioinformatics Research and Applications*, ISBRA 2014, pages 334–345, 2014. doi:10.1007/978-3-319-08171-7_30.

MA18. Estelle Massart and P.-A. Absil. Quotient geometry of the manifold of fixed-rank positive-semidefinite matrices. Technical Report UCL-INMA-2018.06, UCLouvain, November 2018. Preprint: http://sites.uclouvain.be/absil/2018.06.

MBS11. Gilles Meyer, Silvère Bonnabel, and Rodolphe Sepulchre. Regression on fixed-rank positive semidefinite matrices: a Riemannian approach. *Journal of Machine Learning Research*, 12(Feb):593–625, 2011.

MGS+19. Estelle Massart, Pierre-Yves Gousenbourger, Nguyen Thanh Son, Tatjana Stykel, and P.-A. Absil. Interpolation on the manifold of fixed-rank positive-semidefinite matrices for parametric model order reduction: preliminary results. In *Proceedings of the 27th European Symposium on Artifical Neural Networks, Computational Intelligence and Machine Learning (ESANN2019)*, pages 281–286, 2019.

MHB+16. Melissa Marchand, Wen Huang, Arnaud Browet, Paul Van Dooren, and Kyle A. Gallivan. A Riemannian optimization approach for role model extraction. In *Proceedings of the 22nd International Symposium on Mathematical Theory of Networks and Systems*, pages 58–64, 2016.

MMP18. Luigi Malagò, Luigi Montrucchio, and Giovanni Pistone. Wasserstein Riemannian geometry of Gaussian densities. *Information Geometry*, 1(2):137–179, 2018. `doi:10.1007/s41884-018-0014-4`.

MMS11. B. Mishra, G. Meyer, and R. Sepulchre. Low-rank optimization for distance matrix completion. In *Proceedings of the 50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 4455–4460, 2011. `doi:10.1109/CDC.2011.6160810`.

O'N66. Barrett O'Neill. The fundamental equations of a submersion. *The Michigan Mathematical Journal*, 13(4):459–469, 1966. `doi:10.1307/mmj/1028999604`.

O'N83. Barrett O'Neill. *Semi-Riemannian Geometry*, volume 103 of *Pure and Applied Mathematics*. Academic Press Inc. [Harcourt Brace Jovanovich Publishers], New York, 1983.

SASK12. Chafik Samir, P.-A. Absil, Anuj Srivastava, and Eric Klassen. A gradient-descent method for curve fitting on Riemannian manifolds. *Foundations of Computational Mathematics*, 12(1):49–73, 2012. `doi:10.1007/s10208-011-9091-7`.

Smi05. S.T. Smith. Covariance, Subspace, and Intrinsic Cramér–Rao Bounds. *IEEE Transactions on Signal Processing*, 53(5):1610–1630, may 2005. `doi:10.1109/tsp.2005.845428`.

Tak11. A. Takatsu. Wasserstein geometry of Gaussian measures. *Osaka Journal of Mathematics*, 48(4):1005–1026, 2011.

VAV09. Bart Vandereycken, P.-A. Absil, and Stefan Vandewalle. Embedded geometry of the set of symmetric positive semidefinite matrices of fixed rank. In *IEEE/SP 15th Workshop on Statistical Signal Processing*, pages 389–392, 2009. `doi:10.1109/SSP.2009.5278558`.

VAV13. Bart Vandereycken, P.-A. Absil, and Stefan Vandewalle. A Riemannian geometry with complete geodesics for the set of positive semidefinite matrices of fixed rank. *IMA Journal of Numerical Analysis*, 33(2):481–514, 2013. `doi:10.1093/imanum/drs006`.