BIT

Geometric means of structured matrices

**Dario A. Bini · Bruno Iannazzo · Ben Jeuris ·
Raf Vandebril**

Abstract The geometric mean of positive definite matrices is usually identified with the Karcher mean, which possesses all properties—generalized from the scalar case—a geometric mean is expected to satisfy. Unfortunately, the Karcher mean is typically not structure preserving, and destroys, e.g., Toeplitz and band structures, which emerge in many applications. For this reason, the Karcher mean is not always recommended for modeling averages of structured matrices. In this article a new definition of a geometric mean for structured matrices is introduced, its properties are outlined, algorithms for its computation, and numerical experiments are provided. In the Toeplitz case an existing mean based on the Kähler metric is analyzed for comparison.

D.A. Bini
Università di Pisa, Pisa, Italy
e-mail: bini@dm.unipi.it

B. Iannazzo
Università di Perugia, Perugia, Italy
e-mail: bruno.iannazzo@dmi.unipg.it

B. Jeuris · R. Vandebril (✉)
KU Leuven, Leuven, Belgium
e-mail: raf.vandebril@cs.kuleuven.be

B. Jeuris
e-mail: ben.jeuris@cs.kuleuven.be

✡ Springer

1 Introduction

The wish to generalize the concept of a geometric mean to positive definite (*positive* for short) matrices and, on the other hand, the need to average quantities expressed by positive matrices in certain applications have led to the definition and the study of the Karcher mean [8, 9, 33].

Without describing all facets, one can consider the set of $n \times n$ positive matrices, denoted by $\mathcal{P}_n$, as a manifold [1], in particular, there is a diffeomorphism from $\mathcal{P}_n$ to $\mathbb{R}^{n^2}$. In each point of $A \in \mathcal{P}_n$ one can define the tangent space $T_A \mathcal{P}_n$, which can be identified with the space of Hermitian matrices. The Karcher mean can now be defined in terms of a Riemannian geometry defined on $\mathcal{P}_n$ and induced by the inner product

$$g_A(X, Y) := \operatorname{tr}\left(A^{-1} X A^{-1} Y\right), \quad X, Y \in T_A \mathcal{P}_n \tag{1.1}$$

on the tangent space $T_A \mathcal{P}_n$. This inner product g_A makes $\mathcal{P}_n$ a complete Riemannian manifold with non-positive curvature and yields the following distance between two matrices $A, B \in \mathcal{P}_n$:

$$\delta(A, B) = \left(\sum_{k=1}^{n} \log^2 \lambda_k\right)^{1/2}, \tag{1.2}$$

where $\lambda_1, \ldots, \lambda_n$, are the eigenvalues of $A^{-1}B$, which are positive numbers (for all the proofs see [8, Ch. 6]). The Karcher mean of a set of m positive matrices, $A_1, \ldots, A_m \in \mathcal{P}_n$, is defined as the unique positive minimizer $G(A_1, \ldots, A_m)$ of the function

$$f(X; A_1, \ldots, A_m) := \sum_{j=1}^{m} \delta^2(X, A_j). \tag{1.3}$$

Since this mean minimizes the sum of squared intrinsic distances to each of the matrices A_j it is a barycenter of these matrices with respect to the aforementioned metric.

An important feature of the Karcher mean is that it possesses all the properties desired by a geometric mean, like the ten Ando-Li-Mathias (ALM) axioms [2]. For this reason, it is a viable tool in applications requiring some of these properties [7, 34]. A geometric mean should for instance be: permutation invariant, monotone, joint concave, and should satisfy the arithmetic-geometric-harmonic inequality (see [2] for the precise statements of the properties). In particular, one of the most characteristic properties of a geometric mean is its invariance under inversion:

$$G\left(A_1^{-1}, \ldots, A_m^{-1}\right) = G(A_1, \ldots, A_m)^{-1}. \tag{1.4}$$

Prior to having the proofs of all the properties of the Karcher mean, some of which are very elusive [10, 31], other definitions of a matrix geometric mean had been proposed [2, 12, 14, 35], even if nowadays there is large agreement in considering the Karcher mean as the "right" matrix geometric mean.

In certain applications, however, besides the positive definiteness, the data matrices have some further structure in the sense that they belong to some special subset $\mathcal{S}$, say a linear space. For instance, in the design and analysis of certain radar systems, the matrices to be averaged are correlation matrices, which are positive Toeplitz matrices or are positive Toeplitz block matrices with Toeplitz blocks [5, 6, 29, 41]. In these cases, one would like the geometric mean to belong to the same class $\mathcal{S}$ as the data. Unfortunately, the Karcher mean does not preserve many structures, in particular the Karcher mean of Toeplitz and/or band matrices is typically not of Toeplitz and/or band form anymore, as illustrated by the following simple example.

Example 1.1 Let $\mathcal{T}$ be the set of tridiagonal Toeplitz matrices and choose $A_1, A_2 \in \mathcal{T}$ where $A_1 = I$, and $A_2 = \mathrm{tridiag}(1, 2, 1)$ is the matrix with 2's on the main, and 1's appearing on sub- and superdiagonals. We have $A_1 A_2 = A_2 A_1$, thus the Karcher mean equals $(A_1 A_2)^{1/2}$. For $n = 3$ we get

$$(A_1 A_2)^{1/2} = \frac{\sqrt{2}}{4} \begin{bmatrix} \sqrt{2+\sqrt{2}}+2 & \sqrt{2}\sqrt{2-\sqrt{2}} & \sqrt{2+\sqrt{2}}-2 \\ \sqrt{2}\sqrt{2-\sqrt{2}} & \sqrt{2+\sqrt{2}} & \sqrt{2}\sqrt{2-\sqrt{2}} \\ \sqrt{2+\sqrt{2}}-2 & \sqrt{2}\sqrt{2-\sqrt{2}} & \sqrt{2+\sqrt{2}}+2 \end{bmatrix} \tag{1.5}$$

which is neither tridiagonal nor Toeplitz.

In this paper we introduce the concept of a structured geometric mean of positive matrices in such a way that if $A_1, \ldots, A_m \in \mathcal{S}$ also their mean belongs to $\mathcal{S}$. Given a subset $\mathcal{S}$ of $\mathcal{P}_n$ and matrices $A_1, \ldots, A_m \in \mathcal{S}$, we say that $G \in \mathcal{S}$ is a *structured geometric mean* with respect to $\mathcal{S}$ of $A_1, \ldots, A_n$ if the function $f(X; A_1, \ldots, A_m)$ of (1.3) takes its minimum value over $\mathcal{S}$ at G. The set of all structured geometric means of $A_1, \ldots, A_m$ (all minimizers of f) with respect to $\mathcal{S}$ is denoted by $G_{\mathcal{S}} = G_{\mathcal{S}}(A_1, \ldots, A_m)$.

We show that if $\mathcal{S}$ is closed (and nonempty) then $G_{\mathcal{S}}$ is nonempty and the matrices $G \in G_{\mathcal{S}}$ satisfy most of the ALM axioms in a suitably adjusted form. For instance, the invariance under inversion property (1.4) turns into

$$G_{\mathcal{S}}(A_1, \ldots, A_m) = G_{\mathcal{S}^{-1}}\big(A_1^{-1}, \ldots, A_m^{-1}\big)^{-1},$$

where for a set $\mathcal{U} \subseteq \mathcal{P}_n$ we denote $\mathcal{U}^{-1} = \{X^{-1} : X \in \mathcal{U}\}$. That is, the inverse of any structured geometric mean of the matrices $A_1, \ldots, A_m \in \mathcal{S}$ with respect to $\mathcal{S}$ coincides with a structured mean of the inverses $A_1^{-1}, \ldots, A_m^{-1}$ with respect to the set $\mathcal{S}^{-1}$ where these inverses reside.

Moreover, we show that, in many interesting cases, structured geometric means can be characterized in terms of the positive solutions of a suitable vector equation and provide algorithms for their computation.

In the Toeplitz case we also consider a different approach, where the mean is defined as a barycenter for a suitable metric on the manifold [4]. We analyze this barycenter and its properties in detail, obtaining an explicit expression in the real case and a quick algorithm in the complex case.

The article is organized as follows. In Sect. 2 the cost function (1.3) is examined with special focus on the existence of the minimizer over a closed set. The structured matrix mean itself is the subject of study in Sect. 3, where the theoretical properties it should satisfy are examined. Section 4 proposes two algorithms for computing a structured mean G in a linear space together with their convergence analysis. For one algorithm, it is shown that the convergence speed is independent of the condition number of the mean and is faster when the condition numbers of the matrices $A_i^{-1/2} G A_i^{-1/2}$ are smaller, for $i = 1, \ldots, n$. Because of its nature and its convergence properties, this algorithm can be viewed as the natural extension to the structured case of the Richardson-like algorithm introduced and analyzed in [11] for the computation of the Karcher mean of unstructured matrices. In Sect. 5, for Toeplitz matrices, a different structured matrix mean [4] as a barycenter is considered, and an algorithm for computing it is developed. Section 6 shows numerical experiments related to accuracy and speed for computing the structured matrix mean.

Here we recall some basic notation and properties that will be used in the rest of the paper. Given a matrix A, we define $\sigma(A)$ the spectrum of A, that is, the set of all the eigenvalues of A, and $\rho(A) = \max_{\lambda \in \sigma(A)} |\lambda|$ the spectral radius of A. Moreover we denote by $\|A\|_F := (\text{trace}(A^*A))^{1/2} = (\sum_{i,j} |a_{ij}|^2)^{1/2}$ the Euclidean (Frobenius) norm of A, and $\|A\|_s = \rho(A^*A)^{1/2}$ is the spectral norm. By A^* we denote the transposed conjugate of A. Recall that for a positive matrix A there exists a unique positive solution to the equation $X^2 = A$. This solution, denoted by $A^{1/2}$, is called the square root of A [8]. Given a matrix $A \in \mathbb{C}^{n \times n}$, we use the vec-operator to build $\text{vec}(A) \in \mathbb{C}^{n^2}$, a long vector obtained by stacking the columns of A. We will use the Kronecker product $\otimes$ such that $A \otimes B$ is the block matrix whose (i, j)th block is defined as $a_{ij} B$. The vec operator and the Kronecker product interplay in the following way [23]

$$\text{vec}(ABC) = \left(C^T \otimes A\right) \text{vec}(B). \tag{1.6}$$

Finally, we recall a natural partial order in $\mathcal{P}_n$ that will be used in the following: let A and B be positive, we write $A \geq B$ if the matrix $A - B$ is semidefinite positive.

2 Existence of structured geometric means

In this section the existence of a structured geometric mean and its relation to the classical Karcher mean is studied. First some necessities are repeated.

2.1 Uniqueness of the Karcher mean for positive matrices

The Riemannian geometry on $\mathcal{P}_n$ given by the inner product (1.1) turns out to be complete and a parametrization of the geodesic joining two positive matrices A and

B is known to be [8, 30]

$$A\#_t B = A^{1/2}\left(A^{-1/2} B A^{-1/2}\right)^t A^{1/2} = A\left(A^{-1} B\right)^t, \quad t \in [0, 1], \qquad (2.1)$$

where the midpoint $A\#_{1/2}B$ coincides with the geometric mean of the two matrices [24, 30].

Given a set of matrices $A_1, \ldots, A_m \in \mathcal{P}_n$, the function $f(X) = f(X; A_1, \ldots, A_m)$ in (1.3) is strictly geodesically convex, which means that for any two different matrices $X, Y \in \mathcal{P}_n$, we have

$$f(X\#_t Y) < (1 - t)f(X) + tf(Y), \quad 0 < t < 1. \qquad (2.2)$$

This property follows from [8, Exercise 6.1.13], where it is stated that for $m = 1$ the function $f(X)$ is strictly geodesically convex. The case $m > 1$ follows by summing up the m inequalities obtained by applying (2.2) to the functions $f(X) = f(X; A_i)$, for $i = 1, \ldots, m$, respectively.

Geodesical convexity is a key ingredient for the proof of the existence of a unique minimizer of f over $\mathcal{P}_n$ given in [8, Ch. 6]. A different proof is obtained using the fact that $\mathcal{P}_n$, with the inner product (1.1), forms a Cartan–Hadamard manifold [16, 28, 30, 32], which is a Riemannian manifold, complete, simply connected and with non-positive sectional curvature everywhere. On such a Cartan–Hadamard manifold the Karcher mean (the so-called center-of-mass) exists and is unique [17, 27, 28].

The notion of geodesical convexity in $\mathcal{P}_n$ is different from the customary convexity in the Euclidean space where one requires that

$$f\left((1 - t)X + tY\right) \leq (1 - t)f(X) + tf(Y), \quad t \in [0, 1].$$

In fact, the function f is not convex in the traditional sense as the following example shows.

Example 2.1 Consider the set made of the unique matrix $A = 1$, and $x, y \in \mathbb{R}_+^* = \mathcal{P}_1$. We have $f(x) = \delta^2(x, A) = \log^2(x)$ which is not convex. On the other hand the function $\log^2(x)$ is strictly geodesically convex and this can be shown by an elementary argument: in fact, it is continuous and for $x \neq y$

$$\delta^2(\sqrt{xy}, 1) = \log^2(\sqrt{xy}) = \frac{1}{4}\left(\log^2 x + \log^2 y + 2\log x \log y\right)$$

$$= \frac{1}{2}\left(\log^2 x + \log^2 y\right) - \frac{1}{4}(\log x - \log y)^2$$

$$< \frac{1}{2}\left(\log^2 x + \log^2 y\right) = \frac{1}{2}\left(\delta^2(x, 1) + \delta^2(y, 1)\right).$$

Iterative selection of midpoints and a continuity argument completes the proof.

Since f is strictly geodesically convex, it can be proved that it has a unique minimizer over any closed, geodesically convex subset $\mathcal{S}$ of $\mathcal{P}_n$, where we say that a subset $\mathcal{S} \subseteq \mathcal{P}_n$ is geodesically convex if for any $X, Y \in \mathcal{S}$, the entire geodesic $X\#_t Y$,

$t \in [0, 1]$ belongs to $\mathcal{S}$. Indeed, if X_1 and X_2 were two different matrices in $\mathcal{S}$ where f takes its minimum, then from (2.2) it would follow that $f(X_1 \#_t X_2) < f(X_1) = f(X_2)$ for any $0 < t < 1$ which contradicts the assumption.

2.2 Existence of structured geometric means on a closed set

For a generic closed subset $\mathcal{U}$ of $\mathcal{P}_n$, which is not necessarily geodesically convex, we can prove the existence of a minimum point by using the fact that $f(X)$ is continuous.

Theorem 2.1 *Let $\mathcal{U} \subseteq \mathcal{P}_n$ be a closed subset. Then for any $A_1, \ldots, A_m \in \mathcal{P}_n$ the function $f(X) = f(X; A_1, \ldots, A_m)$ has a minimum in $\mathcal{U}$.*

Proof Consider the distance measures $\delta(X, Y)$, given in (1.2), which can be written also as $\delta(X, Y) = \| \log(Y^{-1/2} X Y^{-1/2}) \|_F$, and $d(X, Y)$ (the Thompson metric [39]), given by

$$d(X, Y) = \left\| \log\left(Y^{-1/2} X Y^{-1/2}\right) \right\|_s.$$

From the inequality between the Frobenius and the spectral norm, we have

$$d(X, Y) \leq \delta(X, Y) \leq \sqrt{n} d(X, Y)$$

for all $X, Y \in \mathcal{P}_n$. We also define the order interval $[X, Y]$ as

$$[X, Y] = \{Z \in \mathcal{P}_n : X \leq Z \leq Y\},$$

for all $X, Y \in \mathcal{P}_n$ with $X \leq Y$ [21]. The order interval $[e^{-r} X, e^r X]$ turns out to be the closed ball of radius $r > 0$ around X with respect to the Thompson metric, thus the order interval $[e^{-r} I, e^r I]$ is a compact subset of $\mathcal{P}_n$.

Next, consider the set $Q_t = \{X \in \mathcal{P}_n : f(X) \leq t\}$. Because of the continuity of f, each Q_t is closed in $\mathcal{P}_n$. The definition of f also guarantees the boundedness of Q_t for each t with respect to the Riemannian metric δ. Since we have the above inequality of the distance measures, this implies that each Q_t is also bounded with respect to the Thompson metric. Hence it is possible to find some r such that Q_t is contained in the closed ball of radius r around I with respect to the Thompson metric, that is $Q_t \subset [e^{-r} I, e^r I]$. Since Q_t is now a closed subset of a compact set (in $\mathcal{P}_n$), it is also compact in $\mathcal{P}_n$.

Finally, define $B_t = Q_t \cap \mathcal{U}$, which is a again compact subset in $\mathcal{P}_n$ (as the intersection of a compact and a closed set). The collection of all nonempty B_t is a descending family of nonempty, compact subsets of $\mathcal{P}_n$, which has a nonempty intersection. The elements in this intersection are the minimizers of f in $\mathcal{U}$. $\qquad\square$

In general, uniqueness of the point where $f(X)$ takes its minimum cannot be guaranteed. For instance, if both A and A^{-1} belong to $\mathcal{U}$ while $I = A \#_{1/2} A^{-1}$ does not, then the function $f_1(X) := \delta^2(X, A) + \delta^2(X, A^{-1})$ reaches its minimum at a point $I \neq G \in \mathcal{U}$. Clearly, $f_1(G^{-1}) = f_1(G)$ and if $G^{-1} \neq G$ belongs to $\mathcal{U}$, then we have at least two distinct points of minimum. A more concrete example is the following.
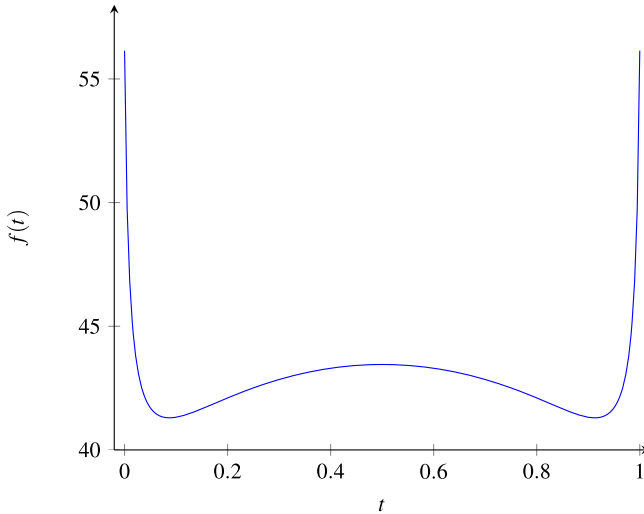
Fig. 1 Graph of $f(t) = \delta^2(G(t), A) + \delta^2(G(t), B)$ for $G(t) = A + t(B - A)$ with $A = I$ and $B = \text{diag}(200, 1/200)$

Example 2.2 Consider the 2×2 matrices $A = I$ and $B = \begin{bmatrix} a & 0 \\ 0 & a^{-1} \end{bmatrix}$, where $a > 1$. Define the segment $\mathcal{U} = \{G(t) = A + t(B - A), t \in [0, 1]\}$, which is closed and convex, but not geodesically convex. The function $f(t) = \delta^2(G(t), A) + \delta^2(G(t), B)$ takes the form $f(t) = \log^2((1 - t)/a + t) + \log^2(a(1 - t) + t) + \log^2((1 - t) + t/a) + \log^2((1 - t) + at)$ and is symmetric with respect to $t = 1/2$. For $a = 200$ the function has the graph shown in Fig. 1 with a local maximum at $t = 1/2$ and two global minima close to the edges of the segment.

3 A theoretical exploration of the structured geometric mean

In this section we discuss the relation between the structured and generic geometric mean, together with the adaptation of the generic properties to the structured setting. We will discuss just the real case, so in this section, the set $\mathcal{P}_n$ stands for the manifold of real positive definite matrices whose tangent space is the set of real and symmetric matrices.

3.1 The geometric and structured geometric mean relation

The properties shown in Sect. 2 imply that a structured geometric mean with respect to $\mathcal{U}$, as defined in the introduction, always exists for any closed subset $\mathcal{U}$ of $\mathcal{P}_n$. In particular, this holds in the cases where $\mathcal{U} = \mathcal{S} \cap \mathcal{P}_n$ for any linear space $\mathcal{S}$ of matrices and also for $\mathcal{U}^{-1} := \mathcal{S}^{-1} \cap \mathcal{P}_n$ where we define $\mathcal{S}^{-1} = \{A^{-1} : A \in \mathcal{S}, \det A \neq 0\}$. This captures a wide class of interesting structures emerging in applications, e.g., Toeplitz and band matrices, as well as their inverses. For simplicity we will restrict our analysis in the remainder of the article to the real case.

More general structures are given in terms of a parametrization $\sigma(t) : \mathcal{V} \to \mathbb{R}^{n \times n}$, with σ a differentiable function defined in the open subset $\mathcal{V}$ of $\mathbb{R}^q$, which we will call the *parameter space*. The set $\mathcal{T} = \sigma(\mathcal{V})$ is the structure determined by σ. If σ is linear and $\mathcal{V} = \mathbb{R}^q$, then $\mathcal{T}$ is a linear space. Examples of sets $\mathcal{T}$ of interest which generally do not form a linear space are the set of matrices with a given displacement rank [13], the set of semiseparable [40], and quasiseparable matrices [19]. For an $n \times n$ symmetric Toeplitz matrix, a possible parametrization is given by

$$\sigma(t) = \sigma([t_0, t_1, \ldots, t_{n-1}]) = \begin{bmatrix} t_0 & t_1 & \ldots & t_{n-1} \\ t_1 & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & t_1 \\ t_{n-1} & \ldots & t_1 & t_0 \end{bmatrix}. \tag{3.1}$$

For a band matrix, one can, e.g., just store the nonzero-elements in a long vector and map them onto their exact locations. In the following, given a closed set $\mathcal{T}$ we let $\mathcal{U} = \mathcal{T} \cap \mathcal{P}_n$.

In Example 2.2 we illustrated that the minimum of the cost function restricted to a closed subset $\mathcal{U} \subseteq \mathcal{P}_n$ is not necessarily unique. For this reason, we consider the structured geometric mean $G_{\mathcal{U}} = G_{\mathcal{U}}(A_1, \ldots, A_m)$ of $A_1, \ldots, A_m \in \mathcal{U}$ as the set of matrices in $\mathcal{U}$ where the function $f(X)$ attains its minimum. Formally speaking, for $A_1, \ldots, A_m \in \mathcal{U}$, let $g \in \mathbb{R}^q$ be such that $\widehat{G} = \sigma(g) \in G_{\mathcal{U}}(A_1, \ldots, A_m)$, then

$$f(\sigma(g); A_1, \ldots, A_m) = \min_{t \in \mathbb{R}^q} f(\sigma(t); A_1, \ldots, A_m).$$

Since $\mathcal{U} \subseteq \mathcal{P}_n$, the minimum over $\mathcal{P}_n$ is less than or equal to the minimum over $\mathcal{U}$. In general it will often happen that $\widehat{G} \neq G(A_1, \ldots, A_m)$ like in (1.5).

3.2 Properties of the geometric mean conveyed to the structured mean setting

Some desired properties for a matrix geometric mean were stated by Ando, Li and Mathias in [2], of which the most noticeable are enlisted here. The other properties include joint concavity, the determinental identity, and continuity from above.

Consistency with scalars If $A_1, \ldots, A_m$ commute, then

$$G(A_1, \ldots, A_m) = (A_1 \cdots A_m)^{1/m}.$$

Permutation invariance For any permutation π of $\{1, \ldots, m\}$,

$$G(A_1, \ldots, A_m) = G(A_{\pi(1)}, \ldots, A_{\pi(m)}).$$

Joint homogeneity

$$G(\alpha_1 A_1, \alpha_2 A_2, \ldots, \alpha_m A_m) = (\alpha_1 \cdots \alpha_m)^{1/m} G(A_1, \ldots, A_m).$$

Monotonicity If $A_i \geq A_i'$, for $i = 1, \ldots, m$, then

$$G(A_1, \ldots, A_m) \geq G(A_1', \ldots, A_m').$$

Invariance under congruence For any nonsingular M,

$$G\big(M^*A_1M, \ldots, M^*A_mM\big) = M^*G(A_1, \ldots, A_m)M.$$

Invariance under inversion

$$G(A_1, \ldots, A_m)^{-1} = G\big(A_1^{-1}, \ldots, A_m^{-1}\big).$$

Arithmetic-geometric-harmonic mean inequality

$$\frac{1}{m}(A_1 + \cdots + A_m) \geq G(A_1, \ldots, A_m) \geq m\big(A_1^{-1} + \cdots + A_m^{-1}\big)^{-1}.$$

Yet another property naturally desired of a geometric mean, but not required in the list of Ando, Li and Mathias, is the *repetition invariance*, that is, for any set of positive matrices $A_1, \ldots, A_m \in \mathcal{P}_n$,

$$G(A_1, \ldots, A_m, A_1, \ldots, A_m) = G(A_1, \ldots, A_m). \tag{3.2}$$

Now, we consider the properties of the structured geometric mean. Some properties such as the permutation invariance trivially hold, others should be restated. In fact, in the generic case the structures we consider are neither invariant under inversion nor under congruence. That is because if $A \in \mathcal{U}$ then it need not necessarily hold that $A^{-1} \in \mathcal{U}$ or $M^*AM \in \mathcal{U}$.

We start with the invariance under inversion as this is one of the most characteristic properties of the geometric mean. To this end we consider the set $\mathcal{T}^{-1} = \{T^{-1} : T \in \mathcal{T}, \det T \neq 0\}$ parametrized with the function $\sigma_{-1}(t) := \sigma(t)^{-1}$. Clearly, the intersection $\mathcal{U}$ of $\mathcal{T}$ with $\mathcal{P}_n$ yields always invertible matrices, so that $\mathcal{T}^{-1} \cap \mathcal{P}_n = \mathcal{U}^{-1}$.

According to our definition, the structured geometric mean of $A_1^{-1}, \ldots, A_m^{-1} \in \mathcal{U}^{-1}$ is given by the set $G_{\mathcal{U}^{-1}}(A_1^{-1}, \ldots, A_m^{-1})$. For any $\widetilde{G} \in G_{\mathcal{U}^{-1}}$, we have $\widetilde{G} = \sigma(\widetilde{g})^{-1}$ such that

$$f\big(\sigma(\widetilde{g})^{-1}; A_1^{-1}, \ldots, A_m^{-1}\big) = \min_{t \in \mathbb{R}^q} f\big(\sigma(t)^{-1}; A_1^{-1}, \ldots, A_m^{-1}\big).$$

Since $\delta(A, B) = \delta(A^{-1}, B^{-1})$, one gets $f(X; A_1, \ldots, A_m) = f(X^{-1}; A_1^{-1}, \ldots, A_m^{-1})$ so that

$$f\big(\sigma(\widetilde{g}); A_1, \ldots, A_m\big) = \min_{t \in \mathbb{R}^q} f\big(\sigma(t); A_1, \ldots, A_m\big)$$

and thus $\widetilde{G}^{-1} \in G_{\mathcal{U}}(A_1, \ldots, A_m)$. Since $\widetilde{G}$ was chosen arbitrarily, and since $\mathcal{U}$ can be interchanged with $\mathcal{U}^{-1}$, we have the analogue of the invariance under inversion for the structured geometric mean:

$$G_{\mathcal{U}}(A_1, \ldots, A_m)^{-1} = G_{\mathcal{U}^{-1}}\big(A_1^{-1}, \ldots, A_m^{-1}\big). \tag{3.3}$$

In a similar manner we can restate the invariance under congruence in a structured style by defining, for any nonsingular M, the set $\mathcal{U}_M := M^*\mathcal{U}M = \{M^*TM : T \in \mathcal{U}\}$.

The invariance under congruence is then understood as

$$G_{\mathcal{U}_M}(M^*A_1M, \ldots, M^*A_mM) = M^*G_{\mathcal{U}}(A_1, \ldots, A_m)M.$$

Joint homogeneity, in order to be defined, requires that the set $\mathcal{T}$ satisfies the following property:

$$A \in \mathcal{T} \quad \Rightarrow \quad \alpha A \in \mathcal{T}$$

for any scalar $\alpha > 0$. This property clearly holds if $\mathcal{T}$ is a linear space or the set formed by the inverses of the nonsingular matrices of a linear space. For these sets, the joint homogeneity holds.

Repetition invariance holds true as well by (1.3), since

$$f(X; A_1, \ldots, A_m, A_1, \ldots, A_m) = 2f(X; A_1, \ldots, A_m),$$

so the minimizers (over a subset) of the functions $f(X; A_1, \ldots, A_m, A_1, \ldots, A_m)$ and $f(X; A_1, \ldots, A_m)$ are the same.

Regarding the remaining properties, we observe that the consistency with scalars is violated, as Example 1.1 shows. Nevertheless, weaker consistency properties hold, such as idempotency, namely $G_{\mathcal{U}}(A, A, \ldots, A) = A$ for each structure $\mathcal{U}$ and $A \in \mathcal{U}$. Moreover, if the set $\mathcal{U}$ is closed and geodesically convex then

$$G_{\mathcal{U}}(A_1, \ldots, A_m) = G(A_1, \ldots, A_m),$$

so the geometric and structured geometric mean coincide. An interesting case of geodesically convex set is given by $\mathcal{U} = \mathcal{T} \cap \mathcal{P}_n$, when $\mathcal{T}$ is an algebra, i.e., a linear space closed under multiplication and inversion.

Finally, the properties related to the ordering of positive matrices such as monotonicity are not true as shown by the following numerical example.

Example 3.1 We consider the four Toeplitz matrices

$$T_1 = \begin{bmatrix} 1 & 1/2 & 1/2 \\ 1/2 & 1 & 1/2 \\ 1/2 & 1/2 & 1 \end{bmatrix}, \qquad T_2 = T_1,$$

$$T_3 = \begin{bmatrix} 3/4 & 1/2 & 0 \\ 1/2 & 3/4 & 1/2 \\ 0 & 1/2 & 3/4 \end{bmatrix}, \qquad S = \begin{bmatrix} 1 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 1 \end{bmatrix},$$

and, using the algorithms presented in the next sections, we compute a structured geometric mean G_ε of the three matrices T_1, T_2 and $T_3 + \varepsilon S$ for various $\varepsilon \geq 0$. The norm of $G_\varepsilon - G_0$ becomes small as ε tends to 0 and we observe that $G_\varepsilon - G_0$ is not positive (semi)definite, while $T_3 + \varepsilon S \geq T_3$. This gives numerical evidence of the lack of monotonicity of a structured geometric mean. On the other hand, computing the arithmetic mean A of T_1, T_2 and T_3, one observes also that the expected inequality $A \geq G_0$ does not hold in this case.

3.3 The structured mean as solution(s) of a vector equation

We start from the Karcher mean, which is obtained as the unique solution in $\mathcal{P}_n$ of the matrix equation

$$\sum_{i=1}^{m} \log\left(XA_i^{-1}\right) = 0. \tag{3.4}$$

Equation (3.4) is obtained using the fact that f is differentiable and has a minimum at the Karcher mean. Thus the Karcher mean satisfies the condition $\nabla f_X = 0$, where $\nabla f_X = 2X^{-1} \sum_{i=1}^{m} \log(XA_i^{-1})$ denotes the (Euclidean) gradient of f with respect to X (see [26, 33]). We remark already that a different metric will be studied in Sect. 4.3.

In the general case, the restriction of f to a structure given by $\sigma(t)$ is investigated. For any minimum g (with corresponding $\sigma(g)$) not located at the boundary of the parameter space, the gradient $\nabla(f \circ \sigma)_t$ of the function with respect to t must be zero, so we are interested in the solutions of the vector equation $\nabla(f \circ \sigma)_t = 0$.

From the chain rule of derivation, one obtains that

$$\nabla(f \circ \sigma)_t = \left(\sum_{i,j} \frac{\partial f(\sigma(t))}{\partial x_{i,j}} \frac{d\sigma_{i,j}(t)}{dt_s}\right)_{s=1,\dots,q} = 0$$

which leads to the vector equation

$$\sum_{i,j} \left(\Gamma\left(\sigma(t)\right)\right)_{i,j} \frac{d\sigma_{i,j}(t)}{dt_s} = 0, \quad s = 1, \dots, q, \tag{3.5}$$

where $\Gamma(X) := \frac{1}{2}\nabla f_X$.

In the case where $\mathcal{T}$ is a linear space, the function $\sigma(t)$ is linear and can be written in matrix form as

$$\text{vec}\left(\sigma(t)\right) = Ut, \quad U \in \mathbb{R}^{n^2 \times q},$$

so that (3.5) turns into

$$U^T \text{vec}\left(\Gamma\left(\sigma(t)\right)\right) = 0, \qquad \Gamma(X) = X^{-1} \sum_{i=1}^{m} \log\left(XA_i^{-1}\right). \tag{3.6}$$

If $\sigma(t)$ is chosen to be orthogonal, i.e. such that $U^T U = I$, then U^T coincides with the Moore–Penrose inverse of U.

When $\mathcal{T}$ denotes the set of symmetric Toeplitz matrices, the parametrization (3.1) leads to a matrix U having orthogonal columns. In fact one has $U^T U = D = \text{diag}(n, 2(n-1), 2(n-2), \dots, 2)$. In particular, for $n = 3$ one has

$$U^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \end{bmatrix}.$$

For $\mathcal{T}$ being the set of symmetric tridiagonal matrices the parametrization

$$\sigma(t) = \begin{bmatrix} t_1 & t_{n+1} & & & & \\ t_{n+1} & t_2 & t_{n+2} & & & \\ & \ddots & \ddots & \ddots & & \\ & & t_{2n-2} & t_{n-1} & t_{2n-1} & \\ & & & t_{2n-1} & t_n \end{bmatrix}$$

also leads to a matrix U having orthogonal columns. Moreover, $U^T U = \operatorname{diag}(I_n, 2I_{n-1})$. For $n = 3$, e.g., one has

$$U^T = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 \end{bmatrix}.$$

4 Algorithms for structured geometric means in the linear case

We will give two algorithms for computing structured geometric means when they are characterized in terms of the solutions g of a vector equation, as, for instance, in the linear case.

We first provide a general definition of a class of algorithms based on a preconditioned functional iteration, then we specialize to two algorithms given by two different preconditioners.

The first, provided in Sect. 4.2 is derived by relying on the projection of the gradient with respect to the Euclidean scalar product. The second, presented in Sect. 4.3, is obtained through projection with respect to the Riemannian metric of $\mathcal{P}_n$ decribed in Sect. 1.

4.1 A preconditioned functional iteration and its convergence

Throughout this section we assume that $A_1, \ldots, A_m \in \mathcal{U}$, where $\mathcal{U} = \mathcal{T} \cap \mathcal{P}_n$ and $\mathcal{T}$ is a linear space with a parametrization $\sigma(t)$ such that $\operatorname{vec}(\sigma(t)) = Ut$, and $D = U^T U$.

The structured geometric mean $G_{\mathcal{U}}$ is defined as the set of minimizers of the function $f(X; A_1, \ldots, A_m)$ over $\mathcal{U}$. These minimizers must be sought among the stationary points of the function f, that is, among the solutions to the vector equation (3.6).

Therefore, a way to design algorithms for computing structured means $G_{\mathcal{U}}$ is to apply numerical techniques to solve the vector equation (3.6). We consider a preconditioned Richardson-like iteration constructed in the spirit of [11]. Let $V(X)$ be a nonsingular and sufficiently differentiable matrix function and define

$$\varphi(t) = t - \theta S(t), \qquad S(t) = V\big(\sigma(t)\big)^{-1} U^T \operatorname{vec}\big(\Gamma\big(\sigma(t)\big)\big),$$
$$t_{\nu+1} = \varphi(t_\nu), \quad \nu = 0, 1, \ldots,$$

(4.1)

where θ is a parameter introduced to enhance convergence, $V(\sigma(t))$ is a preconditioner and t_0 is a given vector such that $\sigma(t_0)$ is positive. Observe that the fixed points of $\varphi(t)$ are the solutions of the vector equation (3.6) and if convergent, the sequence t_ν converges to a solution of the vector equation (3.6).

In the following, given a matrix function $f(X)$, where $X = (x_{i,j})$ and $f(X)$ are $n \times n$ matrices, we denote by $J_f(G)$ the $n^2 \times n^2$ Jacobian matrix of $\mathrm{vec}(f(X))$ with respect to the variable $\mathrm{vec}(X)$ computed at $X = G$, similarly we denote $J_{f \circ \sigma}(t_G)$ the $n^2 \times q$ Jacobian of the composed function $\mathrm{vec}(f(\sigma(t)))$ with respect to the variables $(t_1, \ldots, t_q)$ at $t = t_G$. In this notation, the function in the subscript as well as the variable between parentheses specify if the derivatives are taken w.r.t. the matrix variable X or the vector variable t.

Observe that if $V(\sigma(t))$ is chosen as the Jacobian of $U^T \mathrm{vec}(\Gamma(\sigma(t)))$, then (4.1) coincides with Newton's iteration.

If t_G is a solution of (3.6) and if t_ν is sufficiently near to t_G, then

$$t_{\nu+1} - t_G = J_\varphi(t_G)(t_\nu - t_G) + O\big(\|t_\nu - t_G\|^2\big),$$

so that in order to study the local convergence of this sequence it is sufficient to estimate the spectral radius ρ or any induced norm of $J_\varphi(t_G)$ and determine θ in such a way that $\rho(J_\varphi(t_G)) < 1$. Notice that the Jacobian of $\varphi(t)$ at $t = t_G$ is given by $I - \theta K$ where $K = J_S(t_G)$ is the Jacobian of $S(t)$ at $t = t_G$. Therefore, if we can find a preconditioner $V(t)$ such that K has real positive eigenvalues with minimum and maximum eigenvalues $\kappa_{\min}$ and $\kappa_{\max}$ respectively, then the choice $\theta = 2/(\kappa_{\min} + \kappa_{\max})$ insures local convergence and provides the minimum spectral radius of $J_\varphi(t_G)$ given by

$$\rho\big(J_\varphi(t_G)\big) = \frac{\kappa_{\max} - \kappa_{\min}}{\kappa_{\max} + \kappa_{\min}} = \frac{\mu - 1}{\mu + 1} < 1, \quad \mu = \kappa_{\max}/\kappa_{\min}.$$

Moreover, any values $\hat{\kappa}_{\min} \leq \hat{\kappa}_{\max}$ such that $\hat{\kappa}_{\min} \leq \kappa_{\min} \leq \kappa_{\max} \leq \hat{\kappa}_{\max}$ can be used instead of $\kappa_{\min}$ and $\kappa_{\max}$ to determine a value $\hat{\theta} = 2/(\hat{\kappa}_{\min} + \hat{\kappa}_{\max})$ which insures convergence. Also notice that the closer μ is to 1 the faster is the convergence of the iteration.

Therefore our goal is to perform a spectral analysis of K and to find an upper bound to the ratio $\mu = \kappa_{\max}/\kappa_{\min}$, assuming that all the eigenvalues of K are real positive. From the composition rule of derivatives one finds that

$$K = V\big(\sigma(t_G)\big)^{-1} U^T J_\Gamma(G) U + J_{V(\sigma(t_G))^{-1}}\big(\sigma(t_G)\big) U^T \mathrm{vec}\big(\Gamma\big(\sigma(t_G)\big)\big)$$

and since $U^T \mathrm{vec}(\Gamma(\sigma(t_G))) = 0$, it follows that

$$K = V\big(\sigma(t_G)\big)^{-1} U^T J_\Gamma(G) U. \tag{4.2}$$

To evaluate $J_\Gamma(G)$, we recall that $\Gamma(X) = \sum_{i=1}^m X^{-1} \log(X A_i^{-1})$, so that it is sufficient to determine the formal expression of $J_\psi(G)$ for $\psi(G, A) = G^{-1} \log(G A^{-1})$ for a generic A and then to write $J_\Gamma(G) = \sum_{i=1}^m J_{\psi(G, A_i)}(G)$. In order to evaluate $J_\psi(G)$, we rely on the definition of Fréchet derivative of a matrix function $f(X)$

at X in the direction E

$$Df_X[E] = \lim_{t \to 0} \frac{f(X + tE) - f(X)}{t} = \frac{d}{dt}\bigg|_{t=0} f(X + tE).$$

In fact, the $n^2 \times n^2$ Jacobian matrix $J_f(X)$ of the vector function $\text{vec} \circ f \circ \text{vec}^{-1}$ at $\text{vec}(X)$ is related to the Fréchet derivative by the equation

$$\text{vec}(Df_X[E]) = J_f(X)\,\text{vec}(E). \tag{4.3}$$

We recall also the following properties of the Fréchet derivative [22] where f, g are given matrix functions and $\varphi(X) = X^{-1}$:

$$D(fg)_X[E] = Df_X[E]g(X) + f(X)Dg_X[E], \quad \text{product rule,}$$

$$D(f \circ g)_X[E] = Df_{g(X)}[Dg_X[E]], \quad \text{chain rule,} \tag{4.4}$$

$$D\varphi_X[E] = -X^{-1}EX^{-1}, \quad \text{inversion.}$$

For the derivative of the exponential function we have (see [22, Eq. 10.17a])

$$J_{\exp}(Y) = (I \otimes \exp Y)\beta(Y^T \otimes I - I \otimes Y), \quad \beta(z) = (e^z - 1)/z.$$

Therefore, since $J_{\log}(X) = J_{\exp}(Y)^{-1}$ for $Y = \log X$, we find that

$$J_{\log}(X) = \gamma(\log(X^T) \otimes I - I \otimes \log X)(I \otimes X^{-1}), \quad \gamma(z) = z/(e^z - 1). \tag{4.5}$$

We are now ready to provide an explicit expression of the Fréchet derivative of the function $\psi(X, A) = X^{-1}\log(XA^{-1})$ and of the Jacobian $J_{\psi(X,A)}(X)$.

Lemma 4.1 *Let $\psi(X) = X^{-1}\log(XA^{-1})$. Assume that A, X are positive. For the matrix $J_\psi(X)$ such that $\text{vec}(D\psi_X[E]) = J_\psi(X)\,\text{vec}(E)$ we have*

$$J_\psi(X) = -X^{-1}\log(XA^{-1}) \otimes X^{-1} + (A^{-1} \otimes X^{-1})\gamma(W)(I \otimes AX^{-1}),$$

$$W = \log(XA^{-1}) \otimes I - I \otimes \log(XA^{-1}),$$

with $\gamma(z) = z/(e^z - 1)$.

Proof Since $h(X) := \log(XA^{-1})$ is the composition of $f(X) = \log(X)$ and $g(X) = XA^{-1}$, we get by (4.4)

$$Dh_X[E] = D\log_{XA^{-1}}[EA^{-1}].$$

As $\psi(X)$ is the product of $f(X) = X^{-1}$ and $h(X)$, (4.4) gives us

$$D\psi_X[E] = -X^{-1}EX^{-1}\log(XA^{-1}) + X^{-1}Dh_X[E].$$

Combining the latter two equations yields

$$D\psi_X[E] = -X^{-1}EX^{-1}\log(XA^{-1}) + X^{-1}D\log_{XA^{-1}}[EA^{-1}].$$

By using (4.3) and (1.6) we find that the matrix $J_\psi(X)$ representing $D\psi_X$ is given by

$$J_\psi(X) = -\left(X^{-1}\log\left(XA^{-1}\right)\right)^T \otimes X^{-1} + \left(I \otimes X^{-1}\right)J_{\log}\left(XA^{-1}\right)\left(A^{-T} \otimes I\right).$$

Replacing (4.5) in the equation above and using the fact that $A = A^T$, $X = X^T$ yields

$$J_\psi(X) = -\log\left(A^{-1}X\right)X^{-1} \otimes X^{-1}$$
$$+ \left(I \otimes X^{-1}\right)\gamma\left(\log\left(A^{-1}X\right) \otimes I - I \otimes \log\left(XA^{-1}\right)\right)\left(A^{-1} \otimes AX^{-1}\right).$$

Using the fact that $W\log(V)W^{-1} = \log(WVW^{-1})$, the first term can be written as $-X^{-1}\log(XA^{-1}) \otimes X^{-1}$. The second term can be written as $(I \otimes X^{-1})(A^{-1} \otimes I)\gamma\left(\log(XA^{-1}) \otimes I - I \otimes \log(XA^{-1})\right)(I \otimes AX^{-1})$, which completes the proof. □

Recall that $\Gamma(X) = \sum_{i=1}^m \psi(X, A_i)$ and $G^{-1}\sum_{i=1}^m \log(GA_i^{-1}) = 0$, for $G = \sigma(t_G)$. Then by Lemma 4.1, we obtain the following formula for the Jacobian $J_\Gamma(\sigma(t))$:

$$J_\Gamma(G) = \left(I \otimes G^{-1}\right)H\left(I \otimes G^{-1}\right), \qquad H = \sum_{i=1}^m H_i,$$

$$H_i = \left(A_i^{-1} \otimes I\right)\gamma\left(\log\left(GA_i^{-1}\right) \otimes I - I \otimes \log\left(GA_i^{-1}\right)\right)(I \otimes A_i).$$

Moreover, by using the properties of the Kronecker product and the following relation $\log(GA^{-1}) = A^{1/2}\log(A^{-1/2}GA^{-1/2})A^{-1/2}$, we can write

$$H_i = \left(A_i^{-1/2} \otimes A_i^{1/2}\right)\gamma\left(\log M_i \otimes I - I \otimes \log M_i\right)\left(A_i^{-1/2} \otimes A_i^{1/2}\right),$$

$$M_i = A_i^{-1/2}GA_i^{-1/2}.$$

From this expression it turns out that H_i is positive, and from (4.2) we find that $J_S(t_G)$ is the product of the matrices $V(\sigma(t_G))^{-1}$ and $U^T(I \otimes G^{-1})\sum_{i=1}^m H_i(I \otimes G^{-1})U$, which is a positive matrix.

Thus we may conclude with the following

Theorem 4.1 *The Jacobian K of the function $S(t)$ in (4.2) at $\sigma(t_G) = G$ is given by*

$$K = V^{-1}U^T\left(I \otimes G^{-1}\right)H\left(I \otimes G^{-1}\right)U,$$

$$H = \sum_{i=1}^m H_i, \quad H_i = \left(A_i^{-1/2} \otimes A_i^{1/2}\right)\gamma\left(\log M_i \otimes I - I \otimes \log M_i\right)\left(A_i^{-1/2} \otimes A_i^{1/2}\right),$$

$$M_i = A_i^{-1/2}GA_i^{-1/2},$$

$$\gamma(z) = z/\left(e^z - 1\right).$$

Moreover, the eigenvalues of K are the solutions of the equation

$$\det\left(\kappa V - U^T\left(I \otimes G^{-1}\right)H\left(I \otimes G^{-1}\right)U\right) = 0.$$

4.2 An elementary preconditioner

The simplest choice for the preconditioner $V(t)$ in (4.1) is $V(t) = U^T U = D$. This corresponds to projecting the gradient of the function $f(X, A_1, \ldots, A_p)$ on the set $\mathcal{U}$ according to the Euclidean scalar product. The problem $\det(\kappa I - K) = 0$ turns into the generalized q-dimensional symmetric eigenvalue problem

$$\det\big(U^T\big(\kappa I - \big(I \otimes G^{-1}\big)H\big(I \otimes G^{-1}\big)\big)U\big) = 0.$$

This problem is the projection on the space spanned by the columns of U of the problem $\det(\nu I - (I \otimes G^{-1})H(I \otimes G^{-1})) = 0$ which has real positive solutions.

Now we recall the following result, valid for general positive matrices A, B, which relates the generalized eigenvalues of the pair (A, B) to the ones of the projected pair $(U^T A U, U^T B U)$.

Lemma 4.2 *Let A, B be positive $n \times n$ matrices and U an $n \times m$ matrix. Then the generalized eigenvalues of the pair $(U^T A U, U^T B U)$, which solve the equation $\det(U^T(A - \kappa B)U) = 0$, are real positive and lie in between the maximum and minimum eigenvalues λ of the pair (A, B), such that $\det(A - \lambda B) = 0$. Moreover, the extreme eigenvalues $\lambda_{\min}, \lambda_{\max}$ of the pair (A, B) are such that $\alpha_{\min}/\beta_{\max} \leq \lambda_{\min} \leq \lambda_{\max} \leq \alpha_{\max}/\beta_{\min}$, where $\alpha_{\min}, \alpha_{\max}, \beta_{\min}, \beta_{\max}$ are the minimum and maximum eigenvalues of the matrices A and B, respectively.*

Proof The condition $\det(\lambda B - A) = 0$ is equivalent to $\det(\lambda I - B^{-1/2}AB^{-1/2}) = 0$, which has real positive solutions since $B^{-1/2}AB^{-1/2}$ is positive. The remaining part of the lemma follows from the fact that maximum and minimum eigenvalues of the larger and smaller problems coincide with maximum and minimum value of the Rayleigh quotient $x^T Ax/x^T Bx$ for $x \in \mathbb{R}^n \setminus \{0\}$, and for $x \in \text{span}(U)$, respectively. $\qquad\square$

A first consequence of the above lemma is that the extreme eigenvalues $\kappa_{\min}$ and $\kappa_{\max}$ of K are in between the maximum and the minimum eigenvalue of the n^2-dimensional symmetric matrix $Y = (I \otimes G^{-1})H(I \otimes G^{-1})$, so that the ratio μ between the maximum and minimum eigenvalue of K is less than or equal to the condition number $\mu(Y)$ of the symmetric matrix Y. Moreover since $Y = \sum_{i=1}^m Y_i$ with

$$Y_i = \big(A_i^{-1/2} \otimes A_i^{-1/2}\big)\big(I \otimes M_i^{-1}\big)\gamma\,(\log M_i \otimes I - I \otimes \log M_i)$$
$$\times \big(I \otimes M_i^{-1}\big)\big(A_i^{-1/2} \otimes A_i^{-1/2}\big),$$

one finds that $\widehat{k}_{\min} := \sum_{i=1}^m \lambda_{\min}^{(i)} \leq \kappa_{\min}$ and $\widehat{k}_{\max} := \sum_{i=1}^m \lambda_{\max}^{(i)} \geq \kappa_{\max}$, where $\lambda_{\min}^{(i)}$ and $\lambda_{\max}^{(i)}$ are the minimum and the maximum eigenvalues of Y_i. Moreover, from Lemma 4.2 and from the expression above for Y_i it follows that $\lambda_{\min}^{(i)} \geq \gamma_{\min}^{(i)}/(\alpha_{\max}^{(i)})^2$, $\lambda_{\max}^{(i)} \leq \gamma_{\max}^{(i)}/(\alpha_{\min}^{(i)})^2$, where $\alpha_{\min}^{(i)}, \alpha_{\max}^{(i)}$ are the minimum and the maximum eigenvalues of A_i, respectively, while $\gamma_{\min}^{(i)}$ and $\gamma_{\max}^{(i)}$ are the minimum and maximum eigenvalues of $(I \otimes M_i^{-1})\gamma\,(\log M_i \otimes I - I \otimes \log M_i)(I \otimes M_i^{-1})$, respectively.

From the properties of the matrix function $\gamma(\cdot)$ and from the properties of the Kronecker product one finds that the eigenvalues of the latter matrix can be explicitly given in terms of the eigenvalues $v_r^{(i)}$ of the matrix M_i. In fact, they coincide with $\frac{1}{(v_s^{(i)})^2}(\log t_{r,s}^{(i)})/(t_{r,s}^{(i)} - 1)$ where $t_{r,s}^{(i)} = \frac{v_r^{(i)}}{v_s^{(i)}}$.

Since the function $(\log t)/(t - 1)$ is monotonically decreasing, its minimum and maximum are

$$\eta_{\min}^{(i)} = \left(\log \mu^{(i)}\right)/\left(\mu^{(i)} - 1\right),$$
$$\eta_{\max}^{(i)} = \log\left(1/\mu^{(i)}\right)/\left(1/\mu^{(i)} - 1\right) = \mu^{(i)}\left(\log \mu^{(i)}\right)/\left(\mu^{(i)} - 1\right),$$

for $\mu^{(i)} = \mu(M_i)$ the spectral condition number of M_i. Additionally, taking the factor $(v^{(i)})^{-2}$ into consideration gives

$$\gamma_{\min}^{(i)} \geq \eta_{\min}^{(i)}\left(v_{\max}^{(i)}\right)^{-2},$$
$$\gamma_{\max}^{(i)} \leq \eta_{\max}^{(i)}\left(v_{\min}^{(i)}\right)^{-2} \leq \mu^{(i)}\left(v_{\min}^{(i)}\right)^{-2},$$

where $v_{\min}^{(i)}$ and $v_{\max}^{(i)}$ represent respectively the minimum and maximimum eigenvalue of M_i.

Therefore, we may conclude that the eigenvalues of K are bounded by $\widetilde{\kappa}_{\min} := \sum_{i=1}^{m} \eta_{\min}^{(i)}/(v_{\max}^{(i)}\alpha_{\max}^{(i)})^2$ and $\widetilde{\kappa}_{\max} := \sum_{i=1}^{m} \eta_{\max}^{(i)}/(v_{\min}^{(i)}\alpha_{\min}^{(i)})^2$.

Observe that this bound gets worse when either some matrix is ill-conditioned or if some matrix $A_i^{-1/2}GA_i^{-1/2}$ is ill-conditioned. The latter case cannot occur if the matrices A_i do not differ much from G. The dependence of this bound on the conditioning of A_i makes this algorithm very inefficient as long as some A_i is ill-conditioned. This drawback is overcome in the next section, where we design a more effective preconditioner.

4.3 A preconditioner based on a differential geometric viewpoint

The Karcher mean for positive matrices inherits a beautiful interpretation in terms of differential geometry. It can be considered as the center of mass for a well chosen inner product on the manifold of positive matrices. In this section and in Sect. 5 we consider two approaches inspired by this idea. For more information we refer to the overview in [25], and the articles [15, 20, 36–38].

When considering a manifold optimization approach, the intersection $\mathcal{U}$ of a linear space $\mathcal{T}$ with the manifold of positive matrices $\mathcal{P}_n$ can be viewed as a Riemannian submanifold of $\mathcal{P}_n$ itself, which in turn is called the *enveloping space*. This entails that the inner product from this enveloping space is induced on the submanifold. An immediate consequence is that the gradient of the cost function for the submanifold is given by the orthogonal projection (with respect to the inner product) of the gradient for the enveloping space. Similar to the space of symmetric matrices, being the tangent space to the manifold of positive matrices, the intersection $\mathcal{V}$ of the linear space $\mathcal{T}$ with the space of symmetric matrices is the tangent space to $\mathcal{U}$.

First consider the manifold of positive matrices endowed with the Euclidean inner product $g_X(A, B) = \text{tr}(AB)$, with A and B symmetric, and X a positive matrix. Note

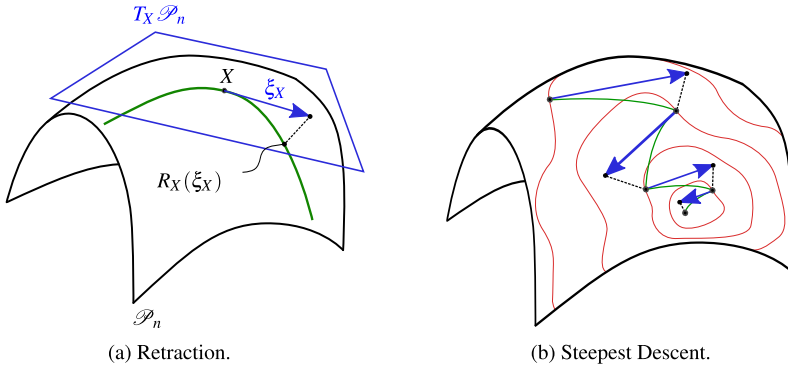(a) Retraction.   (b) Steepest Descent.

Fig. 2 Graphical representation of a retraction and steepest descent flow

that even though this inner product g_X is independent of X, the subscript notation is kept for consistency. In this case, the orthogonal projection of a symmetric matrix A onto $\mathcal{T}$ gives a matrix T, with

$$\text{vec}(T) = U\left(U^T U\right)^{-1} U^T \text{vec}(A),$$

or $\text{vec}(T) = Ut$, with

$$t = \left(U^T U\right)^{-1} U^T \text{vec}(A). \tag{4.6}$$

The expression for the gradient of the Karcher cost function, corresponding to the Euclidean inner product, is known for the manifold of positive matrices and is given by

$$\text{grad}_e f(X; A_1, \ldots, A_m) = 2X^{-1} \sum_{i=1}^{m} \log\left(X A_i^{-1}\right)$$

$$= 2X^{-1/2} \sum_{i=1}^{m} \log\left(X^{1/2} A_i^{-1} X^{1/2}\right) X^{-1/2}. \tag{4.7}$$

The gradient naturally defines the direction of steepest ascent. Nevertheless, the gradient lies in the tangent space, and to build an algorithm from this, a practical way is to follow the gradient and then go back to the manifold through a suitable function, called *retraction*. The precise definition of a retraction, together with general theoretical assumptions it should satisfy, can be found in [1]. Figure 2(a) graphically illustrates the concept of a retraction, where a vector ξ_X in the tangent space $T_X \mathcal{P}_n$ of the positive matrices is retracted to a point $R_X(\xi_X)$ residing on the manifold $\mathcal{P}_n$. On a manifold, the classical steepest descent algorithm is graphically depicted in Fig. 2(b). The thin red lines depict the contour lines, the blue arrows the gradients, and the green curves the retractions to the manifold.

Observe that for $\mathcal{P}_n$ immersed in the set of symmetric matrices, the tangent space at a point is the whole set of symmetric matrices. So one can consider the basic retraction $R_X(A) = X + A$ for a sufficiently small symmetric matrix A.

Entering now the gradient (4.7) in projection (4.6) and applying a gradient descent method with the basic retraction $R_X(A) = X + A$, we arrive exactly at the Richardson-like algorithm for finding the fixed points of function (4.1).

However, since the function f to be minimized is defined through the distance (1.2), it is more natural to consider the manifold of positive matrices endowed with the inner product $g_X(A, B) = \text{tr}(AX^{-1}BX^{-1})$, with A, B and X as before. In this case, the gradient for the enveloping space is known to be

$$\text{grad}_n f(X; A_1, \ldots, A_m) = 2X \sum_{i=1}^{m} \log(A_i^{-1}X)$$

$$= 2X^{1/2} \sum_{i=1}^{m} \log(X^{1/2}A_i^{-1}X^{1/2})X^{1/2}.$$

Note the difference with (4.7).

The orthogonal projection T onto the intersection $\mathcal{V}$ (of $\mathcal{T}$ and the space of symmetric matrices) of this gradient, with respect to the Riemannian scalar product, can be found as the solution of the equations

$$\text{grad}_n f(X) = T + S,$$

$$g_X(S, K) = \text{tr}(SX^{-1}KX^{-1}) = 0, \quad \text{for every } K \in \mathcal{V}.$$

Writing again $\text{vec}(T) = Ut$, we find in parameter space

$$t = \left(U^T(X^{-1} \otimes X^{-1})U\right)^{-1} U^T(X^{-1} \otimes X^{-1}) \text{vec}(\text{grad}_n f(X)). \qquad (4.8)$$

The factor $U^T(X^{-1} \otimes X^{-1})U$ is recurring and is abbreviated as D_X, where the subscript points to the intrinsic variable X. Observe that this Riemannian orthogonal projection can be seen as a Euclidean oblique projection where the two bases of the subspace are the columns of U and $(X^{-1} \otimes X^{-1})U$, respectively.

Using this expression, it is possible to define another gradient descent method where we are now searching the fixed points of the function

$$\varphi(t) = t - \theta D_{\sigma(t)}^{-1} U^T\left(\sigma(t)^{-1} \otimes \sigma(t)^{-1}\right) \text{vec}\left(\sigma(t) \sum_{i=1}^{m} \log(A_i^{-1}\sigma(t))\right). \qquad (4.9)$$

Relying on (1.6) to incorporate the Kronecker product into the vectorization, we find that $(\sigma^{-1} \otimes \sigma^{-1}) \text{vec}(\sigma \sum_{i=1}^{m} \log(A_i^{-1}\sigma)) = \text{vec}(\sum_{i=1}^{m} \log(A_i^{-1}\sigma)\sigma^{-1})$. Applying a property of the matrix logarithm we may rewrite the latter expression as $\text{vec}(\sigma^{-1} \sum_{i=1}^{m} \log(\sigma A_i^{-1}))$. This way, (4.9) takes the form of (4.1) with

$$V(\sigma(t)) = U^T(\sigma(t)^{-1} \otimes \sigma(t)^{-1})U$$

To analyze the convergence of (4.1) with this choice for $V(\sigma(t))$, we have to analyze the eigenvalues of the Jacobian $K = J_S(t_G)$ of $S(t)$ in (4.1) where the equation

$\det(\kappa I - K) = 0$ takes the form of the following generalized eigenvalue problem

$$\det\bigl(U^T\bigl(\kappa\bigl(G^{-1}\otimes G^{-1}\bigr) - \bigl(I\otimes G^{-1}\bigr)H\bigl(I\otimes G^{-1}\bigr)\bigr)U\bigr) = 0. \qquad (4.10)$$

Since the two matrices in (4.10) are positive, in view of Lemma 4.2, the solutions of this generalized eigenvalue problem are real positive and are located in between the maximum and the minimum solution of the larger problem

$$\det\bigl(\lambda\bigl(G^{-1}\otimes G^{-1}\bigr) - \bigl(I\otimes G^{-1}\bigr)H\bigl(I\otimes G^{-1}\bigr)\bigr) = 0,$$

which in turn can be rewritten as a standard eigenvalue problem

$$\det\bigl(\lambda I - \bigl(G^{1/2}\otimes G^{1/2}\bigr)\bigl(I\otimes G^{-1}\bigr)H\bigl(I\otimes G^{-1}\bigr)\bigl(G^{1/2}\otimes G^{1/2}\bigr)\bigr) = 0.$$

Since $H = \sum_{i=1}^{m} H_i$, and the matrices H_i are real symmetric, the eigenvalues of this problem are located in between the sum of the minimum and the sum of the maximum eigenvalues of each subproblem

$$\det\bigl(\lambda I - \bigl(G^{1/2}\otimes G^{1/2}\bigr)\bigl(I\otimes G^{-1}\bigr)H_i\bigl(I\otimes G^{-1}\bigr)\bigl(G^{1/2}\otimes G^{1/2}\bigr)\bigr) = 0, \qquad (4.11)$$

that is $\det(\lambda(G^{-1}\otimes G) - H_i) = 0$, or equivalently $\det(\lambda I - (G\otimes I)\times H_i(I\otimes G^{-1})) = 0$. The matrix in the latter expression is similar to $(A_i^{-1/2}\otimes A_i^{-1/2})\times (G\otimes I)H_i(I\otimes G^{-1})(A_i^{1/2}\otimes A_i^{1/2})$, which, using the expression of H_i provided in Theorem 4.1, can be written as

$$(M_i\otimes I)\gamma(\log M_i\otimes I - I\otimes\log M_i)\bigl(I\otimes M_i^{-1}\bigr).$$

This way, the eigenvalues of (4.11) can be explicitly given in terms of the eigenvalues $v_r^{(i)}$ of the matrix M_i. In fact, they coincide with the $t_{r,s}^{(i)}(\log t_{r,s}^{(i)})/(t_{r,s}^{(i)} - 1)$ where $t_{r,s}^{(i)} = \frac{v_r^{(i)}}{v_s^{(i)}}$.

Since the function $t(\log t)/(t-1)$ is monotone, for the minimum and maximum solution to (4.11) we have

$$\eta_{\min}^{(i)} = \bigl(1/\mu^{(i)}\bigr)\log\bigl(1/\mu^{(i)}\bigr)/\bigl(1/\mu^{(i)} - 1\bigr) = \bigl(\log\mu^{(i)}\bigr)/\bigl(\mu^{(i)} - 1\bigr),$$
$$\eta_{\max}^{(i)} = \mu^{(i)}\bigl(\log\mu^{(i)}\bigr)/\bigl(\mu^{(i)} - 1\bigr),$$

respectively, for $\mu^{(i)} = \mu(M_i)$ the spectral condition number of M_i. Therefore, we may conclude that the eigenvalues of K are in between $\sum_{i=1}^{m}\eta_{\min}^{(i)}$ and $\sum_{i=1}^{m}\eta_{\max}^{(i)}$. This way, we find for the optimal value of θ and for the optimal spectral radius the estimates

$$\theta = \frac{2}{\sum_{i=1}^{m}\frac{\mu^{(i)}+1}{\mu^{(i)}-1}\log\mu^{(i)}},$$

$$\rho = \frac{\sum_{i=1}^{m}\log\mu^{(i)}}{\sum_{i=1}^{m}\frac{\mu^{(i)}+1}{\mu^{(i)}-1}\log\mu^{(i)}}.$$

It is interesting to point out that in this case the convergence speed is related neither to the condition number of the geometric mean G nor to those of the matrices A_i but is related only to the relative distances of G from each A_i measured by the quantities $\mu^{(i)} = \mu(M_i)$, $M_i = A_i^{-1/2} G A_i^{-1/2}$. The closer they are to 1 the faster is the convergence. Therefore, if the matrices to average are not too far from each other, so that the quantities $\mu(M_i)$ are close to 1, then the optimal value of θ is close to $1/m$ and a very fast convergence is expected. This analysis is confirmed by the numerical experiments.

4.4 The case of Toeplitz matrices

From the computational point of view, at each step of the iteration (4.1) one has to compute $U^T \text{vec}(\Gamma(\sigma(t)))$ and then solve a linear system with the matrix $V(\sigma(t))$. The former computation, based on (3.6), requires $O(mn^3)$ arithmetic operations (ops), while the cost of the latter depends on the structure of $V(\sigma(t))$.

In this section we examine the case where $\mathcal{U}$ is the class of symmetric Toeplitz matrices and where $\sigma(t)$ associates t with the Toeplitz matrix having as first column t. We describe a way to make the algorithm of Sect. 4.3 more efficient.

Indeed, for the iteration analyzed in Sect. 4.2, V is the diagonal matrix with diagonal entries $(n, 2n-2, \ldots, 2)$ and the cost of solving a system with matrix V amounts to n divisions.

The iteration examined in Sect. 4.3 has a higher convergence speed but at each step an $n \times n$ system with $V = U^T (X^{-1} \otimes X^{-1}) U$ must be solved, where X is a symmetric positive definite Toeplitz matrix.

We split the computation in two steps. In the first, the n^2 entries of V are computed, in the second step a standard $O(n^3)$ ops linear system solver is used. Concerning the first step we discuss two approaches.

In both approaches the inverse of the Toeplitz matrix X needs to be computed, which can be done efficiently using the Gohberg Semencul formula [13]. Here, vectors v_1, v_2, v_3, v_4 are determined such that $X^{-1} = L(v_1)L(v_2)^T - L(v_3)L(v_4)^T$, where $L(v)$ is the lower triangular Toeplitz matrix whose first column is v. From these, the n^2 entries of X^{-1} can be found. The overall cost is $O(n^2)$ ops.

1. As a first attempt, the entries of V are computed in a straightforward manner using the entries of X^{-1}:

$$V = \begin{bmatrix} \gamma_{1,1} & 2\gamma_{1,2} & \cdots & 2\gamma_{1,n} \\ 2\gamma_{1,2} & 2\gamma_{2,2} & \cdots & 2\gamma_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ 2\gamma_{1,n} & 2\gamma_{2,n} & \cdots & 2\gamma_{n,n} \end{bmatrix},$$

   where

$$\gamma_{1,j} = \sum_{i=1}^{n} \sum_{k=1}^{n-j+1} (X^{-1})_{i,k} (X^{-1})_{i,k+j-1},$$

$$\gamma_{j,p} = \sum_{i=1}^{n-j+1} \sum_{k=1}^{n-p+1} \left((X^{-1})_{i,k} (X^{-1})_{i+j-1,k+p-1} + (X^{-1})_{i,k+p-1} (X^{-1})_{i+j-1,k} \right).$$

The cost of this approach in terms of arithmetic operations is of the order $O(n^4)$.

2. In the second approach, we show that the cost of this computation can be kept at the level of $O(n^3 \log n)$ ops by combining the Gohberg Semencul formula and the FFT. For a given i, the product vector $w_i = (X^{-1} \otimes X^{-1}) U e_i$, where e_i is the ith vector of the canonical basis, is such that $w_i = \text{vec}(X^{-1} E_i X^{-1})$, with E_i being the symmetric Toeplitz matrix whose first column is e_i. Therefore, compute first the columns of $E_i X^{-1}$ by performing $O(n^2)$ additions, and then multiply X^{-1} by these columns, stacking the results to obtain w_i. This computation is performed in $O(n^2 \log n)$ operations for each i by using the Goghberg Semencul formula, since the multiplication of a lower triangular Toeplitz matrix and a vector can be performed in $O(n \log n)$ operations by means of the FFT [13]. Therefore the overall computation of this stage for $i = 1, \ldots, n$ is $O(n^3 \log n)$ ops. Finally, compute for any i the vector $U^T w_i$ for the cost of $O(n^2)$ additions.

The performance of these methods will be compared in Sect. 6.

5 Kähler metric mean for Toeplitz matrices

The Karcher mean of positive definite matrices has the specific interpretation of being the barycenter of the given matrices for the natural metric (1.2) on this manifold. Hence there are in a certain sense two possible generalizations. On the one hand, try to generalize the geometric mean concept, or, on the other hand, try to generalize the barycenter concept. Previously we focused on an extension of the geometric mean. Hereafter we focus on the positive definite Toeplitz matrix manifold itself, denoted by $\mathcal{T}_n$, and consider a barycenter in this case. This mean cannot be called a geometric mean in the sense of satisfying all required properties, but through its intuitive definition, many desirable properties could arise.

The concept of a barycenter is not restricted to the specific metric used to define the Karcher mean. For example, when the set $\mathcal{T}_n$ is endowed with the classical Euclidean inner product, the resulting barycenter is nothing else than the arithmetic mean. Using a probabilty argument, in [3, 4] a metric on $\mathcal{T}_n$ is introduced, called the Kähler metric. This metric results in a complete, simply connected manifold with non-positive sectional curvature everywhere, or a Cartan–Hadamard manifold. Thus, by [17, 27], existence and uniqueness are guaranteed for the barycenter with respect to this metric.

We will recall some known facts about the Kähler metric, and then we will give an explicit formula for the barycenter in the real case and a numerical procedure to compute the barycenter in the complex case.

To construct the Kähler metric, a Toeplitz matrix is first transformed to an n-tuple $(P_0, \mu_1, \ldots, \mu_{n-1})$ in $\mathbb{R}_+^* \times \mathbb{D}^{n-1}$, with $\mathbb{R}_+^*$ the set of strictly positive real numbers and $\mathbb{D}$ the set of complex numbers of modulus less than one. This transformation,

denoted as $\zeta(T) = [P_T, \mu_{T,1}, \ldots, \mu_{T,n-1}]^T$, is performed as follows:

$$P_T = t_0, \qquad \mu_{T,j} = (-1)^j \frac{\det(S_j)}{\det(R_j)},$$

with t_0 the main diagonal element of T, R_j the principal submatrix of size j of T (the upper left $j \times j$ submatrix) and S_j obtained by shifting R_j down one row, or equivalently, by removing the first row and last column of R_{j+1} (the inverse transformation can be found in [42]). In what follows, we use this one-to-one relation between the Toeplitz matrices and the corresponding n-tuple, and when clear by the context, we will neglect the distinction and identify one with the other.

For X and Y being the transformations of two positive Toeplitz matrices $X = [P_X, \mu_{X,1}, \ldots, \mu_{X,n-1}]^T$ and $Y = [P_Y, \mu_{Y,1}, \ldots, \mu_{Y,n-1}]^T$, the metric is given by

$$d(X, Y) = \left(n\sigma(P_X, P_Y)^2 + \sum_{j=1}^{n-1} (n-j)\tau(\mu_{X,j}, \mu_{Y,j})^2 \right)^{1/2},$$

$$\sigma(P_X, P_Y) = \left| \log\left(\frac{P_Y}{P_X}\right) \right|, \qquad \tau(\mu_{X,j}, \mu_{Y,j}) = \operatorname{atanh}\left(\left| \frac{\mu_{Y,j} - \mu_{X,j}}{1 - \mu_{X,j}\mu_{Y,j}} \right| \right),$$

where $\operatorname{atanh}(z) = \frac{1}{2}\log(\frac{1+z}{1-z})$.

The barycenter of the positive Toeplitz matrices T_i, for $i = 1, \ldots, m$, with respect to the Kähler metric will be denoted by $B(T_1, \ldots, T_m) = [P_B, \mu_{B,1}, \ldots, \mu_{B,n-1}]^T$. It is obtained in this transformed space by minimizing the function

$$f(X) = \sum_{i=1}^{m} d^2(X, T_i)$$

over $\mathbb{R}_+^* \times \mathbb{D}^{n-1}$. Notice that the problem of minimizing $f(X)$ can be decoupled into the problems of minimizing $\varphi_0(x) = \sum_{i=1}^{m} \sigma(x, P_{T_i})^2$ over $\mathbb{R}_+^*$, and the $n-1$ scalar functions

$$\varphi_j(z) = \sum_{i=1}^{m} \tau(z, \mu_{T_i,j})^2, \quad j = 1, \ldots, n-1$$

over $\mathbb{D}$. The minimum of $\varphi_0(x)$ is easily obtained as $P_B = (P_{T_1} \cdots P_{T_m})^{1/m}$ by solving the equation $\nabla\varphi_0(x) = 0$. The minimum of $\varphi_j(z)$ is nothing else than the barycenter of $\mu_{T_1,j}, \ldots, \mu_{T_m,j}$ with respect to the customary Poincaré metric on the unit disk and is the point where the gradient

$$\nabla\varphi_j(z) = 2(|z|^2 - 1) \sum_{i=1}^{m} \operatorname{sign}(c_{i,j}) \operatorname{atanh}(|c_{i,j}|), \quad c_{i,j} = \frac{\mu_{T_i,j} - z}{1 - \bar{z}\mu_{T_i,j}}, \qquad (5.1)$$

equals zero.

In the real case we are able to find an explicit expression for this barycenter as well, since $\operatorname{sign}(c)\operatorname{atanh}(|c|) = \operatorname{atanh}(c)$ and after some manipulations we get

$$\mu_{X,j} = \mathcal{C}\left(\left(\mathcal{C}(\mu_{T_1,j}) \cdots \mathcal{C}(\mu_{T_m,j}) \right)^{1/m} \right),$$

where $\mathcal{C}(z) = (1 - z)/(1 + z)$ is the Cayley transform.

In the complex case we were not able to find such an explicit formula but a quick numerical method can be devised using a gradient descent algorithm. We recall that the tangent space to the Poincaré disk can be identified with the complex plane and thus for a sufficiently small tangent vector $v \in \mathbb{C}$, one can consider the retraction $R_z(v) = z + v$, which captures the fact that the manifold is an open subset of the complex plane. The resulting algorithm to find the barycenter of $\mu_1, \ldots, \mu_n \in \mathbb{C}$ is given by the iteration

$$z_{k+1} = z_k + t_k v_k,$$

$$v_k = \left(1 - |z_k|^2\right) \sum_{i=1}^{n} \text{sign}(c_{i,k}) \, \text{atanh}(|c_{i,k}|), \tag{5.2}$$

$$c_{i,k} = \frac{\mu_i - z_k}{1 - \overline{z}_k \mu_i},$$

for a suitable initial value z_0 and a sufficiently small steplength t_k.

Another possibility is to consider the retraction

$$R_z(v) = \frac{z + e^{i\theta} + (z - e^{i\theta})e^{-s}}{1 + \overline{z}e^{i\theta} + (1 - \overline{z}e^{i\theta})e^{-s}}, \quad \theta = \arg v, \ s = \frac{2|v|}{1 - |z|^2},$$

which corresponds to moving along the geodesics of the Poincaré disk. The corresponding gradient descent method is

$$z_{k+1} = R_{z_k}(t_k v_k),$$

with the same v_k as (5.2).

5.1 Properties of the Kähler barycenter

Regarding the properties of this barycenter, it is easily seen that it is permutation invariant, repetition invariant and idempotent (this holds for any barycenter). Moreover, for any $\alpha > 0$, the transformed values of αT are $[\alpha P_T, \mu_{T,1}, \ldots, \mu_{T,m}]^T$ and from the explicit expression of P_B in the real case we get $B(\alpha T_1, T_2, \ldots, T_m) = \alpha^{1/m} B(T_1, \ldots, T_m)$, that is, homogeneity holds.

Unfortunately, this new barycenter does not possess other properties as shown by the following example.

Example 5.1 From the explicit expression for the mean in the real case we get a simple formula for the Kähler barycenter of two 2×2 real matrices

$$T_1 = \begin{bmatrix} x_1 & y_1 \\ y_1 & x_1 \end{bmatrix}, \qquad T_2 = \begin{bmatrix} x_2 & y_2 \\ y_2 & x_2 \end{bmatrix},$$

namely

$$B(T_1, T_2) = \sqrt{x_1 x_2} \begin{bmatrix} 1 & \frac{a-b}{a+b} \\ \frac{a-b}{a+b} & 1 \end{bmatrix}, \quad \text{with} \begin{cases} a = \sqrt{(x_1 + y_1)(x_2 + y_2)} \\ b = \sqrt{(x_1 - y_1)(x_2 - y_2)} \end{cases}.$$

Now consider the following matrices

$$T_1 = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix}, \qquad \widetilde{T}_1 = \begin{bmatrix} 4 & -1 \\ -1 & 4 \end{bmatrix}, \qquad T_2 = \begin{bmatrix} 2 & -1 \\ -1 & 2 \end{bmatrix},$$

with $\widetilde{T}_1 \geq T_1$. By symbolic computation, one gets that

$$B(\widetilde{T}_1, T_2) = \begin{bmatrix} 2\sqrt{2} & \sqrt{2}(\sqrt{5}-3) \\ \sqrt{2}(\sqrt{5}-3) & 2\sqrt{2} \end{bmatrix} \not\geq B(T_1, T_2) = \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix},$$

in fact one eigenvalue of $B(\widetilde{T}_1, T_2) - B(T_1, T_2)$ is $\lambda = \sqrt{10} - 2 - \sqrt{2} < 0$. Thus, we have proved that the Kähler barycenter is not monotonic. Moreover,

$$B(T_1, T_2) \neq (T_1 T_2)^{1/2} = \begin{bmatrix} \sqrt{3} & 0 \\ 0 & \sqrt{3} \end{bmatrix},$$

and hence the Kähler barycenter does not coincide with the Karcher mean for circulant matrices. In particular, it is not a structured geometric mean as defined in Sect. 1.

Observe that in the previous example $B(T_1, T_2)$ surprisingly coincides with the arithmetic mean of T_1 and T_2. It is not difficult to construct examples where it is not true that $B(T_1, T_2) \leq (T_1 + T_2)/2$ as it should be for a geometric mean.

6 Numerical experiments

In this section, the different algorithms proposed in Sects. 4.2 and 4.3 will be compared w.r.t. speed and accuracy. The numerical experiments are confined to Toeplitz matrices, because of applicational interest in computing their structured matrix mean [29]. These matrices are constructed randomly, but with chosen condition number, using the technique described in [18]. Performance, accuracy and computational distance are subjects of the forthcoming investigations. For clarity we remind the reader that the Richardson-iteration corresponds to a projection technique on a manifold, with the classical Euclidean inner product. For all algorithms, the stopping criteria is based on checking the relative size of the gradient and on comparing two consecutive iteration points.

It is worth pointing out that, in spite of the lack of the proof of uniqueness for structured geometric mean in the Toeplitz case, for any fixed set of data matrices used in our experiments, any initial value and any algorithm yielded always the same structured geometric mean. This suggests the conjecture that in the Toeplitz case there is a unique structured geometric mean.

We have also compared the structured geometric mean obtained by our algorithms with the Kähler metric mean, getting in most experiments a relative difference of the order 10^{-1}, which indicates that these two means are relatively far from each other.

6.1 The projection methods

Performance The performance of the projection methods explained in Sect. 4 can be compared by looking at both the number of iterations the methods require and the total amount of computational time they need.
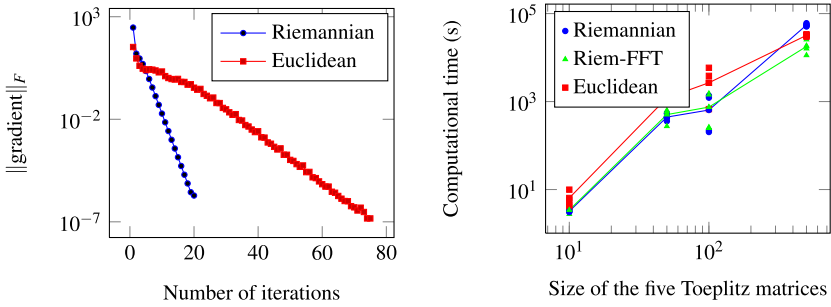
(a) Gradient evolution for five 50×50 matrices.

(b) Computational time for various matrix sizes. The mean of the samples is connected by a line.

Fig. 3 Comparison of the projection methods for Toeplitz matrices. In the legends, Euclidean indicates the method of Sect. 4.2, Riemannian indicates the first approach described in Sect. 4.4, and Riem-FFT the second

In Fig. 3(a), the evolution of the gradient over the iterations is displayed for both techniques (and hence also the number of iterations). Using the projection method introduced in Sect. 4.3 gives a faster decrease of the gradient and results in fewer iteration steps. The number of iterations remains almost constant for this method as the size of the matrices increases. For the projection technique from Sect. 4.2 on the other hand, this number starts to increase when the matrix size grows.

However, comparing expression (4.6) and (4.8), it can be seen that the second one is computationally more expensive and hence the advantage of requiring fewer iterations could be nullified. Therefore, Fig. 3(b) displays the total computational time of both methods for varying sizes of the matrices (both approaches from Sect. 4.4 are shown). The two methods based on Sect. 4.3 maintain an advantage despite their larger computational cost per iteration. Note that for the largest matrix size the computational time of the Euclidean based method appears less than one of the other methods. However, this is caused by the increasing number of iterations required by this Euclidean method. Consequently, the maximum number of iterations is reached before convergence and the algorithms is terminated prematurely. Concerning the operation count in Sect. 4.4, the advantage of the method based on FFT starts to appear when the matrices become sufficiently large.

Accuracy   In order to analyze the accuracy of the projection methods, we implement a high precision version of the first algorithm in Sect. 4.4 using the `vpa` functionality of MATLAB. The relative distance, based on the intrinsic distance (1.2), between this high precision computation and the result of the actual algorithms is shown in Fig. 4. For small condition numbers, the accuracy of all methods is similar in average, but as the condition of the matrices becomes worse, the accuracy of the projection method based on Euclidean geometry deteriorates much faster than that of the method based on the Riemannian geometry. This first method even fails to converge when the condition number of the matrices becomes significantly large. The accuracy of the two approaches in Sect. 4.4 is similar and deteriorates steadily as the condition numbers of the matrices increase.
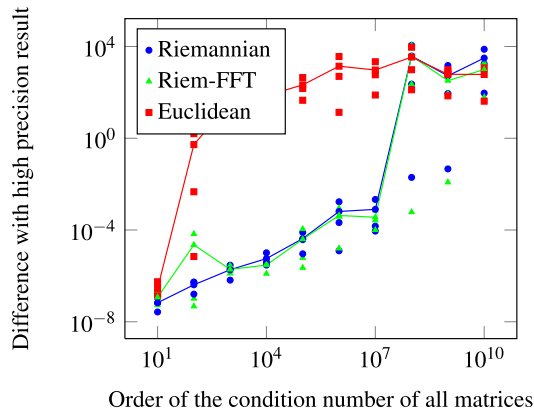
Fig. 4 Accuracy of the projection methods when compared to a high precision version. The mean of the samples is connected by a line. In the legends, Euclidean indicates the method of Sect. 4.2, Riemannian indicates the first approach described in Sect. 4.4, and Riem-FFT the second

7 Conclusions

In this article a generalization of the Karcher mean for positive definite matrices to structured positive definite matrices was proposed. Besides a theoretical investigation and adaptation of the desired properties of such a mean, algorithms were proposed. In the design of the algorithms, two trajectories were put forward, one relying mostly on linear algebra, and one based on differential geometry. A convergence analysis has been performed showing the superiority of the algorithm based on differential geometry. Numerical experiments compared the accuracy and speed of the various techniques and confirmed the theoretical analysis.

In the case of Toeplitz matrices, we have considered also the Kähler metric mean [4], whose properties have been investigated, providing an explicit expression in the real case and a quick algorithm in the complex case. For Toeplitz matrices, both the new structured geometric mean and the Kähler metric mean have not completely satisfying properties. In fact they are neither monotone, nor do they satisfy the arithmetic-geometric inequality. We wonder if it is possible to provide a definition of geometric mean for Toeplitz matrices which behaves well with respect to the ordering of positive matrices.

References

1. Absil, P.-A., Mahony, R., Sepulchre, R.: Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton (2008)
2. Ando, T., Li, C., Mathias, R.: Geometric means. Linear Algebra Appl. **385**, 305–334 (2004)

3. Barbaresco, F.: Information intrinsic geometric flows. In: Mohammad-Djafari, A. (ed.) Bayesian Inference and Maximum Entropy Methods in Science and Engineering. American Institute of Physics Conference Series, vol. 872, pp. 211–218 (2006)
4. Barbaresco, F.: Interactions between symmetric cone and information geometries: Bruhat-tits and Siegel spaces models for high resolution autoregressive Doppler imagery. In: Nielsen, F. (ed.) Emerging Trends in Visual Computing. Lecture Notes in Computer Science, vol. 5416, pp. 124–163. Springer, Berlin (2009)
5. Barbaresco, F.: Robust statistical radar processing in Fréchet metric space: OS-HDR-CFAR and OS-STAP processing in Siegel homogeneous bounded domains. In: Proceedings of IRS'11, International Radar Conference, Leipzig (2011)
6. Barbaresco, F.: Information geometry of covariance matrix: Cartan–Siegel homogeneous bounded domains, Mostow/Berger fibration and Fréchet median. In: Bhatia, R., Nielsen, F. (eds.) Matrix Information Geometry, pp. 199–256. Springer, Berlin (2012)
7. Batchelor, P.G., Moakher, M., Atkinson, D., Calamante, F., Connelly, A.: A rigorous framework for diffusion tensor calculus. Magn. Reson. Med. **53**, 221–225 (2005)
8. Bhatia, R.: Positive Definite Matrices. Princeton Series in Applied Mathematics. Princeton University Press, Princeton (2007)
9. Bhatia, R., Holbrook, J.: Riemannian geometry and matrix geometric means. Linear Algebra Appl. **413**(2–3), 594–618 (2006)
10. Bhatia, R., Karandikar, R.L.: Monotonicity of the matrix geometric mean. Math. Ann. **353**(4), 1453–1467 (2012)
11. Bini, D.A., Iannazzo, B.: Computing the Karcher mean of symmetric positive definite matrices. Linear Algebra Appl. **438**(4), 1700–1710 (2013)
12. Bini, D.A., Iannazzo, B.: A note on computing matrix geometric means. Adv. Comput. Math. **35**(2–4), 175–192 (2011)
13. Bini, D., Pan, V.: Polynomial and Matrix Computations. Birkhäuser, Basel (1994)
14. Bini, D., Meini, B., Poloni, F.: An effective matrix geometric mean satisfying the Ando–Li–Mathias properties. Math. Comput. **79**(269), 437–452 (2010)
15. Bonnabel, S., Sepulchre, R.: Geometric distance and mean for positive semi-definite matrices of fixed rank. SIAM J. Matrix Anal. Appl. **31**(3), 1055–1070 (2009)
16. Bridson, M., Haefliger, A.: Metric Spaces of Non-Positive Curvature. Springer, Berlin (1999)
17. Cartan, E.: Leçons sur la géométrie des espaces de Riemann. Usp. Mat. Nauk **3**(3), 218–222 (1948)
18. Chu, M.T., Golub, G.H.: Structured inverse eigenvalue problems. Acta Numer. **11**, 1–71 (2002)
19. Eidelman, Y., Gohberg, I.C.: On a new class of structured matrices. Integral Equ. Oper. Theory **34**, 293–324 (1999)
20. Fiori, S.: Learning the Fréchet Mean over the Manifold of Symmetric Positive-Definite Matrices. Springer, Media (2009)
21. Gaubert, S., Qu, Z.: The contraction rate in Thompson metric of order-preserving flows on a cone—application to generalized Riccati equations (2012). arXiv:1206.0448v1 [math.MG]
22. Higham, N.J.: Functions of Matrices: Theory and Computation. Society for Industrial and Applied Mathematics, Philadelphia (2008)
23. Horn, R.A., Johnson, C.R.: Topics in Matrix Analysis. Cambridge University Press, Cambridge (1991)
24. Iannazzo, B.: The geometric mean of two matrices from a computational viewpoint (2011). arXiv:1201.0101v1 [math.NA]
25. Jeuris, B., Vandebril, R., Vandereycken, B.: A survey and comparison of contemporary algorithms for computing the matrix geometric mean. Electron. Trans. Numer. Anal. **39**, 379–402 (2012)
26. Karcher, H.: Riemannian center of mass and mollifier smoothing. Commun. Pure Appl. Math. **30**(5), 509–541 (1977)
27. Kendall, W.: Probability, convexity, and harmonic maps with small image I: uniqueness and fine existence. Proc. Lond. Math. Soc. **61**(2), 371–406 (1990)
28. Lang, S.: Fundamentals of Differential Geometry. Graduate Texts in Mathematics. Springer, Berlin (1999)
29. Lapuyade-Lahorgue, J., Barbaresco, F.: Radar detection using Siegel distance between autoregressive processes, application to HF and X-band radar. In: IEEE Radar Conference, 2008. RADAR'08, Rome (2008)
30. Lawson, J., Lim, Y.: The geometric mean, matrices, metrics, and more. Am. Math. Mon. **108**(9), 797–812 (2001)

31. Lawson, J., Lim, Y.: Monotonic properties of the least squares mean. Math. Ann. **351**(2), 267–279 (2011)
32. Lim, Y.: Riemannian and Finsler structures of symmetric cones. Trends Math. **4**(2), 111–118 (2001)
33. Moakher, M.: A differential geometric approach to the geometric mean of symmetric positive-definite matrices. SIAM J. Matrix Anal. Appl. **26**(3), 735–747 (2005)
34. Moakher, M.: On the averaging of symmetric positive-definite tensors. J. Elast. **82**(3), 273–296 (2006)
35. Nakamura, K.: Geometric means of positive operators. Kyungpook Math. J. **49**, 167–181 (2009)
36. Pálfia, M.: The Riemann barycenter computation and means of several matrices. Int. J. Comput. Math. Sci. **3**(3), 128–133 (2009)
37. Qi, C.: Numerical optimization methods on Riemannian manifolds. Ph.D. thesis, Florida State University, College of arts and sciences (2011)
38. Rentmeester, Q., Absil, P.-A.: Algorithm comparison for Karcher mean computation of rotation matrices and diffusion tensors. In: 19th Conference on European Signal Processing, 2011, EUSIPCO (2011)
39. Thompson, A.C.: On certain contraction mappings in a partially ordered vector space. Proc. Am. Math. Soc. **14**(3), 438–443 (1963)
40. Vandebril, R., Van Barel, M., Mastronardi, N.: Matrix Computations and Semiseparable Matrices. Linear Systems, vol. 1. Johns Hopkins University Press, Baltimore (2008)
41. Yang, L.: Medians of probability measures in Riemannian manifolds and applications to radar target detection. Ph.D. thesis, Université de Poitiers (2011)
42. Yang, L., Arnaudon, M., Barbaresco, F.: Geometry of covariance matrices and computation of median. In: AIP Conference Proceedings, 30th International Workshop on Bayesian Inference and Maximum Entropy Methods in Science and Engineering, vol. 1305, pp. 479–486 (2011)