

Computing the Karcher mean of symmetric positive definite matrices

Dario A. Bini^a, Bruno Iannazzo^{b,*}

^a*Dipartimento di Matematica, Università di Pisa, Largo Bruno Pontecorvo 5, 56100, Pisa, Italy. bini@dm.unipi.it*

^b*Dipartimento di Matematica e Informatica, Università di Perugia, Via Vanvitelli 1, 06123, Perugia, Italy. bruno.iannazzo@dm.unipg.it*

Abstract

We present and analyze an iterative method for approximating the Karcher mean of a set of $n \times n$ positive definite matrices A_i , $i = 1, \dots, k$, defined as the unique positive definite solution of the matrix equation $\sum_{i=1}^k \log(A_i^{-1}X) = 0$.

Keywords: Matrix geometric mean; Karcher mean; positive definite matrix

1. Introduction

Averaging real symmetric positive definite matrices is an important problem arising when one has to represent, through a single matrix G , the results of several experiments made up by a set of many $n \times n$ positive matrices A_1, \dots, A_k (throughout the paper when we speak of positive matrices, we refer to real symmetric positive definite matrices). This problem appears, for instance, in applications to elasticity [14], radar signal processing [2, 11], medical imaging [3, 9, 15] and image processing [16].

In these applications, the straightforward choice for G , that would be the arithmetic mean of A_1, \dots, A_k , does not satisfy some expected properties or gives poor results. In elasticity, for instance, one of the properties required from the matrix mean G is that the mean of $A_1^{-1}, \dots, A_k^{-1}$, must coincide with G^{-1} [14]. Among the classical means of positive real numbers a_1, \dots, a_k , this property is satisfied by the geometric mean $g = (a_1 \cdots a_k)^{1/k}$, which justifies the need of defining a suitable geometric mean of positive matrices.

*corresponding author. This work is supported by PRIN 2008 N. 20083KLJEZ, and by GNCS of INdAM.

To this end it has proved useful the Riemannian structure on the set \mathbb{P}_n of positive matrices given by the scalar product $g(M, N) = \text{trace}(A^{-1}MA^{-1}N)$ at the tangent space to \mathbb{P}_n at the point A . This structure makes \mathbb{P}_n a complete Riemannian manifold with negative curvature [5] in which the distance between two positive matrices A and B turns out to be $\delta(A, B) := \|\log(A^{-1/2}BA^{-1/2})\|_F$, where the norm used is the Euclidean or Frobenius norm $\|A\|_F := (\sum_{ij} a_{ij}^2)^{1/2}$.

In such spaces, every compact set has a unique center of mass (see [4, Section 6.1.5]). The geometric mean of k positive matrices, A_1, \dots, A_k , can be defined as their center of mass, that is the unique minimizer (over \mathbb{P}_n) of

$$f(X) = \sum_{i=1}^k \delta^2(A_i, X). \quad (1)$$

This is, for instance, the definition of geometric mean of positive matrices given by Moakher [13], Bhatia and Holbrook [6].

The center of mass of k matrices A_1, \dots, A_k in \mathbb{P}_n has been proved by Moakher [13, Prop. 3.4] to be the unique positive definite solution of the matrix equation

$$\sum_{i=1}^k \log(A_i^{-1}X) = 0, \quad (2)$$

which can be rewritten in the different forms

$$\sum_{i=1}^k \log(XA_i^{-1}) = 0, \quad \sum_{i=1}^k \log(X^{1/2}A_i^{-1}X^{1/2}) = 0, \quad (3)$$

using the formula $M^{-1}\log(K)M = \log(M^{-1}KM)$, valid for any invertible matrix M and any matrix K having real positive eigenvalues (see [10]).

The unique positive solution G of (2) is referred to as “least square geometric mean” [6], or “Riemannian geometric mean” [13], or, more frequently, “Karcher mean” [2, 12]. For this reason we call equation (2) *Karcher equation*. For $k = 2$ we recover $G = A(A^{-1}B)^{1/2}$, that is the usual definition of geometric mean of two matrices. For $k \geq 3$ such an explicit formula is not known.

In this paper, we focus on the numerical solution of the Karcher equation for $k \geq 3$. Some numerical methods have been introduced in the literature: one approach considered in [14] is to solve the Karcher equation by a fixed

point iteration, unfortunately, this iteration lacks convergence in certain examples. Another approach is to use optimization algorithms on the objective function (1), like the Newton method or the steepest descent in a Riemannian manifold (see [1]). In particular, a gradient descent algorithm can be written as

$$X_0 \in \mathbb{P}_n, \quad X_{\nu+1} = X_\nu \exp \left(-\vartheta_\nu \sum_{i=1}^k \log(A_i^{-1} X_\nu) \right), \quad (4)$$

for some step-length $\vartheta_\nu > 0$. The latter iteration has been considered in [12, 15] with $X_0 = A_1$ or $X_0 = I$ and $\vartheta_\nu = \frac{1}{k}$; with this choice iteration (4) lacks convergence for certain data A_1, \dots, A_k . The same iteration has been considered also in [2, 16] but with no comments on the choice of the initial value and of the step-length.

In principle, one could choose ϑ_ν in iteration (4) using a line search strategy, but the computation of the optimal ϑ_ν seems to be too expensive, while heuristic strategies for the step-length [9] may lead to slow convergence in a large number of cases.

Our approach is to consider a linearization of iteration (4) in the spirit of the Richardson iteration, namely,

$$X_{\nu+1} = X_\nu - \vartheta X_\nu \sum_{i=1}^k \log(A_i^{-1} X_\nu), \quad (5)$$

where $\vartheta > 0$ is a suitable parameter. Any solution of equation (3) is a fixed point of equation (5). The above recursion can be equivalently rewritten as

$$X_{\nu+1} = X_\nu - \vartheta X_\nu^{1/2} \sum_{i=1}^k \log(X_\nu^{1/2} A_i^{-1} X_\nu^{1/2}) X_\nu^{1/2}, \quad (6)$$

provided that all the iterates X_ν are positive. Equation (6) shows that if X_ν is real symmetric then $X_{\nu+1}$ is still real symmetric.

We prove that if ϑ is small enough, the sequence $\{X_\nu\}$ generated by (6) starting from X_0 in a suitable neighborhood of G , is formed by positive matrices and converges to G , and through a first order analysis we provide a bound to the convergence rate. In particular, we show that the local convergence of the iteration is linear for small values of the positive parameter ϑ and that if the matrices X_0, A_1, \dots, A_k pairwise commute and $\vartheta = 1/k$, then the convergence is at least quadratic.

In the general case, we provide an optimal value of the parameter ϑ which guarantees local convergence of (6), related to the condition numbers of $M_i = G^{1/2}A_i^{-1}G^{1/2}$, $i = 1, \dots, k$. We show that the closer are the eigenvalues of M_i to 1, the faster is the convergence.

Thus, if X_0 is *close enough* to the solution G and positive then the sequence $\{X_\nu\}$ is well defined, i.e., X_ν is positive for each ν , and converges to G . Indeed, there are cases where starting from a given X_0 , the first iterate X_1 is not positive definite. In order to overcome this drawback it may be convenient to consider also the sequence generated by (4) which shares with (6) the same convergence properties, namely local convergence for the optimal value found for (6). In fact, it can be shown that the matrices X_ν generated by (4) are positive if X_0 is positive.

We implemented the iterations (6) and (4) and tested their performances on some test problems made up by k -tuples of positive definite matrices A_1, \dots, A_k , normalized so that their spectral radius $\rho(A_i) = 1$, with the initial approximation X_0 chosen in the set $\{I, G_{\text{arith}}, G_{\text{cheap}}\}$ where G_{arith} is the arithmetic mean of A_1, \dots, A_k , G_{cheap} is the ‘‘Cheap’’ mean introduced in [8]. Moreover, we provide a strategy for an automatic choice of the parameter ϑ .

The numerical experiments confirm the theoretical analysis and point out the effectiveness of the technique for the automatic choice of the optimal value of the parameter ϑ . In fact, there are cases where the iteration with the value $\vartheta = 1/k$ does not converge or needs a huge number of steps to provide a reasonable approximation to G , while with the optimal value of ϑ , convergence always occurs in practice and the number of steps needed for convergence is reasonably small if starting with $X_0 = G_{\text{cheap}}$.

The paper is organized as follows: in Section 2 we provide the convergence analysis of iteration (6), state our main Theorem 1, analyze the choice of the optimal value of ϑ and comment on iteration (4). In Section 3 we provide the proof of Theorem 1. In Section 4 we discuss some issues related to the implementation and report the results of the numerical experiments that we have performed. Section 5 draws the conclusions.

We complete this section by recalling the definition of functions of diagonalizable matrices which we use frequently in the paper. Let $A \in \mathbb{C}^{n \times n}$ be diagonalizable, that is, $A = MDM^{-1}$, where $D = (d_{ij})$ is diagonal, and $f(x)$ is a function, then $f(A) := Mf(D)M^{-1}$, where $f(D)$ is a diagonal matrix whose diagonal elements are $f(d_{ii})$, for $i = 1, \dots, n$. If $f(x)$ is a multivalued function then the same branch of f must be taken for repeated eigenvalues.

2. Solving the Karcher equation

In this section we design a method for approximating the Karcher mean as a solution of the Karcher equation (2) relying on iteration (6). We provide a convergence analysis of iteration (6) which shows that the sequence X_ν , $\nu = 0, 1, \dots$, generated by our method has global quadratic convergence if the matrices X_0 , A_i , $i = 1, \dots, k$, pairwise commute, and local linear convergence otherwise, if ϑ is suitably chosen. We provide an explicit expression which relates the convergence factor to the condition numbers of the matrices $G^{-1/2}A_iG^{-1/2}$.

For the sake of notational simplicity let us define

$$\varphi(X) = X - \vartheta X^{1/2} \sum_{i=1}^k \log(X^{1/2}A_i^{-1}X^{1/2}) X^{1/2} = X - \vartheta X \sum_{i=1}^k \log(A_i^{-1}X),$$

so that the iteration (6) can be rewritten as $X_{\nu+1} = \varphi(X_\nu)$. For X positive definite, denote $X' = \varphi(X)$.

The scalar case, where the size n of the matrices is 1, can be handled directly, in fact the function $\varphi(X)$ has the unique positive fixed point $G = (A_1 \cdots A_k)^{1/k}$ and $\varphi'(G) = 1 - k\vartheta$. Thus, there is local convergence for $0 < \vartheta < 2/k$, and the convergence is superlinear for $\vartheta = 1/k$.

In order to study the local convergence of (6) in the general case we need to introduce some notation. We denote by G the Karcher mean of A_1, \dots, A_k , i.e., the unique positive solution of equation (2), or, equivalently, the unique positive matrix such that $\varphi(G) = G$. Set $X = G + E$, $X' = \varphi(X) = G + E'$ and denote $\widehat{E} = G^{-1/2}EG^{-1/2}$, $\widehat{E}' = G^{-1/2}E'G^{-1/2}$. Similarly, define the approximation error at step ν as the matrix

$$\widehat{E}_\nu = G^{-1/2}E_\nu G^{-1/2} = G^{-1/2}X_\nu G^{-1/2} - I, \quad E_\nu = X_\nu - G.$$

The vector counterparts of these quantities are defined as $\widehat{e}_\nu = \text{vec}(\widehat{E}_\nu)$, $\widehat{e} = \text{vec}(\widehat{E})$, $\widehat{e}' = \text{vec}(\widehat{E}')$, where $\text{vec}(\cdot)$ is the operator which transforms a matrix into a long vector by stacking its columns.

For $i = 1, 2, \dots, k$, define the following:

$$\begin{aligned} M_i &= G^{1/2}A_i^{-1}G^{1/2}, \\ W_i &= \log(M_i) \otimes I - I \otimes \log(M_i), \\ H_i &= f(W_i), \quad f(x) = x/(e^x - 1), \end{aligned} \tag{7}$$

and $H = \sum_{i=1}^k H_i$, where the function $f(x) = x/(e^x - 1)$ is extended by continuity at $x = 0$ by $f(0) = 1$. Observe that the matrices M_i are positive definite. Moreover, since $f(x)$ is positive and W_i is symmetric, also the matrices H_i are positive definite together with H .

The following result provides the Fréchet derivative of the function $\varphi(X)$. If not differently specified, $\|\cdot\|$ denotes the 2-norm (Euclidean norm) of a vector and the induced matrix norm. Observe that for a matrix $A \in \mathbb{C}^{n \times n}$, it holds that $\|A\|_F = \|\text{vec}(A)\|$.

Theorem 1. *For $H = \sum_{i=1}^k H_i$, the matrix $I - \vartheta H$ represents the Fréchet derivative of $\varphi(X)$ at G , that is,*

$$\hat{e}' = (I - \vartheta H)\hat{e} + O(\|\hat{e}\|^2).$$

The proof of the above theorem is rather technical and is postponed to Section 3. Here we observe that, since the matrix H is positive definite, there exists $\vartheta_0 > 0$ such that the spectral radius $\rho(I - \vartheta H) < 1$ for any $0 < \vartheta < \vartheta_0$.

This property enables us to prove the following convergence result which also provides an explicit representation of the linear part of the error $\hat{e}_{\nu+1}$ as function of \hat{e}_ν .

Theorem 2. *Let ϑ_0 be such that the spectral radius $\rho(I - \vartheta H) \leq \lambda < 1$ for $0 < \vartheta \leq \vartheta_0$. Then there exists $\varepsilon > 0$ such that for any symmetric matrix X_0 with $\|\hat{E}_0\|_F < \varepsilon$ the sequence $\{X_\nu\}$ generated by (6) is such that $\lim_\nu X_\nu = G$ and*

$$\hat{e}_{\nu+1} = (I - \vartheta H)\hat{e}_\nu + O(\|\hat{e}_\nu\|^2),$$

i.e., the iteration (6) is locally convergent to G .

Proof. Since $I - \vartheta H$ is symmetric positive definite for $0 < \vartheta < \vartheta_0$, for the matrix norm induced by the vector norm $\|\cdot\|$ it holds that $\|I - \vartheta H\| = \rho(I - \vartheta H) =: \lambda < 1$. Therefore, in the view of Theorem 1, one has $\|\hat{e}'\| \leq \lambda\|\hat{e}\| + \sigma\|\hat{e}\|^2$ for a suitable constant $\sigma > 0$. In this way, if ε is such that $\varepsilon < (1 - \lambda)/(2\sigma)$, then

$$\lambda\varepsilon + \sigma\varepsilon^2 < \mu\varepsilon < \varepsilon, \quad \mu = \frac{1 + \lambda}{2} < 1,$$

and all the symmetric matrices X_0 in the neighborhood $\mathcal{U}_\varepsilon = \{X : \|G^{-1/2}XG^{-1/2} - I\|_F < \varepsilon\}$ of G are positive definite and such that for any $X \in \mathcal{U}_\varepsilon$ one has

$\varphi(X) \in \mathcal{U}_\varepsilon$. That is, the sequence X_ν is well defined. Moreover, from the inequality $\|\widehat{e}\| \leq \mu \|\widehat{e}\|$, which is valid for any $X \in \mathcal{U}_\varepsilon$, one inductively deduces that $\|\widehat{e}_\nu\| \leq \mu \|\widehat{e}_{\nu-1}\| \leq \mu^\nu \|\widehat{e}_0\|$ which yields $\lim_\nu \|\widehat{e}_\nu\| = 0$. \square

Here we want to comment on the consequences of this result which is at the basis of an adaptive algorithm for the computation of G provided later on.

Observe that the eigenvalues of H_i can be explicitly given in terms of the eigenvalues $\lambda_1^{(i)}, \dots, \lambda_n^{(i)}$, of M_i . In fact, the eigenvalues of W_i are given by $\log(\lambda_r^{(i)}/\lambda_s^{(i)})$, for $r, s = 1, \dots, n$, so that the eigenvalues of H_i are simply

$$\mu_{r,s}^{(i)} = f(\log(\lambda_r^{(i)}/\lambda_s^{(i)})).$$

Since the function $f(x) = x/(e^x - 1)$ is decreasing, then the minimum and the maximum eigenvalues of H_i are

$$\mu_{\min}^{(i)} = \frac{\log(c_i)}{c_i - 1}, \quad \mu_{\max}^{(i)} = c_i \frac{\log(c_i)}{c_i - 1},$$

where $c_i = \lambda_{\max}^{(i)}/\lambda_{\min}^{(i)}$ is the condition number of the matrix M_i .

Observe that, since for symmetric matrices X, Y with eigenvalues ξ_j and η_j , respectively, the eigenvalues of $X + Y$ are in the range $[\min \xi_j + \min \eta_j, \max \xi_j + \max \eta_j]$, one deduces that the eigenvalues of $H = \sum_{i=1}^k H_i$ are in the range

$$[\beta, \gamma] := \left[\sum_{j=1}^k \frac{\log(c_j)}{c_j - 1}, \sum_{j=1}^k c_j \frac{\log(c_j)}{c_j - 1} \right]. \quad (8)$$

The above expression implies that the optimal choice of the parameter ϑ in the iteration (6) is

$$\vartheta = 2/(\gamma + \beta) = 2 / \sum_{j=1}^k \frac{c_j + 1}{c_j - 1} \log c_j \quad (9)$$

and the spectral radius of $I - \vartheta H$ which provides the convergence rate of the iteration is such that

$$\rho(I - \vartheta H) \leq \frac{\gamma - \beta}{\gamma + \beta} = \frac{\sum_{j=1}^k \log c_j}{\sum_{j=1}^k \frac{c_j + 1}{c_j - 1} \log c_j}. \quad (10)$$

Theorem 2 enables us to analyze the convergence speed in the case where the matrices A_i commute, i.e., $A_i A_j = A_j A_i$ or in the case where $\|A_i A_j - A_j A_i\|$ or $\|A_i - A_j\|$ is small.

Consider the case where the matrices commute and where also X_0 commutes with them (e.g., $X_0 = I$). Then, from the fact that $\text{vec}(AEB) = (B \otimes A)\text{vec}(E)$, $\text{vec}(BEA) = (A \otimes B)\text{vec}(E)$, for real symmetric matrices A , B , the condition $AEB = BEA$ implies that we can move from left to right the matrices appearing in a Kronecker product. This fact enables us to write $H_i \hat{e}_\nu = f(W_i) \hat{e}_\nu = f(0) \hat{e}_\nu = \hat{e}_\nu$ so that $H \hat{e}_\nu = k \hat{e}_\nu$. Therefore, from Theorem 2 and (7) one has $\hat{e}_{\nu+1} = (1 - k\vartheta) \hat{e}_\nu + O(\|e_\nu\|^2)$ and the choice $\vartheta = 1/k$ provides the optimal value for ϑ which yields a convergence speed at least quadratic since, in this case

$$\|e_{\nu+1}\| = O(\|e_\nu\|^2).$$

A similar analysis can be carried out if $\|A_i - A_j\|$ is “small”. Here we provide some “qualitative” comments to point out the behavior of the convergence and leave the “quantitative” analysis to the reader. If $\|A_i - A_j\|$ is small, then also $\|A_i - G\|$ is small and the matrices $M_i = G^{1/2} A_i^{-1} G^{1/2}$ are close to the identity matrix. Therefore their condition number c_i is close to 1, $\log c_i$ is close to 0 as well as the bound $\rho(I - \vartheta H)$ of (10), moreover, the values of γ and β are close to k . Therefore, the choice $\vartheta = 1/k$ provides a small bound to the spectral radius $\rho(I - \vartheta H)$. In this case we expect a very fast convergence even though linear.

A different behavior may be encountered in the case where the matrices A_i almost commute, i.e., if $\|A_i A_j - A_j A_i\|$ is small, but they are not close to each other. In fact, in this case the iteration may fail to converge as the following example shows. Choose the k matrices such that A_i differs from a diagonal matrix by a small correction and the diagonal matrices are very different. For instance, consider $k = n$ matrices of size n with diagonal entries $(1, d, d^2, \dots, d^{n-1})$ and all its cyclic permutations, where $0 < d < 1$ is a small positive value. After one step with $\vartheta = 1/n$, the error is expected to be of the order of the perturbation, but in the subsequent steps the error may start to increase and even to diverge. In this case G is close to σI , with $\sigma = d^{(n-1)/2}$, and the matrices $G^{-1/2} A_i G^{-1/2}$ have a large condition number close to $1/d^n$. Therefore, in order to have convergence, a suitable value for $\vartheta \approx 2/(n^2 \log d)$ must be used. The numerical experiments performed in Section 4.1 confirm this analysis.

This proves that iteration (6) with $\vartheta = 1/k$ is not stable for commuting matrices, since a small perturbation in the input may lead to divergence. This example is only of theoretical interest, since for commuting matrices a fair way to compute the Karcher mean is the direct evaluation of the formula $G = (A_1 A_2 \cdots A_k)^{1/k}$.

2.1. Modifying the iteration

In this section we consider iteration (4), equivalently rewritten in a more symmetric form as

$$X_{\nu+1} = X_{\nu}^{1/2} \exp \left(-\vartheta_{\nu} \sum_{i=1}^k \log(X_{\nu}^{1/2} A_i^{-1} X_{\nu}^{1/2}) \right) X_{\nu}^{1/2}, \quad (11)$$

which is a modification of (6) in the sense that the latter can be obtained from the former by replacing the exponential function $\exp(t)$ with its linearization $1 + t$. To be more precise, denoting $R(X) = \sum_{i=1}^k \log(X^{1/2} A_i^{-1} X^{1/2})$, one has $R(G) = 0$ if G solves (6). On the other hand, since the difference

$$\exp(R(X)) - (I + R(X))$$

contains quadratic terms in the entries of $R(X)$, the linear part of the error $E_{\nu+1}$ obtained with (6) has the same form of the linear part obtained with (11), therefore the same comments on the local convergence of (6) apply to the iteration (11).

The iteration (11) still keeps the nice convergence properties of (6) but avoids the possibility of breakdown to which (6) is prone if X_0 is not chosen appropriately. In fact, it is immediate to verify that the matrices X_{ν} generated by (11) are positive. Indeed, the modified iteration has a slightly higher computational cost since at each step a matrix exponential has to be computed.

Observe that in the scalar case, or in the case where X_0 and A_i , $i = 1, \dots, k$, commute with each other then one step of iteration (11) provides the geometric mean.

3. Proof of Theorem 2.1

Our proof relies on the Kronecker notation and on some well known properties of Fréchet derivatives of matrix functions (see [10, Chapter 3]). We

denote the Fréchet derivative of the function $f(x)$ at X in the direction E by the symbol $L_f(X, E)$ and we represent the linear operator $L_f(X, \cdot)$ through the matrix $K_f(X)$ such that $\text{vec}(L_f(X, E)) = K_f(X)\text{vec}(E)$.

By means of the following representation of the Fréchet derivative of the exponential function (obtained using the commutativity on Theorem 10.13 of [10])

$$K_{\exp}(Y) = \psi(Y^T \otimes I - I \otimes Y)(I \otimes \exp(Y)), \quad \psi(t) = \frac{e^t - 1}{t},$$

and applying the rule for the inverse of the Fréchet derivative, we get

$$K_{\log}(X) = (I \otimes X^{-1})f(\log X^T \otimes I - I \otimes \log X), \quad (12)$$

where $f(t) = t/(e^t - 1)$ for $t \neq 0$ and $f(0) = 1$. Recall that for the function \log , equation $f(X + E) = f(X) + L_f(X, E) + o(\|E\|)$, holds with the term $o(\|E\|)$ replaced by $O(\|E\|^2)$.

We are ready to prove the following

Lemma 3. *For the function $\varphi(X) = X \log(A^{-1}X)$ it holds that*

$$L_\varphi(X, E) = E \log(A^{-1}X) + X L_{\log}(A^{-1}X, A^{-1}E).$$

Moreover, $L_\varphi(X, E)$ admits the matrix representation

$$K_\varphi(X) = \log(XA^{-1}) \otimes I + f(\log((A^{-1}X)^T) \otimes I - I \otimes \log(XA^{-1})). \quad (13)$$

Proof. Consider the functions $g(X) = X$ and $h(X) = \log(A^{-1}X)$. Applying to $\varphi = g \cdot h$ the rule for the Fréchet derivative of a product we get

$$L_\varphi(X, E) = E \log(A^{-1}X) + X L_h(X, E), \quad (14)$$

since $L_g(X, E) = E$. We may look at the function $h(X) = \log(A^{-1}X)$ as at the composition of the two functions $\log(X)$ and $A^{-1}X$. In this way, applying the chain rule for the Fréchet derivative, we get $X L_h(X, E) = X L_{\log}(A^{-1}X, A^{-1}E)$ which, together with (14) yields the required expression for $L_\varphi(X, E)$.

Using (12) and the properties $\text{vec}(AB) = (I \otimes A)\text{vec}(B) = (B^T \otimes I)\text{vec}(A)$, we get

$$\begin{aligned} K_\varphi(X)\text{vec}(E) &= (\log(XA^{-1}) \otimes I)\text{vec}(E) + (I \otimes X)\text{vec}(L_{\log}(A^{-1}X, A^{-1}E)) \\ &= (\log(XA^{-1}) \otimes I + (I \otimes A)f(\log(A^{-1}X)^T \otimes I - I \otimes (A^{-1}X))(I \otimes A^{-1})\text{vec}(E), \end{aligned}$$

which, moving $I \otimes A$ and $I \otimes A^{-1}$ inside the argument of f , leads to (13). \square

By applying the above result to the function $\Phi(X) = I - \vartheta \sum_{i=1}^k \varphi_i(X)$ $\varphi_i(X) = X \log(A_i^{-1}X)$, one obtains that the linear mapping $L_\Phi(X, E)$ is represented by the matrix

$$K_\Phi(X) = I - \vartheta \sum_{i=1}^k \log(XA_i^{-1}) \otimes I + \vartheta \sum_{i=1}^k f(\log((A_i^{-1}X)^T)) \otimes I - I \otimes \log(XA_i^{-1}).$$

Since $\sum_{i=1}^k \log GA_i^{-1} = 0$, one deduces that

$$K_\Phi(G) = I - \vartheta \sum_{i=1}^k f(\log(GA_i^{-1})) \otimes I - I \otimes \log(GA_i^{-1}).$$

Whence, replacing GA_i^{-1} by $G^{1/2}M_iG^{-1/2}$ yields

$$K_\Phi(G) = I - \vartheta (G^{1/2} \otimes G^{1/2}) \left(\sum_{i=1}^k f(\log(M_i)) \otimes I - I \otimes \log(M_i) \right) (G^{-1/2} \otimes G^{-1/2}).$$

This completes the proof of Theorem 1.

4. Implementation issues

The Richardson-like iteration (6) can be implemented in formally different ways according to the equivalent formulae:

$$\begin{aligned} X &= X - \vartheta X^{1/2} \sum_{i=1}^k \log(X^{1/2}A_i^{-1}X^{1/2})X^{1/2}, & G &= X, & \text{(a)} \\ X &= X^{1/2} (I - \vartheta \sum_{i=1}^k \log(X^{1/2}A_i^{-1}X^{1/2})) X^{1/2}, & G &= X, & \text{(b)} \\ Y &= Y - \vartheta Y^{1/2} \sum_{i=1}^k \log(Y^{1/2}A_iY^{1/2})Y^{1/2}, & G &= Y^{-1}, & \text{(c)} \\ Y &= Y^{1/2} (I - \vartheta \sum_{i=1}^k \log(Y^{1/2}A_iY^{1/2})) Y^{1/2}, & G &= Y^{-1}. & \text{(d)} \end{aligned} \tag{15}$$

The cases (c) and (d) consist in applying the iteration (a) and (b) respectively, for the computation of the mean of the inverse matrices A_i^{-1} , and then inverting the result. In this approach we reduce the number of matrix

inversion to be performed and improve the numerical stability in the case where the mean G is better conditioned than the matrices A_1, \dots, A_k . In our numerical tests we simply considered the implementation based on (15) (a).

Yet another implementation which is numerically appealing uses the Cholesky factorization of X_ν , say $X_\nu = R_\nu^T R_\nu$,

$$X_{\nu+1} = X_\nu + \vartheta R_\nu^T \left(\sum_{i=1}^k \log(R_\nu^{-T} A_i R_\nu^{-1}) \right) R_\nu. \quad (16)$$

Since the condition number of the Cholesky factor of X_ν in the spectral norm is the square root of the condition number of X_ν , we expect a good accuracy. Moreover, expression (16) can be computed with a smaller computational cost with respect to (15), since forming the Cholesky factor costs less than forming the matrix square root [10].

In order to choose the optimal value of the parameter ϑ according to the results of Section 2 we have to compute the values γ and β in equation (8). This computation can be performed dynamically by means of (8), (9), where $c_j = \text{cond}(G^{1/2} A_j^{-1} G^{1/2})$ can be approximated by replacing the matrix G with the current approximation X_ν . Once again, for the success of this heuristic it is needed that X_0 is a good approximation to G . For this reason, choosing X_0 equal to the Cheap mean is crucial.

A better evaluation of the optimal value of the parameter ϑ could be performed by means of the evaluation of the largest and the smallest eigenvalues of the matrix obtained by H in (7) by replacing G with the current iterate X_ν . This computation can be performed by means of few steps of the Lanczos method, where the computation of the matrix-vector product can be performed with cost $O(kn^3)$ by exploiting the specific Kronecker structure of H . The optimal value of ϑ is then computed as:

$$\vartheta = 2/(\mu_{\max} + \mu_{\min}). \quad (17)$$

In the next section we show by means of several numerical experiments the effectiveness of our algorithms.

4.1. Numerical experiments

The first bunch of tests has been designed to demonstrate that starting the iteration (6) in the form (15) (a), with X_0 equal to the Cheap mean,

defined as the common limit of the sequences $\{A_i^{(\nu)}\}_\nu$, $i = 1, \dots, k$, such that

$$A_i^{(\nu+1)} = A_i^{(\nu)} \exp\left(\frac{1}{k} \sum_{j=1, j \neq i}^k \log((A_i^{(\nu)})^{-1} A_j^{(\nu)})\right), \quad i = 1, \dots, k, \quad (18)$$

we can approximate the Karcher mean with much less iterations than starting with $X_0 = I$ or with $X_0 = \frac{1}{k} \sum_{i=1}^k A_i$. Using the Cholesky factorization of $A_i^{(\nu)}$, iteration (18) can be rewritten in the form

$$A_i^{(\nu)} = R^T R, \quad A_i^{(\nu+1)} = R^T \exp\left(\frac{1}{k} \sum_{j=1, j \neq i}^k \log(R^{-T} A_j^{(\nu)} R^{-1})\right) R, \quad i = 1, \dots, k,$$

which leads to a different implementation where just functions of symmetric matrices are computed. Table 1 reports the number of iterations obtained with the different values of X_0 and with the value of ϑ computed once for all by means of (9) where G has been replaced by X_0 .

The iteration has been halted if the residual correction is less than $1.0\text{e-}11$ or if it is not decreasing. The latter condition has never been encountered in this set of tests. The dataset has been generated randomly with different values of the condition number of the matrices A_i according to the following Matlab commands:

```
n = 10; W = rand(n) - rand(n); X1 = W' * W;
X2 = X1 - eye(n) * t; X = X2/norm(X2, 2);
```

where the parameter t is chosen in such a way that the matrix \mathbf{X} has a given condition number cnd . In fact, the condition number of \mathbf{X} is given by $\text{cnd} = (\lambda_1 - t)/(\lambda_n - t)$, provided that $t < \lambda_n$, where $\lambda_1, \lambda_n \geq 0$ are the largest and the smallest eigenvalues of $\mathbf{X1}$, respectively. This way, it is sufficient to choose $t = (\text{cnd}\lambda_n - \lambda_1)/(\text{cnd} - 1)$. This choice guarantees the nonsingularity of \mathbf{X} .

It is important to point out that the normalization that we have imposed on our data, i.e., $\|X\|_2 = 1$, is no loss of generality. In fact, the Karcher mean $G(A_1, \dots, A_k)$ of a set of positive definite matrices A_1, \dots, A_k is such that

$$G(\alpha_1 A_1, \dots, \alpha_k A_k) = (\alpha_1 \cdots \alpha_k)^{1/k} G(A_1, \dots, A_k)$$

for any $\alpha_1, \dots, \alpha_k > 0$.

$k \backslash X_0$	cond=1.e2			cond=1.e4		
	I	AM	Cheap	I	AM	Cheap
3	74 (0.1)	26 (0.27)	17 (0.24)	114 (0.07)	89 (0.08)	41 (0.16)
4	66 (0.83)	21 (0.19)	17 (0.18)	82 (0.07)	59 (0.09)	37 (0.11)
5	65 (0.07)	19 (0.16)	16 (0.15)	87 (0.05)	58 (0.08)	35 (0.1)
6	62 (0.06)	20 (0.13)	16 (0.13)	81 (0.05)	54 (0.07)	31 (0.09)
7	61 (0.05)	21 (0.11)	15 (0.11)	83 (0.04)	63 (0.05)	29 (0.08)
8	61 (0.05)	20 (0.1)	15 (0.1)	93 (0.03)	55 (0.05)	29 (0.07)
9	58 (0.04)	19 (0.09)	14 (0.09)	89 (0.03)	50 (0.04)	29 (0.06)
10	56 (0.08)	19 (0.08)	14 (0.08)	94 (0.03)	47 (0.05)	28 (0.06)

Table 1: Number of iterations needed to approximate the Karcher mean up to the residual $1.e-11$ by means of the iteration (15) (a), starting with the identity matrix, the arithmetic mean and the Cheap mean. Between parentheses it is reported the value of ϑ computed once for all by means of (9) with G replaced by X_0 .

We reported the case of $n = 10$, $k = 3 : 10$ for different values of the condition number. It is evident the substantial reduction of the number of steps obtained starting with the Cheap mean.

The second bunch of tests concerns an automatization for the choice of ϑ . In this case we considered a set of $n \times n$ matrices A_1, \dots, A_k , with $k = n$, randomly generated in a neighborhood of radius ε for different values of ε and of the condition number of the matrices. More precisely, after generating a diagonal matrix $A_1 = \text{diag}(d_1, \dots, d_n)$, we generated the remaining matrices by adding to A_1 a nonnegative definite random perturbation of norm ε .

We implemented an adaptive version of the iteration with X_0 equal to the Cheap mean, where the parameter ϑ is chosen at each step by means of equation (9). In these formulae, we have replaced the unknown matrix G by the current approximation X_ν .

We have implemented a version where the parameter ϑ is chosen by applying Lanczos method, and finally we have implemented a third version where ϑ is chosen once for all by trying several values of ϑ , sampled at distance 0.01 and choosing the one which provides the minimum number of iterations.

Among the many tests performed we report only two cases which differ for the condition number of the matrices. Table 2 reports the number of iterations for different values of the radius ε in the two cases where $n = 5$ and the center of the neighborhood of radius ε is $A_1 = \text{diag}(1, 10^{-1}, 10^{-2}, 10^{-4}, 10^{-5})$ and $A_1 = \text{diag}(1, 1/5, 1/10, 1/15, 1/20)$, respectively. It is interesting to point

$\varepsilon \backslash \vartheta$	cond=1e5				cond=20			
	1/k	(9)	(17)	samp.	1/k	(9)	(17)	samp.
2e-1	*	39	28	22	6	6	6	5
1e-1	4533	34	30	19	5	5	4	4
1e-2	46	22	17	14	2	2	2	2
1e-3	28	20	14	12	1	1	1	1
1e-4	10	8	7	6	1	1	1	1

Table 2: Number of iterations needed to approximate the Karcher mean for matrices lying in a neighborhood of radius ε with $\vartheta = 1/k$ and with the values for ϑ computed by means of (9), (17) and by sampling.

out that in the first case, where the matrices are ill conditioned, with the value $\vartheta = 1/5$ the iteration fails to converge for $\varepsilon = 0.2$ and requires a large number of iterations for $\varepsilon = 0.1$. A smaller value of ϑ computed with the techniques described in the previous sections leads to a dramatic reduction of the number of iterations.

In the case of well conditioned matrices, the optimal value of ϑ is close to $1/k$ almost independently of the radius ε and the number of iterations is negligible.

As a matter of fact, we have observed that starting with X_0 equal to the Cheap mean the value of ϑ computed by means of (9) does not change much in the iterations. Therefore it is more convenient to compute this value once for all.

Concerning the choice of ϑ , we observe that relying on Lanczos method slightly reduces the number of iterations with respect to the choice based on (9). Moreover, the value of ϑ obtained by sampling the number of iterations and choosing the value which provides the minimum does not dramatically reduce the number of iteration. Therefore, for practical computations, (9) is a good balance between reducing the number of iterations and keeping a low cost.

Finally, we have tested the convergence speed and the value of the parameter ϑ in the case where the matrices A_i , $i = 1, \dots, k$, are almost commuting, i.e., $\|A_i A_j - A_j A_i\|$ is small but the matrices are far from their mean. We have selected the case where $A_1 = \text{diag}(1, 1/10, 1/100, 1/1000, 1/10000)$ and the matrices A_i , for $2 \leq i \leq 5$ are obtained from A_1 by applying a cyclic permutation to the diagonal entries and by adding a positive definite matrix

of norm 10^{-8} . The value of the mean is close to the scalar matrix σI , with $\sigma = 10^{-11/5}$. It is interesting to observe that with the value $\vartheta = 1/k = 1/5$, which is optimal in the commutative case, the iteration fails to converge, while with the choice based on (9) one obtains convergence in just 4 iterations.

5. Conclusions

We have introduced a new iteration depending on a parameter ϑ for computing the Karcher mean of k symmetric positive definite matrices, proved its global quadratic convergence in the commutative case and its local linear convergence in the general case. We have provided a criterion for determining a value of ϑ which guarantees local convergence. By means of a wide numerical experimentation we have shown that choosing as initial approximation the Cheap mean, this iteration converges quickly to the Karcher mean. An adaptive version of this iteration is proposed and implemented in the Matrix Means Toolbox [7].

There are still some questions that need some more investigation. The convergence of the proposed algorithm relies on the global convergence of the Cheap mean iteration and the convergence of iteration (5) to the Karcher mean for $X_0 = G_{\text{cheap}}$ and with the value of ϑ chosen as in (9). Both facts have been observed in a wide set of numerical experiments, however it would be nice to confirm the results with a proof.

Another interesting question concerns the application to a radar problem [11]. In that case the matrices to be averaged are Toeplitz matrices, while their Karcher mean is not. It could be more meaningful, from a physical point of view, to have an averaging procedure which yields a Toeplitz matrix. This is a topic of future research.

References

- [1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization algorithms on matrix manifolds*. Princeton University Press, Princeton, NJ, 2008.
- [2] F. Barbaresco. New Foundation of Radar Doppler Signal Processing based on Advanced Differential Geometry of Symmetric Spaces: Doppler Matrix CFAR & Radar Application. In *International Radar Conference 2009*, Bordeaux, France, October 2009.

- [3] P. G. Batchelor, M. Moakher, D. Atkinson, F. Calamante, and A. Connelly. A Rigorous Framework for Diffusion Tensor Calculus. *Magn. Reson. Med.*, 53:221–225, 2005.
- [4] M. Berger. *A panoramic view of Riemannian geometry*. Springer-Verlag, Berlin, 2003.
- [5] R. Bhatia. *Positive definite matrices*. Princeton Series in Applied Mathematics. Princeton University Press, Princeton, NJ, 2007.
- [6] R. Bhatia and J. Holbrook. Riemannian geometry and matrix geometric means. *Linear Algebra Appl.*, 413(2-3):594–618, 2006.
- [7] D. A. Bini and B. Iannazzo. The Matrix Means Toolbox. <http://bezout.dm.unipi.it/software/mmttoolbox> (Retrieved on July 29, 2011).
- [8] D. A. Bini and B. Iannazzo. A note on computing matrix geometric means. *Adv. Comput. Math.*, 2010. In print, DOI: 10.1007/s10444-010-9165-0.
- [9] P. T. Fletcher and S. Joshi. Riemannian geometry for the statistical analysis of diffusion tensor data. *Signal Processing*, 87:250–262, 2007.
- [10] N. J. Higham. *Functions of Matrices: Theory and Computation*. Society for Industrial and Applied Mathematics, Philadelphia, PA, USA, 2008.
- [11] J. Lapuyade-Lahorgue and F. Barbaresco. Radar detection using Siegel distance between autoregressive processes, application to HF and X-band radar. In *Radar Conference, 2008. RADAR '08, IEEE*, Rome, May 2008.
- [12] J. H. Manton. A globally convergent numerical algorithm for computing the centre of mass on compact Lie groups. In *Eighth International Conference on Control, Automation, Robotics and Vision, 2004. ICARCV 2004 8th*, Kunming, China, December 2004.
- [13] M. Moakher. A differential geometric approach to the geometric mean of symmetric positive-definite matrices. *SIAM J. Matrix Anal. Appl.*, 26(3):735–747, 2005.

- [14] M. Moakher. On the averaging of symmetric positive-definite tensors. *J. Elasticity*, 82(3):273–296, 2006.
- [15] X. Pennec, P. Fillard, and N. Ayache. A Riemannian Framework for Tensor Computing. *Int. J. Comput. Vis.*, 66:41–66, 2006.
- [16] Y. Rathi, O. Michailovich, and A. Tannenbaum. Segmenting images on the Tensor Manifold. In *Proceedings of Computer Vision and Pattern Recognition*, pages 1–8, 2007.