The Riemann barycenter computation and means of several matrices

Miklós Pálfia

Abstract—An iterative definition of any n variable mean function is given in this article, which iteratively uses the two-variable form of the corresponding two-variable mean function. This extension method omits recursivity which is an important improvement compared with certain recursive formulas given before by Ando-Li-Mathias, Petz-Temesi. Furthermore it is conjectured here that this iterative algorithm coincides with the solution of the Riemann centroid minimization problem. Certain simulations are given here to compare the convergence rate of the different algorithms given in the literature. These algorithms will be the gradient and the Newton mehod for the Riemann centroid computation.

Keywords—means, matrix means, operator means, geometric mean, Riemannian center of mass

I. Introduction

The theory of means of two positive matrices was first developed by Kubo and Ando in [14] and since it has found a number of applications in operator theory in [6] and in Diffuse Tensor Magnetic Resonance Imaging (DT-MRI) see ([7], [11]). The development of DT-MRI and also certain approaches in quantum information theory led to the generalization of Lie-Trotter formulas in the Log-Euclidean framework in [1] and beyond. One can see that the general theory of matrix means has some extensive literature as long as we are interested in the means of two matrices. Although the means of two matrices are well established we must say the converse is true for the means of several matrices.

In [2] the geometric mean of matrices was extended to several matrices with a recursive extension method. Although it provides sufficient definition for the n-variable geometric mean, practically it is inapplicable due to its recursivity. This recursive extension was generalized to any mean function in [20], but it suffered from the same problem, and a further one as well, its convergence could only be proven for orderable matrices. In [18] the recursivity was omitted, so the n-variable mean could be constructed directly from the two-variable form, although the proofs still needed the initial orderability.

There is another approach for extending mean functions to several matrices, which is mainly built on differential geometric ideas. It is based on the fact that the geometric mean of two matrices correspond with a minimization problem defined to be the Riemannian centroid with respect to the Thompson metric, which induces the manifold of positive definite self-adjoint matrices. This minimization problem was generalized to several matrices in [16], [17] with respect to the Thompson metric. In these two articles in particular cases, it was possible to solve explicitly the minimization problem and

Miklós Pálfia is with the Department of Automation and Applied Informatics, Budapest University of Technology and Economics, H-1521 Budapest, Hungary, email: palfia.miklos@aut.bme.hu

acctually the formulas yield the corresponding several variable version of the geometric mean.

In this article we will use the non-recursive extension given for any mean function in [18] to the computation of the Riemann barycenter. In the next section basic definitions will be given, which are inherited from previous related work. After that in the following section we will provide some ideas related to the Riemann centroid computation, and a simulation is given which suggests that the iterative algorithm provides the centroid for corresponding metrics.

II. Preliminaries and related work

Let $P(n)$ denote the open convex cone of self adjoint, positive definite $n \times n$ matrices.

Definition II.1. A two-variable function $M : P(n) \times P(n) \mapsto P(n)$ is called a mean function if

(i) $M(X, X) = X$ for every $X \in P(n)$,
(ii) If $X < Y$, then $X < M(X, Y) < Y$,
(iii) If $X < X'$ and $Y < Y'$, then $M(X, Y) < M(X', Y')$,
(iv) $M(X, Y)$ is continuous,
(v) $M(CXC^*, CYC^*) = CM(X, Y)C^*$ $X, Y \in P(n)$ and C is invertible.

Using the above definition 2-variable mean functions can be defined for pairs of positive matrices. This result was carried out by Kubo and Ando who provided certain formulas, for defining and computing a two-variable mean function. First of all they defined the formula for an $M(A, B)$ matrix mean, with the help of the class of normalized operator monotone functions as

$$M(A, B) = A^{1/2} f(A^{-1/2} B A^{-1/2}) A^{1/2}. \qquad (1)$$

Throughout the article we will use the above definitions for 2-variable matrix means.

The next definition provides a class of extension methods using a two-variable mean function. This is a generalized form of the definition given for symmetric means in [18].

Definition II.2. Let $X = (X_1^0 \ldots X_n^0)$ where X_i^0 is self-adjoint positive definite matrix and G be a directed graph, with n vertices and edges given as that, there is one cycle in G, not taking into account the direction of the edges, which contains all vertices and edges (so it is at the same time a *Hamiltonian*- and an *Euler-cycle*). This implies that in G, every vertex has two edges and that G is coherent, which can be seen in Figure 1.

Let us define a one to one correspondence between X_i^0 numbers and G-s vertices. Taking every directed edge in G as an $M(X_j^k, X_l^k)$ (where M is a (not necessarily symmetric)
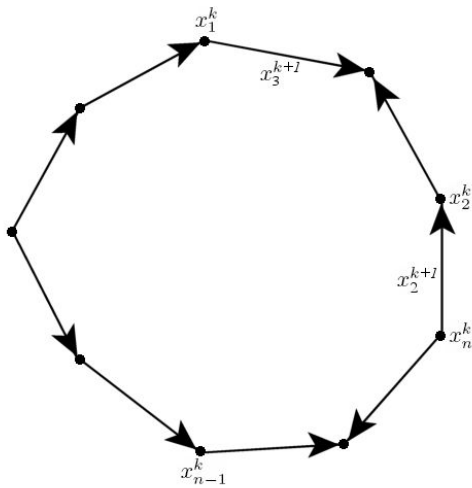
World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:2, No:4, 2008

Fig. 1. This graph represents an iteration scheme, which can be used to extend a mean function to n-variables

mean function and X_j^k, X_l^k are assigned to the two ending points of the edge as previously given), we can define a sequence with an optional n mappings, as a correspondence between every directed edge (X_j^k, X_l^k) and an X_i^{k+1}

$$X_i^{k+1} = M(X_j^k, X_l^k) =$$
$$= \left[X_j^k\right]^{1/2} f\left(\left[X_j^k\right]^{-1/2} X_l^k \left[X_j^k\right]^{-1/2}\right) \left[X_j^k\right]^{1/2} \quad (2)$$
$$i, j, l \in [1, \dots n] \quad j \neq l.$$

The following definition will be our basic extension method for mean functions throughout the article, according to [18].

Definition II.3. Let $X = (X_1^0 \dots X_n^0)$ where X_i^0 is self-adjoint positive definite matrix and $M(X, Y)$ be a two variable mean function. Let us define the following iteration

$$X_i^{k+1} = \begin{cases} M(X_1^k, X_2^k) & \text{if } i = 1, \\ M(X_{n-1}^k, X_n^k) & \text{if } i = n, \\ M(X_{i-1}^k, X_{i+1}^k) & \text{else.} \end{cases} \quad (3)$$

If the initial matrices are orderable, it can be easily proven that the above iterative sequences are convergent and that they must have the same limit point, which is the limit of the recursive extension given in [20]. For further information on the proofs see [18].

Very recently we were able to prove in [19] for certain 2-variable mean functions that are not necessarily symmetric, that the iteration in Definition II.3 for unorderable matrices are convergent and that the sequences have the same limit point. The proof was built on the following theorem and other lemmas, which we do not present here, beacuse they are out of the scope of this article.

Theorem II.1. *The sequences given in Definition II.3, using the arithmetic mean $M(X, Y) = A(X, Y) = (X + Y)/2$, for all n are convergent and have the same limit point which is the n-variable arithmetic mean.*

Proof: We have two different approach for the assertion. According to [18] we know the above to be true for orderable tuples, where it is proven that the iteration in Definition II.3 converge to the same limit for orderable tuples as for the recursive extension method, which extends from n to $(n+1)$ variable. It is easy to see that the arithmetic mean case for matrices can be treated matrix-elementwise, therefor it must be true for unordered matrices.

The second approach is to see, that the iteration in Definition II.3 is matrix-elementwise convergent and the sequences of matrices have the same limit according to [18]. The next step is to check that the n-variable arithmetic mean is invariant under one iterational step given in Definition II.3.

$$\frac{X_1^k + X_2^k + \cdots + X_n^k}{n} \quad (4)$$

Applying one iterational step we get the following equations

$$\frac{X_1^{k+1} + X_2^{k+1} + \cdots + X_n^{k+1}}{n} =$$

$$= \frac{\frac{X_1^k + X_2^k}{2} + \cdots + \frac{X_{i-1}^k + X_{i+1}^k}{2} + \cdots + \frac{X_{n-1}^k + X_n^k}{2}}{n} = \quad (5)$$

$$= \frac{X_1^k + X_2^k + \cdots + X_n^k}{n} = \frac{X_1^0 + X_2^0 + \cdots + X_n^0}{n}$$

for all k.

The above means that the iteration leaves invariant the arithmetic mean of the initial matrices and it is also known that the iteration converges for each sequence, hence their limit must be the arithmetic mean. ∎

In [18] it has been shown, that the iterations in Definition II.2 and Definition II.3 will converge to the same limit as long as the tuples are orderable. This phenomenon, coupled with simulations done in MAPLE, suggest that all iterations with precisely defined mappings according to [18], would converge to the same limit for unorderable tuples as well. For investigating the convergence rate of the iterative, non-recursive extensions and verifing numerically the above mentioned, it is straightforward to present the following simulation. We used the two-variable forms of the *geometric* and the *logarithmic mean*

$$L(x, y) = \frac{x - y}{\log x - \log y} \quad (6)$$

to extend to four variables of matrices. All of the matrices were random positive definite self-adjoint 4×4 matrices. We have started the iteration with different random quadruples for the two different *mean*. We have done two distinct simulations for each of the two *means* with the same starting matrices but with different mappings. We have measured the deviation of the four matrices in each iterational step, which is defined as

$$Deviation(A, B, C, D) =$$
$$= max\{R(X_i, X_j)\}, where\ X_i, X_j \in \{A, B, C, D\}, \quad (7)$$

where $R(A, B) = max\{\rho(A^{-1}B), \rho(B^{-1}A)\}$ and $\rho(T)$ is the spectral radius of T. It is easy to verify that

$$R(A, B) = 1 \Leftrightarrow A = B. \quad (8)$$

TABLE I
EXTENSION OF THE LOGARITHMIC AND GEOMETRIC MEAN TO FOUR
MATRICES.

| Number of Iterations	Geometric mean sequence I.	Geometric mean sequence II.
	Deviation	Deviation
1	16.46115759	18.80397424
2	13.86715043	15.62342582
3	3.972098457	4.182665248
4	3.709805720	2.387243298
5	1.989810482	1.594120461
10	1.177921246	1.186296397
20	1.005130284	1.003401006
Number of Iterations	Logarithmic mean sequence I.	Logarithmic mean sequence II.
	Deviation	Deviation
1	6.637814855	7.533990676
2	4.970743956	5.670547504
3	2.493474496	2.656267728
4	2.188138612	2.095603966
5	1.571889827	1.489249289
10	1.057972042	1.061962393
20	1.003041234	1.002876438
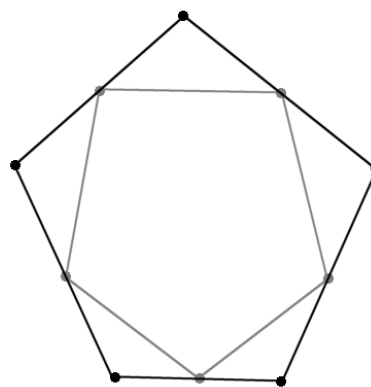
Fig. 2. Iterational Scheme 1 for 5 matrices represented with points.

Now Table I shows that the deviation of the matrices are decreasing exponentially and it appears to be that they are converging to 1. Moreover the differently mapped iterations as well seem to converge to the same limit as the qaudruples do individually.

In the next section we will give some convergence rate considerations according to some topological properties of the matrices.

III. CONVERGENCE RATE CONSIDERATIONS

We have mentioned before that for orderable tuples it was shown in [18] that every iterational scheme in Definition II.2 converges to the same limit point as long as the mean is symmetric. Although it has not been proven for unorderable tuples yet, it implies numerical considerations involved in the actual convergence rate of these algorithms. From now on, we will assume this to be true for unorderable tuples as well.

Firstly we will consider two different mappings for 5 distinct matrices. For the sake of simplicity we will represent these matrices on the next two figures with points as if they were elements of the Euclidean plane. We will apply the two-variable arithmetic mean in the iteration given in Definition II.2. Let us consider the next two mappings given by Figure 2 and 3.

We know that the arithmetic mean for two vectors is given by the halving point of the straight line connecting them together. According to this phenomenon it can be also easily verified that the next five points given by any mapping will be the elements of the set given by the polygon in Figure 2 with black edges. Therefore their limit point will be the element of this set as well. So generally the convergence rate of these iterations is highly dependent from the choice of the mappings in every iterational step. For instance in Figure 2 and 3 the area of the polygons given by gray edges are significantly different. In Figure 3 the mapping was chosen to include the two-variable mean of the farthest points without disobeying the rules given in Definition II.2. It can be easily decided that
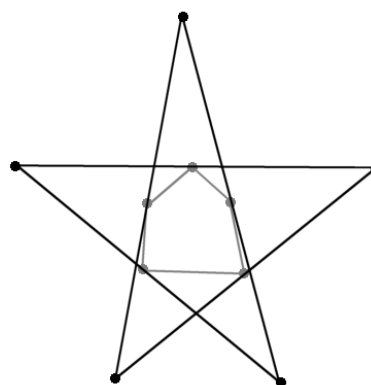

Fig. 3. Iterational Scheme 2 for 5 matrices represented with points.

the second mapping will provide almost the fastest iterational scheme for 5 matrices.

Taking into account the above mentioned, we can modify our iterational scheme with adding some heuristics, making it adaptive to the geometry of the sets given by the points in every iterational step. The function Idealmapping defined by Algorithm 1 returns the array ma which contains the indexes of matrices in such an order, that if we set up one iterational step in such a way that we take the mean of every two matrices that have neighboring indices in ma or have the first and last indices, we can sufficiently reduce the distance between them.

The function defined in Algorithm 1 if applied in every iterational step, it greatly improves the convergence rate of the sequences. Note that a $d(\cdot, \cdot)$ distance function must be provided as well. We suggest choosing distance functions that are not too costly to compute, beacuse it must be carried out

World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:2, No:4, 2008

Algorithm 1 Idealmapping

Require: $x_1, \ldots, x_n$
1: $d \Leftarrow n(n-1)/2$
2: $i \Leftarrow 1, j \Leftarrow 2$
3: **for** $k = 0$ to d **do**
4:     $r[k,1] \Leftarrow d(x_i, x_j)$
5:     $r[k,2] \Leftarrow i, r[k,3] \Leftarrow j$
6:     **if** $j = n$ **then**
7:        $j \Leftarrow n - i + 2, i \Leftarrow 1$
8:     **else**
9:        $i \Leftarrow i + 1, j \Leftarrow j + 1$
10:    **end if**
11: **end for**
12: sort r by $r[k,1]$ descending
13: $ma[1] \Leftarrow r[1,2]$
14: $j \Leftarrow r[1,3]$
15: **for** $k = 2$ to n **do**
16:    find largest $r[i,1]$ for such i that ($r[i,2] = j$ or $r[i,2] = j$) and $r[i,2] \notin ma$ and $r[i,3] \notin ma$
17:    **if** $r[i,2] = j$ **then**
18:      $j \Leftarrow r[i,3], ma[k] \Leftarrow r[i,2]$
19:    **else**
20:      $j \Leftarrow r[i,2], ma[k] \Leftarrow r[i,3]$
21:    **end if**
22: **end for**
23: **return** ma

$n(n-1)/2$ times per iterational step. We have run several simulations with the above heuristics, but before providing the results, we will move on to the Riemann centroid problem.

IV. THE RIEMANN CENTROID

The Riemann centroid is the solution of the following minimization problem

$$argmin_{X>0} \sum_{i=1}^{n} d(X, X_i)^2. \qquad (9)$$

The above has a unique solution if the sectional curvature of the manifold corresponding to the metric induced by the distance function $d(a,b)$ is non-positive [13]. If we reduce this problem to the manifold of $k \times k$ positive definite matrices, we can define the solution of the minimization problem as the average of the X_i matrices corresponding to the metric induced by the distance function $d(a,b)$. The solution of the above defined problem can be given as the extremal points of the cost function defined as

$$C(X) = \sum_{i=1}^{n} d(X, X_i)^2. \qquad (10)$$

To find the these points we have to compute the gradient of $C(X)$ and solve the equation

$$gradC(X) = 0. \qquad (11)$$

The gradient can be written as

$$gradC(X) = -\sum_{i=1}^{n} \log_X(X_i), \qquad (12)$$

where $\log_p(q)$ denotes the inverse of the exponential map given at p, see [13].

Firstly one can start examining the above problem by taking only two points of the manifold and trying to solve (9) for this two points

$$argmin_{X>0} d(X, X_1)^2 + d(X, X_2)^2. \qquad (13)$$

The solution of the above is the middle point of the distance minimizing geodesic curve $\gamma_{X_1,X_2}(t)$ connecting X_1 and X_2, so if this curve is parametrized in a way that $\gamma_{X_1,X_2}(0) = X_1$ and $\gamma_{X_1,X_2}(1) = X_2$, then $\gamma_{X_1,X_2}(1/2) = argmin_{X>0} d(X, X_1)^2 + d(X, X_2)^2$. Let us suppose that the metric $d(\cdot, \cdot)$ has such properties that if $X_1 < X_2$ then $X_1 < \gamma_{X_1,X_2}(1/2) < X_2$. In this case this minimization problem corresponds with a symmetric mean function $M(A,B)$ defined as

$$M(A,B) = argmin_{X>0} d(X, A)^2 + d(X, B)^2 = \gamma_{A,B}(1/2). \qquad (14)$$

One might arrive at the problem, that if $d(\cdot, \cdot)$ induces an $M(A,B)$ mean function in the above described way, then the extension of the mean function to n-variables can be defined as the solution of the minimization problem (9). The arising question is that whether the solution of the minimization problem, if unique, equals with the iterative extensions given in the second section.

Let us study the case of the geometric mean. The geometric mean for two positive definite matrices is given as

$$G(A,B) = A^{1/2} \left(A^{-1/2} B A^{-1/2} \right)^{1/2} A^{1/2}. \qquad (15)$$

The above matrix mean has been studied extensively, see [2], [16], [17], [10]. This mean is induced by the metric defined on the manifold of positive definite matrices by the inner product at $p \in M$

$$\langle X, Y \rangle_p = tr \left\{ X^* p^{-1} Y p^{-1} \right\}, \qquad (16)$$

where X and Y are elements of the tangent bundle $T_p M$ given at $p \in M$. For this metric the exponential map and its inverse can be written explicitly as

$$\begin{aligned} \exp_p(X) &= p^{1/2} \exp(p^{-1/2} X p^{-1/2}) p^{1/2} \\ \log_p(X) &= p^{1/2} \log(p^{-1/2} X p^{-1/2}) p^{1/2}. \end{aligned} \qquad (17)$$

Note that in this case $gradC(X)$ given in (12) can expressed in a closed form with matrix functions $\exp, \log$ and the square root. Therefor if the initial X_i matrices commute, then (11) can be solved in closed form as

$$X = X_1^{1/n} \cdots X_n^{1/n}. \qquad (18)$$

The above formula is the same as the geometric mean for positive numbers, actually the above minimization problem formulated for numbers yields the same solution, the geometric mean of n numbers.

There is another metric where the solution of (9) can be expressed in closed form for matrices, that is the standard Euclidean metric for matrices with 0 curvature. The solution will be the arithmetic mean. For these two special cases we know that the iterative extension defined before yields

World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:2, No:4, 2008

also the solution of the minimization problem, because the solutions can be expressed in closed form, which correspond with the arithmetic and geometric mean of matrices. But we know that the iterational sequences, if the initial matrices are orderable, converge to the same extension formulas given for the arithmetic and the geometric mean, see [18]. This coincidence motivated the further study of the iterative extensions of mean functions. The simulations provided in the next section will be given for the geometric mean, because it is the only matrix mean function for which it is nontrivial that the iterative sequences' limit is the corresponding solution of the minimization problem, but in certain situations, for example when the initial matrices commute, we know this to be true. Furthermore it is relatively easy to compute the needed differential forms of the manifold induced by the Froebenius inner product (16).

V. THE COMPUTATION OF THE RIEMANN CENTROID AND SIMULATIONS

There are several methods which provide an approximate solution of the minimization problem given in (9). Two of them should be considered here, the first is the gradient descent algorithm, the second is the Newton method. Both methods have much in common, but generally the Newton method is considered to be faster then the gradient method, but it requires the computation of the Hessian of the cost function (10). Both algorithms requires the computation of the gradient of the cost function. Generally expressing the gradient in closed form requires solutions of the Euler-Lagrange equations related to the geodesic equations. To solve the geodesic equations in closed form respect to the metric and the ending points of the geodesic can be extremely difficult.

Taking into account the above mentioned problems, we will only provide a simulation which is related to the metric induced by the inner product (16) given on the tangent bundle $T_q M$. We already know that this metric induces the geometric mean if the matrices commute. The other good properties of this metric are that the exponential map, the logarithmic map and the curvature endomorphism can be relatively easiy found and expressed in closed form, see [10].

We have used the following two algorithms for computing the centroid:

Algorithm 2 Gradient Method

  set tolerance $\epsilon > 0, x_t \Leftarrow x_0$
  $x \Leftarrow x_t$
  repeat
    $g \Leftarrow gradC(x)$
    obtain α_k by applying the Armijo line-search rule
    $x_t \Leftarrow exp_x(-\alpha_k g)$
  until $R(x, x_t) > \epsilon$

We have been comparing these two methods with the ones given in the second section and modified by Algorithm 1, mostly we were interested in smaller number of starting matrices. The simulations showed that the iterative methods

---

Algorithm 3 Newton Method

  set tolerance $\epsilon > 0, x_t \Leftarrow x_0$
  $x \Leftarrow x_t$
  repeat
    $g \Leftarrow gradC(x)$
    Solve $HessC(x) \cdot d = -g$ for d
    if $\langle d, g \rangle_x \geq 0$ **then**
      $d = -g$
    end if
    obtain α_k by applying the Armijo line-search rule
    $x_t \Leftarrow exp_x(-\alpha_k g)$
  until $R(x, x_t) > \epsilon$

---

presented here tend to converge to the approximate solutions found by the gradient and the Newton method.

If one is only interested in the number of iterations needed for reaching a tolerance level, one might find that the Newton method is the best choice next to the gradient. But taking into account the computational costs, quite astonishingly the best choice is the iterational method given in the second section modified by Algorithm 1, although Figure 4,6 would suggest the opposite. In Figure 4, 6 one can compare the distinct methods convergence rate. Here Iteration 1 was defined by Definition II.3 and Iteration 3 was given by Definition II.2 modified by Algorithm 1. Actually in the case of the Newton method, the evaluation of the Hessian was so computationally costly that it cancelled out the rapid convergence of the algorithm. Due to the adaptive properties of both of the gradient and Newton methods, there were certain distributions of the matrices, when the algorithms did not converge at all or the convergence was quite slow.

The iterative processes given in this article were more robust then both methods, the computations were easier to carry out, therefore the overall performance was better than both methods, when there were relatively small number of matrices.

VI. CONCLUSIONS

In this article iterational algorithms were presented, which provide an extension of a two-variable mean function. It has been proven that the iterative methods converge to the same limit point. These methods provide more sophisticated extensions than the others given before, because they omit recursivity and they use the two-variable form directly.

These methods appear to provide the solution of the Riemann barycenter problem, according to the simulations given here in this context. Furthermore these methods provide solutions faster then the original gradient or Newton methods. There are no adaptive parameters in the algorithms, so the speed of convergence can be more precisely predicted.

REFERENCES

[1] E. Ahn, S. Kim and Y. Lim, *An extended Lie-Trotter formula and its applications*, Linear Algebra and its Appl., 427 (2007), pp. 190–196.
[2] T. Ando, C-K. Li and R. Mathias, *Geometric means*, Linear Alg. Appl., 385 (2004), pp. 305–334.
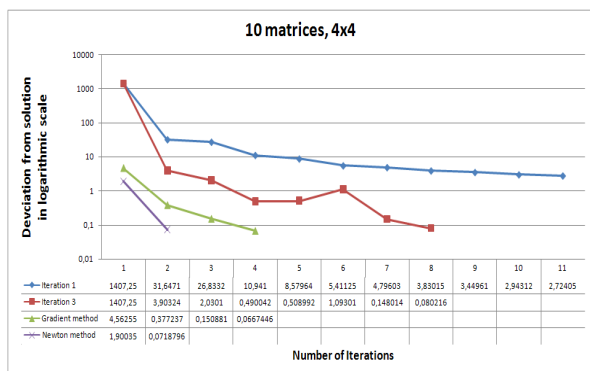[3] W. N. Anderson Jr., *Shorted operators*, SIAM J. Appl. Math., 20 (1971), pp. 520–525.

World Academy of Science, Engineering and Technology
International Journal of Mathematical and Computational Sciences
Vol:2, No:4, 2008

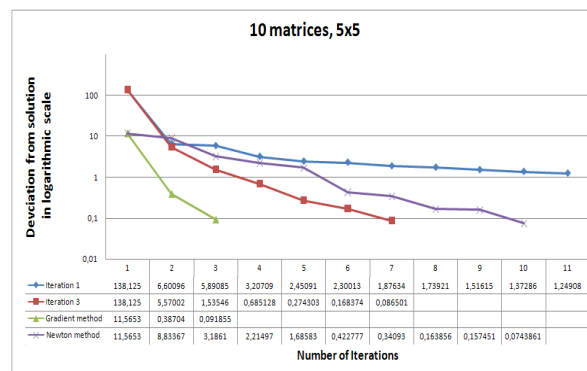Fig. 4.   Convergence rate results for 10, 4×4 matrices.

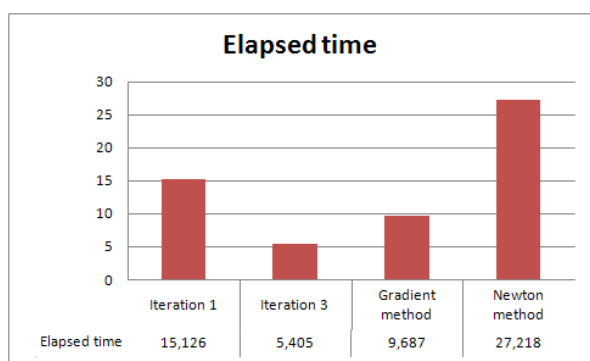Fig. 6.   Convergence rate results for 10, 5×5 matrices.

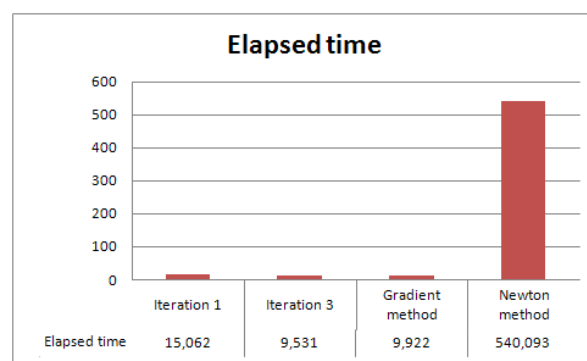Fig. 5.   Execution time results for the methods given in Figure 4.

Fig. 7.   Execution time results for the methods given in Figure 6.

[4] W. N. Jr. Anderson and R. J. Duffin, *Series and parallel addition of matrices*, J. Math. Anal. Appl., 26 (1969), pp. 576-594.

[5] W. N. Jr. Anderson and G. E. Trapp, *Shorted Operators II*, SIAM J. Appl. Math., 28 (1975), pp. 60-71.

[6] J. Antezanaa, G. Corachb and D. Stojanoff, *Bilateral shorted operators and parallel sums*, Linear Alg. Appl., 414 (2006), pp. 570–588.

[7] V. Arsigny, P. Fillard, X. Pennec and N. Ayache, *Geometric Means in a Novel Vector Space Structure on Symmetric Positive-Definite Matrices*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 328–347.

[8] R. Bhatia, *Matrix Analysis*, Springer-Verlag, New York, 1996.

[9] R. Bhatia and J. Holbrook, *Riemannian geometry and matrix geometric means*, Linear Alg. Appl., 413 (2006), pp. 594618.

[10] R. Ferreira and J. Xavier, *Hessian of the Riemannian squared-distance function on connected locally symmetric spaces with applications*, Controlo 2006, 7th Portuguese Conference on Automatic Control , special session on control, optimization and computation, September 2006.

[11] P. Fillard, V. Arsigny, X. Pennec and N. Ayache, *Joint estimation and smoothing of clinical DT-MRI with a Log-Euclidean metric*. Research Report RR-5607, INRIA, Sophia-Antipolis, France, June 2005.

[12] F. Hiai and H. Kosaki, *Means of Hilbert space operators*, Lecture Notes In Maths. 1820 (2003), Springer, .

[13] H. Karcher, *Riemannian center of mass and mollifier smoothing*, Comm. Pure Appl. Math., vol. 30 (1977), pp. 509541.

[14] F. Kubo and T. Ando, *Means of positive linear operators*, Math. Ann., 246 (1980), pp. 205–224.

[15] 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, Kunming, China, December 2004.

[16] M. Moakher, *A Differential Geometric Approach to the Geometric Mean of Symmetric Positive-Definite Matrices*, SIAM J. Matrix Anal. Appl., 26 (2005), pp. 735–747.

[17] M. Moakher, *Means and averaging in the group of rotations*, SIAM J. Matrix Anal. Appl., vol. 24, no. 1 (2002), pp. 116.

[18] M. Pálfia, *Algorithmic definition of means acting on positive numbers and operators*, submitted (arXiv:math/0503510).

[19] M. Pálfia, *Means of unorderable matrices and the Riemann centroid*, submitted to Linear Algebra and its Applications.

[20] D. Petz and R. Temesi, *Means of positive numbers and matrices*, SIAM J. Matrix Anal. Appl., 27 (2006), pp. 712–720.