数理解析研究所講究録
第 1737 巻 2011 年 65-70

STOPPING CRITERIA FOR MULTIVARIABLE GEOMETRIC MEANS

YONGDO LIM

ABSTRACT. We provide an upper bound for the number of iterations necessary to achieve a desired level of accuracy for multivariable geometric means obtained via symmetrization procedures and induction. It is shown that the upper bound for the number of iterations depends only on the diameter of the the set of n matrices and the desired convergence tolerance, and is quite accurate for the Alm mean.

1. INTRODUCTION

The problem of averaging a date of positive definite matrices invariant under the "matrix inversion" (self-duality) is very important and is an active current research area. One of two-variable matrix means of positive definite matrices satisfying the self-duality is the geometric mean $A\#B=A^{-1/2}(A^{-1/2}BA^{-1/2})^{1/2}A^{1/2}$, which appears as a unique positive definite solution of the Riccati equation $XA^{-1}X=B$. Also it appears as the unique Nesterov-Todd primal-dual scaling points of A and B for the logarithmic barrier functional $F(X)=-\log\det(X)$ which has played an important role in their development of a theoretical foundation for efficient primal-dual interior-point algorithms for problems of minimizing linear functionals over the intersection of an affine subspace on the cone [10]. In the Riemannian manifold of positive definite matrices equipped with the Riemannian trace metric (it coincides with the Hessian metric of the logarithmic barrier functional) $ds=||A^{-1/2}dA\,A^{-1/2}||_2=(\mathrm{tr}(A^{-1}dA)^2)^{1/2}$ ([2, 3, 6]), the curve $t\mapsto A\#_tB:=A^{1/2}(A^{-1/2}BA^{-1/2})^tA^{1/2}$ is the unique geodesic line between A and B and its geodesic middle (midpoint) $A\#_{1/2}B$ is the geometric mean of A and B. The geometric mean $A\#B$ is also appeared as a midpoint of the Thompson metric, which is obtained from the Finsler structure of spectral norms at each tangent space [11, 5], given by

$$d(A,B)=\max\{\log M(A/B),\log M(B/A)\}=||\log A^{-1/2}BA^{-1/2}||$$

where $M(A/B) := \inf\{\lambda > 0 : A \leq \lambda B\}$, the largest eigenvalue of $B^{-1/2}AB^{-1/2}$.

Two recent approaches for extending two-variable geometric mean of positive definite matrices to higher order have been given by Ando, Li and Mathias [1] and Bini, Meini and Poloni [4] via "symmetrization methods" and induction. Let A_1, A_2, A_3 be positive definite matrices. Starting with $(A_1^{(0)}, A_2^{(0)}, A_3^{(0)}) = (A_1, A_2, A_3)$ define

$$\begin{aligned}
(A_1^{(1)}, A_2^{(1)}, A_3^{(1)}) &= (A_1\#_t(A_2\#A_3), A_2\#_t(A_1\#tA_3), A_3\#_t(A_1\#A_2)),\\
&\vdots\\
(A_1^{(r+1)}, A_2^{(r+1)}, A_3^{(r+1)}) &= (A_1^{(r)}\#_t(A_2^{(r)}\#A_3^{(r)}), A_2^{(r)}\#_t(A_1^{(r)}\#A_3^{(r)}), A_3^{(r)}\#_t(A_1^{(r)}\#A_2^{(r)})).
\end{aligned}$$

It is shown [4] that the sequences $\{A_i^{(r)}\}_{r=0}^{\infty}, i = 1, 2, 3$, converge to a common limit (depending on $t \in (0, 1]$), yielding a geometric mean of 3-positive definite matrices, denoted by $G(t; A_1, A_2, A_3)$. The ALM and BMP symmetrization procedures are the case $t = 1$ and $t = 2/3$, respectively. Inductively, letting $G = G(t_1, \ldots, t_{n-2}; A_1, \ldots, A_n)$ the symmetrization procedure of $n+1$-positive definite matrices is defined by

$$\beta_t(\mathbb{A}) = (A_1\#_t G(\mathbb{A}_{\neq 1}), A_2\#_t G(\mathbb{A}_{\neq 2}), \ldots, A_{n+1}\#_t G(\mathbb{A}_{\neq n+1})), t = t_{n-1} \in (0, 1]$$

where $\mathbb{A} = (A_1, \ldots, A_{n+1})$ and $\mathbb{A}_{\neq i} = (A_1, \ldots, A_{i-1}, A_{i+1}, \ldots, A_{n+1})$. Then its iteration $\beta_t^r(\mathbb{A}) = (A_1^{(r)}, \ldots, A_n^{(r)})$ approaches to a common limit $\lim_{r\to\infty} A_i^{(r)} = X^*$ for all i, yielding $n+1$-geometric mean $G(t_1, \ldots, t_{n-2}, t_{n-1}; \mathbb{A})$; in particular $\mathrm{Alm}_{n+1}(\mathbb{A})$ with $t_i = 1$ and $\mathrm{Bmp}_{n+1}(\mathbb{A})$ with $t_i = \frac{i}{i+1}$ for all i, respectively. This shows that any sequence $\mathbf{t} = \{t_n\}_{n=3}^{\infty}$ in the interval $(0, 1]$ provides a family of means

$$G_{\mathbf{t}} = \{G_{\mathbf{t},n}\}_{n=2}^{\infty}, \quad G_{\mathbf{t},n} = G(t_3, \ldots, t_n; \cdot).$$

For instances for $\mathbf{1} = (1, 1, \ldots)$ and $\mathbf{s} = (\frac{2}{3}, \frac{3}{4}, \ldots, \frac{n-1}{n}, \ldots)$,

$$G_{\mathbf{1}} = \mathrm{ALM} = \{\mathrm{Alm}_n\}_{n=2}^{\infty}, \quad G_{\mathbf{s}} = \mathrm{BMP} = \{\mathrm{Bmp}_n\}_{n=2}^{\infty}.$$

It turns out that the Bmp_{n+1} mean is much more computationally efficient than Alm_{n+1} mean [4].

The main purpose of this paper is to study a stopping criteria for the symmetrization procedures of $G_{\mathbf{t},n}$ in terms of the Thompson metric. It is shown that our upper bound is quite accurate for the Alm mean and the upper bound decreases as n increases on any bounded region of positive definite matrices.

2. Multivariable geometric means

We have the following functional characterization and properties of $G_{\mathbf{t}}$.

Theorem 2.1. *For a given sequence* $\mathbf{t} = \{t_n\}_{n=3}^{\infty}$ *in* $(0,1]$, *there exists a unique family* $G_{\mathbf{t}} = \{G_{\mathbf{t},n}\}_{n=2}^{\infty}$ *of continuous functions* $G_n : P(k)^n \to P(k)$ *that satisfies*

(i) $G_2(A,B) = A\#B$;

(ii) $G_n(A, A, \dots, A) = A$;

(iii) $G_n(\mathbb{A}) = G_n(A_1 \#_{t_n} G_{n-1}(\mathbb{A}_{\neq 1}), \dots, A_n \#_{t_n} G_{n-1}(\mathbb{A}_{\neq n}))$

for all $n \geq 2$. *Each* $G_{\mathbf{t},n}$ *satisfies the following properties*

(P1) $g(\mathbb{A}) = (A_1 \cdots A_n)^{1/n}$ *if* A_i*'s commute;*

(P2) (*Joint homogeneity*) $g(\mathbf{a} \cdot \mathbb{A}) = (a_1 \cdots a_n)^{1/n} g(\mathbb{A})$;

(P3) (*Permutation invariance*) $g(\mathbb{A}_\sigma) = g(\mathbb{A})$, *where* $\mathbb{A}_\sigma = (A_{\sigma(1)}, \dots, A_{\sigma(n)})$;

(P4) (*Monotonicity*) *If* $B_i \leq A_i$ *for all* $1 \leq i \leq n$, *then* $g(\mathbb{B}) \leq g(\mathbb{A})$;

(P5) (*Continuity*) g *is continuous;*

(P6) (*Congruence invariance*) $g(M^*\mathbb{A}M) = M^* g(\mathbb{A}) M$ *for invertible matrix* M;

(P7) (*Joint concavity*) $g(\lambda\mathbb{A} + (1-\lambda)\mathbb{B}) \geq \lambda g(\mathbb{A}) + (1-\lambda) g(\mathbb{B})$ *for* $0 \leq \lambda \leq 1$;

(P8) (*Self-duality*) $g(\mathbb{A}^{-1})^{-1} = g(\mathbb{A})$;

(P9) (*Determinantal identity*) $\mathrm{Det} g(\mathbb{A}) = \prod_{i=1}^{n} (\mathrm{Det} A_i)^{1/n}$; *and*

(P10) (*AGH mean inequalities*) $n(\sum_{i=1}^{n} A_i^{-1})^{-1} \leq g(\mathbb{A}) \leq \frac{1}{n}\sum_{i=1}^{n} A_i$.

Remark 2.2. The property (*ii*) can be replaced by the monotonicity (cf. [1]).

Remark 2.3. In [1] and [4], the authors actually obtained a stronger version of ($P5$) for the Thompson metric; it is shown that ($g = \mathrm{Alm}_n, \mathrm{Bmp}_n$)

$$R(g(A_1, \dots, A_n), g(B_1, \dots, B_n)) \leq \prod_{i=1}^{n} R(A_i, B_i)^{1/n},$$

where $R(A,B) = \max(\rho(A^{-1}B), \rho(B^{-1}A))$ and $\rho(X)$ denotes the spectral radius of X. In terms of the Thompson metric, we can rewrite it as follows:

(P11) $d(g(A_1, \dots, A_n), g(B_1, \dots, B_n)) \leq \frac{1}{n}\sum_{i=1}^{n} d(A_i, B_i)$.

One can see that this holds true for $G_{\mathbf{t},n}$.

Remark 2.4. The $(n+1)$-geometric mean $G_{\mathbf{t},n+1}(\mathbb{A})$ appears as the common limit of the symmetrization $\beta : \mathrm{P}(k)^{n+1} \to \mathrm{P}(k)^{n+1}$ by

$$\beta(\mathbb{A}) := \big(A_1 \#_{t_{n+1}} G_{\mathbf{t},n}(\mathbb{A}_{\neq 1}), \dots, A_{n+1} \#_{t_{n+1}} G_{\mathbf{t},n}(\mathbb{A}_{\neq n+1})\big).$$

For $r \in \mathbb{N}_0$, we let $\beta^r(\mathbb{A}) = (A_1^{(r)}, \dots, A_{n+1}^{(r)}) \in \mathrm{P}(k)^{n+1}$. Then $A_i^{(0)} = A_i$ and

$$A_i^{(r+1)} = A_i^{(r)} \#_{t_{n+1}} G_{\mathbf{t},n}(\mathbb{A}_{\neq i})$$

for all i.

Set

$$\mu_n := 1 - \frac{(n-1)t_{n+1}}{n}, \quad \nu_n := 1 - \frac{(n-1)t_{n+1}}{n^2}.$$

Then $0 < \mu_n < \nu_n < 1$ for all $n \geq 2$.

Using the non-positive curvature property of the Thompson metric [5]

$$d(A\#_t B, C\#_t D) \leq (1-t)d(A,C) + td(B,D), \quad t \in [0,1],$$

Theorem 2.5. *We have*

$$\Delta(\beta^r(\mathbb{A})) \leq \mu_n^r \Delta(\mathbb{A}), \tag{2.1}$$

$$d(A_i^{(r)}, A_i^{(r+1)}) \leq t_{n+1}\mu_n^r \Delta(\mathbb{A}), \tag{2.2}$$

where $\Delta(\mathbb{A}) = \max\{d(A_i, A_j) : 1 \leq i, j \leq n+1\}$.

Given $\mathbb{A} = (A_1, \dots, A_{n+1})$ and a convergence tolerance $\epsilon > 0$, consider

$$N_{\mathbf{t}}(\epsilon) := \min\{r : d(A_1^{(r)}, G_{\mathbf{t},n+1}(\mathbb{A})) \leq \epsilon\}.$$

Our main result is the following.

Theorem 2.6. *We have*

$$N_{\mathbf{t}}(\epsilon) \leq \left\lfloor \frac{\log \frac{n}{n-1} + \log \frac{\Delta(\mathbb{A})}{\epsilon}}{\log \frac{n}{n(1-t_{n+1})+t_{n+1}}} \right\rfloor + 1.$$

Proof. By Theorem 2.5 and by the triangular inequality,

$$d(A_1^{(r)}, G_{\mathbf{t},n+1}(\mathbb{A})) \leq \frac{n}{n-1}\left(1 - \frac{(n-1)t_{n+1}}{n}\right)^r \Delta(\mathbb{A}).$$

Solving the inequality RHS $\leq \epsilon$ for r, we have

$$N_{\mathbf{t}}(\epsilon) \leq \left\lfloor \frac{\log \frac{\epsilon(n-1)}{n\Delta(\mathbb{A})}}{\log \frac{n(1-t_{n+1})+t_{n+1}}{n}} \right\rfloor + 1 = \left\lfloor \frac{\log \frac{n\Delta(\mathbb{A})}{\epsilon(n-1)}}{\log \frac{n}{n(1-t_{n+1})+t_{n+1}}} \right\rfloor + 1.$$

□

Remark 2.7. We note that the upper bounds of $N_{\mathbf{t}}(\epsilon)$ is decreasing for the variable $t_{n+1} \in (0,1]$ and also decrease as n increases when $\mathbb{A}$ varies over a bounded subset of positive definite matrices.

Example 2.8. Number of iterations needed from our upper bound is reported in Table 1 as $t_{n+1} = 1, \frac{n}{n+1}, \frac{2}{3}, \frac{1}{2}$, with $\epsilon = 10^{-10}$ and $\Delta(\mathbb{A}) = 2\log 500$.

TABLE 1

t_{n+1}	Iter.($n=3$)	Iter.($n=5$)	Iter.($n=10$)	Iter.($n=10^2$)	Iter.($n=10^3;10^5$)
1	24	17	12	6	4(3)
$\frac{n}{n+1}$	38	24	16	7	5(3)
$\frac{2}{3}$	45	34	28	24	24(24)
$\frac{1}{2}$	65	51	43	38	37(37)

Example 2.9. For $n = 3, 5$, we consider $G_{\mathbf{1},n+1}(A_1, \dots, A_{n+1})$ and $G_{\mathbf{s},n+1}(A_1, \dots, A_{n+1})$ where

$$A_1 = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 3 & -2 \\ 0 & -2 & 2 \end{pmatrix}, \quad A_2 = \begin{pmatrix} 2 & 1 & 0 \\ 1 & 3 & 2 \\ 0 & 2 & 2 \end{pmatrix}, \quad A_3 = \begin{pmatrix} 1 & 0 & 1 \\ 0 & 10 & 0 \\ 1 & 0 & 50 \end{pmatrix},$$

$$A_4 = \begin{pmatrix} 1 & 0 & -1 \\ 0 & 10 & 0 \\ -1 & 0 & 50 \end{pmatrix}, \quad A_5 = \begin{pmatrix} 4 & -1 & 3 \\ -1 & 9 & 4 \\ 3 & 4 & 9 \end{pmatrix}, \quad A_6 = \begin{pmatrix} 9 & -6 & 0 \\ -6 & 9 & 2 \\ 0 & 2 & 8 \end{pmatrix}.$$

We note that $\Delta(A_1, \dots, A_6) \leq 5$ for the Thompson metric. As a stopping criterion for each computed mean, we chose

$$d(A_i^{(r+1)}, A_i^{(r)}) \leq 2\epsilon, \tag{2.3}$$

with $\epsilon = 10^{-10}$. In Table 3, we report the number of iterations needed. The result shows that our upper bound is quite accurate for the Alm mean, but rather poor for the Bmp mean.

TABLE 2

n	$g = \mathrm{Alm}_n$	$g = \mathrm{Bmp}_n$	$N_1(\epsilon)$	$N_s(\epsilon)$
3	$23(t=1)$	$4(t=\frac{3}{4})$	23	37
5	$16(t=1)$	$4(t=\frac{5}{6})$	16	23

Remark 2.10. A general method of finding the number of iteration for multivariable geometric means from non-symmetric and non-expansive means has recently been developed in [9].

REFERENCES

[1] T. Ando, C. K. Li and R. Mathias, Geometric means, Linear Algebra Appl. **385** (2004), 305–334.

[2] R. Bhatia, On the exponential metric increasing property, Linear Algebra Appl. **375** (2003), 211-220.

[3] R. Bhatia, *Positive definite matrices*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ, 2007.

[4] D. Bini, B. Meini and F. Poloni, An effective matrix geometric mean satisfying the Ando-Li-Mathias properties, Math. Comp. **79** (2010), 437-452.

[5] G. Corach, H. Porta and L. Recht, Convexity of the geodesic distance on spaces of positive operators, Illinois J. Math. **38**(1994), 87-94.

[6] J. D. Lawson and Y. Lim, The geometric mean, matrices, metrics, and more, Amer. Math. Monthly **108** (2001), 797-812.

[7] J. D. Lawson and Y. Lim, Symmetric spaces with convex metrics, Forum Math. **19** (2007), 571-602.

[8] J. D. Lawson and Y. Lim, A general framework for extending means to higher orders, Colloq. Math. **113** (2008), 191-221.

[9] Y. Lim, Stopping criteria for the Ando-Li-Mathias and Bini-Meini-Poloni geometric means, Linear Algebra Appl., to appear.

[10] Y. E. Nesterov and M. J. Todd, Self-scaled barriers and interior-point methods for convex programming, Mathematics of Operations Research **22** (1997), 1-42.

[11] A. C. Thompson, On certain contraction mappings in a partially ordered vector space, Proc. Amer. Math. Soc. **14** (1963), 438-443.

DEPARTMENT OF MATHEMATICS, KYUNGPOOK NATIONAL UNIVERSITY, TAEGU 702-701, KOREA

E-mail address: ylim@knu.ac.kr