Hindawi Publishing Corporation
Journal of Applied Mathematics
Volume 2014, Article ID 507175, 7 pages
http://dx.doi.org/10.1155/2014/507175

*Research Article*

# Riemannian Gradient Algorithm for the Numerical Solution of Linear Matrix Equations

**Xiaomin Duan,[1,2] Huafei Sun,[1] and Xinyu Zhao[3,4]**

[1] *School of Mathematics, Beijing Institute of Technology, Beijing 100081, China*
[2] *School of Science, Dalian Jiaotong University, Dalian 116028, China*
[3] *School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China*
[4] *School of Materials Science and Engineering, Dalian Jiaotong University, Dalian 116028, China*

Correspondence should be addressed to Huafei Sun; huafeisun@bit.edu.cn

Received 6 August 2013; Revised 10 December 2013; Accepted 10 December 2013; Published 6 January 2014

Academic Editor: Zhi-Hong Guan

Copyright © 2014 Xiaomin Duan et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

A Riemannian gradient algorithm based on geometric structures of a manifold consisting of all positive definite matrices is proposed to calculate the numerical solution of the linear matrix equation $Q = X + \sum_{i=1}^{m} A_i^T X A_i$. In this algorithm, the geodesic distance on the curved Riemannian manifold is taken as an objective function and the geodesic curve is treated as the convergence path. Also the optimal variable step sizes corresponding to the minimum value of the objective function are provided in order to improve the convergence speed. Furthermore, the convergence speed of the Riemannian gradient algorithm is compared with that of the traditional conjugate gradient method in two simulation examples. It is found that the convergence speed of the provided algorithm is faster than that of the conjugate gradient method.

## 1. Introduction

The linear matrix equation

$$Q = X + \sum_{i=1}^{m} A_i^T X A_i, \tag{1}$$

where $A_1, A_2, \ldots, A_m$ are arbitrary $n \times n$ real matrices, $m$ is a nonnegative integer, and $A_i^T$ denotes the transpose of the matrix $A_i$, arises from many fields, such as the control theory, the dynamic programming, and the stochastic filtering [1–4]. In the past decades, there has been increasing interest in the solution problems of this equation. In the case of $m = 1$, some numerical methods, such as Bartels-Stewart method, Hessenberg-Schur method, and Schur and QR decompositions method, were proposed in [5, 6]. Based on the Kronecker product and the fixed point theorem in partially ordered sets, some sufficient conditions for the existence of a unique symmetric positive definite solution are given in [7, 8]. Ran and Reurings ([7, Theorem 3.3] and [9, Theorem 3.1]) pointed out that if $Q - \sum_{i=1}^{m} A_i^T Q A_i$ is a positive definite matrix, then there exists a unique solution and it is symmetric positive definite. Recently, under the condition that (1) is consistent, Su and Chen presented an efficient numerical iterative method based on the conjugate gradient method (CGM) [10].

In addition, based on geometric structures on a Riemannian manifold, Duan et al. proposed a natural gradient descent algorithm to solve algebraic Lyapunov equations [11, 12]. Following them, we investigate the solution problem of (1) in the view of Riemannian manifolds. Note that this solution of (1) is a symmetric positive definite matrix and the set of all the symmetric positive definite matrices can be considered as a manifold. Thus, it is more convenient to investigate the solution problem with the help of these geometric structures on this manifold. To address such a need, a new framework is presented in this paper to calculate the numerical solution, which is based on the geometric structures on the Riemannian manifold of positive definite symmetric matrices. We will first describe briefly some

fundamental knowledge on manifold. Then a Riemannian gradient algorithm is proposed, where the geodesic distance on the curved Riemannian manifold is taken as an objective function and the geodesic curve is treated as the convergence path. Also the optimal variable step sizes corresponding to the minimum value of the objective function are provided in order to improve the convergence speed. Finally, the behavior of the provided algorithm and the traditional CGM is compared and demonstrated in two simulation examples.

## 2. A Survey of Some Geometrical Concepts

In the present section, we briefly survey the geometry of smooth manifolds by recalling concepts as tangent spaces, the Riemannian gradient, geodesic curves, and geodesic distance. More details can be found in [13, 14]. Specially, these concepts of manifold consisting of all symmetric positive definite matrices have also been introduced in [15–20].

*2.1. Riemannian Manifolds.* Let us denote a Riemannian manifold by $\mathbb{M}$. Its tangent space at point $Y \in \mathbb{M}$ is denoted by $T_Y\mathbb{M}$. Given any pair of points $V, W \in T_Y\mathbb{M}$, an inner product $\langle V, W\rangle_Y \in \mathbb{R}$ is defined. The specification of the inner product for a Riemannian manifold turns it into a metric space. In fact, the length of a curve $\phi_{Y,V} : [0,1] \to \mathbb{M}$, such that $\phi_{Y,V}(0) = Y \in \mathbb{M}$ and $\dot{\phi}_{Y,V}(0) = V \in T_Y\mathbb{M}$, is given by

$$\ell(\phi_{Y,V}) := \int_0^1 \sqrt{\left\langle \dot{\phi}_{Y,V}(t), \dot{\phi}_{Y,V}(t)\right\rangle_{\phi_{Y,V}(t)}}\,\mathrm{d}t. \tag{2}$$

Given arbitrary $Y \in \mathbb{M}$ and $V \in T_Y\mathbb{M}$, the curve $\gamma_{Y,V} : [0,1] \to \mathbb{M}$ of shortest length is called geodesic curve. Such minimal length $\ell(\gamma_{Y,V})$ is called geodesic distance between two points, namely, $\gamma_{Y,V}(0) = Y$ and $\gamma_{Y,V}(1)$. The Riemannian distance between two points is denoted by

$$\mathrm{d}(\gamma_{Y,V}(0), \gamma_{Y,V}(1)) := \ell(\gamma_{Y,V}). \tag{3}$$

It is obvious that if any pair of points on manifold $\mathbb{M}$ can be connected by a geodesic curve, then it is possible to measure the distance between any given pair of points on it. Given a regular function $f : \mathbb{M} \to \mathbb{R}$, its Riemannian gradient $\nabla_Y f$ in the direction of the vector $V \in T_Y\mathbb{M}$ measures the rate of change of the function $f$ in the direction $V$. Namely, given any smooth curve $\phi_{Y,V} : [0,1] \to \mathbb{M}$, such that $\phi_{Y,V}(0) = Y \in \mathbb{M}$ and $\dot{\phi}_{Y,V}(0) = V$, the Riemannian gradient $\nabla_Y f$ is the unique vector in $T_Y\mathbb{M}$ such that

$$\langle V, \nabla_Y f\rangle_Y = \frac{\mathrm{d}}{\mathrm{d}t} f(\phi_{Y,V}(t))\big|_{t=0}. \tag{4}$$

Instead of the Euclidean space, by using the Riemannian gradient the optimization method may be readily extended to smooth manifolds [21, Chapter 7]. For this purpose, let us consider the differential equation on manifold $\mathbb{M}$

$$\dot{y} = -\nabla_{y(t)} f, \qquad y(0) = \overline{y} \in \mathbb{M}. \tag{5}$$

If the function $f : \mathbb{M} \to \mathbb{R}$ is bounded and the smooth manifold $\mathbb{M}$ is compact, then the solution of such differential equation tends to a local minimum of the function $f$ in $\mathbb{M}$, depending on the initial point $\overline{y}$. In fact, it is achieved by definition of the Riemannian gradient

$$\frac{\mathrm{d}}{\mathrm{d}t} f(y(t)) = \left\langle \dot{y}(t), \nabla_{y(t)} f\right\rangle_{y(t)} = -\left\langle \nabla_{y(t)} f, \nabla_{y(t)} f\right\rangle_{y(t)} \le 0. \tag{6}$$

*2.2. Manifold of Symmetric Positive Definite Matrices.* Let us denote the general linear group by $\mathbb{GL}(n)$, which consists of all nonsingular real matrices. And $X > 0$ means that $X$ is a symmetric positive definite matrix. For a different notation $X - Y > 0$, we will use $X > Y$. And the manifold consisting of all symmetric positive definite matrices is defined by $\mathbb{S}^+(n) = \{X \in \mathbb{R}^{n\times n} \mid X^T = X, X > 0\}$. The tangent space at a point $X \in \mathbb{S}^+(n)$ is given by $T_X\mathbb{S}^+(n) = \{V \in \mathbb{R}^{n\times n} \mid V^T = V\}$. On manifold $\mathbb{S}^+(n)$, the Riemannian metric at the point $X$ is defined by

$$\langle V_1, V_2\rangle_X = \mathrm{tr}\left[X^{-1}V_1X^{-1}V_2\right], \tag{7}$$

where $V_1, V_2 \in T_X\mathbb{S}^+(n)$ and tr denotes the trace of the matrix. Then, manifold $\mathbb{S}^+(n)$ with the Riemannian metric (7) becomes a Riemannian manifold. Moreover, it is a curved manifold since this metric is invariant under the group action of $\mathbb{GL}(n)$. The geodesic curve at the point $X \in \mathbb{S}^+(n)$ in the direction $V \in T_X\mathbb{S}^+(n)$ can be expressed as

$$\gamma_{X,V}(t) = X^{1/2}\exp\left(tX^{-1/2}VX^{-1/2}\right)X^{1/2}. \tag{8}$$

Obviously, this geodesic curve is entirely contained in manifold. The Hopf-Rinow theorem implies that manifold $\mathbb{S}^+(n)$ with the Riemannian metric (7) is geodesically complete ([22, Theorem 2.8]). This means, for any given pair $X, Y \in \mathbb{S}^+(n)$, we can find a geodesic curve $\gamma_{X,V}(t)$, such that $\gamma_{X,V}(0) = X$ and $\gamma_{X,V}(1) = Y$, by taking the initial velocity $\dot{\gamma}_{X,V}(0) = X^{1/2}\log(X^{-1/2}YX^{-1/2})X^{1/2} \in T_X\mathbb{S}^+(n)$. According to (2), the geodesic distance $\mathrm{d}(X, Y)$ can be computed explicitly by

$$\mathrm{d}^2(X, Y) = \mathrm{tr}\left[\log^2\left(X^{-1/2}YX^{-1/2}\right)\right]. \tag{9}$$

## 3. Riemannian Gradient Algorithm

Some researches show that the solution equation (1) is unique positive definite if $A_i$ and $Q$ satisfy ([7, Theorem 3.3] and [9, Theorem 3.1])

$$Q > \sum_{i=1}^{m} A_i^T Q A_i. \tag{10}$$

For convenience, we set $\mathscr{L}(X) := X + \sum_{i=1}^{m} A_i^T X A_i$. And it is obvious that $\mathscr{L}(X)$ is positive definite for any $X \in \mathbb{S}^+(n)$. Our purpose is to seek a matrix $X \in \mathbb{S}^+(n)$ so that $\mathscr{L}(X)$ is as close as possible to the given matrix $Q$. We figure out some key points in designing such an algorithm as follows.

(1) For some $X$, the difference between the given positive definite matrix $Q$ and the matrix $\mathscr{L}(X)$ should be determined and calculated firstly.

(2) Then the Riemannian gradient descent algorithm can be used to adjust $X$ so that the difference will be as small as possible.

Note that the Riemannian manifold $\mathbb{S}^+(n)$ is geodesically complete; hence we take the geodesic distance between $Q$ and $\mathscr{L}(X)$ as the objective function $J(X)$; namely,

$$J(X) = \mathrm{d}^2(Q, \mathscr{L}(X)). \tag{11}$$

Then the exact solution of (1) may be defined as

$$X_* = \arg\min_{X\in\mathbb{S}^+(n)} J(X). \tag{12}$$

In order to find a minimizer for the criterion $\mathrm{d}^2(Q, \mathscr{L}(X))$, we may make use of the differential equation (5). The minimum is apparently achieved for a value $X_*$ such that

$$\nabla_X J(X)\big|_{X=X_*} = 0. \tag{13}$$

It is necessary to solve the optimization problem (12) by a numerical optimization algorithm that takes the geometry of the Riemannian manifold $\mathbb{S}^+(n)$ into account. In particular, starting from an initial guess $X_0$, it is possible to provide an iterative learning algorithm that generates a pattern $X_t \in \mathbb{S}^+(n)$ at any learning step $t \in \mathbb{N}$. Following [21, 23], such sequence may be generated by moving from each point $X_t \in \mathbb{S}^+(n)$ to the next point $X_{t+1}$ along a short geodesic curve in the opposite direction of the Riemannian gradient of the objective function $J(X_t)$. Namely, if we set

$$V_t = \nabla_{X_t} J(X_t), \tag{14}$$

then the Riemannian gradient algorithm may be expressed as

$$X_{t+1} = \gamma_{X_t,-V_t}(\eta), \tag{15}$$

where $\eta$ denotes any suitable step-size schedule that drives the iterative algorithm (15) to be convergent. Moreover, with the aim to employ the learning algorithm (14) and (15), we first need to compute the Riemannian gradient of the objective function $J(X_t)$ to be optimized. In the definition of the Riemannian gradient (4), the generic smooth curve $\phi_{Y,V}$ may be replaced with a geodesic curve. We obtain the following theorem.

**Theorem 1.** *The Riemannian gradient of the objective function $J(X)$ at point $X$ is given by*

$$\nabla_X J(X) = 2X\left(\log\left(Q^{-1}\mathscr{L}(X)\right)\mathscr{L}(X)^{-1} + \sum_{j=1}^{m} A_j \log\left(Q^{-1}\mathscr{L}(X)\right)\mathscr{L}(X)^{-1}A_j^T\right)X. \tag{16}$$

*Proof.* First, we compute the derivative of the real-valued function

$$J(s(t)) = \mathrm{tr}\left[\log^2\left(Q^{-1/2}\mathscr{L}(s(t))Q^{-1/2}\right)\right] \tag{17}$$

with respect to $t$, where $s(t) = X^{1/2}\exp(tX^{-1/2}VX^{-1/2})X^{1/2}$ is the geodesic curve emanating from $X$ in the direction $V = \dot{s}(0)$. Using Proposition 2.1 in [16] and some properties of the trace, we have

$$\begin{aligned}
&\frac{\mathrm{d}}{\mathrm{d}t}J(s(t))\big|_{t=0} \\
&= 2\,\mathrm{tr}\left[\left(\log\left(Q^{-1}\mathscr{L}(X)\right)\mathscr{L}(X)^{-1} + \sum_{j=1}^{m} A_j \log\left(Q^{-1}\mathscr{L}(X)\right)\mathscr{L}(X)^{-1}A_j^T\right)V\right] \\
&= \left\langle V, 2X\left(\log\left(Q^{-1}\mathscr{L}(X)\right)\mathscr{L}(X)^{-1} + \sum_{j=1}^{m} A_j \log\left(Q^{-1}\mathscr{L}(X)\right)\mathscr{L}(X)^{-1}A_j^T\right)X\right\rangle_X.
\end{aligned} \tag{18}$$

From (4), it is shown that formula (16) is valid. □

*Remark 2.* In fact, note that $\log(A^{-1}BA) = A^{-1}(\log B)A$, for all $A \in \mathbb{GL}(n)$ and for all $B \in \mathbb{S}^+(n)$; then $\nabla_X J(X)$ can be rewritten as

$$\begin{aligned}
\nabla_X J(X) &= 2X\left(\mathscr{L}(X)^{-1/2}\log\left(\mathscr{L}(X)^{1/2}Q^{-1}\mathscr{L}(X)^{1/2}\right)\mathscr{L}(X)^{-1/2} \right. \\
&\quad + \sum_{j=1}^{m} A_j \mathscr{L}(X)^{-1/2} \\
&\quad \left. \times \log\left(\mathscr{L}(X)^{1/2}Q^{-1}\mathscr{L}(X)^{1/2}\right)\mathscr{L}(X)^{-1/2}A_j^T\right)X.
\end{aligned} \tag{19}$$

Therefore it can be shown that $\nabla_X J(X)$ belongs to $T_X\mathbb{S}^+(n)$.

**Corollary 3.** *The Riemannian gradient $\nabla_X J(X)$ has the unique zero point $X_*$. Furthermore, for the suitable step-size $\eta$, the sequence $X_t$ generated by* (15) *converges to the solution $X_*$ of* (1).

*Proof.* Obviously, $X_*$ is a zero point of the Riemannian gradient $\nabla_X J(X)$. If there exists $X'_* \neq X_*$ such that the iterative process (15) stops, then it means $\nabla_{X'_*} J(X'_*) = 0$, which is equivalent to

$$\log\left(Q^{-1}\mathscr{L}\left(X'_*\right)\right)\mathscr{L}\left(X'_*\right)^{-1} + \sum_{j=1}^{m} A_j \log\left(Q^{-1}\mathscr{L}\left(X'_*\right)\right)\mathscr{L}\left(X'_*\right)^{-1}A_j^T = 0. \tag{20}$$

On (20), the matrices are multiplied by $X_*'$ and the traces of the matrices are taken; then we have

$$\operatorname{tr}\left[\log\left(Q^{-1}\mathscr{L}\left(X_*'\right)\right)\right] = 0. \tag{21}$$

Also, if we replace $X_*'$ with $X_*$ and follow the similar procedure as the above equation, then the following equalities are valid:

$$\begin{aligned}
&\operatorname{tr}\left[\log\left(Q^{-1}\mathscr{L}\left(X_*'\right)\right)\mathscr{L}\left(X_*'\right)^{-1}\left(X_* + \sum_{i=1}^{m} A_i^T X_* A_i\right)\right] \\
&\quad = \operatorname{tr}\left[\log\left(Q^{-1}\mathscr{L}\left(X_*'\right)\right)\mathscr{L}\left(X_*'\right)^{-1} Q\right] \\
&\quad = 0.
\end{aligned} \tag{22}$$

Combining (21) and (22), it can be obtained that

$$\begin{aligned}
&\operatorname{tr}\left[\log\left(Q^{-1}\mathscr{L}\left(X_*'\right)\right)\right] - \operatorname{tr}\left[\log\left(Q^{-1}\mathscr{L}\left(X_*'\right)\right)Q\right] \\
&\quad = \operatorname{tr}\left[\log\left(Q^{-1}\mathscr{L}\left(X_*'\right)\right)\left(I - \mathscr{L}\left(X_*'\right)^{-1}Q\right)\right] \\
&\quad = \operatorname{tr}\left[\log\left(Q^{-1/2}\mathscr{L}\left(X_*'\right)Q^{-1/2}\right)\right. \\
&\qquad \left.\times\left(I - Q^{1/2}\mathscr{L}\left(X_*'\right)^{-1}Q^{1/2}\right)\right] \\
&\quad = 0.
\end{aligned} \tag{23}$$

Note that the matrix $Q^{1/2}\mathscr{L}(X_*')^{-1}Q^{1/2}$ is positive definite; hence its orthogonal similarity diagonalization can be expressed as $U\Lambda U^T$, where $U$ is an orthogonal matrix, $\Lambda = \operatorname{diag}(\lambda_1, \ldots, \lambda_n)$, and $\lambda_i > 0$ $(1 \le i \le n)$. Thus, we have

$$\operatorname{tr}\left[(\Lambda - I)\log\Lambda\right] = \sum_{i=1}^{n}(\lambda_i - 1)\log\lambda^i = 0. \tag{24}$$

Note that the second equality of (24) is equivalent to

$$\prod_{i=1}^{n}\lambda_i^{\lambda_i - 1} = 1. \tag{25}$$

Thus we have $\lambda_1 = \lambda_2 = \cdots = \lambda_n = 1$. That means that $X_*'$ is also the solution of (1), which is contrary to the uniqueness of the solution. Therefore, the Riemannian gradient $\nabla_X J(X)$ has the unique zero point $X_*$. This completes the proof of Corollary 3. □

Now, the Riemannian gradient algorithm with a constant step-size (RGACS) becomes

$$X_{t+1} = X_t^{1/2}\exp\left(-\eta X_t^{-1/2}\nabla_{X_t}J\left(X_t\right)X_t^{-1/2}\right)X_t^{1/2}, \tag{26}$$

where the initial value $X_0 \in \mathbb{S}^+(n)$. The step-size $\eta$ corresponding to the best convergence speed can be obtained experimentally and the details will be described in our following simulations.

*3.1. Riemannian Gradient Algorithm with Variable Step Sizes.* In order to improve the convergence speed of the RGACS, an optimal variable step-size schedule $\eta_t$ will be considered and defined as follows. The step-size $\eta_t$ should be evaluated in such a way that the objective function $J(X_{t+1})$ at step $t+1$ is reduced as much as possible with respect to $J(X_t)$. We may regard $J(X_{t+1})$ as a function of the step-size $\eta_t$ that

$$J\left(X_{t+1}\right) = \mathrm{d}^2\left(Q, \mathscr{L}\left(\gamma_{X_t,-V_t}\left(\eta_t\right)\right)\right). \tag{27}$$

Thus, we can optimize the value of $\eta_t$ to ensure that the difference $J(X_t) - J(X_{t+1})$ could be as large as possible. Since it is difficult to solve this nonlinear problem exactly, we use a suboptimal approximation in practice. Under the hypothesis that the step-size value $\eta_t$ is small enough, we may invoke the expansion of the function $J(X_{t+1})$ around the point $\eta_t = 0$ as follows:

$$J\left(X_{t+1}\right) \approx J\left(X_t\right) + C_{1,t}\eta_t + \frac{1}{2}C_{2,t}\eta_t^2. \tag{28}$$

Then the optimal step-size $\eta_t^*$, corresponding to the maximum of the difference $J(X_t) - J(X_{t+1})$, can be expressed as

$$\eta_t^* \approx -\frac{C_{1,t}}{C_{2,t}}. \tag{29}$$

The coefficients $C_{1,t}$, and $C_{2,t}$ can be calculated using the first-order and the second-order derivatives of the function $J(X_{t+1})$ with respect to the parameter $\eta_t$ at the point $\eta_t = 0$. On the basis of the above discussion, we may obtain the optimal step-size

$$\eta_t^* \approx \frac{\operatorname{tr}\left[\left(X\mathscr{L}\left(\log\left(Q^{-1}\mathscr{L}\left(X_t\right)\right)\mathscr{L}\left(X_t\right)^{-1}\right)\right)^2\right]}{\operatorname{tr}\left[\left(\mathscr{L}\left(X_t\right)^{-1}\mathscr{L}\left(\nabla_{X_t}J\left(X_t\right)\right)\right)^2\right]}. \tag{30}$$

Then, the Riemannian gradient algorithm with variable step-sizes (RGAVS) can be written explicitly as

$$X_{t+1} = X_t^{1/2}\exp\left(-\eta_t^* X_t^{-1/2}\nabla_{X_t}J\left(X_t\right)X_t^{-1/2}\right)X_t^{1/2}, \tag{31}$$

where $X_0 \in \mathbb{S}^+(n)$ is the initial value.

*3.2. Volume Feature of Riemannian Gradient Algorithm.* The RGACS shows an interesting volume feature that the ratio $\det(X_{t+1})/\det(X_t)$ always keeps a certain relationship. In order to prove the volume feature, let us recall two properties of determinant operator det; namely, for each $A, B \in \mathbb{GL}(n)$, there hold

$$\begin{aligned}
&\det(AB) = \det(A)\det(B), \\
&\det\left(\exp(A)\right) = \exp\left(\operatorname{tr}(A)\right).
\end{aligned} \tag{32}$$

By applying the determinant operator to both sides of (31), we obtain

$$\begin{aligned}
&\det\left(X_{t+1}\right)\\
&\quad=\det\left(X_t\right)\exp\left(\operatorname{tr}\left[-\eta X_t^{-1/2}\nabla_{X_t}J\left(X_t\right)X_t^{-1/2}\right]\right)\\
&\quad=\det\left(X_t\right)\exp\left(\operatorname{tr}\left[\eta\left(\log\left(\mathscr{L}(X_t)^{-1}Q\right)\mathscr{L}(X_t)^{-1}X_t\right.\right.\right.\\
&\qquad\qquad+\log\left(\mathscr{L}(X_t)^{-1}Q\right)\mathscr{L}(X_t)^{-1}\\
&\qquad\qquad\left.\left.\left.\times\sum_{j=1}^{m}A_j^TX_tA_j\right)\right]\right)\\
&\quad=\det\left(X_t\right)\exp\left(\operatorname{tr}\left[\eta\log\left(\mathscr{L}(X_t)^{-1}Q\right)\right]\right)\\
&\quad=\det\left(X_t\right)\det\left(\left(\mathscr{L}(X_t)^{-1}Q\right)^{\eta}\right).
\end{aligned}\tag{33}$$

Therefore, the following property is valid:

$$\frac{\det\left(X_{t+1}\right)}{\det\left(X_t\right)}=\left(\frac{\det(Q)}{\det(\mathscr{L}\left(X_t\right)}\right)^{\eta}.\tag{34}$$

This equation shows that the determinant of (1) is closely linked with that of the sequence $X_t$. Moreover, the ratio $\det(X_{t+1})/\det(X_t)$ is directly proportional to $\det(Q)/\det(\mathscr{L}(X_t))$ for the fixed step-size $\eta$. In the case of the RGAVS, the similar conclusion can also be obtained.

## 4. Simulations

Two simulation examples are given to compare the convergence speed of the Riemannian gradient algorithm (the RGACS and the RGAVS) with those of the traditional CGM. The error criterion used in these simulations is $\rho(\mathscr{L}(X)-Q)<10^{-15}$, where $\rho(\cdot)$ denotes the spectral radius of the matrix.

*Example 4.* First we consider the matrices equation

$$Q=X+A_1^TXA_1+A_2^TXA_2+A_3^TXA_3\tag{35}$$

in the case of $n=5$. And $Q$, $A_1$, $A_2$, $A_3$ are as follows:

$$Q=\begin{pmatrix}4.6350&0.0698&0.3104&0.1990&0.2918\\0.0698&4.0376&0.0742&0.0694&0.0690\\0.3104&0.0742&4.2242&0.1554&0.1826\\0.1990&0.0694&0.1554&4.1472&0.1572\\0.2918&0.0690&0.1826&0.1572&4.2134\end{pmatrix},$$

$$A_1=\begin{pmatrix}0.0834&0.0007&0.0293&0.0142&0.0139\\0.0141&0.2975&0.0154&0.0289&0.0339\\0.0244&0.0152&0.0317&0.0270&0.0117\\0.0251&0.0097&0.0140&0.0135&0.0240\\0.0158&0.1499&0.0275&0.0077&0.0157\end{pmatrix},$$

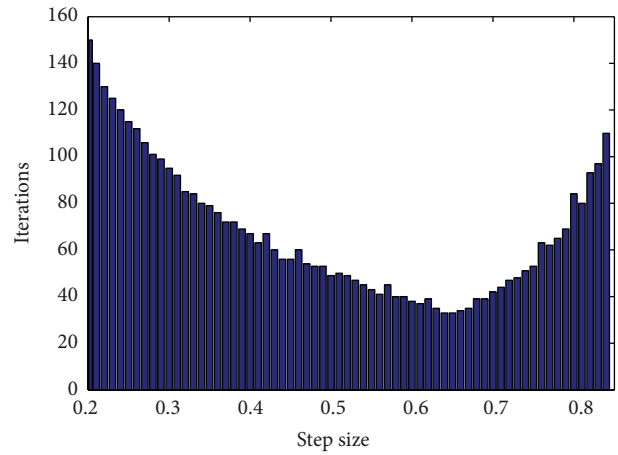


Figure 1: Step sizes versus the iterations in the RGACS.

$$A_2=\begin{pmatrix}0.2624&0.0166&0.0454&0.0172&0.0160\\0.0893&0.0608&0.0242&0.0029&0.0061\\0.0054&0.0190&0.0589&0.0036&0.0160\\0.0278&0.0050&0.0255&0.0689&0.0048\\0.0319&0.0064&0.0103&0.0219&0.0018\end{pmatrix},$$

$$A_3=\begin{pmatrix}8.1622&0.6020&0.4505&0.8258&0.1067\\2.7943&7.0630&0.0838&0.5383&0.9619\\0.3112&0.6541&6.2290&0.9961&0.0046\\0.5285&0.6892&0.9133&1.0782&0.7749\\0.1656&0.7482&0.1524&0.4427&9.0073\end{pmatrix},\tag{36}$$

which satisfy the condition (10). Both the RGACS and RGAVS are used and compared with the traditional CGM, and $Q$ is taken as the initial value $X_0$ in the following simulation. In the RGACS case, the step sizes versus the iterations are shown in Figure 1. It can be found that the iterations will reduce gradually as the step-size changes from zero to a critical value; then the iterations will increase gradually above this critical value. Therefore, the critical value about 0.64 can be selected experimentally as an optimal step-size and used in the RGACS.

Figure 2 shows the convergence comparison between the RGACS and the RGAVS and the traditional CGM. It can be found that the RGAVS possesses the most steady convergence and the fastest convergence speed, and it needs 23 iterations to obtain the numerical solution of (35) as follows:

$$\begin{pmatrix}0.0856&-0.0265&0.0151&-0.1065&0.0148\\-0.0265&0.0970&0.0166&-0.1581&-0.0023\\0.0151&0.0166&0.1591&-0.3608&0.0283\\-0.1065&-0.1581&-0.3608&2.3761&-0.1723\\0.0148&-0.0023&0.0283&-0.1723&0.0654\end{pmatrix}.\tag{37}$$

The case of the CGM realizes the given error through 33 iterations and the slowest one among them is the RGACS with 31 steps. By comparison, it is shown that the convergence

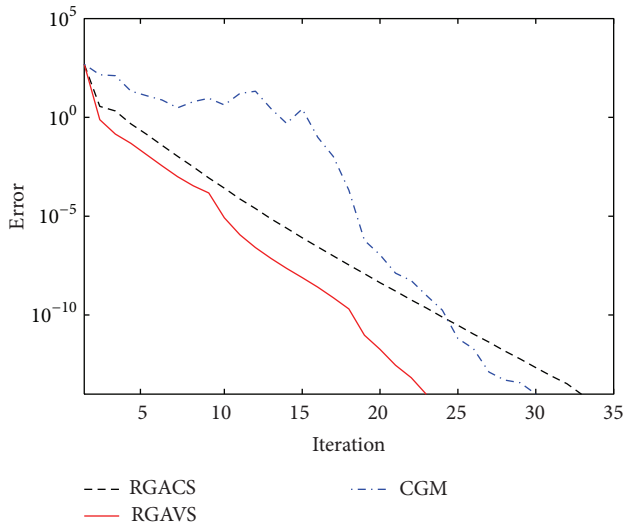


Figure 2: Comparison of convergence speeds (log scale).

speed of the RGAVS is faster than both those of the RGACS and the traditional CGM in solving (35).

*Example 5.* A 20-order linear matrix equation with $m = 2$ is also considered here following the similar procedure as Example 4, where $Q \in \mathbb{S}^{+}(20)$, and $A_1, A_2 \in \mathbb{R}^{20\times 20}$ satisfy the condition (10). The simulation indicates that the numerical solution needs 26, 45, and 37 iterations in the RGAVS, the RGACS ($\eta = 0.74$), and the traditional CGM, respectively. Therefore, the convergence speed of the RGAVS is also better than those of the RGACS and the CGM.

## 5. Conclusion

Using geometric structures of manifold consisting of all symmetric positive definite matrices, the Riemannian gradient algorithm is developed to calculate the numerical solution for the linear matrix equation $Q = X + \sum_{i=1}^{m} A_i^T X A_i$. In this algorithm, the geodesic distance on the curved Riemannian manifold is taken as an objective function and the geodesic curve is treated as the convergence path. Also, the variable step sizes algorithm corresponding to the minimum value of the objective function is provided in order to improve the convergence speed in the constant step-size case. Finally, the convergence speed of the Riemannian gradient algorithm has been compared with the traditional conjugate gradient method by two simulation examples in 5-order and 20-order linear matrix equations, respectively. It is shown that the convergence speed of the Riemannian gradient algorithm with variable step sizes is faster than that of the conjugate gradient method. Although the Riemannian gradient algorithm is used here for the linear matrix equation in the real number field, it may be generalized to the complex case as well.

## Conflict of Interests

The authors declare that there is no conflict of interests regarding the publication of this paper.

## Acknowledgments

The authors would like to express their sincere thanks to the referees for their valuable suggestions, which greatly improved the presentation of this paper. This subject is supported by the National Natural Science Foundations of China (nos. 61179031, 10932002, and 51105033).

## References

[1] M. Berzig, "Solving a class of matrix equations via the Bhaskar-Lakshmikantham coupled fixed point theorem," *Applied Mathematics Letters*, vol. 25, no. 11, pp. 1638–1643, 2012.

[2] J. C. Engwerda, "On the existence of a positive definite solution of the matrix equation $X + A^T X^{-1} A = I$," *Linear Algebra and Its Applications*, vol. 194, pp. 91–108, 1993.

[3] W. L. Green and E. W. Kamen, "Stabilizability of linear systems over a commutative normed algebra with applications to spatially-distributed and parameter-dependent systems," *SIAM Journal on Control and Optimization*, vol. 23, no. 1, pp. 1–18, 1985.

[4] W. Pusz and S. L. Woronowicz, "Functional calculus for sesquilinear forms and the purification map," *Reports on Mathematical Physics*, vol. 8, no. 2, pp. 159–170, 1975.

[5] C.-Y. Chiang, E. K.-W. Chu, and W.-W. Lin, "On the $\star$-Sylvester equation $AX \pm X^{\star} B^{\star} = C$," *Applied Mathematics and Computation*, vol. 218, no. 17, pp. 8393–8407, 2012.

[6] Z. Gajic and M. T. J. Qureshi, *Lyapunov Matrix Equation in System Stability and Control*, vol. 195, Academic Press, London, UK, 1995.

[7] A. C. M. Ran and M. C. B. Reurings, "The symmetric linear matrix equation," *Electronic Journal of Linear Algebra*, vol. 9, pp. 93–107, 2002.

[8] M. C. B. Reurings, *Symmetric matrix equations [Ph.D. thesis]*, Vrije Universiteit, Amsterdam, The Netherlands, 2003.

[9] A. C. M. Ran and M. C. B. Reurings, "A fixed point theorem in partially ordered sets and some applications to matrix equations," *Proceedings of the American Mathematical Society*, vol. 132, no. 5, pp. 1435–1443, 2004.

[10] Y. Su and G. Chen, "Iterative methods for solving linear matrix equation and linear matrix system," *International Journal of Computer Mathematics*, vol. 87, no. 4, pp. 763–774, 2010.

[11] X. Duan, H. Sun, L. Peng, and X. Zhao, "A natural gradient descent algorithm for the solution of discrete algebraic Lyapunov equations based on the geodesic distance," *Applied Mathematics and Computation*, vol. 219, no. 19, pp. 9899–9905, 2013.

[12] X. Duan, H. Sun, and Z. Zhang, "A natural gradient descent algorithm for the solution of Lyapunov equations based on the geodesic distance," *Journal of Computational Mathematics*, vol. 219, no. 19, pp. 9899–9905, 2013.

[13] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, vol. 1, Publish or Perish, Berkeley, Calif, USA, 3rd edition, 1999.

[14] S. Lang, *Fundamentals of Differential Geometry*, vol. 191, Springer, 1999.

[15] F. Barbaresco, "Interactions between symmetric cones and information geometrics: Bruhat-tits and Siegel spaces models for high resolution autoregressive Doppler imagery," in *Emerging Trends in Visual Computing*, vol. 5416 of *Lecture Notes in Computer Science*, pp. 124–163, 2009.

[16] M. Moakher, "A differential geometric approach to the geometric mean of symmetric positive-definite matrices," *SIAM Journal on Matrix Analysis and Applications*, vol. 26, no. 3, pp. 735–747, 2005.

[17] M. Moakher, "On the averaging of symmetric positive-definite tensors," *Journal of Elasticity*, vol. 82, no. 3, pp. 273–296, 2006.

[18] A. Schwartzman, *Random ellipsoids and false discovery rates: statistics for diffusion tensor imaging data [Ph.D. thesis]*, Stanford University, 2006.

[19] J. D. Lawson and Y. Lim, "The geometric mean, matrices, metrics, and more," *The American Mathematical Monthly*, vol. 108, no. 9, pp. 797–812, 2001.

[20] R. Bhatia, *Positive definite Matrices*, Princeton Series in Applied Mathematics, Princeton University Press, Princeton, NJ, USA, 2007.

[21] C. Udrişte, *Convex Functions and Optimization Methods on Riemannian Manifolds*, vol. 297 of *Mathematics and its Applications*, Kluwer Academic Publishers, Dordrecht, Germany, 1994.

[22] M. P. D. Carmo, *Riemannian Geometry*, Springer, 1992.

[23] D. G. Luenberger, "The gradient projection method along geodesics," *Management Science*, vol. 18, pp. 620–631, 1972.

Hindawi

- Impact Factor **1.730**
- **28 Days** Fast Track Peer Review
- All Subject Areas of Science
- Submit at http://www.tswj.com