 CMC, vol.54, no.2, pp.181-195, 2018

Solution of Algebraic Lyapunov Equation on Positive-Definite Hermitian Matrices by Using Extended Hamiltonian Algorithm

Muhammad Shoaib Arif [1], Mairaj Bibi [2] and Adnan Jhangir [3]

Abstract: This communique is opted to study the approximate solution of the Algebraic Lyapunov equation on the manifold of positive-definite Hermitian matrices. We choose the geodesic distance between $-A^HX - XA$ and P as the cost function, and put forward the Extended Hamiltonian algorithm (EHA) and Natural gradient algorithm (NGA) for the solution. Finally, several numerical experiments give you an idea about the effectiveness of the proposed algorithms. We also show the comparison between these two algorithms EHA and NGA. Obtained results are provided and analyzed graphically. We also conclude that the extended Hamiltonian algorithm has better convergence speed than the natural gradient algorithm, whereas the trajectory of the solution matrix is optimal in case of Natural gradient algorithm (NGA) as compared to Extended Hamiltonian Algorithm (EHA). The aim of this paper is to show that the Extended Hamiltonian algorithm (EHA) has superior convergence properties as compared to Natural gradient algorithm (NGA). Upto the best of author's knowledge, no approximate solution of the Algebraic Lyapunov equation on the manifold of positive-definite Hermitian matrices is found so far in the literature.

Keywords: Information geometry, algebraic lyapunov equation, positive-definite hermitian matrix manifold, natural gradient algorithm, extended hamiltonian algorithm.

1 Introduction

It is well known that many engineering and mathematical problems, say, signal processing, robot control and computer image processing [Cafaro (2008); Cohn and Parrish (1991); Barbaresco (2009); Brown and Harris (1994)], can be reduced as obtaining the numerical solution of the following algebraic Lyapunov equation

$$A^HX + XA + P = 0, \tag{1}$$

where P is a positive-definite Hermitian matrix, H denotes the conjugate transpose of a Hermitian matrix.

The solution of the algebraic Lyapunov equation is gaining more and more attention in the field of computational mathematics [Datta (2004); Golub, Nash and Vanloan (1979)]. Several algorithms are used to get the approximate solution of the above-mentioned e-quation. For instance, Ran et al. [Ran and Reurings (2004)] put forward the fixed point

[1] Department of Mathematics, Air University, PAF Complex, E-9, Islamabad, Pakistan.
E-mail: shoaib.arif@mail.au.edu.pk
[2] Department of Mathematics, Comsats Institute of Information Technology, Park Road, Chak-Shahzad, Islamabad, Pakistan. E-mail: mairaj_maths@comsats.edu.pk
[3] Department of Mathematics, Comsats Institute of Information Technology, Wah Cantt, Pakistan.
* Corresponding Author: Adnan Jhangir. Email: adnan.jahangir@ciitwah.edu.pk.

algorithm, the Cholesky decomposition algorithm was presented by Li et al. [Li and White (2002)], and a preconditioned Krylov method to get the solution of the Lyapunov equation was given by Jbilou [Jbilou (2010)]. Vandereycken et al. [Vandereycken and Vandewalle (2010)] provided a Riemannian optimization approach to compute the low-rank solution of the Lyapunov matrix equation. Deng et al. [Deng, Bai and Gao (2006)] designed iterative orthogonal direction methods according to the fundamental idea of the classical conjugate direction method for the standard system of linear equations to obtain the Hermitian solutions of the linear matrix equations $AXB = C$ and $(AX, XB) = (C, D)$. Recently, Su et al. [Su and Chen (2010)] proposed a modified conjugate gradient algorithm (MCGA) to solve Lyapunov matrix equations and some other linear matrix equations, which seemed to be the generalized results. The traditional method like modified conjugate gradient algorithm (MCGA) are first order learning algorithms, hence the convergence speed of MCGA is very slow.

Another interesting approach to solve algebraic Lyapunov equation is by considering the set of matrices as a manifold and applying the techniques from differential geometry and information geometry. Recently Arif et al. [Arif, Zhang and Sun (2016)] solved the algebraic Lyapunov equation on matrix manifold by information geometric algorithm. Duan et al. [Duan, Sun and Zhang (2014); Duan, Sun, Peng et al. (2013)] solved continuous algebraic Lyapunov equation and discrete Lyapunov equation on the space of positive-definite symmetric matrices by using natural gradient algorithm. Also, Luo et al. [Luo and Sun (2014)] gives the solution of discrete algebraic Lyapunov equation on the space of positive-definite symmetric matrices by using Extended Hamiltonian algorithm. In both the papers, the authors have considered the set of positive-definite symmetric matrices as a matrix manifold and used the geodesic distance between $A^H X + XA$ and $-P$ to find the solution matrix X.

Up to date, however, there has been few investigation on the solution problem of the Lyapunov matrix equation in the view of Riemannian manifolds. Chein [Chein (2014)] gives the numerical solution of ill posed positive linear system he combines the methods from manifold theory and sliding mode control theory and develop an affine nonlinear dynamical system. This system is proven asymptotically stable by using argument from Lyapunov stability theory.

In this article, a new frame work is proposed to calculate the numerical solution of continuous algebraic Lyapunov matrix equation on the space of positive-definite Hermitian matrices by using natural gradient algorithm and Extended Hamiltonian algorithm. Moreover, we present the comparison of the solution obtained by the two algorithms.

Note that this solution is a positive definite Hermitian matrices is a global asymptotically stable linear system and the set of all the positive definite Hermitian matrices can be taken as a manifold. Thus, it is more convenient to investigate the solution problem with the help of these attractive features on the manifold. To address such a need, we focus on a numerical method to solve the Lyapunov matrix equation on the manifold.

The gradient is usually adapted to minimize the cost function by adjusting the parameters of the manifold. However, the convergence speed can be seen to be slow if a small change in the parameters changes largely the cost function. In order to overcome this problem of poor convergence, the work has been done in two different directions. Firstly, Amari et al. [Amari (1998); Amari and Douglas (2000); Amari (1999)] introduced the Natural Gradient Algorithm (NGA) which employed the Fisher information matrix on the Riemannian structure of manifold based on differential geometry. Another approach is based on the inclusion of momentum term in the ordinary gradient method to enhance the convergence speed. This is a second-order learning algorithm that was developed by Fiori et al. [Fiori (2011, 2012)], which is called the Extended Hamiltonian Algorithm (EHA).

Although, both the natural gradient algorithm and extended Hamiltonian algorithm defines the steepest descent direction, but we must compute explicitly the Fisher information matrix in the natural gradient algorithm and the steepest descent direction in the extended Hamiltonian algorithm at each iterative step. So the computational cost of both the algorithms are comparatively high. Moreover, the trajectory of the parameters obtained by the implementation of extended Hamiltonian algorithm is closer to the geodesic as compared to one obtained by natural gradient algorithm.

Rest of the paper is organized as follows. Section 2 is a preliminary survey on the manifolds of positive-definite Hermitian m atrices. Third section presents the solution of algebraic Lyapunov matrix equation by Extended Hamiltonian algorithm and Natural gradient algorithm and also illustrates the convergence speed of EHA compared with NGA using numerical examples. Section 4 concludes the results presented in section 3.

2 The Riemannian structure on the manifold of positive-definite Hermitian matrices

In this paper, we denote the set of $n \times n$ Positive-definite Hermitian matrices by $H(n)$. This set can be considered as a Riemannian manifold by defining the Riemannian metric on it. Moakher [Moakher (2005)] in his paper, gives the concept of geodesic connecting two matrices on $H(n)$. Observing that the geodesic distance represents the infimum of lengths of the curves connecting any two matrices. Here, we take geodesic distance as the cost function to minimize the distance between two matrices in $H(n)$. The following background material and important results are taken from Zhang [Zhang (2004)], Moakher et al. [Moakher and Batcherlor (2006)].

All $n \times n$ positive-definite Hermitian matrices forms an n^2-dimensional manifold, which is denoted by $H(n)$. Also denote the space of all $n \times n$ Hermitian matrices by $H'(n)$. The exponential map from $H'(n)$ to $H(n)$, given by:

$$\exp(X) = \sum_{m=0}^{\infty} \frac{X^m}{m!},$$

is one-to-one and onto. Its inverse i.e., the logarithmic map from $H(n)$ to $H'(n)$, defined by

$$\ln(X) = \sum_{m=1}^{\infty} (-1)^{m+1} \frac{(X-I)^m}{m},$$

for X in a neighbourhood of the identity I of $H(n)$.

Let E_{kl} denotes matrix whose all entries are zero except the $k-th$ line and $l-th$ column which is 1, then the basis of the manifold $H(n)$ can be given as

$$E_p = \begin{cases} E_{kl}, & k = l, \\ E_{kl} + E_{lk}, & k < l, \\ i(E_{kl} - E_{lk}), & k > l \end{cases} \tag{2}$$

where $i^2 = -1$, p is obtained by some rule from the pair (k, l). Hence, any positive-definite Hermitian matrix $Q \in H(n)$ can be shown as

$$Q = \sum_{i=1}^{n^2} \theta^i E_i,\ \theta^i \in \mathbb{R},$$

where $\{\theta^i\}$ satisfy positive-definite and belong to some open subset of $\mathbb{R}^{n^2}$. Therefore, $\{\theta^i\}$ form a coordinate of the manifold $H(n)$. As $H(n)$ is an open subset of $H'(n)$, so for each $Q \in H(n)$, the tangent space $T_Q H(n)$ is identified by $H'(n)$ and $\{\frac{\partial}{\partial \theta^i}\}_{i=1}^{n^2}$ can serve as the basis of the tangent space.

Definition 2.1 (Duan et al. [Duan, Sun, Peng and Zhao (2013)]). Let g be the Riemannian metric on the positive-definite Hermitian matrix manifold $H(n)$, for $Q \in H(n)$ the inner product on $T_Q H(n)$ can be defined as

$$g_Q(M, N) = \frac{1}{2} tr(Q^{-1} M Q^{-1} N), \tag{3}$$

where $M, N \in T_Q H(n)$.

Obviously, the metric defined above satisfies the fundamental properties of Riemannian metric and keeps invariant under base transformation on the tangent space.

Definition 2.2 (Duan et al. [Duan, Sun and Zhang (2014); Luo and Sun (2014)]). Let $\gamma : [0, 1] \to M$ be a piecewise smooth curve on manifold M, we define the length of γ as

$$l(\gamma) = \int_0^1 \sqrt{< \dot{\gamma}(t), \dot{\gamma}(t) >_{\gamma(t)}} dt = \int_0^1 \sqrt{tr(\gamma^{-1}(t)\dot{\gamma}(t))^2} dt, \tag{4}$$

then the distance between any two point $x, y \in M$ can be defined as

$$d(x, y) = \inf\{l(\gamma) | \gamma : [0, 1] \to M, \gamma(0) = x, \gamma(1) = y\}. \tag{5}$$

Proposition 2.1 (Duan et al. [Duan, Sun, Peng et al. (2013); Luo and Sun (2014)]). For the defined Riemannian metric (3) on the positive-definite Hermitian matrix manifold $H(n)$. We get the geodesic originating from Q along X direction as follows

$$\gamma(t) = Q^{\frac{1}{2}} \exp(t Q^{-\frac{1}{2}} X Q^{-\frac{1}{2}}) Q^{\frac{1}{2}}. \tag{6}$$

Hence, the geodesic distance between Q_1, Q_2 is shown as

$$d(Q_1, Q_2) = \left\| \log(Q_1^{-\frac{1}{2}} Q_2 Q_1^{-\frac{1}{2}}) \right\|_F. \tag{7}$$

According to Hopf-Rinow theorem, the positive-definite Hermitian matrix manifold is complete, which means we can always find a geodesic that connects any two points $Q_1, Q_2 \in H(n)$.

In our case, the geodesic curve $\gamma(t)$ is given by

$$\gamma(t) = x^{1/2}(x^{-1/2} y x^{-1/2})^t x^{1/2} \in M$$

with $\gamma(0) = x; \gamma(1) = y$ and $\dot{\gamma(0)} = x^{1/2} \ln(x^{-1/2} y x^{-1/2}) x^{1/2} \in H(n)$ then the midpoint of x and y is defined by $x \circ y = x^{1/2}(x^{-1/2} y x^{-1/2}) x^{1/2}$ and the geodesic distance $d(x, y)$ can be computed explicitly by

$$d(x, y) = \left\| \log(x^{-\frac{1}{2}} y x^{-\frac{1}{2}}) \right\|_F = \left(\sum_{i=1}^{n} \ln(\lambda_i)^2 \right)^{1/2} \tag{8}$$

where λ_i are eigenvalues of $x^{-1/2} y x^{-1/2}$. or $x^{-1} y$,.

3 Solution of Algebraic Lyapunov matrix equation

Suppose the state of the system $X(t)$ is $\dot{x}(t) = Ax(t)$. Consider the Lyapunov function

$y(t) = x(t)^H X x(t)$

on the complex field, we have

$\dot{y}(t) = \dot{x}(t)^H X x(t) + x(t)^H X \dot{x}(t),$

then

$$\begin{aligned} \dot{y}(t) &= (Ax(t))^H X x(t) + x(t)^H X(Ax(t)), \\ &= x(t)^H (A^H X + XA) x(t) \end{aligned} \tag{9}$$

In order to make the system stable, the state Eq. (9) must be negative definite, which yields

$A^H X + XA = -P,$

where P is a positive-definite Hermitian matrix.

The uniqueness of the solution of Algebraic Lyapunov Eq. (1) is a well-known fact, stated below (see Davis et al. [Davis, Gravagne, Robert et al. (2010)]):

Theorem 3.1. Given a positive-definite Hermitian matrix $P > 0$, there exists a unique positive-definite Hermitian $X > 0$ satisfying (1) if and only if the linear system $\dot{x} = Ax$ is globally asymptotically stable i.e. the real part of eigenvalues of A is less than 0.

3.1 Extended Hamiltonian algorithm

Considering the algebraic Lyapunov Eq. (1) on the positive-definite Hermitian matrix manifold, its solution can be described as finding a positive-definite Hermitian matrix X on $H(n)$ such that the matrix $-A^H X - XA$ is as close as P (see Fig. 1).

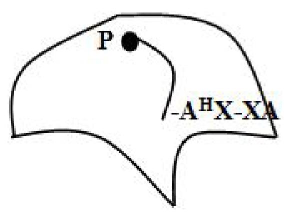

Figure 1: Geodesic distance on positive-definite hermitian matrix manifold

To describe the distance between $-A^H X - XA$ and P, we choose the geodesic distance between them as the measure, that is to say the target function is

$$J(X) = d^2(-A^H X - XA, P) = \left\| \log(P^{-\frac{1}{2}}(-A^H X - XA)P^{-\frac{1}{2}}) \right\|_F^2, \tag{10}$$

then the optimal solution of the Eq. (1) is

$$X^* = \underset{X \in H(n)}{\arg} \min J(X). \tag{11}$$

Let $X = (\zeta^1, \zeta^2, \ldots, \zeta^m) \in \mathbf{R}^m$ be a parameter space of the matrix manifold $H(n)$ on which cost function $J(X_t)$ is defined. Following Lemma can be used to find the gradient of the cost function $J(X_t)$:

Lemma 3.1 (Zhang [Zhang (2004)]). Let $f(X)$ be the scalar function of the matrix X, if $df(X) = \text{tr}(WdX)$ holds, then the gradient of $f(X)$ with respect to X is

$$\partial_X f(X) = W^H. \tag{12}$$

Theorem 3.2. Let $J(X)$ be the function in (10), then the gradient of $J(X)$ with respect to the positive-definite Hermitian matrix X is

$$\begin{aligned}\partial_x J(X) = {} & P^{-\frac{1}{2}}Y^H P^{\frac{1}{2}}(A^H X + XA)^{-1}A^H + AP^{-\frac{1}{2}}Y^H P^{\frac{1}{2}}(A^H X + XA)^{-1} \\ & + A(A^H X + XA)^{-1}P^{\frac{1}{2}}Y^H P^{-\frac{1}{2}} + (A^H X + XA)^{-1}P^{\frac{1}{2}}YP^{-\frac{1}{2}}A^H.\end{aligned} \tag{13}$$

Proof of the above Theorem see the Appendix.

Theorem 3.3. On the positive definite Hermitian matrix system, if the i-th iteration matrix and direction matrix are X_i, V_i respectively, then $(i+1)$-th iteration matrix and direction matrix X_{i+1}, V_{i+1} satisfy

$$\begin{cases} X_{i+1} &= X_i^{\frac{1}{2}} \exp(\eta X_i^{-\frac{1}{2}} V_i X_i^{-\frac{1}{2}}) X_i^{\frac{1}{2}}, \\ V_{i+1} &= \eta(V_i X_i^{-1} V_i - \nabla_{X_i} J(X_i)) + (1-\eta\mu)V_i, \end{cases} \tag{14}$$

where

$$\nabla_{X_i} J(X_i) = X_i \partial_{X_i} J(X_i) X_i,$$

the sufficient small number η is the learning rate, μ satisfies $\sqrt{2\lambda_m} < \mu < \frac{1}{\eta}$, λ_m is the largest eigenvalue of the Hessian matrix of the cost function. The X and V iteration continue until the stopping criterion is met. See Fiori [Fiori (2011, 2012)] for more details.

By these discussion, we present the extended Hamiltonian algorithm to find the solution of the algebraic Lyapunov Eq. (1) on the positive-definite Hermitian matrix manifold $H(n)$.

Algorithm 3.1. For the manifold $H(n)$ the algorithm is given as follows. Here $J(X)$ is the cost function (10).

1. Input initial matrix X_0, initial direction V_0, step size η and error tolerance $\varepsilon > 0$;
2. Calculate the gradient $\partial_{X_i} J(X_i)$ by (13);
3. If $J(X_i) < \varepsilon$, then halt;
4. Update X, V according to (14) and go back to step 2.

3.1.1 Numerical experiment

Consider the submanifold $PH(2)$ of $H(2)$ defined by:

$$PH(2) = \left\{ \begin{bmatrix} \zeta^1 & \zeta^2 + i\zeta^3 \\ \zeta^2 - i\zeta^3 & \zeta^4 \end{bmatrix}; \zeta^i \in \mathbb{R}, \zeta^1 > 0, \zeta^1\zeta^4 - (\zeta^2)^2 - (\zeta^3)^2 > 0 \right\}. \tag{15}$$

Now we consider the algebraic Lyapunov equation on the manifold of positive-definite Hermitian matrices.

$$A^H X + XA + P = 0,$$

where

$$A = \begin{bmatrix} -2 & -1-i \\ i & -1 \end{bmatrix}$$

is any matrix with real part of its eigenvalues negative by Theorem 3.1,
and

$$P = \begin{bmatrix} 3 & 1+\frac{3}{2}i \\ 1-\frac{3}{2}i & 3 \end{bmatrix} \in PH(2).$$

In this experiment, we choose initial guess X_0 and initial direction V_0 as

$$X_0 = \begin{bmatrix} 0.8 & -0.3i \\ 0.3i & 2.2 \end{bmatrix} \in PH(2)$$

$$V_0 = \begin{bmatrix} -0.5 & -0.1+0.2i \\ -0.1-0.2i & -0.5 \end{bmatrix} \in H'(2)$$

Taking the step size $\eta = 0.1$ and $\mu = 6$, then after 41 iterations, we obtain the optimal solution under the error tolerance $\varepsilon = 10^{-3}$ as follows,

$$\begin{bmatrix} 0.9968 & 0.0013-0.4947i \\ 0.0013+0.4947i & 1.9895 \end{bmatrix}.$$

In fact, the exact solution of (1) on the positive-definite Hermitian matrix manifold in this example is

$$\begin{bmatrix} 1 & -0.5i \\ 0.5i & 2 \end{bmatrix}.$$

In Fig. 2, $\zeta^1, \zeta^2, \zeta^3, \zeta^4$ represent coordinates of the manifold $PH(2)$, S and A denote the initial matrix and the goal matrix respectively. The coordinates $\zeta^1, \zeta^3, \zeta^4$ are taken along coordinate axes and ζ^2 is represented by colour bar. The curve from S to A shows us the optimal trajectory by EHA. Fig. 2 also shows the geodesic connecting S and A obtained by (6).

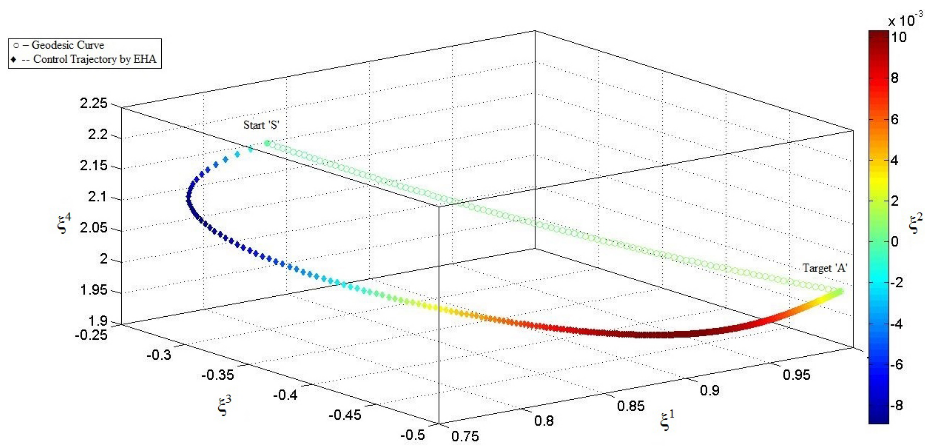

Figure 2: The optimal trajectory of EHA where $\eta = 0.1$, $\mu = 6$ and $\varepsilon = 10^{-3}$

Futhermore, we compare the efficiency of the algorithm with different step sizes. In Fig. 3, the descent curves corresponding to $\eta = 0.1, 0.15, 0.2$ show us the relation between $J(X)$ and iterations.

From the Fig. 3, we can find that if η is too small, the iterations are many and the algorithm converges slowly. However, the step size can not be too large and may result in divergence of this algorithm. Therefore, we need to adjust the step size to obtain the best convergence speed.

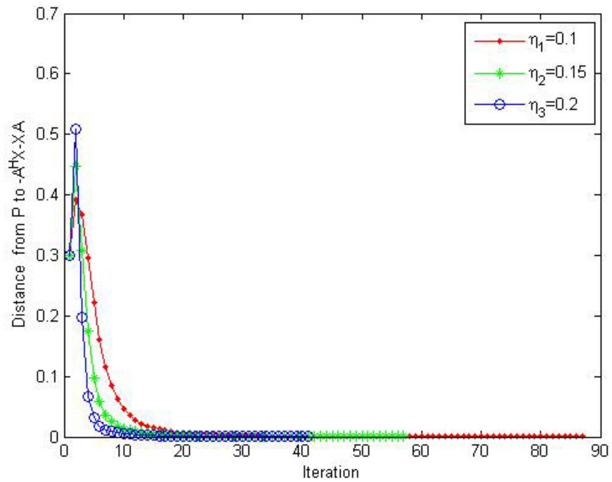

Figure 3: The efficiency of the Algorithm with different step size

3.2 Natural Gradient Algorithm

Since $H(n)$ is a Riemannian manifold, not a Euclidean space, therefore, it is non optimal to make use of the classical Frobenius inner product:

$$\langle A, B\rangle = \mathrm{tr}(A^T B) \tag{16}$$

as a flat metric on manifold $H(n)$ for this geometric problem **?**. Moreover, since the geodesic $A + t(B - A)$ is a negative metric for some values of t, so it is not appropriate to apply the ordinary gradient methods on the manifold $H(n)$ with metric (16). Observing that the geodesic is the shortest path between two points on a manifold, therefore we take geodesic distance as the cost function, denoted by:

$$J(X) = d^2(P, -A^H X - XA), \tag{17}$$

then the optimal solution of Algebric Lyapunov equation is obtained by

$$X^* = \underset{X\in H(n)}{\arg} \min J(X). \tag{18}$$

As stated above, the ordinary gradient can not give the steepest descent direction of the cost function $J(X_t)$ on manifold $H(n)$, whereas the natural gradient algorithm (NGA) does. Below we state an important Lemma, which gives the iterative step in the natural gradient algorithm.

Lemma 3.2 (Amari [Amari (1998)])**.** Let $X = (\zeta^1, \zeta^2, \dots \zeta^m) \in \mathbb{R}^m$ be a parameter space on the Riemannian manifold $H(n)$, and consider a function $L(\zeta)$. Then the natural gradient algorithm is given by:

$$\zeta_{t+1} = \zeta_t - \eta_t G^{-1} \nabla L(\zeta_t) \tag{19}$$

where $G^{-1} = (g^{ij})$ is the inverse of the Riemannian metric $G = (g_{ij})$ and

$$\frac{\partial}{\partial X^i} J(X_t) = \tag{20}$$

$$2tr\left(P^{-\frac{1}{2}}\log(P^{-\frac{1}{2}}(-A^H X_t - X_t A)P^{-\frac{1}{2}})P^{\frac{1}{2}}(A^H X_t + X_t A)^{-1}(A^H\frac{\partial X_t}{\partial X^i} + \frac{\partial X_t}{\partial X^i}A)\right),$$

Now, we will give the natural gradient descent algorithm for the considered Eq. (1), taking the geodesic distance $J(X_t)$ as the cost function and the negative of the gradient of the cost function $J(X_t)$ about X_t to give the descent direction in the iterative equation.

Theorem 3.4. The iteration on manifold $H(n)$ is given by

$$X_{t+1} = X_t - \eta G^{-1} \nabla J(X_t), \tag{21}$$

where the component of gradient $\nabla J(X_t)$ satisfies

$$\frac{\partial J(X_t)}{\partial X^i} = 2tr(P^{-\frac{1}{2}} \log(P^{-\frac{1}{2}}(-A^H X_t - X_t A)P^{-\frac{1}{2}}$$
$$(A^H X_t + X_t A)^{-1}(A^H \frac{\partial X_t}{\partial X^i} + \frac{\partial X_t}{\partial X^i} A), \tag{22}$$

where $i = 1, 2, \ldots, m$.

For Proof of above Theorem See the Appendix.

By these discussion, we present the natural gradient algorithm to find the solution of the algebraic Lyapunov Eq. (1) on the manifold $H(n)$ of positive-definite Hermitian matrices.

Algorithm 3.2. For the coordinate $X = (\zeta^1, \zeta^2, \ldots, \zeta^m)$ of the considered manifold H(n), the natural gradient algorithm is given by;

1. Set $X_\circ = (\zeta_\circ^1, \zeta_\circ^2, \ldots, \zeta_\circ^m)$ as the initial input matrix X and choose required tolerance $\epsilon_\circ > 0$.
2. Compute $J(X_t) = d^2(P, -A^H X_t - X_t A)$
3. If $\|\nabla J(X_t)\|_F < \epsilon_\circ$, then halt.
4. Update the vector X by $X_{t+1} = X_t - \eta G^{-1} \nabla J(X_t)$, where $X_t = (\zeta_t^1, \zeta_t^2, \ldots, \zeta_t^m)$, η is learning rate and go back to step 2.

3.2.1 Numerical Simulations

Consider the submanifold $PH(2)$ of $H(2)$ defined by:

$$PH(2) = \left\{ \begin{bmatrix} \zeta^1 & \zeta^2 + i\zeta^3 \\ \zeta^2 - i\zeta^3 & \zeta^4 \end{bmatrix}; \zeta^i \in \mathbb{R}, \zeta^1 > 0, \zeta^1\zeta^4 - (\zeta^2)^2 - (\zeta^3)^2 > 0 \right\}. \tag{23}$$

Now we consider the algebraic Lyapunov equation on the manifold of positive-definite Hermitian matrices:

$$A^H X + XA + P = 0,$$

where

$$A = \begin{bmatrix} -2 & -1-i \\ i & -1 \end{bmatrix}$$

is any matrix with real part of its eigenvalues negative by Theorem 3.1,
and

$$P = \begin{bmatrix} 3 & 1 + \frac{3}{2}i \\ 1 - \frac{3}{2}i & 3 \end{bmatrix} \in PH(2).$$

In this experiment, we choose initial guess X_0 as

$$X_0 = \begin{bmatrix} 0.8 & -0.3i \\ 0.3i & 2.2 \end{bmatrix} \in PH(2)$$

Taking the step size $\eta = 0.035$, then after 44 iterations, we obtain the optimal solution under the error tolerance $\varepsilon = 10^{-2}$ as follows,

$$\begin{bmatrix} 1.0046 & -0.0002 - 0.5048i \\ -0.0002 + 0.5048i & 2.0146 \end{bmatrix}.$$

In fact, the exact solution of (1) in this example is:

$$\begin{bmatrix} 1 & -0.5i \\ 0.5i & 2 \end{bmatrix}.$$

In Fig. 4, $\zeta^1, \zeta^2, \zeta^3, \zeta^4$ represent coordinates of the manifold $PH(2)$, S and A denote the initial matrix and the goal matrix respectively. The coordinates $\zeta^1, \zeta^3, \zeta^4$ are taken along coordinate axes and ζ^2 is represented by colour bar. The curve from S to A shows us the optimal trajectory by NGA. Fig. 4 also shows the geodesic connecting S and A obtained by (6).

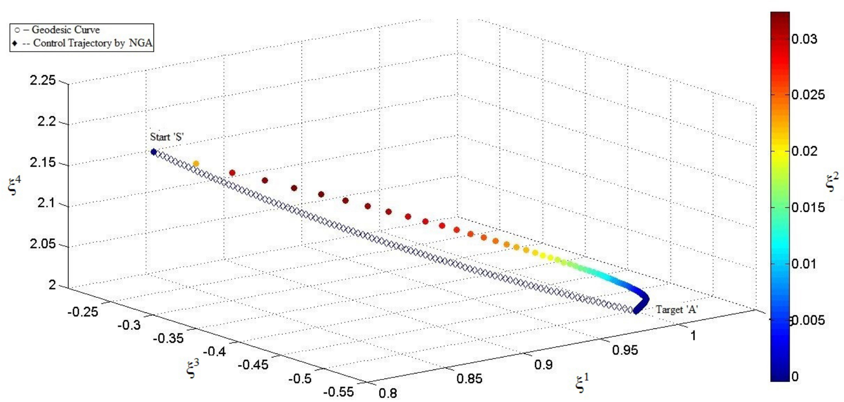

Figure 4: The optimal trajectory of NGA where $\eta = 0.035$ and $\varepsilon = 10^{-2}$

Futhermore, we compare the efficiency of the algorithm with different step sizes. In Fig. 5, the descent curves corresponding to $\eta = 0.015, 0.025, 0.035$ show us the relation between $J(X)$ and iterations.

From the Fig. 5, we can find that if η is too smaller, the iterations are many and the algorithm convergent slowly. However, the step size can not be too large, which may result in divergence in this algorithm. Therefore, we need to adjust the step size in order to obtain the best convergence speed.

3.3 Comparison of NGA and EHA

We apply the natural gradient algorithm 3.2 and extended Hamiltonian algorithm 3.1 to solve the algebraic Lyapunov Eq. (1). From the following example, we can see the efficiency of the two proposed algorithms.

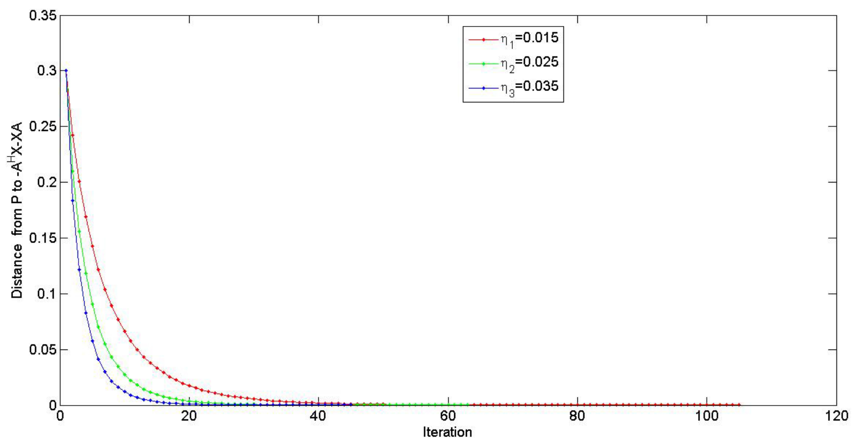

Figure 5: The efficiency of the Natural gradient Algorithm with different step size

Consider the submanifold $PH(2)$ of $H(2)$ defined by:

$$PH(2) = \left\{ \begin{bmatrix} \zeta^1 & \zeta^2 + i\zeta^3 \\ \zeta^2 - i\zeta^3 & \zeta^4 \end{bmatrix}; \zeta^i \in \mathbb{R}, \zeta^1 > 0, \zeta^1\zeta^4 - (\zeta^2)^2 - (\zeta^3)^2 > 0 \right\}. \tag{24}$$

Now we consider the algebraic Lyapunov equation on the manifold of positive-definite Hermitian matrices.

$A^H X + XA + P = 0,$

where

$$A = \begin{bmatrix} -2 & -1-i \\ i & -1 \end{bmatrix}$$

is any matrix with real part of its eigenvalues negative by Theorem 3.1, and

$$P = \begin{bmatrix} 3 & 1+\frac{3}{2}i \\ 1-\frac{3}{2}i & 3 \end{bmatrix} \in PH(2).$$

In this experiment, we choose initial guess X_0 and initial direction V_0 as

$$X_0 = \begin{bmatrix} 0.8 & -0.3i \\ 0.3i & 2.2 \end{bmatrix} \in PH(2)$$

$$V_0 = \begin{bmatrix} -0.5 & -0.1+0.2i \\ -0.1-0.2i & -0.5 \end{bmatrix} \in H'(2)$$

According to algorithm 3.1, we get the solution of algebraic Lyapunov equation with $\eta = 0.07, \mu = 4$ and error tolerance $\epsilon = 10^{-3}$ as

$$\begin{bmatrix} 0.9975 & 0.0006 - 0.4971i \\ 0.0006 - 0.4971i & 1.9930 \end{bmatrix}.$$

According to algorithm 3.2, we get the solution of the algebraic Lyapunov equation with $\eta = 0.07$ and error tolerance $\epsilon = 10^{-3}$ as

$$\begin{bmatrix} 1.0005 & -0.0001 - 0.5006i \\ -0.0001 + 0.5006i & 2.0016 \end{bmatrix}.$$

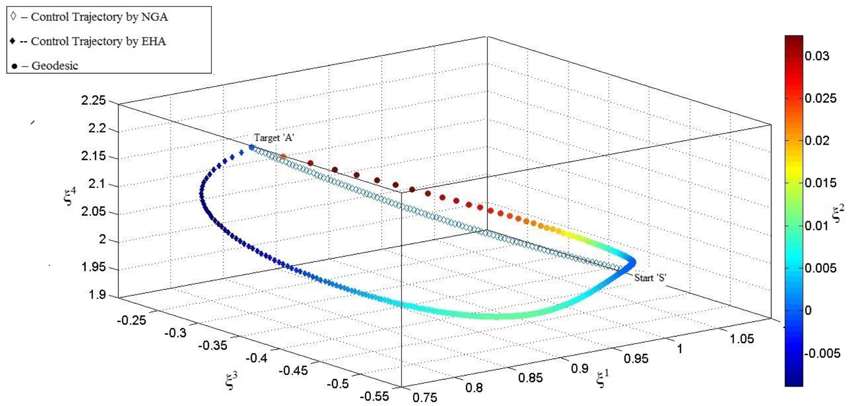

Figure 6: The optimal trajectory of $X(t)$ by NGA and EHA where $\eta = 0.07, \mu = 4$ and $\epsilon = 10^{-3}$

Besides, the optimal trajectory of $X(t)$ from the initial input to the target matrix is shown in Fig. 6.

In Fig. 6, $\zeta^1, \zeta^2, \zeta^3, \zeta^4$ represent parameters of the vector $X(t)$, S and A denote the initial matrix and the goal matrix respectively. The parameters $\zeta^1, \zeta^2, \zeta^3$ are taken along coordinate axes and ζ^4 is represented by colour bar. The curves from S to A shows us the optimal trajectory of $X(t)$ by NGA and EHA. Fig. 6 also shows the geodesic connecting S and A obtained by (6). In addition, although the trajectory of the input $X(t)$ given by EHA is not optimal, but the convergence is faster than NGA.

The EHA and NGA are respectively applied to get the solution of the algebraic Lyapunov equation. In particular, the behaviour of the cost function is shown in Fig. 7. In early stages of learning, the EHA decreases much faster than NGA with the same learning rate. The result shows that the EHA has faster convergence speed and need 95 iterations to obtain optimal solution of Algebraic Lyapunov equation as compared to NGA which converges after 155 iterations.

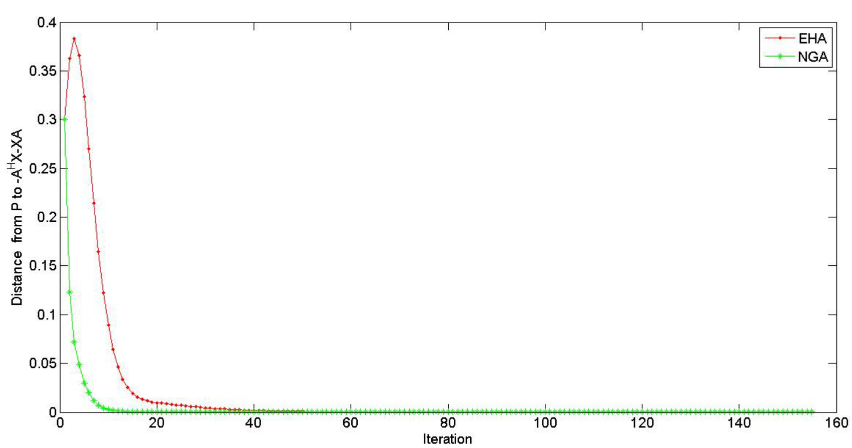

Figure 7: Comparison of convergence speed of EHA an NGA

4 Conclusion

We studied the solution of continuous algebraic Lyapunov equation by considering the positive-definite Hermitian matrices as a Riemannian manifold and used geodesic distance to find the solution. Here we used two algorithms, the extended Hamiltonian algorithm and the natural gradient algorithm to get the approximate solution of algebraic Lyapunov matrix equation. Finally, several numerical experiments give you an idea about the effectiveness of the proposed algorithms. We also show the comparison between these two algorithms EHA and NGA. Henceforth we conclude that the extended Hamiltonian algorithm has better convergence speed than the natural gradient algorithm, whereas the trajectory of the solution matrix is optimal in case of NGA as compared to EHA.

5 Appendix

Proof of Theorem 3.2

Proof. Since $-A^H X - XA$ is Hermitian. Let $Y = log(P^{-\frac{1}{2}}(-A^H X - XA)P^{-\frac{1}{2}})$

$$\begin{aligned} dY &= dlog(P^{-\frac{1}{2}}(-A^H X - XA)P^{-\frac{1}{2}}) \\ &= (P^{-\frac{1}{2}}(-A^H X - XA)P^{-\frac{1}{2}})^{-1}P^{-\frac{1}{2}}(-A^H dX - dXA)P^{-\frac{1}{2}}) \\ &= P^{\frac{1}{2}}(-A^H X - XA)^{-1}(-A^H dX - dXA)P^{-\frac{1}{2}} \end{aligned}$$

then, we have

$$dJ(X) = d(tr(Y^H Y)) = tr(dY^H Y + Y^H dY)$$

According to Lemma 4.3.1, the geodesic of $J(X)$ with respect to X is

$$\begin{aligned} \partial_x J(X) = {} & P^{-\frac{1}{2}}Y^H P^{\frac{1}{2}}(A^H X + XA)^{-1}A^H + AP^{-\frac{1}{2}}Y^H P^{\frac{1}{2}}(A^H X + XA)^{-1} \\ & + A(A^H X + XA)^{-1}P^{\frac{1}{2}}Y^H P^{-\frac{1}{2}} + (A^H X + XA)^{-1}P^{\frac{1}{2}}YP^{-\frac{1}{2}}A^H. \end{aligned}$$

Proof of Theorem 3.4

Proof. According to above Lemma, we can get the iterative process

$$X_{t+1} = X_t - \eta G^{-1}\nabla J(X_t),$$

where the Fisher metric matrix G is obtained by 3. Let $X(t) = \log((-A^H X_t - X_t A)^{-\frac{1}{2}} P(-A^H X_t - X_t A)^{-\frac{1}{2}})$, it is easy to show that $X(t)$ is Hermitian from (20) in Lemma 3.2 and the properties of the trace of a matrix, we have the components of the gradient $\nabla J(X_t)$:

$$\begin{aligned} \frac{\partial}{\partial X^i} J(X_t) &= 2\mathrm{tr}\left(\log(P^{-\frac{1}{2}}(-A^H X_t - X_t A)P^{-\frac{1}{2}})\frac{\partial}{\partial X^i}(\log(P^{-\frac{1}{2}}(-A^H X_t - X_t A)P^{-\frac{1}{2}}))\right) \\ &= 2\mathrm{tr}\left(P^{-\frac{1}{2}}\log(P^{-\frac{1}{2}}(-A^H X_t - X_t A)P^{-\frac{1}{2}})P^{\frac{1}{2}}(-A^H X_t - X_t A)^{-1}(A^H\frac{\partial X_t}{\partial X^i} + \frac{\partial X_t}{\partial X^i}A)\right), \end{aligned}$$

where $i = 1, 2, \ldots, m$. ∎

Acknowledgement: The authors wish to express their appreciation to the reviewers for their helpful suggestions which greatly improved the presentation of this paper.

References

Amari, S. (1998): Natural gradient works efficiently in learning. *Neural Computation*, vol. 10, no. 2, pp. 251-276.

Amari, S. (1999): Natural gradient for over-and under-complete bases in ica. *Neural Computation*, vol. 11, no. 8, pp. 1875-1883.

Amari, S.; Douglas, S. C. (2000): Why natural gradient, acoustics. *Speech and Signal Processing*, vol. 2, pp. 1213-1216.

Arif, M. S.; Zhang, E. C.; Sun, H. (2016): An information geometric algorithm for algebraic lyapunov equation on positive definite matrix manifolds. *Transaction of Beijing institute of Technology*, vol. 36, no. 2, pp. 205-208.

Barbaresco, F. (2009): Interactions between symmetric cone and information geometries. *Bruhat-Tits and Siegel Spaces Models for High Resolution Autoregressive Doppler Imagery, ETVC*, vol. 5416, pp. 124-163.

Brown, M.; Harris, C. (1994): *Neurofuzzy Adaptive Modelling and Control*. Prentice Hall New York.

Cafaro, C. (2008): Information-geometric indicators of chaos in gaussian models on statistical manifolds of negative ricci curvature. *International Journal of Theoretical Physics*, vol. 47, no. 11, pp. 2924-2933.

Chein, S. L. (2014): A sliding mode control algorithm for solving an ill-posed positive linear system. *Computers, Materials & Continua*, vol. 39, no. 2, pp. 153-178.

Cohn, S. E.; Parrish, D. F. (1991): The behavior of forecast covariances for a kalman filter in two dimensions. *Neural Computation*, vol. 119, no. 8, pp. 1757-1785.

Datta, B. N. (2004): *Numerical Methods for Linear Control Systems*. Elsvier Academic-Press.

Davis, J. M.; Gravagne, I. A.; Robert, J.; Marks, I. (2010): Algebraic and dynamic lyapunov equations on time scales. *42nd South Eastern Symposium on System Theory at University of Texas, USA*, pp. 329-334.

Deng, Y. B.; Bai, Z. Z.; Gao, Y. H. (2006): Iterative orthogonal direction methods for hermitian minimum norm solutions of two consistent matrix equations. *Numerical Linear Algebra with Applications*, vol. 13, no. 10, pp. 801-823.

Duan, X.; Sun, H.; Peng, L.; Zhao, X. (2013): 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.

Duan, X.; Sun, H.; Zhang, Z. (2014): A natural gradient algorithm for the solution of lyapunov equations based on the geodesic distance. *Journal of Computational Mathematics*, vol. 32, no. 1, pp. 93-106.

Fiori, S. (2011): Extended hamiltonian learning on riemannian manifolds, theoretical aspects. *IEEE Transactions on Neural Networks*, vol. 22, no. 5, pp. 687-700.

Fiori, S. (2012): Extended hamiltonian learning on riemannian manifolds, numerical aspects. *IEEE Transactions on Neural Networks and Learning Systems*, vol. 23, no. 1, pp. 7-21.

Golub, G. H.; Nash, S.; Vanloan, C. (1979): A hessenberg-schur method for the problem $ax + xb = c$. *IEEE Transactions on Automatic Control*, vol. 24, no. 6, pp. 909-913.

Jbilou, K. (2010): Adi preconditioned krylov methods for large lyapunov matrix equations. *Linear Algebra & Its Applications*, vol. 432, no. 10, pp. 2473-2485.

Li, J. R.; White, J. (2002): Low-rank solution of lyapunov equations. *SIAM Journal on Matrix Analysis and Applications*, vol. 24, no. 1, pp. 260-280.

Luo, Z.; Sun, H. (2014): Extended hamiltonian algorithm for the solution of discrete algebraic lyapunov equations. *Applied Mathematics and Computation*, vol. 234, pp. 245-252.

Moakher, M. (2005): 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.

Moakher, M.; Batcherlor, P. G. (2006): *Symmetric Positive-Definite M atrices: From Geometry to Applications and Visualizatin, Visualization and Processing of Tensor Fields*. Springer.

Ran, A. C. M.; Reurings, M. C. B. (2004): A fixed point theorem in partially ordered sets and some applications to matrix equations. *Proceedings of the American Mathematical Society*, vol. 132, pp. 1435-1443.

Su, Y. F.; Chen, G. L. (2010): Iterative methods for solving linear matrix equation and linear matrix system. *Numerical Linear Algebra with Applications*, vol. 87, no. 4, pp. 763-774.

Vandereycken, B.; Vandewalle, S. (2010): A riemannian optimization approach for computing low-rank solutions of lyapunov equations. *SIAM Journal on Matrix Analysis and Applications*, vol. 31, no. 5, pp. 2553-2579.

Zhang, X. D. (2004): *Matrix Analysis and Application*. Springer, Beijing.