Applied Mathematical Sciences, Vol. 4, 2010, no. 1, 21 - 30

Approximate Low Rank Solution of Generalized Lyapunov Matrix Equations via Proper Orthogonal Decomposition

Amer Kaabi

Department of Basic Science
Khoramshahr Marine Science and Technology University
Khoramshahr, Iran
kaabi_amer@yahoo.com

Abstract

We generalized a direct method for generalized Lyapunov matrix equation, using proper orthogonal decomposition (POD). Such equations arise in model reduction of descriptor systems.

Keywords: Generalized Lyapunov matrix equations, approximation theory, infinite dimensional problems, matrix algebra, proper orthogonal decomposition

1 Introduction

Lyapunov matrix equations play a essential role in control theory [4, 5, 6, 11]. As we know, The necessary and sufficient condition for $AX + XA^T = -DD^T$ to have a unique solution is that

$$\lambda(A) \cap \lambda(-A^T) = \emptyset,$$

where $\lambda(A)$ and $\lambda(-A^T)$ are the spectrums of A and $-A^T$, respectively. Moreover, for the symmetric right hand side, this solution is also symmetric. There are a number of direct methods for solving the Lyapunov matrix equations numerically, the most important of which are the Bartels-Stewart [1], Hessenberg-Schur [8] and the Hammarling methods [9]. But these methods is not suitable for big and sparse Lyapunov matrix equations. In [12], Y saad proposed Krylov subspace methods of Galerkin type for computing low rank solutions of large and sparse Lyapunov matrix equations. His methods are used for such large and sparse Lyapunov matrix equations which have stable matrix coefficient

A, i.e., all eigenvalues of matrix A have negative real parts, and B be a vector in $\mathbb{R}^n$. The preconditioned Krylov subspace methods for large and sparse lypunov matrix equations is presented by M. Hochbruck and G. Starke [10]. They constructed SSOR and ADI(r) preconditioners for accelerating the rate of convergence of Krylov subspace methods (see [13]). In addition, for solving Lyapunov matrix equations, they used coupled two-term recurrence version of QMR iterative method without look-ahead which have proposed by Freund and Nachtigal [7].

Recently J. R. Singler [14] proposed low rank solution of Lyapunov matrix equations via proper orthogonal decomposition (POD). In this paper, we use this idea for obtaining the numerical solution of the generalized continuous -time Lyapunov equation (GCLE)

$$A^T X E + E^T X A = -BB^T, \tag{1}$$

where $A, E \in R^{n\times n}$ and $B \in R^{n\times m}$ are given matrices.

The POD-based algorithm discussed here also computes an approximate low rank solutions of stable Lyapunov equations.Unlike many other large-scale algorithms, the POD-based approach is not iterative; instead, the solution is constructed by simulating m linear differential equations, where m is the rank of B, and then computing POD eigenvalues and mods. The main computational cost of algorithm is approximating the solutions of the linear differential equations. Thus,this algorithm is applicable to large-scale systems when the rank of B is relatively small.

Throughout this paper, we let H be Hilbert space with inner product $(.,.)$ and corresponding norm $\| \,.\, \|_H = (.,.)^{\frac{1}{2}}$. For the matrix Lyapunov equation, H is taken to be $\mathbb{R}^n$ and inner product can be taken as standard as the standard dot product, $(a,b) = a^T b$, or a weighted dot product, $(a,b) = a^T M b$,where $M \in R^{n\times n}$ is symmetric positive definite.

This paper is organized as follows. In Section 2, a brief description of basic properties of generalized Lyapunov matrix equations is given, and the POD based algorithm for generalized Lyapunov matrix equations is presented in Section 3. In Section 4 the approximation theory and error bounds are summarized. Finally, Section 5 summarizes the main conclusion of this paper.

2 Basic Properties of Generalized Lyapunov Matrix Equations

In the past few years its generalizations, the generalized continuous -time Lyapunov equation (1). The right hand side $-BB^T$ is symmetric and so is the solution matrix X if the equation has a unique solution. We can consider the

GCLE as special of generalized Sylvester equation

$$R^T XS + U^T XV = -Y, \tag{2}$$

where in general, X and Y are $n \times m$ matrices. As we know, the generalized Sylvester matrix equation (2) can be written as a big linear system of equations

$$(S^T \otimes R^T + V^T \otimes U^T)vec(X) = -vec(Y)$$

where $\otimes$ denote the Kronecker product, g_{ij}, are entries of the given $n \times m$ matrix G, and

$$vec(G) = (g_{11}, ..., g_{n1}, g_{12}, ..., g_{n2}, ..., g_{1m}, ..., g_{nm})^T.$$

The solvability of (2) depends on the generelized eigenstructure of the matrix pairs (R, U) and (V, S). A matrix pencil $\alpha R - \beta U$ is called regular iff there exists a pair of complex number $(\acute{\alpha}, \acute{\beta})$ such that $\acute{\alpha} R - \acute{\beta} U$ is nonsingular. If $\alpha Rx = \beta Ux$ holds for a vector $x \neq 0$, the pair $(\alpha, \beta) \neq (0, 0)$ is a generalized eigenvalue. Two generalized eigenvalue (α, β) and (γ, δ) are considered to be equal iff $\alpha\delta = \beta\gamma$. The set of all generalized eigenvalues of the pencil $\alpha R - \beta U$ is designated by $\sigma(R, U)$.

Theorem 1. The matrix equation (2) has a unique solution if and only if
1.$\alpha R - \beta U$ and $\alpha V - \beta S$ are pencils and
2.$\sigma(R, U) \bigcap \sigma(-V, S) = \emptyset$.

Proof : See [2].

Let (α_i, β_i) denote the generalized eigenvalues of the matrix pencil $\alpha A - \beta E$. For simplicity, we switch to the more conventional form of a matrix pencil $A - \lambda E$, whose eigenvalues are given by $\lambda_i = \frac{\beta_i}{\alpha_i}$ with $\lambda_i = \infty$ when $\alpha_i = 0$. Applying the above theorem to (1) give the following corollary.

corollary 1. Let $A - \lambda E$ be a regular pencil. Then the GCLE (1) has a unique solution if and only if all eigenvalues of $A - \lambda E$ are finite and $\lambda_i + \lambda_j$ for any two eigenvalues λ_i and λ_j of $A - \lambda E$.

Proof : See [2].

As a consequence, singularity of one of the matrices A and E implies singularity of the GCLE (1). Since A and E play a symmetric role, we expect E to be invertible. Under this assumption equations (1) is equivalent to Lyapunov equations

$$(AE^{-1})^T X + X(AE^{-1}) = -(E^{-T}B)(E^{-T}B)^T \tag{3}$$

and the classical result about the positive definite solution of the stable Lyapunov equation [15] remains valid for the generalized equations.

Theorem 2. Let E be nonsingular, BB^T be positive definite (semidefinite) and $Re(\lambda_i) < 0$ for all eigenvalue λ_i of $A - \lambda E$, then the solution matrix X of the GCLE (1) is positive definite (semidefinite).

Proof : See [15].

3 POD Based Algorithm For Generalized Lyapunov Matrix Equations

Let $L^2(0,\infty;H)$ be the set of all functions w such that $w(t) \in H$ for all $t \geq 0$ and whose H norm is square integrable, i.e.,

$$\| w \|_{L^2(0,\infty;H)} = (\int_0^\infty \| w(t) \|_H^2 \, dt)^{\frac{1}{2}} < \infty.$$

A sequence of functions $\{w_k\} \subset L^2(0,\infty;H)$ converge to $w \in L^2(0,\infty;H)$ if $\| w_k - w \|_{L^2(0,\infty;H)} \longrightarrow 0$ as $k \longrightarrow \infty$.

Definition 1: The continuous POD operator $X : H \longrightarrow H$ for a dataset $\{w_j\}_{j=1}^m \subset L^2(0,\infty;H)$ is defined by

$$Xz = \int_0^\infty \sum_{j=1}^m (z, w_j(t)) w_j(t) dt.$$

The continuous POD operator is self adjoint, compact, and nonnegative; thus, the eigenvalues of X may be ordered $\lambda_1 \geq \lambda_2 \geq ... \geq 0$ and the corresponding orthonormal eigenvectors $\{\varphi_k\} \subset H$ form a compact set.

Definition 2: The eigenvalues $\{\lambda_k\}$ the continuous POD operator X are called the POD eigenvalues of $\{w_j\}$ and the orthonormal eigenvectors of $\{\varphi_k\} \subset H$ are called the POD modes of $\{w_j\}$.

An important feature of proper orthogonal decomposition is that the POD eigenvalues and modes of a time varying dataset $\{w_j\}_{j=1}^m \subset L^2(0,\infty;H)$ can computed using the method on snapshots. The main idea is to approximate each w_j with functions whose POD eigenvalues and modes are easily computable.

Theorem 3. Let $w_j^N \in L^2(0,\infty;H)$ be a sequence functions converging to $w_j \in L^2(0,\infty;H)$ for $j = 1, ..., m$. Let $\{\lambda_k^N, \varphi_k^N\}$ and $\{\lambda_k, \varphi_k\}$ denote the POD eigenvalues and modes of $\{w_j^N\}_{j=1}^m$ and $\{w_j\}_{j=1}^m$, respectively. Then for each k,

$$lim_{N\longrightarrow\infty} \mid \lambda_k^N - \lambda_k \mid = 0, \quad lim_{N\longrightarrow\infty} \| \varphi_k^N - \varphi_k \|_H = 0.$$

Furthermore, as $N \longrightarrow \infty$,

$$\sum_{k\geq 1} \lambda_k^N \longrightarrow \sum_{k\geq 1} \lambda_k.$$

Proof : See [14].

A popular approach to the method of snapshots is to use piecewise constant functions (in time) to approximate the functions w_j. For simplicity, we focus on the case $m = 1$, i.e., there is only one function in the dataset. The algorithm is similar for $m > 1$.

Method of Snapshots (for $m = 1$**):**
1. Let $a_j \approx w(t_j)$ be approximate of $w(t)$ at times $0 = t_0 < t_1 < ... < t_N = T$ for $j = 0, ..., N$.
2. Let $v_j = \frac{a_j + a_{j-1}}{2}$ be the approximate average value of $w(t)$ over the jth time interval for $j = 0, ..., N$.
3. Let $\delta_j = t_j - t_{j-1}$ be the jth time step for $j = 0, ..., N$.
4. Let Γ be the symmetric $N \times N$ matrix whose entries are the inner products $\Gamma_{ij} = (\delta_j^{\frac{1}{2}} v_j, \delta_i^{\frac{1}{2}} v_i)$.
5. Let $\lambda_1 \geq \lambda_2 \geq ... \geq \lambda_N \geq 0$ be the ordered eigenvalues of Γ with corresponding orthonormal eigenvectors $\{\gamma_k\}_{k=1}^{N}$.
6. The (approximate) POD eigenvalues are given by $\{\lambda_k\}$ and, if $\lambda_k \neq 0$, the (approximate) kth POD mode is

$$\varphi_k = \lambda_k^{\frac{-1}{2}} \sum_{j=1}^{N} \delta_j^{\frac{1}{2}} [\gamma_k]_j v_j,$$

where $[\gamma_k]_j$ is the jth component of γ_k.
We suppose AE^{-1} and $E^{-T}B$ have the following properties. In the matrix case, $AE^{-1} \in R^{n \times n}$ is exponentially stable and $E^{-T}B \in R^{n \times m}$. In the infinite dimensional case, $AE^{-1} : D(AE^{-1}) \subset H \longrightarrow H$ generate an exponentially stable C_0- semigroup $e^{AE^{-1}t}$ over H and $E^{-T}B : R^m \longrightarrow H$ is finite rank and bonded. This assumption implies that $E^{-T}B$ must be take the form

$$E^{-T}Bu = \sum_{j=1}^{m} eb_j u_j,$$

where each eb_j in H and $u = [u_1, ..., u_m]^T \in R^m$ (see [16]). Note that this representation for$E^{-T}B$ also holds for the matrix problem; in this case, eb_j is the jth column of $E^{-T}B$.
The exact solution $X : H \longrightarrow H$ of the Lyapunov equation (3) is given by [3]

$$Xz = \int_0^{\infty} e^{(AE^{-1})^T t} (E^{-T}B)(E^{-T}B)^T e^{AE^{-1}t} z dt.$$

We now show that the Lyapunov solution equals the continues POD operator for dataset $\{w_j\}$ given in the main algorithm.
Proposition 1: The unique solution $X : H \longrightarrow H$ of the Lyapunov equation (3) takes the form

$$Xz = \int_0^{\infty} \sum_{j=1}^{m} (z, w_j(t)) w_j(t) dt, \tag{4}$$

where each w_j is exact solution of the linear evolution equation

$$\dot{w}_j(t) = AE^{-1} w_j(t), \quad w_j(0) = eb_j, \quad j = 1, ..., m. \tag{5}$$

Proof :
The solution may be factored as $X = PP^*$, where $P:L^2(0,\infty;R^m) \longrightarrow H$ is defined by

$$Pu = \int_0^\infty e^{(AE^{-1})^T t}(E^{-T}B)u(t)dt.$$

and $X = P^* : H \longrightarrow L^2(0,\infty;R^m)$, the adjoint of P, is given by $P^*z = (E^{-T}B)^* e^{(AE^{-1})^* t}$. Again, given the assumptions above on $E^{-T}B$, the operator must have the form

$$E^{-T}Bu = \sum_{j=1}^{m} eb_j u_j,$$

where $u = [u_1, ..., u_m]^T \in R^m$, and each eb_j is in H. then we have
P u$= \int_0^\infty e^{(AE^{-1})^T t}(E^{-T}B)u(t)dt = \int_0^\infty \sum_{j=1}^m u_j(t)w_j(t)dt$, where $w_j(t) = e^{(AE^{-1})^T t}eb_j$ for $j = 1, ..., m$. This implies that each $w_j \in L^2(0,\infty;H)$ is the solution of the linear evolution equation (5). The adjoint operator P^*: $H \longrightarrow L^2(0,\infty;R^m)$, is easily computed to be

$$[Pz](t) = [(z, w_1(t)), ..., (z, w_m(t))]^T$$

.
again using $X = PP^*$ gives the expression (4).
□
Since the Lyapunov equation X equal the continuous POD operator for $\{w_j\}$, the POD eigenvalues and modes equal, by definition, the eigenvalues and orthonormal eigenvectors of X.
Corollary 2: Let $w_j \in L^2(0,\infty;H)$ be the exact solution of the linear evolution equation (5), for $j = 1, ..., m$. The POD eigenvalues $\{\lambda_k\}$ and modes $\{\varphi_k\} \subset H$ of the dataset $\{w_j\}$are the eigenvalues and orthonormal eigenvectors of the unique solution $X : H \longrightarrow H$ of the Lyapunov equation (3).
The truncated eigenvalue expansion of X is given by

$$X_r z = \sum_{k=1}^{r} \lambda_k (z, \varphi_k(t))\varphi_k. \tag{6}$$

The algorithm to approximate the solution $X : H \longrightarrow H$ of the Lyapunov equation (3) can be briefly summarized as follows.

POD Based Algorithm
1. Let w_j^N be an approximation to the solution w_j of the linear differential equation (5).
2. Compute $\{\lambda_k^N\}$ and $\{\varphi_k^N\}$, the POD eigenvalues and modes of the dataset $\{w_k^N\}_{j=1}^m$,e.g., by method of snapshots.
3. Choose r and form the rth order approximate Lyapunov solution X_r^N :

$H \longrightarrow H$ given by

$$X_r^N z = \sum_{j=1}^{m} \lambda_k^N (z, \varphi_k^N)\varphi_k^N, \tag{7}$$

where $(.,.)$ is inner product over the Hilbert space.
The choice of the order r and the approximation level N will discuss in Section 4.

4 Approximation Theory And Error Bounds

Let K be a compact linear operator from a Hilbert space H_1 to a Hilbert space H_2. The operator norm of K is defined by

$$\| K \| = sup_{0 \neq x \in H_1} \frac{\| Kx \|}{\| x \|} = \sigma_1,$$

where $\sigma_1 \geq \sigma_2 \geq ... \geq 0$ denote the singular value of K,i.e.,

$$\| K \|_{tr} = \sum_{k \geq 1} \sigma_k.$$

A best rank r approximating, K_r, to K is given by a solution of the following problem: find the minimnizer over all rank operator F_r of the operator norm error $\| K - F_r \|$. A solution of this problem (which may not be unique)is given by the rth order truncated singular value decomposition of K. the best value of the operator norm error is σ_{r+1}, the first neglected singular value. The truncated singular value decomposition also gives a best rank r approximation of K if the norm is taken to be the trace norm. in this case, the best trace norm is given by $\sum_{k>r} \sigma_k$, the sum of the neglected singular values. In this work, many of the operators we consider map a Hilbert space into itself and are compact, self adjoin, and nonnegative. The eigenvalues of such an operator can be ordered $\lambda_1 \geq \lambda_2 \geq ... \geq 0$. Furthermore, the eigenvalues are equal to the singular values and the truncated eigenvalue expansion is equal to the truncated singular value decomposition. Thus, the truncated eigenvalue expansion provides the best low rank approximation in this case.
We let $\{w_j^N\}_{j=1}^m$ be approximations of solutions $\{w_j\}_{j=1}^m$ of solutions of the differential equations (5).We let $\{\lambda_k^N, \varphi_k^N\}$ and $\{\lambda_k, \varphi_k\}$ denote the POD eigenvalues and modes of $\{w_j^N\}$ and $\{w_j\}$, respectively. Also, $X_r^N : H \longrightarrow H$ as define in (7) denote the approximate Lyapunov solution and . $X_r : H \longrightarrow H$ define in (6) denote the rth order truncated eigenvalue expansion of the Lyapunov solution.
Theory 4: Let r be given. Suppose for $j = 1, ..., m$, $w_j^N \longrightarrow w_j$ in $L^2(0, \infty; H)$ as $N \longrightarrow \infty$. Then $\lambda_k^N \longrightarrow \lambda_k$ and $\varphi_k^N \longrightarrow \varphi_k$ for $1 \leq k \leq r$. Also, as the POD eigenvalues and modes converge,$X_r^N \longrightarrow X_r$ in the operator norm.

Proof : See [14].
Theory 5: The operator norm error between X_r^N and X, the exact solution of the Lyapunov equation(3), is bounded as follows:

$$\| X - X_r^N \| \leq \lambda_{r+1} + \sum_{k=1}^{r} (| \lambda_k - \lambda_k^N |) + 2\lambda_k^N \| \varphi_k - \varphi_k^N \|_H).$$

Proof : See [14].
Theory 5: The trace norm error between X_r^N and X, the exact solution of the Lyapunov equation(1), is bounded as follows:

$$\| X - X_r^N \|_{tr} \leq \sum_{k>r} \lambda_k^N + C^N (\sum_{j=1}^{m} \| w_j - w_j^N \|_{L^2_{(0,\infty;H)}})^{\frac{1}{2}},$$

where

$$C^N = (\sum_{k\geq 1} \lambda_k^N)^{\frac{1}{2}} + (\sum_{k\geq 1} \lambda_k)^{\frac{1}{2}}.$$

Proof : See [14].
As each $w_j^N \longrightarrow w_j$ in $L^2_{(0,\infty;H)}$, the last term in the error bounds tend to zero, and also,

$$\sum_{k>r} \lambda_k^N \longrightarrow \sum_{k>r} \lambda_k, \quad C^N \longrightarrow 2(\sum_{k\geq 1} \lambda_k)^{\frac{1}{2}}.$$

We note that both terms in the error bound can be approximately computed or estimated.

5 Conclusion

We generalized the POD-based algorithm to to compute approximate low rank solutions of generalized Lyapunov equations. The algorithm is applicable to large-scale matrix problems as well as a class of infinite dimensional problems. The quality of the approximate solution can be ascertained by simple, computable a priori error bounds.

References

[1] R. H. Bartels and G. W. Stewart. Algorithm 432. Solution of matrix equation AX+XB=C. Comm. ACM 15:820-826, 1972.

[2] K. W. E. Chu, The solution of the matrix equation $AXB-CXD = Y$ and $(YA - DZ, YC - BZ) = (E, F)$, Linear Algebra Appl.,93:93-105,1987.

[3] R. F. Curtain and H. j. Zwart, An Introduction to Infinite-Dimensional Linear System Theory. New York:Springer-Verlag, 1995.

[4] B. N. Datta and K. Datta, Theoretical and computational aspects of some linear algebra problems in control theory. In computational and combinatorical methods in systems Theory (C. I. Byrnes and A. Lindquist,Eds.), Elsevier, Amsterdam. 201-212, 1986.

[5] B. A. Francis, A Course in H_∞ Contorol Theory. New York: Springer-Verlag, 1987.

[6] K. Glover, All optimal Hankel-norm approximations of linear multivariable system and thier $L^\infty-$ error bounds, Int. J. Control, vol. 39, pp. 1115-1193, 1984.

[7] R. W. Freund and N. M. Nachtigal. An implemetation of the QMR methodbased on coupled two-term recurrences. Technical Report 92.15,RIACS, June 1992.

[8] G. H. Golub and C. F. Van Loan. A Hessenberg-Schur method for the problem $AX - XB = C$. IEE Trans. Automat. Control., 19:167-188, 1979.

[9] S. Hammarling, Numerical solution of the stable, non-negetive definite Lyapunov equation, IMA J. Numer. Anal., 2 pp. 303-323, 1982.

[10] M. Hockhbruck, G. starke. Preconditioned Krylov subspace methods for Lyapunov matrix equations,Siam J. Matrix anal appl.16(1), (1995) 156-171.

[11] A. J. Laub, M. T. Heath, C. C. Paige, and R. C. Ward, Computation of sysrem balancing transformations and other applications of simultaneous diagonalization algorithms. IEEE Trans. Automat. Contr., vol AC-32, pp 115-122, 1987.

[12] Y. Saad, Numerical solution of large Lyapanov equations, in Signal Processing, Scattering, Operator Theory and Numerical Methods(M. A. Kaashoek, J. H. van Schuppen and A.C.M. ran, eds.), Birkhauser, Boston, 1990 pp.503-511.

[13] V. Simoncini, A new iterative method for solving Large-scale Lyapunov matrix equations, SIAM J. Sci, Compute., vol. 29, no. 3, pp. 1268-1288,2007.

[14] J. R. Singler, Approximate low rank solutions of Lyapunov equations via proper orthogonal decomposition, Proceedings of the 2008 American Control Conference, June 2008, pp. 267-272.

[15] J. Snyder, M. Zakai, On non-negetive solution of the equation $AD + DA^T = -C$, SIAM J. Appl. Math., 18:704-714,1970.

[16] J. Weidmann, Linear operator in Hilbert spaces. New York:Springer-Verlag, 1980.

Received: May, 2009