Proceedings of the 19th International Symposium on Mathematical Theory of Networks and Systems – MTNS 2010 • 5–9 July, 2010 • Budapest, Hungary

Numerical Algorithm for Structured Low Rank Approximation Problem

S. R. Khare*, H. K. Pillai, M. N. Belur

***Abstract*— In this paper we discuss an important problem of Structured Low Rank Approximation (SLRA) of linearly structured matrices. This is a very important problem having many applications like computation of approximate GCD, model order reduction to name a few. In this paper we formulate SLRA problem as an unconstrained optimization problem on a smooth matrix manifold. We use Armijo line search algorithm on the matrix manifold to compute the nearest SLRA of the given matrix.**

***Index Terms*— Singular Value Decomposition (SVD), Structured Low Rank Approximation (SLRA), Optimization on Matrix Manifolds**

I. INTRODUCTION

The problem of Structured Low Rank Approximation (SLRA) of a given structured matrix has found a lot of applications in the literature: approximate GCD of univariate polynomials, model order reduction to mention a few among many. The SLRA problem deals with finding the nearest rank deficient matrix with the same structure. Clearly the usual approach of singular value decomposition fails as the structure of the matrix is not preserved. Several formulations are discussed in the literature to solve this problem, for instance optimization formulation and lift and project algorithm as in [1], structured total least squares formulation as in [2], [3], [4]. Several algorithms can be found in [5], [6], [7], [8] based on techniques of structured total least squares to compute the nearest SLRA of the Sylvester matrix to compute an approximate GCD of given univariate polynomials. Further similar results about Hankel matrices can be found in [9]. An exhaustive literature survey of this field can be found in [10].

In this paper we use a formulation of the SLRA problem as in [2], [3]. However we use a different approach to solve the optimization problem. The paper is organized as follows. In Section II we formulate the problem formally. In Section III we discuss a procedure to compute the nearest SLRA based on Lagrange Multiplier approach. Further some numerical examples are discussed in Section IV. Finally we conclude in Section V.

II. PROBLEM FORMULATION

A subset $\Omega \subset \mathbb{R}^{p\times q}$ is called a linear structure of matrices if Ω is a linear subspace of $\mathbb{R}^{p\times q}$. Let $B = \{B_1, B_2, \dots, B_N\}$ be a basis of Ω. In general, $N < pq$. However $N = pq$ implies that $\Omega = R^{p\times q}$, that is there is no structure involved. Now we define SLRA problem as it is defined in [1].

Problem Statement 2.1: Given $\Omega \subset \mathbb{R}^{p\times q}$, a linear structure, and $X \in \Omega$ such that $\mathrm{rank}(X) = k$ for $k \le \min\{p,q\}$, find a matrix Y such that

$$\min_{Y\in\Omega,\mathrm{rank}(Y)=k-1} \|X-Y\|_F .$$

In this paper, we consider the Frobenius norm as the matrix norm. However in the problem definition above, one can use any matrix norm. We now give another optimization formulation of the SLRA problem as discussed in [2], [3]. Let $X \in \Omega$ be given as $X = \sum_{i=1}^N x_i B_i$. WLOG we assume $p \ge q$. Let $\mathrm{rank}(X) = n$. Then to find $Y \in \Omega$ such that $Y = \sum_{i=1}^N y_i B_i$ such that

$$\min_{y_i,v} \sum_{i=1}^{N} c(B_i)(x_i - y_i)^2 \quad (1)$$

subject to

$$\left(\sum_{i=1}^{N} y_i B_i\right) v = 0 \ , \ v^T v = 1.$$

where $c : B \longrightarrow \mathbb{R}_+$ is a function. This function c relates the cost function in terms of vectors $x = [x_1, x_2, \dots, x_N]^T$ and $y = [y_1, y_2, \dots, y_N]^T$ to the Frobenius norm of the difference of the matrices X and Y. We illustrate this in the following example. Let $\Omega \subset \mathbb{R}^{3\times 3}$ be the set of all symmetric matrices. Then $\dim(\Omega) = 6$. A basis $B = \{B_1, B_2, \dots, B_6\}$ that we choose is as follows.

$$B_1 = \begin{bmatrix}1&0&0\\0&0&0\\0&0&0\end{bmatrix},\ B_2 = \begin{bmatrix}0&1&0\\1&0&0\\0&0&0\end{bmatrix},\ B_3 = \begin{bmatrix}0&0&1\\0&0&0\\1&0&0\end{bmatrix},$$

$$B_4 = \begin{bmatrix}0&0&0\\0&1&0\\0&0&0\end{bmatrix},\ B_5 = \begin{bmatrix}0&0&0\\0&0&1\\0&1&0\end{bmatrix},\ B_6 = \begin{bmatrix}0&0&0\\0&0&0\\0&0&1\end{bmatrix}$$

Here the function c is defined as follows:

$$c(B_1) = 1,\ c(B_2) = 2,\ c(B_3) = 2,$$
$$c(B_4) = 1,\ c(B_5) = 2,\ c(B_6) = 1.$$

Then it is clear that $\sum_{i=1}^N c(B_i)(x_i - y_i)^2 = \|X - Y\|_F^2$.

We now explain lift and project algorithm as explained in [1]. The output of this algorithm is used as an initial stage for Armijo line search algorithm to be used in the next section. The lift and project algorithm does not yield the nearest SLRA, however we use the output of this algorithm as an initial guess for the method that we propose. Before writing the algorithm formally, we describe the idea behind the algorithm briefly. Let $X \in \mathbb{R}^{p\times q}$ be a structured matrix

*Corresponding Author

The authors are with the Department of Electrical Engineering, Indian Institute of Technology Bombay, Powai, Mumbai, India 400076

E-mail ids: S. R. Khare swanand@ee.iitb.ac.in, H. K. Pillai hp@ee.iitb.ac.in, M. N. Belur belur@ee.iitb.ac.in

ISBN 978-963-311-370-7

of rank r. Then the nearest rank $r-1$ approximation of X is computed using the SVD of X. However this destroys the structure of the matrix. So this low rank approximation is projected back onto the space of structured matrices. This procedure is iterated until one gets the structured low rank approximation. This procedure can shown to be a descent method and hence the convergence of this procedure is guaranteed. Now we write the algorithm formally. Let $\mathbf{P}$ be the projection operator defined on the subspace of matrices with given structure, that is $\mathbf{P}: \mathbb{R}^{p\times q} \to \Omega$.

Algorithm 2.2: Lift and Project Algorithm for SLRA

Input: $X \in \Omega$, a rank r matrix
Output: $\tilde{X} \in \Omega$, a rank $r-1$ matrix

Initialize $\tilde{X} = X$
while rank $(\tilde{X}) == r$
  Compute the SVD of X as $X = \sum_{i=1}^{r} \sigma_i u_i v_i^T$
  Compute the low rank approximation as
  $\hat{X} = \sum_{i=1}^{r-1} \sigma_i u_i v_i^T$
  $\tilde{X} = \mathbf{P}\hat{X}$
end while

III. Algorithm to compute the nearest SLRA

We now propose a method to solve this optimization problem. Though the formulation is same as in the paper of [2], the approach to the solve the problem are different. In this reference the optimization formulation is shown to be equivalent to the generalized nonlinear eigenvalue problem. We break the optimization problem (1) into two nested optimization problems as follows:

$$\min_{v\in\mathbb{R}^q} \left\{ \begin{array}{c} \min_{y_i} \sum_{i=1}^{N} c(B_i)(x_i-y_i)^2 \\ \text{subject to} \\ \left(\sum_{i=1}^{N} y_i B_i\right) v = 0 \end{array} \right\}$$
$$\text{subject to}$$
$$v^T v = 1.$$

The optimization problem inside the braces is called the inner optimization problem. This inner optimization is shown to have a closed form solution. This solution can be completely expressed in terms of the optimization variable v of the outer optimization. In order to solve the inner optimization problem we use Lagrange multiplier approach. Thus the inner optimization problem, I, becomes:

$$I: \min_{y_i,\ell} \sum_{i=1}^{N} c(B_i)(x_i-y_i)^2 + \ell^T \left(\sum_{i=1}^{N} y_i B_i\right) v$$

where ℓ is a vector of Lagrange multipliers. Now differentiating with respect to y_i and ℓ we get,

$$\text{wrt } y_i: \; y_i = x_i - \frac{1}{2c(B_i)} \ell^T B_i v \text{ for } i = 1,2,\ldots,N \quad (2)$$

$$\text{wrt } \ell: \; \left(\sum_{i=1}^{N} y_i B_i\right) v = 0. \quad (3)$$

Substituting for y_i from equation (2) in equation (3), we get,

$$\left(\sum_{i=1}^{N} (x_i - \frac{1}{2c(B_i)} \ell^T B_i v) B_i\right) v = 0$$
$$\Rightarrow \quad \left(\sum_{i=1}^{N} x_i B_i\right) v = \sum_{i=1}^{N} \frac{1}{2c(B_i)} \ell^T B_i v B_i v$$
$$\Rightarrow \quad Xv = D_v \ell \quad (4)$$

where D_v is defined as

$$D_v \;=\; \sum_{i=1}^{N} \frac{1}{2c(B_i)} B_i v (B_i v)^T. \quad (5)$$

Note that D_v is a symmetric nonnegative definite matrix. In order to compute ℓ we solve linear system (4) to get

$$\ell = D_v^{-1} X v. \quad (6)$$

Substituting ℓ from (6) in equation (2), we get,

$$y_i = x_i - \frac{1}{2c(B_i)} v^T X^T D_v^{-1} B_i v \text{ for } i = 1,2,\ldots,N \quad (7)$$

Thus the optimal value for the inner optimization problem I is given by,

$$\sum_{i=1}^{N} c(B_i)(x_i-y_i)^2 = \sum_{i=1}^{N} (v^T X^T D_v^{-1} B_i v)^2. \quad (8)$$

The outer optimization can be stated completely in the optimization variable v as follows:

$$\min_{v\in\mathbb{R}^q} \sum_{i=1}^{N} \left(v^T X^T D_v^{-1} B_i v\right)^2 \quad (9)$$
$$\text{subject to}$$
$$v^T v = 1.$$

Notice that the constraint set of this optimization problem is S^{q-1}, a unit sphere in $\mathbb{R}^q$, a smooth manifold. Alternatively we can view this constrained optimization problem as a non constrained optimization problem on the manifold S^{q-1}. We use the gradient search algorithm (see [11]) to solve this problem. Before proceeding further, we state the unconstrained optimization problem and expression for gradient of the cost function that is required in the gradient search algorithm.

$$\min_{v\in S^{q-1}} f(v) \quad (10)$$

where

$$f(v) = \sum_{i=1}^{N} \left(v^T X^T D_v^{-1} B_i v\right)^2. \quad (11)$$

For a given matrix $M(v)$ of size $m\times m$, consider $M(k,j) = m_{kj}(v)$ is a function of $v = [v_1, v_2, \ldots, v_n]^T$. Then the notation $\frac{d}{dv_i}M$ means the following:

$$\frac{d}{dv_i} M = \left[\frac{d}{dv_i} m_{kj}(v)\right]_{k,j=1,\ldots,m} \quad (12)$$

The gradient of the cost function, denoted by $\nabla_v f$, is given by,

$$\nabla_v f = \quad 2\sum_{i=1}^{N}\left(v^T X^T D_v^{-1} B_i v\right) \times \left\{X^T D_v^{-1} B_i v + B_i^T D_v^{-1} X v - \begin{bmatrix} v^T X^T D_v^{-1}\left(\frac{d}{dv_1} D_v\right) D_v^{-1} B_i v \\ v^T X^T D_v^{-1}\left(\frac{d}{dv_2} D_v\right) D_v^{-1} B_i v \\ \vdots \\ v^T X^T D_v^{-1}\left(\frac{d}{dv_n} D_v\right) D_v^{-1} B_i v \end{bmatrix}\right\} \tag{13}$$

The gradient search algorithm is implemented as in [11, Chapter 4].

IV. Examples

Example 4.1: In this example, we consider a Hankel matrix of size 6×6. We compute the nearest SLRA to the given Hankel matrix. The dimension of space of Hankel matrices of size $n \times n$ in the space of $\mathbb{R}^{n\times n}$ is $2n-1$. Hence any Hankel matrix of size $n \times n$ can be represented by a vector in $\mathbb{R}^{2n-1}$. In this case, consider the given Hankel matrix, H, be represented as a vector $h \in \mathbb{R}^{11}$. Let $h = [\ 3.1472\ \ 4.0579\ \ -3.7301\ \ 4.1338\ \ 1.3236\ \ -4.0246\ \ -2.2150\ \ 0.4688\ \ 4.5751\ \ 4.6489\ \ -3.4239\ ]^T$ denote the given Hankel matrix. The nearest rank deficient Hankel matrix, $\tilde{H}$, is computed and corresponding vector $\tilde{h}$ obtained as $\tilde{h} = [\ 3.1489\ \ 4.0560\ \ -3.7309\ \ 4.1323\ \ 1.3283\ \ -4.0284\ \ -2.2071\ \ 0.4674\ \ 4.6121\ \ 4.6183\ \ -3.3530\ ]^T$. Note that $\|H - \tilde{H}\|_F = 0.1074$ and the smallest singular value of $\tilde{H}$ is 9.6469×10^{-16}.

Example 4.2: In this example, we consider a matrix $A \in \mathbb{R}^{5\times 5}$. Since we do not consider any structure on A, the dimension of the space under consideration is $5 \times 5 = 25$. The nearest low rank approximation of A is computed using the proposed algorithm. The results of the proposed algorithm match with the results obtained from singular value decomposition of A.

V. Concluding Remarks

In this paper we proposed a numerical algorithm to compute the nearest SLRA of a given structured low rank approximation of the given linearly structured matrix. In order to compute the nearest SLRA we formulated the problem as an unconstrained optimization problem on a matrix manifold, namely the unit sphere, S^N, in $\mathbb{R}^N$. The optimization problem on the sphere is solved using a line search algorithm as proposed in [11]. The numerical results are obtained for Hankel structured, a special linearly structured and unstructured matrices.

References

[1] M. T. Chu, R. E. Funderlic, and R. J. Plemmons, "Structured low rank approximation," *Linear Algebra and Its Applications*, vol. 366, pp. 157–172, 2003.

[2] B. D. Moor, "Structured total least squares and l_2 approximation problems," *Linear Algebra and its Applications*, vol. 188,189, pp. 163–205, 1993.

[3] B. D. Moor, "Total least squares for affinely structured matrices and the noisy realization problem," *IEEE Transactions on Signal Processing*, vol. 42, pp. 3104–3114, 1994.

[4] J. B. Rosen, H. Park, and J. Glick, "Total least norm formulation and solution for structured problems," *SIAM Journal of Matrix Analysis and Applications*, vol. 17, pp. 110–126, 1996.

[5] B. Li, Z. Yang, and L. Zhi, "Fast low rank approximation of a sylvester matrix by structured total least norm," *Japan Society for Symbolic and Algebraic Computation*, vol. 11, pp. 165–174, 2005.

[6] E. Kaltofen, Z. Yang, and L. Zhi, "Structured low rank approximation of a sylvester matrix," *Symbolic Numeric Computation, D. Wang and L. Zhi Editors, Trends in Mathematics*, pp. 69–83, 2006.

[7] J. Winkler and J. Allan, "Structured low rank approximations of the sylvester resultant matrix for approximate gcds of bernstein basis polynomials," *Electronic Transctions on Numerical Analysis*, vol. 31, pp. 141–155, 2008.

[8] J. Winkler and J. Allan, "Structured total least norm and approximate gcds of inexact polynomials," *Journal of Computational and Applied Mathematics*, vol. 215, pp. 1–13, 2008.

[9] M. A. Hitz, "On computing nearest singular hankel matrices," in *Proceedings of the 2005 international symposium on Symbolic and algebraic computation, Beijing, China*, pp. 171–176, 2005.

[10] I. Markovski, "Structured low rank approximation and its applications," *Automatica*, vol. 44, pp. 891–909, 2008.

[11] P. A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princetion University Press, 2008.