Lanczos Algorithm on the Grassmann manifold

Ana Paula Lopes[1], Ana Julia Viamonte[2], and Antonio Jose Pascoal[2]

[1] ISCAP, Porto, Portugal
aplopes@iscap.ipp.pt
[2] Universidade Portucalense, Porto, Portugal
ajs@upt.pt; ajfp@upt.pt

Abstract. In this paper we developed a new Lanczos algorithm on the Grassmann manifold. This work comes in the wake of the article by A. Edelman, T. A. Arias and S. T. Smith, The geometry of algorithms with orthogonality constraints [EAS98], where they presented a new conjugate gradient algorithm on the Grassmann and Stiefel manifolds. These manifolds which are based on orthogonality constraints, yields penetrating insight into many numerical algorithms of linear algebra. They have developed an approach to numerical algorithms involving orthogonality constraints. As the Lanczos method and the method of conjugate gradients are closely related, and one of the main problems of the Lanczos method is the loss of orthogonality, arose the idea of checking whether it would be possible to get a Lanczos algorithm on the Grassmann manifold.

1 Introduction

The problem of computing eigenvalues, eigenvectors and invariant subspaces is always present in areas as diverse as Engineering, Physics, Computer Science and Mathematics. Considering the importance of these problems in many practical applications, it is not surprising that has been and continues to be the subject of intense research. Lately, it has been verified that the iterations of eigenvalue and eigenvector problems are best analyzed in some special spaces. This geometric approach has helped us to analyse methods such as the QR algorithm, but however, it hasn't led to new algorithms largely due to the difficulty in establishing a bridge between the geometry of the abstract spaces and the well known algorithms of numerical linear algebra. The optimization problem of the estimative of the invariant subspaces is made explicit with a geometric approach. However a geometrical treatment of Grassmann manifold appropriate for numerical linear algebra is not present in standard references. The iterations start from an initial estimative and produce a sequence of p-dimensional subspaces which become increasingly better estimative of the invariant subspaces. Thus successive iterates belong to a particular space, formed by all the p-dimensional subspaces of R^n, called Grassmann manifold, denoted by $Gr(p, n)$.

The concept of the fiber bundle structure that allows us to describe the relationship between subspaces and bases and also allows us to explain how objects

in geometric abstract of the Grassmann manifold can be translated into numerical formulas. One of the objectives of this work is to prove that a geometric approach could lead to a new efficient algorithm, even in problems as classics as computing invariant subspaces, due to the fiber bundle structure on the Grassmann manifold. We develop a Restarted Block Lanczos algorithm, and we show that this algorithm is an iteration on Gr(p,n) and is competitive when compared with other algorithms.

2 Grassmann and Stiefel Manifolds in Numerical Linear Algebra

In standard differential geometry references we cannot found an appropriate geometrical treatment of the Stiefel and Grassmann manifolds for numerical linear algebra. It is essential to understand these manifolds, which represent orthogonality constraints, as in the symmetric eigenvalue problem and for that we must understand the relationship between the geometric entities and the numerical representation of these quantities. If we wish to represent a subspace we can pick a basis of vectors that span the subspace. The relationship between subspaces and bases is done by the principal fiber bundle structure on the Grassmann manifold. The Grassmann manifold is the set of all p-dimensional subspaces of R^n aand is represented by $Gr(p,n)$. One possible way to represent numerically an element y of $Gr(p,n)$ i e, a p-dimensional subspace of R^n , consists of specifying an $n \times p$ full column rank matrix Y whose columns span the space y, and we can write $Y = span(y)$. The span of full rank $n \times p$ matrix is an element of $Gr(p,n)$ if and only if Y has full rank.

The set of $n \times p$ matrices with full rank is called the non compact Stiefel manifold and is denoted by $ST(p,n)$, i.,

$$ST(p,n) = \{Y \in R^{n \times p} : r(Y) = p\} \tag{1}$$

We denote by $St(p,n)$, termed the Stiefel manifold, the set of all $n \times p$ orthogonal real matrices

$$St(p,n) = \{Y \in R^{n \times p} : Y^T Y = I_p\} \tag{2}$$

The set of $n \times n$ invertible matrices is denoted by GL_n,

$$GL_n = \{A \in R^{n \times n} : det(a) \neq 0\} \tag{3}$$

The Grassmann manifold can be identified with the quotient space

$$Gr(p,n) = ST(p,n)/GL_p \tag{4}$$

For details concerning Stiefel manifolds, Grassmann manifolds, please refer to standard texts such as [Boo75].

3 Lanczos on Grassmann Manifold

The Lanczos algorithm is a method for computing some eigenvalues of a large symmetric matrix A and their eigenvectors. The idea consists in building a sequence of nested subspaces span$\{\nu, A\nu, A^2\nu, \ldots\}$ and solving the eigenproblem reduced to these subspaces. We will generalize this method and its variants to the Grassmann manifold and show that such generalization is not useless as a numerical algorithm. The block restarted Lanczos is a method on Grassmann because if we record the subspaces we get just after restarting, then we obtain a sequence of subspaces with the same dimension. Until now, as far as we know, people haven't though of Lanczos as a subpace iteration, so it may be interesting to point out that restarted Lanczos is a Grassmann iteration. We intend to obtain new convergence results and make extensive comparisons with other Grassmannian methods like the ones mentioned in ([1998]). We propose a new algorithm called Grassmann-Lanczos, which induces a subspace iteration, i.e, an iteration on the Grassmann manifold, that can be written as follows.

4 Algorithm Grassmann-Lanczos (GL)

Let A be an symmetric matrix $n \times n$

Consider y_0 an p-dimensional subspace of IR^n, i.e.,$y_0 \in Gr(p,n)$. The algorithm produce a sequence of subspaces

$$\begin{array}{ccc} Gr(p,n) & \rightarrow & Gr(p,n) \\ y & \rightarrow & y_+ \end{array}$$

defined by:

- Input(A,Y_1, m)
- For i=1 to m do
  - For j=1 to i do
    - $B_{i\ j} = Y_i^t A Y_j$
  - $M_i = Y_i^t A Y_i$
  - $Z_{i+1} = AY_i - Y_i\ M_i - \sum_{j=1}^{i-1} Y_i B_{i\ j}$
  - Compute QR Decomposition of Z_{i+1}: $Z_{i+1} = Q_j R_j = Y_{i+1} R_{i+1}$
- $Q = (Y_1, Y_2, \ldots, Y_m)$
- $M = Q^t A Q$
- $X = Z_{i+1}$
- $y_+ = span(QX)$

The matrix Y_1, $n \times p$, is that $y \in span(Y_1)$. The algorithm is defined as a mapping $Gr(p,n) \rightarrow Gr(p,n)$ and this algorithm is well-defined, in other words, the output $Y_+ \in Gr(p,n)$ (a p-dimensional subspace) is not affected by the choices made in the computation process. In order to prove this, we assume

that make another choice in step (1.). We obtain a matrix $\ddot{Y} = Y * R$, for some non-singular $p \times p$ matrix R. It is easy to see that at point (2.), we will get a Krylov subspace $K_m(\ddot{Y}) = K_n(Y)$ .Then we take an orthogonal basis, for instance we will get a basis $\ddot{Q} = Q * R$ for some orthogonal $p \times p$ matrix R. At point (3.), you compute

$$\ddot{M} = \ddot{Q}^T A \ddot{Q} \tag{5}$$

$$= (QR)^T A(QR) \tag{6}$$

$$= R^T Q^T AQR \tag{7}$$

$$= R^T MR \tag{8}$$

Notice that M and $\ddot{M}$ have the same spectrum, and the eigenvectores of $\ddot{M}$ are those of M after left multiplication by R^T . Therefore, the dominated eigenspace of $\ddot{M}$ is equal to R^T multiplication by the dominated eigenspace of M. Consequently, at point (4.), your $\ddot{X}$ satisfies $\ddot{X} = R^T * X * S$ where S is another $p \times p$ orthogonal matrix. Finally, in (5.), we have

$$\ddot{Q} * \ddot{X} = Q * R * R^T * X * S = Q * X * S \tag{9}$$

Consequently, $\ddot{Q} * \ddot{X}$ and $Q * X$ have the same column space.

This shows that none of the choices we made affected the output of the iteration mapping.

Observe that the subspaces

$$S_k = span\{Y, AY, A^2Y, \ldots, A^{k-1}Y\} \tag{10}$$

and

$$S'_k = span\{Y, G, AG, A^2G, \ldots, A^{k-1}G\} \tag{11}$$

with

$$G = AY - Y(Y^T AY) \tag{12}$$

$$= (I - YY^T)AY \tag{13}$$

$$= (I - Y(Y^T Y)^{-1}Y^T)AY \tag{14}$$

are identical. If we consider the orthogonal projector onto the orthogonal complement to the span of the Y defined by

$$\Pi_{Y_+} = I - Y(Y^T Y)^{-1}Y^T \tag{15}$$

Then, we can write equation (14) as

$$G = \Pi_{Y_+} AY \tag{16}$$

and

$$S'_k = span\{Y, \Pi_{Y_+} AY, A\Pi_{Y_+} AY, \ldots, A^{k-1}\Pi_{Y_+} AY\} \tag{17}$$

5 Parallel Grassmann-Lanczos Algorithm (PGL)

Analyzing the GL algorithm, and as this method is well-defined because the output $Y_* \in Gr(p, n)$ (a p-dimensional subspace) is not affected by the choices made in the computation process, we conclude that this method might be implemented in parallel computing. In a message passing programming model we can distribute the large vectors and/or matrices on the processors while the small ones kept on just one processor. Each Y_i code can be Computed independently from the others, we only need do the comunication to calculate $AY - I$.

6 Conclusions and Future Works

This paper offers a new approach to the Lanczos algorithm. Lanczos algorithm is very competitive method whose main problem is the loss of orthogonality, but the introduction of Grassman manifolds seems to resolve this problem ([1998])]. Analyzing the algorithm that we have proposed, and as this method is well-defined because the output $Y_* \in Gr(p, n)$ (a p-dimensional subspace) is not affected by the choices made in the computation process, we conclude that this method might be easily implemented by blocks and in parallel computing. Thus, all the data has lead us to conclude that this method is not only competitive in sequential computation, but also in parallel computing. Further experimentation may be needed in practice. If one is really considering the pure linear symmetric eigenvalue problem, then we think that pure conjugate gradient procedures must be inferior to Lanczos procedures. In order to remove losses of orthogonality we try an idealized Grassman algorithm and we hope to prove this is a better way, because we think that these algorithms should be understood in the correct geometrical framework following the ideas of Demmel and Parlett.

References

[2004] Absil, P. A., Sepulchre, R., Van Dooren P., Mahony,R. : Cubically convergent iterations for invariant subspace computation. SIAM J. Matrix Analysis vol 26, 1 (2004) 70–96

[2003] Baglama, J. , Calvetti D. , Reichel, L. : IRBL: An Implicitly Restarted Block Lanczos Method for Large-Scale Hermitian Eigenproblems. SIAM J. Sci. Comput. Vol. 24, n5 (2003) 1650–1677

[1975] Boothby, W. M. : An introduction to differentiable manifolds and Riemannian geometry. Academic Press (1975)
[1985] Cullum, J. C.,Willoughby,R. A. : Lanczos Algoritms for Large Symmetric Eigenvalue Computations. Vol.I and Vol. II, Birkhauser.
[1987] Demmel, J.W.: Three Methods for finding estimates of invariant subspaces. Computing. 38, 43–57.
[1998] Edelman, A., Arias, T.A., Smith,S. T.: The Geometry of Algorithms with Orthogonality Constraints. SIAM J. Matrix Anal. Appl. 20, n2(1998) 303–353
[2006] Lopes, A.P.: O mtodo de Lanczos na Variedade de Grassman. PhD theses Universidade Portucalense (2006)
[1998] Parlett, B. N. : The symmetric eigenvalue problem. Prentice Hall series in Computational Mathematics.