The common invariant subspace problem: an approach via Gröbner bases

Donu Arapura a, Chris Peterson b, *

a Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA
b Department of Mathematics, Colorado State University, Fort Collins, CO 80523, USA
Received 24 March 2003; accepted 17 November 2003
Submitted by R.A. Brualdi

Abstract

Let $A$ be an $n \times n$ matrix. It is a relatively simple process to construct a homogeneous ideal (generated by quadrics) whose associated projective variety parametrizes the one-dimensional invariant subspaces of $A$. Given a finite collection of $n \times n$ matrices, one can similarly construct a homogeneous ideal (again generated by quadrics) whose associated projective variety parametrizes the one-dimensional subspaces which are invariant subspaces for every member of the collection. Gröbner basis techniques then provide a finite, rational algorithm to determine how many points are on this variety. In other words, a finite, rational algorithm is given to determine both the existence and quantity of common one-dimensional invariant subspaces to a set of matrices. This is then extended, for each $d$, to an algorithm to determine both the existence and quantity of common $d$-dimensional invariant subspaces to a set of matrices.

© 2004 Published by Elsevier Inc.

Keywords: Eigenvector; Invariant subspace; Grassmann variety; Gröbner basis; Algorithm

1. Introduction

Let $A$ and $B$ be two $n \times n$ matrices. In 1984, Dan Shemesh gave a very nice procedure for constructing the maximal $A, B$ invariant subspace, $\mathcal{V}$, on which $A$ and $B$ commute. He then showed that if $\mathcal{V} \neq 0$ then $A$ and $B$ share a common eigenvector [6].

* Corresponding author.
E-mail addresses: dvb@math.purdue.edu (D. Arapura), peterson@math.colostate.edu (C. Peterson).

0024-3795/$ - see front matter © 2004 Published by Elsevier Inc.
doi:10.1016/j.laa.2003.03.001
Theorem 1 (Shemesh). Let $A$ and $B$ be two $n \times n$ matrices. Let $[A, B] = AB - BA$. Then $A$ and $B$ share a common eigenvector if and only if $\mathcal{N} \neq 0$ where

$$\mathcal{N} = \bigcap_{k,l=1}^{n-1} \ker[A^k, B^l].$$

It is clear from the theorem that one can determine if two matrices share a common eigenvector using a finite number of rational operations.

Several attempts were made to extend Shemesh’s result to a procedure for detecting when two matrices, $A$ and $B$, share a common invariant subspace [1,4,7]. In [1], the problem was completely solved when the rank of the invariant subspace was restricted to be two. In [4], there were no restrictions on the rank of the invariant subspace but the eigenvalues of $A$ were required to be distinct. In [7], an algorithm was proposed that could be carried through in a wide variety of circumstances but certain aspects of the algorithm remained incomplete. In each of these papers, one reduces the common invariant sub-space problem to the common eigenvector problem by considering $r$th exterior powers of the matrices (i.e. $r$th compound matrices). This is a very natural approach but some difficulties arise due to the method of detecting common eigenvectors. Namely, in the method given by Shemesh’s Theorem, it is difficult to determine if the detected eigenvectors satisfy the Plücker conditions.

In this paper, we take a completely different approach to the common eigenvector problem. In extending the method to the $r$th exterior powers of the matrices, enough information is retained to allow a complete solution to the problem of detecting the existence of common invariant subspaces and to the problem of determining the number of common invariant subspaces.

Our approach is as follows. Let $A_1, A_2, \ldots, A_t$ be a collection $n \times n$ matrices. We construct a homogeneous ideal $I_{A_1, A_2, \ldots, A_t}$ with the property that the projective variety associated to the ideal parametrizes the one-dimensional subspaces which are invariant subspaces for every member of the collection. Gröbner basis techniques can then be used to determine various properties of this variety including its degree and dimension. To determine if $A_1, A_2, \ldots, A_t$ share a common eigenvector one only needs to determine if the projective variety (in $\mathbb{P}^{n-1}$) defined by the homogeneous ideal $I_{A_1, A_2, \ldots, A_t}$ contains any points. To extend our method to the common invariant subspace problem, we first find the one-dimensional subspaces which are invariant spaces for each of the $r$th exterior powers of $A_1, A_2, \ldots, A_t$. Then we determine which of these common one-dimensional invariant subspaces satisfy the Plücker conditions. This is done by intersecting the projective variety parametrizing all common one-dimensional invariant subspaces of the $r$th exterior powers with an appropriate Grassmann variety.

2. Construction of $I_{A_1, A_2, \ldots, A_t}$ and the Grassmann variety

Let $A$ be an $n \times n$ matrix. Let $x$ be an $n \times 1$ matrix of indeterminates. $x$ is an eigenvector of $A$ (or the zero vector) if and only if the two vectors $Ax$ and $x$ are
linearly dependent. In other words, \( x \) is an eigenvector of \( A \) (or the zero vector) if and only if the matrix \( M = [Ax|x] \) has rank less than 2.

**Definition 2.** Let \( A \) be an \( n \times n \) matrix. Let \( x \) be an \( n \times 1 \) matrix of indeterminates. \( I_A \) is defined to be the ideal generated by the \( 2 \times 2 \) minors of the matrix \( M = [Ax|x] \).

Let \( \mathbb{A}^n_k \) denote affine \( n \)-space over an algebraically closed field \( k \) and let \( \mathbb{P}^n_k \) denote projective \( n \)-space over \( k \). If all the entries of the matrix \( A \) lie in \( k \) then the affine variety defined by \( I_A \) lies in \( \mathbb{A}^n_k \).

**Proposition 3.** The affine variety defined by \( I_A \) is the union of all the eigenvectors of \( A \) (and the zero vector).

**Proof.** \( x \) is an eigenvector of \( A \) (or the zero vector) if and only if all of the \( 2 \times 2 \) minors of \( M \) are zero. \( \Box \)

**Remark 4.** Since the ideal, \( I_A \), is generated by homogeneous polynomials, we can also view the ideal as defining a projective variety in \( \mathbb{P}^n_k \). Each one-dimensional invariant subspace will correspond to a point on this variety.

**Example 5.** If

\[
A = \begin{bmatrix}
0 & 0 & 3 \\
1 & 4 & 5 \\
0 & 6 & 2
\end{bmatrix}
\quad \text{and} \quad
x = \begin{bmatrix}
x \\
y \\
z
\end{bmatrix},
\]

then

\[
M = [Ax|x] = \begin{bmatrix}
3z & x \\
x + 4y + 5z & y \\
6y + 2z & z
\end{bmatrix}
\]

and

\[
I_A = (3yz - x^2 - 4xy - 5xz, 3z^2 - 6xy - 2xz, xz + 2yz + 5z^2 - 6y^2) \subset \mathbb{C}[x, y, z].
\]

\( I_A \) defines three lines through the origin in \( \mathbb{A}^3_k \) or equivalently, defines 3 points in \( \mathbb{P}^2_k \).

**Definition 6.** Let \( A_1, A_2, \ldots, A_t \) be a collection of \( n \times n \) matrices. \( I_{A_1,A_2,\ldots,A_t} \) is defined to be the ideal generated by the union of the generators of \( I_{A_1}, I_{A_2}, \ldots, I_{A_t} \).

In a manner similar to before, if the entries of \( A_1, A_2, \ldots, A_t \) lie in an algebraically closed field \( k \) then \( I_{A_1,A_2,\ldots,A_t} \) determines an affine variety in \( \mathbb{A}^n_k \) and a projective variety in \( \mathbb{P}^n_k \).
Proposition 7. The variety defined by $I_{A_1, A_2, \ldots, A_t}$ is the intersection of the varieties defined by $I_{A_1}, I_{A_2}, \ldots, I_{A_t}$.

Proof. This follows immediately from the definition of $I_{A_1, A_2, \ldots, A_t}$. See for instance [2, Chapter 4.3, pp. 181–189]. □

Thus, the set of matrices $A_1, A_2, \ldots, A_t$ have a common eigenvector if and only if the projective variety defined by $I_{A_1, A_2, \ldots, A_t}$ is not the empty set. Equivalently, $A_1, A_2, \ldots, A_t$ have a common eigenvector if and only if the affine variety defined by $I_{A_1, A_2, \ldots, A_t}$ has points other than the origin. Standard Gröbner basis techniques can be used to determine if the projective variety defined by $I_{A_1, A_2, \ldots, A_t}$ is the empty set. The variety can be decomposed into equidimensional components by a finite rational algorithm [3]. Since both the operation of taking the radical and the operation of determining an equidimensional decomposition can be carried out via a finite rational algorithm, if the coefficients of the generators of $I_{A_1, A_2, \ldots, A_t}$ all lie in a field $k' \subseteq k$ then the ideal of each of the equidimensional components can be generated by forms all of whose coefficients also lie in $k'$.

Definition 8. Let $V$ be a vector space consisting of the zero vector and eigenvectors common to a set of matrices $A_1, A_2, \ldots, A_t$. Then $V$ is a maximal space of common eigenvectors if it is not a proper subspace of any other vector space consisting of the zero vector and eigenvectors common to $A_1, A_2, \ldots, A_t$.

The degree of the pure $d$-dimensional component of the projective variety is equal to the number of maximal spaces of common eigenvectors of dimension $d + 1$. The degree of each component can be determined by standard Gröbner basis techniques. In summary, there is a finite rational algorithm for determining if a set of matrices share a common eigenvector. Furthermore, for each $d$ there is a finite rational algorithm for determining the number of $d$-dimensional maximal spaces of common eigenvectors to a set of matrices.

Let $A$ be a matrix and let $V$ be an invariant subspace of $A$ (i.e. $AV \subseteq V$). It is well known that the operation of taking the $d$th exterior power (i.e. the $d$th compound) commutes with matrix multiplication. Thus, if we take the $d$th exterior power of the equation $AV \subseteq V$ then we obtain $(AV)^{(d)} = A^{(d)}V^{(d)} \subseteq V^{(d)}$. If the dimension of $V$ is $d$ then this equation simplifies to $A^{(d)}L \subseteq L$ where $L$ is a one-dimensional subspace. In other words, $L$ will be a one-dimensional invariant subspace of $A$. Hence, a necessary condition for two matrices, $A$ and $B$, to have an invariant subspace of dimension $d$ is for $A$ and $B$ to have a common eigenvector. However, even if $A$ and $B$ have a common eigenvector, we still cannot conclude that $A$ and $B$ have a common invariant subspace of dimension $d$. We need the common eigenvector of $A$ and $B$ to be a $d$th exterior power of a $d$-dimensional space. In other words, we need the point in projective space determined by $L$ to lie on an appropriate Grassmann Variety.
Definition/Theorem 9. Let \( G(n, r) \) denote the Grassmann variety of all \( r \)-dimensional linear subspaces of the vector space \( V = k^n \). Let \( N = \binom{n}{r} - 1 \). There is a natural inclusion of \( G(n, r) \) as a projective variety in \( \mathbb{P}^N \) via the Plücker embedding. In this embedding, the Grassmann variety is the common zero locus of a set of quadratic polynomials (known as the Plücker relations).

The Plücker relations can be computed by hand or they can be computed symbolically by the following method. Let \( X \) be the \( \binom{n}{r} \times 1 \) matrix whose \((i, 1)\)th entry is the variable \( A_i \). Let \( M \) be the \( n \times r \) matrix whose \((i, j)\)th entry is the variable \( a(i,j) \). Let \( M^{(r)} \) denote the \( r \)th exterior power of \( M \). Let \( P = X - M^{(r)} \) and let \( I \) be the ideal whose generators are the entries of \( P \). Thus \( I \subseteq k[A_1, A_2, \ldots, A_{\binom{n}{r}}, a(1,1), \ldots, a(n,r)] \). Let \( J_P = I \cap k[A_1, A_2, \ldots, A_{\binom{n}{r}}] \). Then \( J_P \) is the homogeneous ideal defining \( G(n, r) \). The generators of \( J_P \) are the Plücker relations. For more information about Grassmann varieties, see [5], Chapter 6.

3. Algorithms

Algorithm for determining the number of maximal common eigenspaces of dimension \( d \).

- Input: \( A_1, A_2, \ldots, A_t \).
- Compute \( I = I_{A_1,A_2,\ldots,A_t} \).
- Compute the radical of \( I \), call it \( \text{Rad}(I) \).
- If \( \text{Rad}(I) \neq m \) (\( m \) denotes the maximal ideal) then \( A_1, A_2, \ldots, A_t \) have a common eigenvector.
- Decompose \( \text{Rad}(I) \) into its equidimensional components, \( C_0, C_1, \ldots \) where \( C_i \) defines an equidimensional projective variety of dimension \( i \).
- For each \( i \), let \( d_i = \deg(C_i) \).
- \( d_i \) is the number of \( i + 1 \)-dimensional maximal common eigenspaces.

Algorithm for determining the existence of common invariant subspaces of dimension \( d \).

- Input: \( A_1, A_2, \ldots, A_t \).
- Compute the \( d \)th exterior powers \( A_1^{(d)}, A_2^{(d)}, \ldots, A_t^{(d)} \) of the matrices.
- Compute \( I = I_{A_1^{(d)},A_2^{(d)},\ldots,A_t^{(d)}} \).
- Compute the generators of the homogeneous ideal of \( GR(n, d) \), call them \( G_1, G_2, \ldots \).
- Let \( J \) be the ideal whose generators consist of the generators of \( I \) together with the equations \( G_1, G_2, \ldots \).
- Compute the radical of \( J \), call it \( \text{Rad}(J) \).
• If \( \text{Rad}(J) \neq m \) (\( m \) denotes the maximal ideal) then \( A_1, A_2, \ldots, A_t \) have a common invariant subspace of dimension \( d \).

**Remark 10.** In the proceeding algorithm, the projective variety defined by \( \text{Rad}(J) \) gives information as to the number of different common invariant subspaces of dimension \( d \). For instance, if the variety is zero-dimensional, then the number of different common invariant subspaces of dimension \( d \) is equal to the degree of the variety. If the dimension of the variety is larger than zero then there is a family of common invariant subspaces of dimension \( d \). If one is only interested in detecting the existence of common invariant subspaces of dimension \( d \) then it is not necessary to compute the radical of \( J \). It is enough to determine if the projective scheme defined by \( J \) consists of any points. This is a less expensive computation.

**Example 11.** Let

\[
A = \begin{bmatrix} 2 & 1 & 0 & 1 \\ 0 & 3 & 0 & 1 \\ 1 & 0 & 3 & 0 \\ 1 & 0 & 1 & 2 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} -1 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 1 & 0 & -1 \end{bmatrix}.
\]

Using the first algorithm, one finds that \( I_{A,B} \) defines a projective variety of degree 2 and dimension 0 in \( \mathbb{P}^3 \) thus there are two one-dimensional maximal common eigenspaces. Using the second algorithm, one finds that \( I_{A(2),B(2)} \) defines a projective variety of which two points lie on GR(4,2) thus there are 2 common invariant subspaces of dimension 2. Using again the second algorithm, one finds that \( I_{A(3),B(3)} \) defines a projective variety of which two points lie on GR(4, 3) thus there are two common invariant subspaces of dimension 3.

**4. Final remarks**

The algorithms outlined in the previous section are all finite rational algorithms. As a consequence, if the field generated by the entries of \( A_1, A_2, \ldots, A_t \) is \( k' \) then every ideal appearing in the algorithms has a generating set whose coefficients all lie in \( k' \). This has implications for when a basis can be chosen for an invariant subspace such that the entries in the basis vectors all lie in \( k' \). For instance, if there is a unique invariant subspace of dimension \( d \) then a basis for this subspace can be chosen inside of \( k' \). In general, one should not expect to find a basis inside of \( k' \) but rather in some algebraic extension. It should also be remarked that actual bases for the subspaces involved cannot, in general, be determined from these algorithms. To determine bases for the subspaces involved would require carrying out a primary decomposition of the ideals that appear in the algorithms. There are various methods for carrying out partial primary decompositions (via symbolic methods) and these
can be applied to the problem at hand. One case where a full primary decomposition can be attained is when the degrees of the varieties involved are less than 5. This is related to the problem of finding exact representations of the roots of a degree \(d\) polynomial, sometimes it can be done and sometimes it cannot be done. It can always be done for a polynomial of degree less than 5.

In the first algorithm, if one is only interested in establishing the existence of a common eigenvector then it is not necessary to compute the radical of \(I\). It is enough to determine the dimension of the scheme defined by \(I\). Since \(I\) will be generated by quadrics, this computation can frequently be done in a timely manner. Similarly, in the second algorithm, it is not necessary to compute the radical of \(J\). It is enough to determine the dimension of the scheme defined by \(J\). Since \(J\) will also be generated by quadrics, again, this computation can frequently be done in a timely manner.

Acknowledgement

The authors would like to thank the NSF for partial support of this project.

References