

A criterion for the existence of common invariant subspaces of matrices

Michael Tsatsomeros *

April 16, 2001

Abstract

It is shown that square matrices A and B have a common invariant subspace W of dimension $k \geq 1$ if and only if for some scalar s , $A + sI$ and $B + sI$ are invertible and their k -th compounds have a common eigenvector, which is a Grassmann representative for W . The applicability of this criterion and its ability to yield a basis for the common invariant subspace are investigated.

Key words: invariant subspace, compound matrix, decomposable vector, Grassmann space

AMS subject classifications: 15A75, 47A15

1 Introduction

The lattice of invariant subspaces of a matrix is a well studied concept with several applications. In particular, there is substantial information known about matrices having the same invariant subspaces, or having chains of common invariant subspaces (see [7]). The question of existence of a non-trivial invariant subspace common to two or more matrices is also of interest. It arises, among other places, in the study of irreducibility of decomposable positive linear maps on operator algebras [3]. It also comes up in what might be called Burnside's Theorem, namely that if two square complex matrices do not have a common non-trivial invariant subspace, then the algebra that they generate is the entire matrix algebra. The practical aspects of determining whether two matrices have a common invariant subspace of dimension $k \geq 1$ are considered in [5] where a necessary condition is provided; this condition is also shown to be sufficient under certain assumptions. In what follows, we extend the above results by identifying a necessary and sufficient condition for the existence of a common invariant subspace of dimension k . We also describe a strategy for checking this condition and for determining a basis for a common invariant subspace.

*Department of Mathematics and Statistics, University of Regina, Regina, Saskatchewan, Canada S4S 0A2. Research partially supported by NSERC research grant.

We begin with some basic facts from multilinear algebra that can be found in [4, 11]. For positive integers $k \leq n$, let $\langle n \rangle = \{1, \dots, n\}$ and denote by $Q_{k,n}$ the k -tuples of $\langle n \rangle$ with elements in increasing order. The members of $Q_{k,n}$ are considered ordered lexicographically.

For any matrix $X \in \mathbf{C}^{m \times n}$ and nonempty $\alpha \subseteq \langle m \rangle$, $\beta \subseteq \langle n \rangle$, let $X[\alpha | \beta]$ denote the submatrix of X in rows and columns indexed by α and β , respectively. Given an integer $0 < k \leq \min(m, n)$, the k -th compound of X is defined as the $\binom{m}{k} \times \binom{n}{k}$ matrix $X^{(k)} = (\det X[\alpha | \beta])_{\alpha \in Q_{k,m}, \beta \in Q_{k,n}}$. Matrix compounds satisfy $(XY)^{(k)} = X^{(k)}Y^{(k)}$. The exterior product of $x_i \in \mathbf{C}^n$, $i = 1, \dots, k$, denoted by $x_1 \wedge \dots \wedge x_k$, is the $\binom{n}{k}$ -component vector equal to the k -th compound of $X = [x_1 | \dots | x_k]$; that is,

$$x_1 \wedge \dots \wedge x_k = X^{(k)} = (\det X[\alpha | \langle k \rangle])_{\alpha \in Q_{k,n}}.$$

Consequently, if $A \in \mathbf{C}^{n \times n}$ and $0 < k \leq n$, the first column of $A^{(k)}$ is precisely the exterior product of the first k columns of A . Exterior products satisfy the following:

$$(1) \quad x_1 \wedge \dots \wedge x_k = 0 \iff x_1, \dots, x_k \text{ are linearly dependent.}$$

$$(2) \quad \mu_1 x_1 \wedge \dots \wedge \mu_k x_k = \prod_{i=1}^k \mu_i (x_1 \wedge \dots \wedge x_k) \quad (\mu_i \in \mathbf{C}).$$

$$(3) \quad A^{(k)}(x_1 \wedge \dots \wedge x_k) = Ax_1 \wedge \dots \wedge Ax_k.$$

When a vector $x \in \mathbf{C}^{\binom{n}{k}}$ is viewed as a member of the k -th Grassmann space over \mathbf{C}^n , it is called *decomposable* if $x = x_1 \wedge \dots \wedge x_k$ for some $x_i \in \mathbf{C}^n$, $i = 1, \dots, k$; we refer to x_1, \dots, x_k as the *factors* of x . By conditions (2) and (3), those decomposable vectors whose factors are linearly independent eigenvectors of $A \in \mathbf{C}^{n \times n}$ are eigenvectors of $A^{(k)}$. The spectrum of $A^{(k)}$ coincides with the set of all possible k -products of the eigenvalues of A . In general, not all eigenvectors of a matrix compound are decomposable (see [13] for an example).

Consider now a k -dimensional subspace $W \subseteq \mathbf{C}^n$ spanned by $\{x_1, \dots, x_k\}$. By (1) we have that

$$W = \{x \in \mathbf{C}^{\binom{n}{k}}; x \wedge x_1 \wedge \dots \wedge x_k = 0\}.$$

The vector $x_1 \wedge \dots \wedge x_k$ is known as a *Grassmann representative for W* . It follows that two k -dimensional subspaces spanned by $\{x_1, \dots, x_k\}$ and $\{y_1, \dots, y_k\}$, respectively, coincide if and only if for some nonzero $\mu \in \mathbf{C}$,

$$(4) \quad x_1 \wedge \dots \wedge x_k = \mu (y_1 \wedge \dots \wedge y_k);$$

that is, Grassmann representatives for a subspace differ only by a nonzero scalar factor.

2 The criterion

Recall that W is an *invariant subspace* of $A \in \mathbf{C}^{n \times n}$ if $Ax \in W$ for all $x \in W$. As $\{0\}$ and \mathbf{C}^n are trivially invariant subspaces of every matrix in $\mathbf{C}^{n \times n}$, we restrict our attention to non-trivial invariant subspaces. First is a slight refinement of the necessary condition for the existence of

a common non-trivial invariant subspace shown in [5]; its proof is included for completeness and in order to note that the common eigenvector of the compounds mentioned is decomposable, its factors being basis vectors of the common invariant subspace.

Theorem 2.1 *Let W be a common invariant subspace of $A, B \in \mathbf{C}^{n \times n}$ with basis $\{x_1, \dots, x_k\}$, $1 \leq k < n$. Then the following hold.*

- (i) [5] *For all $s \in \mathbf{C}$, $(A + sI)^{(k)}, (B + sI)^{(k)}$ have a common eigenvector x .*
- (ii) *The eigenvector x in (i) is a Grassmann representative for W , i.e., $x = x_1 \wedge \dots \wedge x_k$.*

Proof. As A and B have a common invariant subspace W if and only if $A + sI$ and $B + sI$ do, we need only show the conclusions when $s = 0$. Let $Q \in \mathbf{C}^{n \times n}$ be an invertible matrix whose first k columns are the basis vectors of W , $\{x_1, \dots, x_k\}$. Then

$$R = Q^{-1}AQ = \begin{pmatrix} R_{11} & R_{12} \\ 0 & R_{22} \end{pmatrix} \quad \text{and} \quad S = Q^{-1}BQ = \begin{pmatrix} S_{11} & S_{12} \\ 0 & S_{22} \end{pmatrix},$$

where $S_{11}, R_{11} \in \mathbf{C}^{k \times k}$. It follows that

$$R^{(k)} = (Q^{(k)})^{-1}A^{(k)}Q^{(k)} \quad \text{and} \quad S^{(k)} = (Q^{(k)})^{-1}B^{(k)}Q^{(k)}.$$

Because of the block upper triangular structure of R , in the first column of $R^{(k)}$ only the diagonal entry may be nonzero. Hence the first column of $Q^{(k)}$, call it x , is an eigenvector of $A^{(k)}$ and, similarly, an eigenvector of $B^{(k)}$. By definition of Q and the k -th compound, $x = x_1 \wedge \dots \wedge x_k$. ■

We can now prove the following necessary and sufficient conditions for the existence of a common non-trivial invariant subspace.

Theorem 2.2 *Let $A, B \in \mathbf{C}^{n \times n}$ and W be a subspace of \mathbf{C}^n of dimension k , $1 \leq k < n$. The following are equivalent.*

- (i) *W is a common invariant subspace of A and B .*
- (ii) *There exists decomposable $x \in \mathbf{C}^{\binom{n}{k}}$ such that for all $s \in \mathbf{C}$, x is a common eigenvector of $(A + sI)^{(k)}$ and $(B + sI)^{(k)}$.*
- (iii) *There exist decomposable $x \in \mathbf{C}^{\binom{n}{k}}$ and $\hat{s} \in \mathbf{C}$ such that $A + \hat{s}I$ and $B + \hat{s}I$ are invertible and x is a common eigenvector of $(A + \hat{s}I)^{(k)}$ and $(B + \hat{s}I)^{(k)}$.*

The vector x in (ii) and (iii) is a Grassmann representative for W .

Proof. The implication (i) \Rightarrow (ii) follows from Theorem 2.1. The implication (ii) \Rightarrow (iii) is trivial. To show (iii) \Rightarrow (i), let \hat{s} be such that $A + \hat{s}I, B + \hat{s}I$ are invertible. It follows that their k -th compounds are also invertible. Thus there exist nonzero $\lambda \in \mathbf{C}$ and $x_1, \dots, x_k \in \mathbf{C}^n$ such that $x = x_1 \wedge \dots \wedge x_k \neq 0$ and

$$(A + \hat{s}I)^{(k)}x = (A + \hat{s}I)x_1 \wedge \dots \wedge (A + \hat{s}I)x_k = \lambda (x_1 \wedge \dots \wedge x_k).$$

That is, by (4) and the discussion preceding it,

$$\text{span}\{x_1, \dots, x_k\} = \text{span}\{(A + \hat{s}I)x_1, \dots, (A + \hat{s}I)x_k\}.$$

Similarly,

$$\text{span}\{x_1, \dots, x_k\} = \text{span}\{(B + \hat{s}I)x_1, \dots, (B + \hat{s}I)x_k\}.$$

It follows that $W = \text{span}\{x_1, \dots, x_k\}$ is a common invariant subspace of $A + \hat{s}I$ and $B + \hat{s}I$, and thus a common invariant subspace of A and B , having x as its Grassmann representative. ■

Notice that once the existence of a common invariant subspace for two matrices is established, this subspace is left invariant by all analytic functions of the two matrices, as well as all words formed from the two matrices. It is also clear that the above theorem can be extended to the case of any number of matrices having a common invariant subspace.

In the sufficient condition for the existence of a common invariant subspace of A and B of dimension $2 \leq k < n$ provided in [5, Theorem 3.1], it is assumed that B is invertible and $A^{(k)}$ has distinct eigenvalues. When $2 \leq k < n$, the latter assumption implies indeed that the eigenvalues of A are also nonzero and distinct. In turn, all the eigenvectors of $A^{(k)}$ must be decomposable. In other words, the sufficient condition in [5] is superseded by the ‘only if’ part of the following necessary and sufficient condition.

Corollary 2.3 *Let $A, B \in \mathbf{C}^{n \times n}$ be invertible. Then A, B have a common invariant subspace of dimension k , $1 \leq k < n$, if and only if $A^{(k)}, B^{(k)}$ have a common decomposable eigenvector.*

Proof. The necessity of the compounds having a common decomposable eigenvector follows from Theorem 2.1 and the sufficiency from (iii) \Rightarrow (i) of Theorem 2.2 with $\hat{s} = 0$. ■

A related problem that has recently arisen in mathematical biology [2] poses a natural matrix theoretic question, namely, when do two matrices have a common invariant convex cone? By a *convex cone* we mean a subset of \mathbf{R}^n that contains all nonnegative linear combinations of given vectors. The *dimension* of a cone Γ is the dimension of its span, $\Gamma - \Gamma$. In connection with the common invariant subspace problem, we have the following result.

Corollary 2.4 *Let Γ be cone of dimension k and suppose that $A\Gamma \subseteq \Gamma$, $B\Gamma \subseteq \Gamma$, where $A, B \in \mathbf{C}^{n \times n}$. Then $A^{(k)}$ and $B^{(k)}$ have a common decomposable eigenvector that is a Grassmann representative of $\Gamma - \Gamma$.*

Proof. Notice that $W = \Gamma - \Gamma$ is a subspace of dimension k left invariant by A and B . The conclusions follow readily from Theorem 2.2. ■

Certainly, as every subspace is a convex cone, the interest in the above corollary is when Γ is not a subspace but rather a *pointed* cone (i.e., $\Gamma \cap (-\Gamma) = \{0\}$) or a *proper* cone (i.e., closed, pointed, with non-empty interior). *Conditions that are necessary and sufficient for two matrices to have a common pointed or proper cone invariant are not known.*

3 Implementation of the criterion

The algorithm provided in [5] resolves the problem of existence of a common invariant subspace when one of the matrices has distinct eigenvalues (requiring in the process a finite number of rational computations). Our results lead to a plausible strategy for all matrices, including the opportunity of providing a basis for the common invariant subspace. As in [5], use will be made of the following criterion for the existence of a common eigenvector.

Theorem 3.1 ([12]) *Let $X, Y \in \mathbf{C}^{p \times p}$ and*

$$K = \sum_{m, \ell=1}^{p-1} [X^m, Y^\ell]^* [X^m, Y^\ell],$$

where $[X^m, Y^\ell]$ denotes the commutator $X^m Y^\ell - Y^\ell X^m$. Then X and Y have a common eigenvector if and only if K is not invertible.

Consider now the following plan for discovering a k -dimensional ($1 \leq k < n$) common invariant subspace of given matrices $A, B \in \mathbf{C}^{n \times n}$.

1. Find s such that $A + sI$ and $B + sI$ are invertible.
2. Compute $X = (A + sI)^{(k)}$, $Y = (B + sI)^{(k)}$ and K as in Theorem 3.1.
3. If K is invertible, A and B do not have a common invariant subspace of dimension k .
4. Otherwise, compute bases for the intersections of eigenspaces of X and Y .
5. If the intersecting eigenspaces of X and Y contain a nonzero decomposable vector, then A and B have a common invariant subspace of dimension k . Otherwise, no such subspace exists.
6. Find a decomposable vector belonging to an eigenspace intersection of Step 4 and factor it; its factors form a basis for a common invariant subspace of A and B .

Steps 1-3 of the above plan are straightforward. Steps 2-3 can be computationally expensive, depending on n and k . Step 4 can be performed using [8, Algorithm 12.4.3]. Steps 5 and 6 can be theoretically and practically challenging; we will return to them after the following illustrative example.

Example 3.2 Let us consider whether

$$A = \begin{pmatrix} 3 & -3 & 1 \\ 0 & 4 & 0 \\ -1 & -3 & 5 \end{pmatrix}, \quad B = \begin{pmatrix} -1 & -3 & 5 \\ -2 & 6 & 2 \\ -7 & -1 & 11 \end{pmatrix}$$

have a common invariant subspaces of dimension $k = 2$ or not. The eigenvalues of A are all equal to 4 and of B are 4 (double) and 8. We compute the second compounds of A and B to be

$$X = A^{(2)} = \begin{pmatrix} 12 & 0 & -4 \\ -12 & 16 & -12 \\ 4 & 0 & 20 \end{pmatrix}, \quad Y = B^{(2)} = \begin{pmatrix} -12 & 8 & -36 \\ -20 & 24 & -28 \\ 44 & -8 & 68 \end{pmatrix}.$$

The matrix K of Theorem 3.1 is a scalar multiple of

$$\begin{pmatrix} 1 & 0 & 1 \\ 0 & 0 & 0 \\ 1 & 0 & 1 \end{pmatrix}$$

and thus it is not invertible. Using Matlab's *null* routine (which computes orthogonal bases, either rationally via the echelon form or via the singular value decomposition), we find that

$$\text{Nul}(X - 16I) = \text{span}\{\alpha, \beta\}, \quad \text{where } \alpha = (-1 \ 0 \ 1)^t, \beta = (0 \ 1 \ 0)^t,$$

$$\text{Nul}(Y - 16I) = \text{span}\{\gamma\} \quad \text{where } \gamma = (1 \ -1 \ -1)^t, \text{ and}$$

$$\text{Nul}(Y - 32I) = \text{span}\{\delta\} \quad \text{where } \delta = (1 \ 1 \ -1)^t.$$

Since $\gamma, \delta \in \text{span}\{\alpha, \beta\}$, we have that γ, δ are common eigenvectors of X and Y . Moreover, γ and δ are decomposable as $\gamma = x_1 \wedge x_2$ and $\delta = x_1 \wedge x_3$, where

$$x_1 = (1 \ 0 \ 1)^t, \quad x_2 = (0 \ 1 \ -1)^t, \quad x_3 = (0 \ 1 \ 1)^t.$$

It follows that $\text{span}\{x_1, x_2\}$ and $\text{span}\{x_1, x_3\}$ are common invariant subspaces of A and B . Notice also that, in agreement with Theorem 2.2 when $k = 1$, A and B have a common invariant subspace of dimension 1, namely, the span of x_1 , which is a common eigenvector of A and B . ■

Step 5 of our plan raises in essence the following question: *Given a subspace U of the k -th Grassmann space over \mathbf{C}^n , when does U contain a nonzero decomposable vector?* A similar problem for subspaces of the tensor product of finite dimensional vector spaces is considered in [1] and is associated with the existence of a quasi-positive semidefinite operator of negative inertia k .

There are of course subspaces U that do not contain a nonzero decomposable vector, e.g., the span of a non-decomposable vector. If the subspace in question is one dimensional, the quadratic Plücker relations for decomposability can be used (see [11, Vol. II, §4.1, Definition 1.1]). For example,

$$x = (x_1, \dots, x_6)^t \in \mathbf{C}^{\binom{4}{2}} \text{ is decomposable if and only if } x_1x_6 - x_2x_5 + x_3x_4 = 0.$$

As it can be seen in the next proposition, the answer to the above question for a general subspace is related to the dimension of the *Grassmann variety*, identified here with the set $G = \{z \in \mathbf{C}^{\binom{n}{k}} : z \neq 0 \text{ is decomposable}\}$.

Proposition 3.3 *Let U be a subspace of $\mathbf{C}^{\binom{n}{k}}$. If $\dim U > \binom{n}{k} - k(n - k)$, then U contains a decomposable vector.*

Proof. The Grassmann variety G on projective space $\mathbf{P}^{\binom{n}{k}-1}$ has dimension $k(n - k)$ [9, Chapter XIV]. The subspace U is also an algebraic variety of dimension $\dim U$. In addition, U and G are homogeneous sets (i.e., closed under scalar multiplication). By [10, Corollary 3.30], if $\dim U + k(n - k) > \binom{n}{k}$, then the intersection of U and G is nonempty and thus U contains a nonzero decomposable vector. ■

In certain cases, all vectors in the intersecting subspaces of step 4 are decomposable. For example, if A is assumed to have distinct eigenvalues, then for an appropriate s , $X = (A + sI)^{(k)}$ also has distinct eigenvalues [5]. As a consequence, all eigenvectors of X are exterior products of eigenvectors of A and are therefore decomposable. Another classical case leads to the following result.

Proposition 3.4 *Let $A, B \in \mathbf{C}^{n \times n}$ and W be a subspace of \mathbf{C}^n of dimension $n - 1$. The following are equivalent.*

- (i) W is a common invariant subspace of A and B .
- (ii) There exists $s \in \mathbf{C}$ such that $(A + sI)^{(n-1)}$ and $(B + sI)^{(n-1)}$ are invertible and have a common eigenvector.
- (iii) A^* and B^* have a common eigenvector u such that $W^\perp = \text{span}\{u\}$.

Proof. We first notice that any common eigenvector of $(A + sI)^{(n-1)}$ and $(B + sI)^{(n-1)}$ is an element of the $(n - 1)$ -st Grassmann space over \mathbf{C}^n . By [11, Vol. II, §4.1, Theorem 1.3] the Grassmann variety in this case coincides with the whole space, i.e., every vector is decomposable. Thus the equivalence of (i) and (ii) follows from Theorem 2.2. The equivalence of (i) and (iii) is elementary. ■

Next we comment on step 6 of our plan, that is, finding a decomposable vector and its factors. This problem is related to problems that arise in linear control theory (e.g., pole assignment by feedback). They are referred to as determinantal assignment problems and their solution involves finding a minimal parameterization of the quadratic Plücker relations (see e.g., [6]).

We conclude with one more example for which Proposition 3.4 is not relevant and, as the matrices do not have distinct eigenvalues, the results in [5] do not apply.

Example 3.5 Let

$$A = \begin{pmatrix} 1 & 0 & 1 & 1 \\ 1 & 2 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 1 & 1 & 0 & 1 \end{pmatrix}, \quad B = A^t.$$

The eigenvalues of A and B are 3 and 1 (triple). The second compounds of A and B are

$$X = A^{(2)} = \begin{pmatrix} 2 & -1 & -1 & -2 & -2 & 0 \\ 0 & 2 & 1 & 0 & 0 & -1 \\ 1 & -1 & 0 & -1 & -1 & 1 \\ 0 & 2 & 1 & 4 & 2 & 0 \\ -1 & 0 & 1 & 0 & 2 & 0 \\ 0 & -2 & -1 & -2 & -1 & 2 \end{pmatrix}, \quad Y = B^{(2)} = X^t.$$

The matrix K of Theorem 3.1 is not invertible and thus X and Y have common eigenvectors. We find that $V_1 = \text{Nul}(X - I) = \text{span}\{\alpha, \beta\}$ and $V_2 = \text{Nul}(Y - I) = \text{span}\{\gamma, \delta\}$, where

$$\alpha = (0 \ 1 \ -1 \ -1 \ 1 \ 0)^t, \quad \beta = (-1 \ 2 \ -1 \ -1 \ 0 \ 1)^t,$$

$$\gamma = (2 \ -1 \ -1 \ 1 \ 1 \ 0)^t, \delta = (1 \ 0 \ -1 \ 1 \ 0 \ 1)^t.$$

Also $\dim(V_1 \cap V_2) = 2 + 2 - \dim(V_1 \cup V_2) = 1$. In fact,

$$\alpha - \beta = \gamma - \delta = (1 \ -1 \ 0 \ 0 \ 1 \ -1)^t$$

spans $V_1 \cap V_2$ and is decomposable by the quadratic Plücker relations. Its factors are

$$x_1 = (1 \ 0 \ 0 \ -1)^t, \quad x_2 = (0 \ 1 \ -1 \ 0)^t$$

and thus $\text{span}\{x_1, x_2\}$ is a common invariant subspace of A and $B = A^t$. A similar analysis for the eigenspaces corresponding to 3 gives that $\text{span}\{y_1, y_2\}$ is also a common invariant subspace, where

$$y_1 = (1 \ 0 \ 0 \ 1)^t, \quad y_2 = (0 \ 1 \ 1 \ 0)^t.$$

■

References

- [1] D. Choudhury, R. A. Horn, and S. J. Pierce. Quasi-positive definite operators and matrices. *Linear Algebra and Its Applications*, 99:161-176, 1988.
- [2] R. Edwards. Analysis of continuous-time switching networks. *Physica D*, to appear.
- [3] D. R. Farenick. Irreducible positive linear maps on operator algebras. *Proceedings of the American Mathematical Society*, 124(11):3381-3390, 1996.
- [4] M. Fiedler. *Special Matrices And Their Applications In Numerical Mathematics*. Martinus Nijhoff Publishers, 1986.
- [5] A. George and K. D. Ikramov. Common invariant subspaces of two matrices. *Linear Algebra and Its Applications*, 287:171-179, 1999.
- [6] C. Giannakopoulos, G. Kalogeropoulos, and N. Karcianas. The Grassmann variety of non-dynamic compensators: the determinantal assignment problem of linear systems. *Bull. Soc. Math. Grèce (N.S.)* 24:33-58, 1983.
- [7] I. Gohberg, P. Lancaster, and L. Rodman. *Invariant Subspaces of Matrices with Applications*. Wiley-Interscience, 1986.
- [8] G. H. Golub and C. F. Van Loan. *Matrix Computations*. The John Hopkins University Press, Baltimore, 1989.
- [9] W. V. D. Hodge and P. D. Pedoe. *Methods of Algebraic Geometry* Vol. 2. Cambridge University Press, 1952.
- [10] D. Mumford. *Algebraic Geometry I: Complex Projective Varieties*. Grundlehren der math. Wissenschaften, Vol. 221, Springer, New York, 1976.

- [11] M. Marcus. *Finite Dimensional Multilinear Algebra*. Two Volumes. Marcel Dekker, 1973-1975.
- [12] D. Shemesh. Common eigenvectors of two matrices. *Linear Algebra and Its Applications*, 62:11-18, 1984.
- [13] M. Tsatsomeros, J. S. Maybee, D. D. Olesky, and P. van den Driessche. Nullspaces of matrices and their compounds. *Linear and Multilinear Algebra*, 34(3/4) : 291-300, 1993.