Common Eigenvectors of Two Matrices

Dan Shemesh
Department of Mathematics
Technion — Israel Institute of Technology
32000 Haifa, Israel

Submitted by Hans Schneider

ABSTRACT

A computable criterion is given for two square matrices to possess a common eigenvector, as well as a criterion for one matrix to have an eigenvector lying in a given subspace. Some applications are discussed.

1. INTRODUCTION

A nonzero vector x in $\mathbf{C}^n$ is a common eigenvector of the n-square, complex matrices A and B if there exist complex numbers λ and μ such that

$$Ax = \lambda x,$$

$$Bx = \mu x.$$

Whenever the two matrices A and B commute, they possess at least one common eigenvector.

In 1935 McCoy [3] proved that the matrices A and B have simultaneous triangularization (i.e. there exists a nonsingular matrix P such that $P^{-1}AP$ and $P^{-1}BP$ are triangular) iff for every polynomial $p(x, y)$ of the noncommutative variables x, y the matrix $p(A, B)[A, B]$ ($[A, B]$ being the commutator $AB - BA$) is nilpotent. Consequently if the McCoy condition ($p(A, B)[A, B]$ is nilpotent for every $p(x, y)$ as above) holds, then A and B have a common eigenvector.

The McCoy condition is sufficient, though not very easy to check, for the existence of a common eigenvector. In Section 3 a computable condition is given which is necessary and sufficient for two matrices to have a common eigenvector.

Another way of attacking the above problem is to find whether or not there is an eigenvector of A corresponding to a fixed eigenvalue λ, which is

52 Vanderbilt Ave., New York, NY 10017 0024-3795/84/$3.00

also an eigenvector of B. We can restate this problem as: Does there exist an eigenvector of B which belongs to the kernel of $A-\lambda I$? Or, more generally, when does an n-square, complex matrix have an eigenvector lying in a given subspace of $\mathbf{C}^n$? This problem is solved in Section 2; its solution is related to system theory.

In Section 4 some applications are given. The problems of existence of common roots of two polynomials and two matrix polynomials are treated there, as well as the question of when two matrices have a common eigenvector corresponding to the same eigenvalue (i.e. $Ax=\lambda x$ and $Bx=\lambda x$).

2. CONSTRAINED EIGENVECTOR PROBLEM

Let A be an n-square complex matrix, and let U be a subspace of $\mathbf{C}^n$. Under what conditions does the matrix A have an eigenvector which belongs to the subspace U? Since U can be characterized as the kernel of a (not necessarily square) matrix B, the above problem can be rewritten as follows: Is there a nonzero vector x satisfying

$$\begin{aligned} &Ax=\lambda x, \\ &Bx=0 \qquad (B \text{ being an } m\times n \text{ matrix}), \end{aligned} \tag{2.1}$$

or, in a more compact form,

$$\begin{pmatrix} A \\ B \end{pmatrix} x=\lambda \begin{pmatrix} x \\ 0 \end{pmatrix}? \tag{2.2}$$

The following theorem settles this problem.

Theorem 2.1. *There is an eigenvector x of A which satisfies $Bx=0$ if and only if $\bigcap_{k=0}^{q-1}\ker(BA^k)\neq\{0\}$, where q is any integer greater than or equal to the degree of the minimal polynomial of A.*

Observation. All powers of A of order $\geqslant q$ are linear combinations of the first q powers $I, A, A^2,\dots,A^{q-1}$. Thus, denoting by $\mathscr{M}$ the intersection of the kernels in the theorem, we get $\mathscr{M}=\bigcap_{k=0}^{q-1}\ker(BA^k)=\bigcap_{k=0}^{\infty}\ker(BA^k)$.

Proof. Let $x\neq 0$ be a vector in $\mathscr{M}$. Then $BA^kx=0$, $k=0,1,2,\dots,$ and hence also $BA^{k+1}x=BA^k(Ax)=0$ for $k=0,1,2,\dots,$ which means that Ax is also in $\mathscr{M}$.

The subspace $\mathscr{M}\neq\{0\}$ is an invariant subspace of A. Thus A has an eigenvector in $\mathscr{M}\subset\ker B$, and this eigenvector is a solution of (2.1).

The other direction is obvious. If $x \neq 0$ is a solution of (2.1), then x satisfies $A^k x = \lambda^k x \in \ker B$, so $x \in \bigcap_{k=0}^{\infty} \ker(BA^k) = \mathcal{M}$. ∎

REMARK 2.1. $\mathcal{M}$ is the kernel of the positive semidefinite matrix $K = \sum_{k=0}^{q-1} (BA^k)^* (BA^k)$. Therefore (2.1) has a nontrivial solution iff K is singular. (The singularity of K can be easily checked.)

REMARK 2.2. $\mathcal{M}$ is also the kernel of the $mq \times n$ matrix

$$L = \begin{pmatrix} B \\ BA \\ \vdots \\ BA^{q-1} \end{pmatrix}.$$

Now Theorem 2.1 can be restated: the problem (2.1) is solvable iff $\operatorname{rank}(L) < n$.

The last remark has the following consequences in system theory (for definitions and relevant results see [2]):

THEOREM 2.2. *The system $\dot{x}(t) = Ax(t) + Bu(t)$ is completely controllable if there is no eigenvector of A^* lying in the kernel of B^*.*

THEOREM 2.3. *The system $\dot{x}(t) = Ax(t)$, $y(t) = Bx(t)$ is completely observable iff no eigenvector of A is contained in* $\ker(B)$.

REMARK 2.3. Although, in the general case, the exact number of linearly independent eigenvectors of A which are in $\ker(B)$ cannot be figured out by Theorem 2.1, an upper bound can be given:

(a) The number of linearly independent solutions of (2.1) is $\leqslant \dim(\mathcal{M})$.

(b) If A is similar to a diagonal matrix, then $\dim(\mathcal{M})$ is the exact number of independent eigenvectors of A lying in $\ker(B)$. This follows from the fact that every invariant subspace of a diagonalizable matrix is spanned by eigenvectors of this matrix.

3. COMMON EIGENVECTORS

The results of the previous section can be used to check the solvability of

$$\begin{aligned} Ax &= \lambda x \\ Bx &= \mu x, \qquad x \neq 0, \end{aligned} \tag{3.1}$$

where A and B are complex, n-square matrices and λ, μ are unknown eigenvalues of A and B respectively.

For each fixed eigenvalue μ of B, one can use Theorem 2.1 to check if the problem

$$Ax = \lambda x,$$
$$(B - \mu I)x = 0, \qquad x \neq 0$$

is solvable.

This method requires the knowledge of all the eigenvalues of B (or A), which might be very difficult to achieve.

The following theorem gives a criterion for the solvability of (3.1), i.e. for the existence of common eigenvectors of A and B. This criterion does not require knowledge of the eigenvalues of A or B.

THEOREM 3.1. *The matrices A and B have a common eigenvector iff*

$$\bigcap_{k,l=1}^{n-1} \ker[A^k, B^l] \neq \{0\}$$

(where n, in the intersection, can be replaced by p and q, the degrees of the minimal polynomials of A and B).

Proof. First suppose that $x \neq 0$ is a common eigenvector of A and B. Then $Ax = \lambda x$ and $Bx = \mu x$ imply that $A^k B^l x = B^l A^k x = \lambda^k \mu^l x$. Thus $[A^k, B^l]x = A^k B^l x - B^l A^k x = 0$, i.e.

$$0 \neq x \in \bigcap_{k,l=1}^{\infty} \ker[A^k, B^l] = \bigcap_{k,l=1}^{n} \ker[A^k, B^l]$$

(as in the observation following Theorem 2.1).

In order to prove the second direction let us denote by $\mathcal{N}$ the subspace $\bigcap_{k,l=1}^{n} \ker[A^k, B^l] = \bigcap_{k,l=1}^{\infty} \ker[A^k, B^l]$.

We suppose first that $\mathcal{N}$ (which is supposed to be $\neq \{0\}$) is an invariant subspace of both A and B. Now note that A and B commute on $\mathcal{N}$ ($[A, B]x = 0$ for every $x \in \mathcal{N}$). Thus the restrictions of the operators represented by A and B to the subspace $\mathcal{N}$ are commutative and hence they have a common eigenvector, which is, of course, a common eigenvector of A and B. In order to finish the proof the invariance of $\mathcal{N}$ should be proved.

Choose a nonzero vector x in $\mathcal{N}$, and define $\mathcal{N}_x = \{p(A, B)x | p \in \mathcal{P}\}$, where $\mathcal{P}$ is the set of all complex polynomials in two noncommutative variables. $\mathcal{N}_x$ is, obviously, an invariant subspace of A and B. Now, since $x \in \mathcal{N}$, $A^kB^lx = B^lA^kx$, and therefore every monomial in A and B operating on x has the form A^rB^sx [e.g. $A^2BAB^3x = A^2B(AB^3x) = A^2B(B^3Ax) = A^2(B^4Ax) = A^2(AB^4x) = A^3B^4x$]. This means that any element of $\mathcal{N}_x$ is of the form $y = (\sum_{i,j} a_{ij} A^iB^j)x$, and using the above argument we get $ABy = BAy = (\sum a_{ij} A^{i+1}B^{j+1})x$. Thus A and B commute on $\mathcal{N}_x$, and the proof is finished if $\mathcal{N}_x$ replaces $\mathcal{N}$ in the previous argument. Moreover, $\mathcal{N}_x \subset \mathcal{N}$, so $\mathcal{N} = \bigcup_{x \in \mathcal{N}} \mathcal{N}_x = \sum_{x \in \mathcal{N}} \mathcal{N}_x$. Thus $\mathcal{N}$, as a sum of invariant subspaces, is invariant for both A and B. ■

Note that $\mathcal{N}_x$ is a minimal A, B invariant subspace which contains x and on which A and B commute. On the other hand $\mathcal{N}$ is the maximal A, B invariant subspace on which A and B commute.

REMARK 3.1. As in the previous section, we can construct the matrices

$$K = \sum_{k,l=1}^{n-1} [A^k, B^l]^*[A^k, B^l] \quad \text{and} \quad L = \begin{pmatrix} [A, B] \\ [A, B^2] \\ \vdots \\ [A, B^{n-1}] \\ [A^2, B] \\ \vdots \\ [A^{n-1}, B^{n-1}] \end{pmatrix},$$

each of which has $\mathcal{N}$ as its kernel. Now we can restate Theorem 3.1: A and B possess a common eigenvector if and only if the matrix K (or the matrix L) has rank less than n.

REMARK 3.2. Again, as in Section 2, one can observe that the number of independent common eigenvectors of A and B is less than or equal to $\dim(\mathcal{N})$. If A and B are both diagonalizable, then $\dim(\mathcal{N})$ is the exact number of linearly independent common eigenvectors of A and B.

4. APPLICATIONS

Let $p(t)=a_0+a_1t+\cdots+a_{n-1}t^{n-1}+t^n$ be a monic complex polynomial, and let C_p denote the companion matrix of p, i.e.

$$C_p=\begin{pmatrix} 0 & 1 & & & \\ & 0 & 1 & & \\ & & \ddots & \ddots & \\ & & & 0 & 1 \\ -a_0 & -a_1 & & & -a_{n-1} \end{pmatrix}.$$

Let p and q be monic polynomials of the same degree. Then p and q have a common root iff C_p and C_q have a common eigenvector. This result follows from the fact that the only eigenvectors of C_p are multiples of $(1,\lambda,\lambda^2,\dots,\lambda^{n-1})^t$, where λ is a root of p. Thus Theorem 3.1 applied to C_p, C_q gives a criterion for the existence of common roots of the polynomials p and q.

Now let $P(t)=A_0+tA_1+t^2A_2+\cdots+t^{n-1}A_{n-1}+t^nI$ be a matrix polynomial, where $A_0,A_1,\dots,A_{n-1},I$ are complex square matrices of the same order, say m.

We say that λ is an eigenvalue of P if there exists a nonzero vector x satisfying $P(\lambda)x=0$; x is then a corresponding eigenvector. Construct now the block companion matrix of P:

$$C_P=\begin{pmatrix} 0 & I & & & \\ & 0 & I & & \\ & & \ddots & \ddots & \\ & & & 0 & I \\ -A_0 & -A_1 & \cdots & -A_{n-2} & -A_{n-1} \end{pmatrix}.$$

Suppose y is an eigenvector of C_P. Then $y=(x_0^t,x_1^t,\dots,x_{n-1}^t)^t$, where x_i, $i=0,1,\dots,n-1$, are n-tuples and

$$C_py=\lambda y,$$

i.e.

$$\begin{pmatrix} 0 & I & & & \\ & 0 & I & & \\ & & \ddots & \ddots & \\ & & & 0 & I \\ -A_0 & -A_1 & \cdots & -A_{n-2} & -A_{n-1} \end{pmatrix}\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}=\lambda\begin{pmatrix} x_0 \\ x_1 \\ \vdots \\ x_{n-1} \end{pmatrix}.$$

This implies that $x_1 = \lambda x_0$, $x_2 = \lambda x_1, \ldots, x_{n-1} = \lambda x_{n-2}$. So $y = (x_0^t, \lambda x_0^t, \lambda^2 x_0^t, \ldots, \lambda^{n-1} x_0^t)^t$, and comparing the nth components of $C_p y = \lambda y$, we get

$$(-A_0 - \lambda A_1 - \cdots - \lambda^{n-1} A_{n-1}) x_0 = \lambda(\lambda^{n-1} x_0), \qquad \text{i.e.,} \quad P(\lambda) x_0 = 0.$$

The conclusion of the above discussion is that the eigenvector of the companion matrix C_P must be of the form $(x^t, \lambda x^t, \lambda^2 x^t, \ldots, \lambda^{n-1} x^t)^t$ where $P(\lambda)x = 0$ and $x \neq 0$.

Consider now two monic matrix polynomials P and Q of the same degree. Then, if C_P and C_Q have a common eigenvector, it has to be $(x^t, \lambda x^t, \ldots, \lambda^{n-1} x^t)^t$ where $P(\lambda)x = Q(\lambda)x = 0$. Therefore we have proved

THEOREM 4.1. *The two matrices C_P and C_Q have a common eigenvector iff P and Q have common eigenvector corresponding to the same eigenvalue λ.*

REMARK 4.1. Here and in the previous discussion, the limitation that the two polynomials should be of the same degree can be handled by multiplying one of these polynomials by a proper power of t (or $t - \alpha$, when α is not a root of the two polynomials).

A simple application of Theorem 4.1 is the following

REMARK 4.2. The two n-square complex matrices, A, B have a common eigenvector corresponding to the same eigenvalue (i.e. $Ax = \lambda x$ and $Bx = \lambda x$) iff the matrices

$$\begin{pmatrix} 0 & I \\ A & 0 \end{pmatrix}, \qquad \begin{pmatrix} 0 & I \\ B & 0 \end{pmatrix}$$

have a common eigenvector (I is the $n \times n$ identity matrix).

Denote the above two matrices by $\tilde{A}, \tilde{B}$ respectively. It follows from the last remark and Theorem 3.1 that A and B have a common eigenvector corresponding to the same eigenvalue iff $\bigcap_{k,l=1}^{2n-1} \ker[\tilde{A}^k, \tilde{B}^l] \neq \{0\}$ (n is the order of the square matrices A and B).

Checking the commutators $[\tilde{A}^k, \tilde{B}^l]$ and their kernels yields the conclusion that there exists a nonzero vector x which satisfies

$$Ax = \lambda x,$$
$$Bx = \lambda x$$

iff $\mathscr{N} \cap \ker(A - B) \neq \{0\}$ (where $\mathscr{N}$ is the subspace $\bigcap_{k,l=1}^{n-1} \ker[A^k, B^l]$ mentioned in Theorem 3.1).

We conclude with a generalization of this result.

THEOREM 4.2. *Let p be a complex polynomial in two variables. The problem*

$$\begin{aligned} Ax &= \lambda x, \\ Bx &= \mu x, \qquad x \neq 0, \\ p(\lambda, \mu) &= 0 \end{aligned} \tag{4.1}$$

is solvable iff $\mathscr{N} \cap \ker p(A, B) \neq \{0\}$.

Proof. Since A and B commute on $\mathscr{N}$, they commute with $p(A, B)$ on the subspace $\mathscr{N}$, and hence $\ker p(A, B) \cap \mathscr{N}$ is again an A- and B-invariant subspace. On this subspace A and B commute and thus possess a common eigenvector x. This vector x satisfies $Ax = \lambda x$, $Bx = \mu x$, and $p(A, B)x = p(\lambda, \mu)x = 0$ [$x \in \ker p(A, B)$]. x is a nonzero vector; hence $p(\lambda, \mu)$ must vanish.

Conversely, if x is a solution of (4.1), then x belongs to $\mathscr{N}$ and again $p(A, B)x = p(\lambda, \mu)x$; but $p(\lambda, \mu) = 0$ implies that x is in $\ker p(A, B)$ and $0 \neq x \in \mathscr{N} \cap \ker P(A, B)$. ∎

REFERENCES

1 M. P. Drazin, J. W. Dungey, and K. W. Gruenberg, Some theorems on commutative matrices, *J. London Math. Soc.* 26:221–228 (1951).
2 D. G. Luenberger, *Introduction to Dynamic Systems*, Wiley, New York, 1979.
3 N. H. McCoy, On the characteristic roots of matric polynomials, *Bull. Amer. Math. Soc.* 42:592–600 (1936).
4 H. Shapiro, Simultaneous block triangularization and block diagonalization of sets of matrices, *Linear Algebra Appl.* 25:129–137 (1979).

Received 14 April 1982; revised 2 May 1983