Available online at www.sciencedirect.com
ScienceDirect
Linear Algebra and its Applications 430 (2009) 1–6
LINEAR ALGEBRA AND ITS APPLICATIONS
www.elsevier.com/locate/laa

Generators of matrix algebras in dimension 2 and 3

Helmer Aslaksen [a,*], Arne B. Sletsjøe [b]

[a] Department of Mathematics, National University of Singapore, Singapore 117543, Singapore
[b] Department of Mathematics, University of Oslo, P.O. Box 1053, Blindern, 0316 Oslo, Norway

Received 14 June 1995; accepted 8 May 2006
Available online 8 October 2008
Submitted by T.J. Laffey

Abstract

Let K be an algebraically closed field of characteristic zero and consider a set of 2×2 or 3×3 matrices. Using a theorem of Shemesh, we give conditions for when the matrices in the set generate the full matrix algebra.

Keywords: Generator; Matrix; Algebra

1. Introduction

Let K be an algebraically closed field of characteristic zero, and let $M_n = M_n(K)$ be the algebra of $n \times n$ matrices over K. Given a set $S = \{A_1, \ldots, A_p\}$ of $n \times n$ matrices, we would like to have conditions for when the A_i generate the algebra M_n. In other words, determine whether every matrix in M_n can be written in the form $P(A_1, \ldots, A_p)$, where P is a noncommutative polynomial. (We identify scalars with scalar matrices so the constant polynomials give the scalar matrices.) The case $n = 1$ is of course trivial, and when $p = 1$, the single matrix A_1 generates a commutative subalgebra. We therefore assume that $n, p \geqslant 2$. This question has been studied by many authors, see for example the extensive bibliography in [2]. We will give some results in the case of $n = 2$ or 3. We would like to thank the referees and the editor for making nontrivial improvements to the paper.

* Corresponding author.
E-mail addresses: aslaksen@math.nus.sg (H. Aslaksen), arnebs@math.uio.no (A.B. Sletsjøe).
URLs: http://www.math.nus.edu.sg/aslaksen/ (H. Aslaksen), http://www.math.uio.no/arnebs/ (A.B. Sletsjøe).

0024-3795/$ - see front matter © 2008 Published by Elsevier Inc.
doi:10.1016/j.laa.2006.05.022

2. General observations

Let $\mathscr{A}$ be the algebra generated by S. If we could show that the dimension of $\mathscr{A}$ as a vector space is n^2, it would follow that $\mathscr{A} = M_n$. This can sometimes be done when we know a linear spanning set $\mathscr{B} = \{B_1, \ldots, B_q\}$ of $\mathscr{A}$. Let M be the $n^2 \times q$ matrix obtained by writing the matrices in $\mathscr{B}$ as column vectors. We would like to show that rank $M = n^2$. Since M is an $n^2 \times n^2$ matrix and rank $M =$ rank (MM^*), it suffices to show that $\det(MM^*) \neq 0$. Unfortunately, the size of $\mathscr{B}$ may be big [4]. In this paper we will combine this method with results of Shemesh and Spencer and Rivlin to get some simple results for $n = 2$ or 3.

The starting point is the following well-known consequence of Burnside's Theorem.

Lemma 1. *Let $\{A_1, \ldots, A_p\}$ be a set of matrices in M_n where $n = 2$ or 3. The A_i's generate M_n if and only if they do not have a common eigenvector or a common left-eigenvector.*

We can therefore use the following theorem due to Shemesh [5].

Theorem 2. *Two $n \times n$ matrices, A and B, have a common eigenvector if and only if*

$$\sum_{k,l=1}^{n-1} [A^k, B^l]^*[A^k, B^l]$$

is singular.

Adding scalar matrices to the A_i's does not change the subalgebra they generate, so we sometimes assume that our matrices lie in $\mathfrak{sl}_n = \{M \in M_n | \operatorname{tr} M = 0\}$. We also sometimes identify matrices in M_n with vectors in K^{n^2}, and if $N_1, \ldots, N_{n^2} \in M_n$, then $\det(N_1, \ldots, N_{n^2})$ denotes the determinant of the $n^2 \times n^2$ matrix whose jth column is N_j, written as $(N_{j1}, \ldots, N_{jn})^t$, where N_{jk} is the kth row of N_j for $k = 1, 2, \ldots, n$. We write the scalar matrix aI as a. When we say that a set of matrices generate M_n, we are talking about M_n as an algebra, while when we say that a set of matrices form a basis of M_n, we are talking about M_n as a vector space.

3. The 2 × 2 case

The following theorem is well-known, but we include a proof since it illustrated a technique we will use in the 3×3 case. Notice that the proof gives us an explicit basis for M_2.

Theorem 3. *Let $A, B \in M_2$. A and B generate M_2 if and only if $[A, B]$ is invertible.*

Proof. A direct computation shows that

$$\det(I, A, B, AB) = -\det(I, A, B, BA) = \det[A, B].$$

Hence

$$\det(I, A, B, [A, B]) = 2\det[A, B]. \tag{1}$$

But if $I, A, B, [A, B]$ are linearly independent, then the dimension of $\mathscr{A}$ as a vector space is 4, so A and B generate M_2. □

We call $[M, N, P] = [M, [N, P]]$ a double commutator. The characteristic polynomial of A can be written as

$$x^2 - x \operatorname{tr} A + ((\operatorname{tr} A)^2 - \operatorname{tr} A^2)/2.$$

It follows that the discriminant of the characteristic polynomial of A can be written as

$$\operatorname{disc}(A) = 2 \operatorname{tr} A^2 - (\operatorname{tr} A)^2.$$

Lemma 4. *Let $A, B, C \in M_2$ and suppose that no two of them generate M_2. Then A, B, C generate M_2 if and only if the double commutator $[A, B, C] = [A, [B, C]]$ is invertible.*

Proof. A direct computation shows that

$$\det(I, A, B, C)^2 = -\det[A, [B, C]] - \operatorname{disc}(A)\det[B, C]. \tag{2}$$

But if I, A, B, C are linearly independent, then A, B and C generate M_2. □

Notice that the above proof gives us an explicit basis for M_2. We can now give a complete solution for the case $n = 2$.

Theorem 5. *The matrices $A_1, \ldots, A_p \in M_2$ generate M_2 if and only if at least one of the commutators $[A_i, A_j]$ or double commutators $[A_i, A_j, A_k] = [A_i, [A_j, A_k]]$ is invertible.*

Proof. If $p > 4$, the matrices are linearly dependent, so we can assume that $p \leqslant 4$. Suppose that A_1, A_2, A_3, A_4 generate M_2, but that no proper subset of them generates M_2. Then the four matrices are linearly independent, and we can write the identity I as a linear combination of them. If the coefficient of A_4 in this expression is nonzero, then A_1, A_2, A_3, I span and therefore generate M_2, so A_1, A_2, A_3 generate M_2. Thus, if $A_1, \ldots, A_p$ generate M_2, we can always find a subset of three of these matrices that generate M_2. The result now follows from Theorem 3 and Lemma 4. □

4. Two 3 × 3 matrices

In the case of two 3×3 matrices, we have the following well-known theorem.

Theorem 6. *Let $A, B \in M_3$. If $[A, B]$ is invertible, then A and B generate M_3.*

For $M \in M_3$, we define $H(M)$ to be the linear term in the characteristic polynomial of M. Hence

$$H(M) = ((\operatorname{tr} M)^2 - \operatorname{tr} M^2)/2,$$

which is equal to the sum of the three principal minors of degree two of M. Notice that $H(M)$ is invariant under conjugation, and that if $[A, B]$ is singular, then $[A, B]$ is nilpotent if and only if $H([A, B]) = 0$.

The following theorem shows that if $[A, B]$ is invertible and $H([A, B]) \neq 0$, then we can give an explicit basis for M_3.

Theorem 7. *Let $A, B \in M_3$. Then*

$$\det(I, A, A^2, B, B^2, AB, BA, [A, [A, B]], [B, [B, A]]) = 9 \det[A, B] H([A, B]), \tag{3}$$

so if $\det[A, B] \neq 0$ *and* $H([A, B]) \neq 0$, *then*

$$\{I, A, A^2, B, B^2, AB, BA, [A, [A, B]], [B, [B, A]]\}$$

form a basis for M_3.

The proof of (3) is by direct computation. Notice that this can be thought of as a generalization of (1) and (2).

We can also use Shemesh's Theorem to characterize pairs of generators for M_3.

Theorem 8. *The two* 3×3 *matrices* A *and* B *generate* M_3 *if and only if both*

$$\sum_{k,l=1}^{2} [A^k, B^l]^*[A^k, B^l] \qquad \text{and} \qquad \sum_{k,l=1}^{2} [A^k, B^l][A^k, B^l]^*$$

are invertible.

5. Three or more 3 × 3 matrices

We start with the following theorem due to Laffey [1].

Theorem 9. *Let* $\mathscr{S}$ *be a set of generators for* M_3. *If* $\mathscr{S}$ *has more than four elements*, *then* M_3 *can be generated by a proper subset of* $\mathscr{S}$.

It is therefore sufficient to consider the cases $p = 3$ or 4. Following the approach outlined earlier, we start by finding a linear spanning set. Using the polarized Cayley–Hamilton Theorem, Spencer and Rivlin [6,7] deduced the following theorem.

Theorem 10. *Let* $A, B, C \in M_3$. *Define*

$$\begin{aligned}
S(A) &= \{A, A^2\} \\
T(A, B) &= \{AB, A^2B, AB^2, A^2B^2, A^2BA, A^2B^2A\} \\
S(A_1, A_2) &= T(A_1, A_2) \cup T(A_2, A_1) \\
T(A, B, C) &= \{ABC, A^2BC, BA^2C, BCA^2, A^2B^2C, CA^2B^2, ABCA^2\} \\
S(A_1, A_2, A_3) &= \bigcup_{\sigma \in S_3} T(A_\sigma(1), A_\sigma(2), A_\sigma(3)).
\end{aligned}$$

1. *The subalgebra generated by* A *and* B *is spanned by*

   $I \cup S(A) \cup S(B) \cup S(A, B)$.

2. *The subalgebra generated by* A, B *and* C *is spanned by*

   $I \cup S(A) \cup S(B) \cup S(A, B) \cup S(A, B, C)$.

These spanning sets are not optimal. They include words of length 5. Paz [3] has proved that M_n can be generated by words of length $\lceil (n^2+2)/3 \rceil$. For M_3 this gives words of length 4. The general bound has been improved by Pappacena [4].

We next give a version of Shemesh's Theorem for three 3×3 matrices.

Theorem 11. *The matrices $A, B, C \in M_3$ have a common eigenvector if and only the matrix*

$$\begin{aligned} M(A,B,C) = & \sum_{\substack{M\in S(A),\\ N\in S(B)}} [M,N]^*[M,N] + \sum_{\substack{M\in S(A),\\ N\in S(C)}} [M,N]^*[M,N] \\ & + \sum_{\substack{M\in S(B),\\ N\in S(C)}} [M,N]^*[M,N] + \sum_{\substack{M\in S(A,B),\\ N\in S(C)}} [M,N]^*[M,N] \end{aligned}$$

is singular.

Proof. Let $\mathscr{A}$ be the algebra generated by A, B, C. Set

$$V = \bigcap_{\substack{M\in S(A),\\ N\in S(B)}} \ker[M,N] \bigcap_{\substack{M\in S(A),\\ N\in S(C)}} \ker[M,N] \bigcap_{\substack{M\in S(B),\\ N\in S(C)}} \ker[M,N] \bigcap_{\substack{M\in S(A,B),\\ N\in S(C)}} \ker[M,N].$$

We claim that V is invariant under $\mathscr{A}$. Let $v \in V$ and consider $\mathscr{A}v$. We know from Theorem 10 that any element of $\mathscr{A}$ is a linear combination of terms of the form

$$p(A,B)C^i q(A,B)C^j r(A,B)$$

with $p(A,B), q(A,B), r(A,B) \in I \cup S(A) \cup S(B) \cup S(A,B)$. Since

$$v \in \ker[S(A,B),S(C)] \cap \ker[S(A),S(C)] \cap \ker[S(B),S(C)],$$

we get

$$\begin{aligned} & p(A,B)C^i q(A,B)C^j r(A,B)v = p(A,B)C^i q(A,B)r(A,B)C^j v \\ & \quad = p(A,B)C^{i+j} q(A,B)r(A,B)v \\ & \quad = p(A,B)q(A,B)r(A,B)C^{i+j} v = C^{i+j} p(A,B)q(A,B)r(A,B)v. \end{aligned}$$

In the same way we use the fact that $v \in [S(A), S(B)]$ to sort the terms of the form $p(A,B)q(A,B)r(A,B)v$, so that we finally get

$$\mathscr{A}v = \left\{ \sum a_{ijk} C^i B^j A^k v \mid 0 \leqslant i,j,k \leqslant 2, a_{ijk} \in K \right\}.$$

Using the above technique, it follows easily that $\mathscr{A}v \subset V$ and that V is $\mathscr{A}$ invariant. Hence we can restrict $\mathscr{A}$ to V, but since the elements of $\mathscr{A}$ commute on V, they have a common eigenvector, and we can finish as in the proof of Theorem 2. $\square$

From this we deduce the following theorem.

Theorem 12. *Let $A, B, C \in M_3$. Then A, B, C generate M_3 if and only if both $M(A,B,C)$ and $M(A^t, B^t, C^t)$ are invertible.*

For the case of four matrices, we can prove the following theorem.

Theorem 13. *The matrices $A_1, A_2, A_3, A_4 \in M_3$ have a common eigenvector if and only the matrix*

$$M(A_1, A_2, A_3, A_4) = \sum_{\substack{i,j=1,\\ i<j}}^{4} \left(\sum_{\substack{M \in S(A_i),\\ N \in S(A_j)}} [M, N]^*[M, N] \right) + \sum_{\substack{i,j=1,\\ i<j}}^{3} \left(\sum_{\substack{M \in S(A_i, A_j),\\ N \in S(A_4)}} [M, N]^*[M, N] \right) + \sum_{\substack{M \in S(A_1, A_2),\\ N \in S(A_3)}} [M, N]^*[M, N] + \sum_{\substack{M \in S(A_1, A_2, A_3),\\ N \in S(A_4)}} [M, N]^*[M, N].$$

is singular.

Proof. Similar to the proof of Theorem 11. □

From this we deduce the following theorem.

Theorem 14. *Let $A, B, C, D \in M_3$. Then A, B, C, D generate M_3 if and only if both $M(A, B, C, D)$ and $M(A^t, B^t, C^t, D^t)$ are invertible.*

References

[1] T.J. Laffey, Irredundant generating sets for matrix algebras, Linear Algebra Appl., 52 (1983) 457–478.
[2] T.J. Laffey, Simultaneous reduction of sets of matrices under similarity, Linear Algebra Appl. 84 (1986) 123–138.
[3] A. Paz, An application of the Cayley–Hamilton theorem to matrix polynomials in several variables, Linear and Multilinear Algebra 15 (1984) 161–170.
[4] C.J. Pappacena, An upper bound for the length of a finite-dimensional algebra, J. Algebra 197 (1997) 535–545.
[5] D. Shemesh, Common eigenvectors of two matrices, Linear Algebra Appl. 62 (1984) 11–18.
[6] A.J.M. Spencer, R.S. Rivlin, The theory of matrix polynomials and its application to the mechanics of isotropic continua, Arch. Rational Mech. Anal. 2 (1959) 309–336.
[7] A.J.M. Spencer, R.S. Rivlin, Further results in the theory of matrix polynomials, Arch. Rational Mech. Anal. 4 (1959) 214–230.