CORE
Metadata, citation and similar papers at core.ac.uk
Provided by Repositório Institucional da Universidade de Aveiro

Linear spanning sets for matrix spaces

G. Micheli[a,1,*], J. Rosenthal[a,1], P. Vettori[b,1,2]

[a] *University of Zurich, Winterthurstrasse 190, CH-8057 Zürich, Switzerland*

[b] *University of Aveiro, Campus de Santiago, 3810-193 Aveiro, Portugal*

Abstract

Necessary and sufficient conditions are given on matrices A, B and S, having entries in some field $\mathbb{F}$ and suitable dimensions, such that the linear span of the terms A^iSB^j over $\mathbb{F}$ is equal to the whole matrix space.

This result is then used to determine the cardinality of subsets of $\mathbb{F}[A]S\mathbb{F}[B]$ when $\mathbb{F}$ is a finite field.

Keywords: Matrices, linear span, cyclic matrices, finite fields.
2010 MSC: 15A03,15A69

1. Introduction

We start by stating a purely linear algebra problem:

Problem 1.1. *Let m,n be integers and $\mathbb{F}$ be any field. Let A,S,B be matrices having entries in $\mathbb{F}$ of dimensions $m\times m$, $m\times n$ and $n\times n$ respectively. Give necessary and sufficient conditions for the $\mathbb{F}$-linear span of $\{A^iSB^j\}_{i,j\in\mathbb{N}_0}$ to be equal to the whole matrix space $\mathbb{F}^{m\times n}$.*

A solution to this problem will be provided in Section 3.

*Corresponding author
Email addresses: giacomo.micheli at math.uzh.ch (G. Micheli), rosenthal at math.uzh.ch (J. Rosenthal), pvettori at ua.pt (P. Vettori)

[1] Authors supported in part by Swiss National Science Foundation grant SNF no. 149716.

[2] This work was supported by Portuguese funds through the CIDMA (Center for Research and Development in Mathematics and Applications) and the Portuguese Foundation for Science and Technology ("FCT–Fundação para a Ciência e a Tecnologia"), within project UID/MAT/04106/2013.

Notice that the previous problem has also an impact in Cryptography since it gives necessary and sufficient conditions for the attack in [1, Section 3] to be performed in *provable* polynomial time.

Starting with Section 4 we will assume that the base field $\mathbb{F}$ represents the finite field $\mathbb{F} = \mathbb{F}_q$ having cardinality q. Under these conditions and the conditions that $\gcd(m,n) = 1$ and the characteristic polynomials of the matrices A and B are irreducible we are able to show in Section 4 that $\{A^i S B^j\}_{i,j\in\mathbb{N}_0}$ spans the whole vector space $\mathbb{F}^{m\times n}$ as soon as $S \neq 0$.

In Section 5 we will prove that whenever the set $\{A^i S B^j\}_{i,j\in\mathbb{N}_0}$ spans the whole matrix ring as a vector space over the finite field $\mathbb{F}$, the cardinality of subsets of $\mathbb{F}[A]S\mathbb{F}[B]$ can be explicitly computed. A particular instance of this computation (i.e. when S is the identity matrix and A, B have irreducible characteristic polynomial) has already been approached via inequalities in [2].

2. Notation and Preliminaries

Let $\mathbb{F}$ be a field and denote by $\langle \mathcal{S} \rangle_{\mathbb{F}}$ the linear span over $\mathbb{F}$ of a set $\mathcal{S}$ of elements in some $\mathbb{F}$-vector space. Entries, rows and columns of matrices are indexed by integers starting from zero; I_n and, respectively, $0_{m\times n}$ denote the $n \times n$ identity matrix and the $m \times n$ zero matrix — indices may be omitted when no ambiguity arises.

Moreover, given $M \in \mathbb{F}^{n\times n}$,

- the minimal polynomial μ_M of M is the monic generator of the ideal $\{p(s) \in \mathbb{F}[s] : p(M) = 0\}$;
- the characteristic polynomial of M is $\chi_M(s) = \det(sI - M)$;
- $\mathcal{E}_M$ is the set of eigenvalues of M, i.e., the zeros of χ_M in some field extension of $\mathbb{F}$;
- $\mathcal{L}_M^\lambda$ and $\mathcal{R}_M^\lambda$ are the left and, respectively, right eigenspaces of M associated with $\lambda \in \mathcal{E}_M$;
- $\mathcal{L}_M = \bigcup_{\lambda\in\mathcal{E}_M} \mathcal{L}_M^\lambda \setminus \{0\}$ and $\mathcal{R}_M = \bigcup_{\lambda\in\mathcal{E}_M} \mathcal{R}_M^\lambda \setminus \{0\}$ are the sets of left and, respectively, right eigenvectors of M.
- M is cyclic (or non-derogatory) if one of the following equivalent conditions holds true:

- $\mu_M = \chi_M$;
- M is similar to a companion matrix;
- each eigenspace of M has dimension 1, i.e., every eigenvalue has geometric multiplicity 1.

The definition of the Kronecker product and some of its properties are given next. More details may be found, for instance, in [3, Section 12.1].

The Kronecker product of matrices $M \in \mathbb{F}^{m\times p}$ and $N \in \mathbb{F}^{n\times q}$ is the block matrix

$$M \otimes N = [m_{i,j}N]_{0\le i<m, 0\le j<p} \in \mathbb{F}^{mn\times pq},$$

representing the tensor product of the linear maps corresponding to M and N. Therefore, it satisfies the property

$$(M \otimes N)(P \otimes Q) = MP \otimes NQ, \tag{1}$$

whenever the matrix products on the right side can be computed.

The (column) vectorization of M is the (column) vector $\mathrm{v}(M) \in \mathbb{F}^{mp}$ formed by stacking the columns of M. Note that $\mathrm{v} : \mathbb{F}^{m\times p} \to \mathbb{F}^{mp}$ is an isomorphism of $\mathbb{F}$-vector spaces, establishing a correspondence between entry (i, j) of M and entry $i + mj$ of $\mathrm{v}(M)$.

Using this notation, given three matrices M, X, N of suitable dimensions,

$$\mathrm{v}(MXN) = (N^\top \otimes M)\,\mathrm{v}(X). \tag{2}$$

3. A basis for the vector space of $m \times n$ matrices

Let matrices A, B, and S as in Problem 1.1 and define

$$\mathcal{V}_{A,B;S} = \langle \{A^i S B^j\}_{i,j\ge 0} \rangle_{\mathbb{F}}.$$

In this and in the following section, conditions will be given, which ensure that the dimension of $\mathcal{V}_{A,B;S}$ is maximal, i.e., equal to mn.

Theorem 3.1. *Let $A \in \mathbb{F}^{m\times m}$, $B \in \mathbb{F}^{n\times n}$, and $S \in \mathbb{F}^{m\times n}$. Then, the following conditions are equivalent:*

$$\mathcal{V}_{A,B;S} = \mathbb{F}^{m\times n}; \tag{3}$$

$$uSv \neq 0,\ \forall u \in \mathcal{L}_A, v \in \mathcal{R}_B. \tag{4}$$

The proofs of the implications of Theorem 3.1 will be shown separately. In particular, "(3) $\Rightarrow$ (4)" will be demonstrated later on as a consequence of formula (14), concluding the proof of the theorem, while the converse implication will be stated as an independent proposition after two preparatory lemmas.

The first one provides a necessary condition for (4) and, as a consequence of Theorem 3.1, for (3).

Lemma 3.2. *If condition* (4) *holds, then both A and B are cyclic.*

Proof. Let $\mathbb{E}$ be an extension field containing all eigenvalues of A and B. Given any left eigenvector $u \in \mathcal{L}_A$, consider the linear map $\gamma_u : \mathbb{E}^n \to \mathbb{E}$, $x \mapsto uSx$, whose kernel has at least dimension $n-1$. If B is not cyclic, it has a right eigenspace $\mathcal{R}_B^\alpha \subseteq \mathbb{E}^n$ of dimension greater than one. Therefore, there exists a nonzero vector $v \in \mathcal{R}_B^\alpha \cap \ker \gamma_u$ such that $\gamma_u(v) = uSv = 0$.

The same reasoning may be applied exchanging the role of A and B, thus showing that if either A or B is not cyclic, condition (4) cannot be satisfied. □

The second lemma is well known in the case $\mathbb{F} = \mathbb{C}$ (see [4, 5]). For completeness, a self-contained proof for any field $\mathbb{F}$ will be given here.

Lemma 3.3. *Let $H \in \mathbb{F}^{p\times p}$, $K \in \mathbb{F}^{p\times q}$ and assume that $\mathcal{E}_H \subseteq \mathbb{E}$, extension field of $\mathbb{F}$. Then, for any $d \geq \deg \mu_H$,*

$$\operatorname{rank}_{\mathbb{F}} \begin{bmatrix} K & HK & \cdots & H^{d-1}K \end{bmatrix} = p \Leftrightarrow \operatorname{rank}_{\mathbb{E}} \begin{bmatrix} \lambda I - H & K \end{bmatrix} = p,\ \forall \lambda \in \mathcal{E}_H.$$

Proof. Observe that for any matrix M with entries in $\mathbb{F}$, $\operatorname{rank}_{\mathbb{F}} M = \operatorname{rank}_{\mathbb{E}} M$, since the rank depends only on the invertibility (in $\mathbb{F}$) of square submatrices of M. So, this equivalent statement will be proved:

$$\operatorname{rank}_{\mathbb{E}} \begin{bmatrix} K & HK & \cdots & H^{d-1}K \end{bmatrix} < p \Leftrightarrow \exists \lambda \in \mathcal{E}_H : \operatorname{rank}_{\mathbb{E}} \begin{bmatrix} sI - H & K \end{bmatrix} < p.$$

"$\Rightarrow$": Be $u \in \mathbb{E}^{1\times p}$ a nonzero vector such that $u \begin{bmatrix} K & HK & \cdots & H^{d-1}K \end{bmatrix} = 0$ and be $a \in \mathbb{E}[s]$ any generator of the principal ideal $\mathcal{I} = \{f \in \mathbb{E}[s] : uf(H) = 0\}$. Since $\mu_H \in \mathcal{I}$, $\deg a \leq \deg \mu_H \leq d$ and $a(\lambda) = 0$ for some $\lambda \in \mathcal{E}_H$. Write $a(s) = (\lambda - s)b(s)$, with $b(s) = \sum_{i=0}^{d-1} b_i s^i \notin \mathcal{I}$, and let $v = ub(H)$. Then, $v \neq 0$,

$$vK = ub(H)K = \sum_{i=0}^{d-1} b_i u H^i K = \sum_{i=0}^{d-1} b_i 0 = 0,$$

and $v(\lambda I - H) = u(\lambda I - H)b(H) = ua(H) = 0$. Thus, $v \begin{bmatrix} \lambda I - H & K \end{bmatrix} = 0$.

"⇐": There exist $\lambda \in \mathcal{E}_H$ and a nonzero $u \in \mathbb{E}^{1\times p}$ such that $u \begin{bmatrix} \lambda I - H & K \end{bmatrix} = 0$, i.e., $uH = \lambda u$ and $uK = 0$. Hence,

$$u \begin{bmatrix} K & HK & \cdots & H^{d-1}K \end{bmatrix} = u \begin{bmatrix} K & \lambda K & \cdots & \lambda^{d-1}K \end{bmatrix} = 0.$$

□

Proposition 3.4. *Under the assumptions of Theorem 3.1,* (4) ⇒ (3).

Proof. Assuming that condition (4) is satisfied, a sequence of implications will be established, which prove that also condition (3) holds true.

First of all, note that matrices $\{A^i S B^j\}$ generate $\mathbb{F}^{m\times n}$ if and only if the corresponding vectors $\{\mathbb{v}(A^i S B^j)\}$ generate $\mathbb{F}^{mn}$. Therefore, we get that

$$(3) \Leftrightarrow \langle \{\mathbb{v}(A^i S B^j)\}_{i,j\geq 0} \rangle_{\mathbb{F}} = \mathbb{F}^{mn}. \tag{5}$$

By (2) and (1), it follows that

$$\mathbb{v}(A^i S B^j) = \mathbb{v}(A^i S B^j I_n) = (I_n \otimes A^i)\, \mathbb{v}(S B^j) = (I_n \otimes A)^i\, \mathbb{v}(S B^j).$$

Let $F = I_n \otimes A \in \mathbb{F}^{mn\times mn}$, which is a block diagonal matrix, and be G the $mn \times n$ matrix whose columns are $\mathbb{v}(SB^j)$, $0 \leq j < n$. The (right) image of G, i.e., its column span, corresponds through $\mathbb{v}$ to the span of SB^j, $0 \leq j < n$. Analogously, for any $0 \leq i < m$, the image of $F^i G$ corresponds to the span of $A^i S B^j$, $0 \leq j < n$. Hence, by the Cayley-Hamilton Theorem,

$$\langle \{\mathbb{v}(A^i S B^j)\}_{i,j\geq 0} \rangle_{\mathbb{F}} = \mathrm{img}_{\mathbb{F}} \begin{bmatrix} G & FG & \cdots & F^{m-1}G \end{bmatrix}. \tag{6}$$

Observe that the degree of the minimal polynomial $\mu_F = \mu_{I\otimes A} = \mu_A$ cannot be greater than m and so, by (5), (6), and Lemma 3.3, we can state that

$$\begin{aligned} (3) &\Leftrightarrow \mathrm{img}_{\mathbb{F}} \begin{bmatrix} G & FG & \cdots & F^{m-1}G \end{bmatrix} = \mathbb{F}^{mn} \\ &\Leftrightarrow \mathrm{rank}_{\mathbb{E}} \begin{bmatrix} \lambda I - F & G \end{bmatrix} = mn,\ \forall \lambda \in \mathcal{E}_A, \end{aligned} \tag{7}$$

being $\mathbb{E}$ an extension field of $\mathbb{F}$ containing the eigenvalues of F, i.e., of A.

In order to determine the conditions that guarantee that the rank of the polynomial matrix $C(s) = \begin{bmatrix} sI - F & G \end{bmatrix}$ does not drop as $s \in \mathcal{E}_A$, it is necessary to analyze the structure of $C(s)$ with greater detail.

Denote by G_i, $0 \leq i < n$, the $m \times n$ blocks forming matrix G. Then

$$C(s) = \begin{bmatrix} sI - F & G \end{bmatrix} = \begin{bmatrix} sI - A & & & & G_0 \\ & sI - A & & & G_1 \\ & & \ddots & & \vdots \\ & & & sI - A & G_{n-1} \end{bmatrix}. \tag{8}$$

For any $\alpha \in \mathcal{E}_A$, the rank of $C(\alpha)$ is mn if and only if $wC(\alpha) \neq 0$ for every $w \neq 0$. In particular, we shall consider only nonzero vectors w such that $w(\alpha I - F) = 0$, since otherwise condition $wC(\alpha) \neq 0$ would be obviously satisfied. As $\alpha I - F = I_n \otimes (\alpha I - A)$, it turns out that $w(\alpha I - F) = 0$ if and only if $w = \begin{bmatrix} u_0 & u_1 & \cdots & u_{n-1} \end{bmatrix}$, with $u_i \in \mathcal{L}_A^\alpha$, $0 \leq i < n$. Under this condition,

$$\begin{aligned} wC(\alpha) &= \begin{bmatrix} u_0 & u_1 & \cdots & u_{n-1} \end{bmatrix} \begin{bmatrix} \alpha I - A & & & G_0 \\ & \alpha I - A & & G_1 \\ & & \ddots & \vdots \\ & & \alpha I - A & G_{n-1} \end{bmatrix} \\ &= \begin{bmatrix} 0 & u_0 G_0 + u_1 G_1 + \cdots + u_{n-1} G_{n-1} \end{bmatrix}. \end{aligned} \tag{9}$$

By Lemma 3.2, A is cyclic. It follows that, since the eigenspace $\mathcal{L}_A^\alpha$ has dimension 1, it is generated by one (eigen)vector, say $u \neq 0$, whence $u_i = \gamma_i u$, $\gamma_i \in \mathbb{E}$ for $0 \leq i < n$, not all zero. Summing up, the rank of $C(\alpha)$ is mn if the linear combination

$$\gamma_0 u G_0 + \gamma_1 u G_1 + \cdots + \gamma_{n-1} u G_{n-1}$$

is not zero for any choice of the (not all zero) coefficients γ_i, $i = 0, \ldots, n-1$, i.e., if the vectors $\{uG_i\}_{0 \leq i < n}$ are linearly independent. Hence, by equivalence (7), it follows that

$$(3) \Leftarrow \{uG_i\}_{0 \leq i < n} \text{ are } \mathbb{E}\text{-linearly independent, } \forall u \in \mathcal{L}_A. \tag{10}$$

Consider now any $u \in \mathbb{E}^{1 \times m}$ and define the matrix

$$D = (I_n \otimes u)G = \begin{bmatrix} uG_0 \\ uG_1 \\ \vdots \\ uG_{n-1} \end{bmatrix} \in \mathbb{E}^{n \times n}.$$

Moreover, let $(SB^j)_i$ be the i-th column of SB^j for every $0 \leq i < n$ and $0 \leq j < n$.

By definition, the j-th column of G is $\mathrm{v}(SB^j)$, which contains, stacked, vectors $(SB^j)_i$. Therefore, in particular, the j-th column of G_i, is $(SB^j)_i$. Consequently, the j-th component of uG_i, which is the entry at (i, j) of D, is $u(SB^j)_i$. At the same time, this value is the i-th component (column) of uSB^j, i.e, the entry at (j, i) of the matrix whose rows are uSB^j. In other words,

$$D^\top = \begin{bmatrix} uSB^0 \\ uSB^1 \\ \vdots \\ uSB^{n-1} \end{bmatrix}.$$

Since D is square, its rows are linearly independent if and only if its columns share the same property. Applying again Lemma 3.3 with $H = B^\top$ and $K = (uS)^\top$, we get that

$$\begin{aligned}
&\{uG_i : 0 \le i < n\} \text{ are } \mathbb{E}\text{-linearly independent} && \Leftrightarrow \qquad (11)\\
&\{uSB^j : 0 \le j < n\} \text{ are } \mathbb{E}\text{-linearly independent} && \Leftrightarrow\\
&\operatorname{rank}_{\mathbb{E}} \begin{bmatrix} (uS)^\top & B^\top(uS)^\top & \cdots & (B^\top)^{n-1}(uS)^\top \end{bmatrix} = n && \Leftrightarrow\\
&\operatorname{rank}_{\mathbb{E}} \begin{bmatrix} \lambda I - B^\top & (uS)^\top \end{bmatrix} = \operatorname{rank}_{\mathbb{E}} \begin{bmatrix} \lambda I - B \\ uS \end{bmatrix} = n,\ \forall \lambda \in \mathcal{E}_B.
\end{aligned}$$

As before, define $E_u(s) = \left[\begin{smallmatrix} sI - B \\ uS \end{smallmatrix}\right] \in \mathbb{E}^{(n+1)\times n}[s]$ and consider any $\beta \in \mathcal{E}_B$.

By Lemma 3.2, also matrix B is cyclic, being $\operatorname{rank}_{\mathbb{E}} \begin{bmatrix} \beta I - B \end{bmatrix} = n - 1$. Therefore, the rank of $E_u(\beta)$ is actually n if $E_u(\beta)v \neq 0$ for any $v \in \mathcal{R}_B^\beta$. In this case, condition $E_u(\beta)v \neq 0$ reduces to $uSv \neq 0$ and, consequently,

$$\operatorname{rank}_{\mathbb{E}} \begin{bmatrix} \lambda I - B \\ uS \end{bmatrix} = n,\ \forall \lambda \in \mathcal{E}_B. \Leftarrow uSv \neq 0,\ \forall v \in \mathcal{R}_B. \qquad (12)$$

In particular, considering only $u \in \mathcal{L}_A$, as in condition (4), the sequence of implications (12), (11) and (10) concludes the proof. □

In order to prove the converse implication of Proposition 3.4 we introduce the necessary notation and state a fundamental result.

Given $A \in \mathbb{F}^{m\times m}$, $B \in \mathbb{F}^{n\times n}$, and $S \in \mathbb{F}^{m\times n}$, let $r_{i,j} = \mathbb{v}(A^i S B^j)$ and define

$$R_{A,B;S} = \begin{bmatrix} r_{0,0} & r_{1,0} & \cdots & r_{m-1,0} & r_{0,1} & r_{1,1} & \cdots & r_{m-1,n-1} \end{bmatrix} \in \mathbb{F}^{mn\times mn}. \qquad (13)$$

Then, given $v \in \mathbb{F}^n$, $\operatorname{diag}(v) \in \mathbb{F}^{n\times n}$ is the diagonal matrix defined by the components of v. Moreover, let $\operatorname{diag}(M) = \operatorname{diag}(\mathbb{v}(M))$ for any matrix M.

Finally, let $\overline{x}^n = \begin{bmatrix} 1 & x & \cdots & x^{n-1} \end{bmatrix}$ and be $\mathcal{V}^n_{x_0,x_1,\ldots,x_l}$ the matrix whose rows are ${\overline{x}_0}^n, {\overline{x}_1}^n, \ldots, {\overline{x}_l}^n$.

Proposition 3.5. *Let $A \in \mathbb{F}^{m\times m}$, $B \in \mathbb{F}^{n\times n}$, and $S \in \mathbb{F}^{m\times n}$. Suppose that $u_h \in \mathcal{L}_A^{\alpha_h}$, $0 \le h < s$, and $v_k \in \mathcal{L}_B^{\beta_k}$, $0 \le k < t$, are the rows and, respectively, columns of matrices $U \in \mathbb{E}^{s\times m}$ and $V \in \mathbb{E}^{n\times t}$ in a suitable extension field $\mathbb{E}$ of $\mathbb{F}$. Then,*

$$(V^\top \otimes U) R_{A,B;S} = \operatorname{diag}(USV)(\mathcal{V}^n_{\beta_0,\ldots,\beta_{t-1}} \otimes \mathcal{V}^m_{\alpha_0,\ldots,\alpha_{s-1}}). \qquad (14)$$

Proof. Observe that, for any row u_h of U and column v_k of V, the following equalities hold true: $u_h A^i = \alpha_h^i u_h$ and $B^j v_k = \beta_k^j v_k$. Thus, by (2),

$$(v_k^\top \otimes u_h)\, \mathbb{v}(A^i S B^j) = u_h A^i S B^j v_k = u_h S v_k\, \alpha_h^i \beta_k^j$$

and from (13) it follows that

$$(v_k^\top \otimes u_h) R_{A,B;S} = u_h S v_k (\overline{\beta}_k^n \otimes \overline{\alpha}_h^m).$$

Stacking up all these equalities, we get equation (14). □

Using Proposition 3.5, we are finally in a position to complete the proof of Theorem 3.1.

Proof (of Theorem 3.1). Given Proposition 3.4, it only remains to show that (3) ⇒ (4).

Suppose that the nonzero left-eigenvector $u \in \mathcal{L}_A^\alpha$ and right-eigenvector $v \in \mathcal{R}_B^\beta$ satisfy $uSv = 0$. Then, taking $U = u$ and $V = v$ in formula (14), we get

$$(v^\top \otimes u) R_{A,B;S} = (uSv)(\overline{\beta}^n \otimes \overline{\alpha}^m) = 0,$$

showing that $R_{A,B;S}$ does not have full rank. Therefore, its columns $\mathbb{v}(A^i S B^j)$ are linearly dependent and the set of matrices $A^i S B^j$ cannot generate $\mathbb{F}^{m\times n}$. □

Example 3.6. *Consider the following matrices, with $m, n \geq 2$:*

$$A = \begin{bmatrix} 0 & 0 \\ I_{m-1} & 0 \end{bmatrix} \in \mathbb{F}^{m\times m}, \quad B = \begin{bmatrix} 0 & I_{n-1} \\ 0 & 0 \end{bmatrix} \in \mathbb{F}^{n\times n}, \quad S = \begin{bmatrix} 1 & 0 \\ 0 & 0_{(m-1)\times(n-1)} \end{bmatrix} \in \mathbb{F}^{m\times n}.$$

Note that A and B are the left and, respectively, right companion matrices of $\mu_A(s) = s^m$ and $\mu_B(s) = s^n$. Therefore, they are cyclic, their only eigenvalue is $\lambda = 0$, they are nilpotent, and their eigenspaces are generated by $u = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}$ (left eigenvector of A) and $v = \begin{bmatrix} 1 & 0 & \cdots & 0 \end{bmatrix}^\top$ (right eigenvector of B).

Even though S has rank 1, $uSv = 1 \neq 0$, whence condition (4) *of Theorem 3.1 is satisfied. Therefore, $\mathbb{F}$–linear combinations of matrices $E_{i,j} = A^i S B^j$, with $0 \leq i < m$ and $0 \leq j < n$, generate $\mathbb{F}^{m\times n}$ for any field $\mathbb{F}$.*

Indeed, it is straightforward to check that each $E_{i,j}$ is one of the mn elements of the canonical basis of $\mathbb{F}^{m\times n}$, having its unique nonzero entry, equal to 1*, at position (i, j). In other words, $\mathbb{v}(E_{i,j})$ is the $i + mj$-th vector of the canonical basis of $\mathbb{F}^{mn}$.*

To the authors' knowledge, equality (3) and the equivalent condition that was presented in Theorem 3.1 have not been considered in the literature before (not even

when $m = n$: see, for instance, the survey [6] containing a small section about spanning sets of matrix algebras).

A comparison with previous results can be made only in the case $m = n = 2$ and $S = I$, verifying that $\mathbb{F}^{2\times 2}$ is spanned by linear combinations of $A^i B^j$, $i, j = 0, 1$, if and only if it can be generated by A and B as a matrix algebra (see, for example, [7], where this problem is thoroughly investigated). Indeed, in the following it is shown that a well-known criterium for the latter problem, the invertibility of the commutator of A and B, is equivalent to condition (4) presented in Theorem 3.1.

Proposition 3.7. *Let $A, B \in \mathbb{F}^{2\times 2}$. Then, the commutator $[A, B] = AB - BA$ is invertible if and only if $uv \neq 0$. for any $u \in \mathcal{L}_A$ and $v \in \mathcal{R}_B$.*

Proof. We will show that $[A, B]$ is singular if and only if there exist vectors $u \in \mathcal{L}_A$ and $v \in \mathcal{R}_B$ such that $uv = 0$.

Notice that if A is not cyclic, i.e., it is a multiple of the identity, or $B = 0$, both conditions are satisfied, since $[A, B] = 0$ and $\mathcal{L}_A = \mathbb{F}^{1\times 2} \setminus \{0\}$ or $\mathcal{R}_B = \mathbb{F}^2 \setminus \{0\}$. Hence, we may assume that A is cyclic and, without loss of generality, in Jordan form and that $B \neq 0$. Since by adding a scalar matrix cI, c in any field extension of $\mathbb{F}$, to A or B or multiplying them by any nonzero value does not change their commutator's rank, the general situation can be represented by the following two simplified cases (in which matrix A has only one or two different eigenvalues):

$$A \text{ is equal to } A_1 = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \text{ or to } A_2 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \text{ and } B = \begin{bmatrix} \alpha & \gamma \\ \beta & 0 \end{bmatrix} \neq 0.$$

Observe that A_1 has only one independent left eigenvector, e.g., $u_1 = \begin{bmatrix} 0 & 1 \end{bmatrix}$ and A_2 has independent left eigenvectors u_1 and $u_2 = \begin{bmatrix} 1 & 0 \end{bmatrix}$. Moreover, under the condition $\beta\gamma = 0$, B has eigenvalues 0 and α and its right eigenvectors are nonzero multiples of the vectors

$$v_1 = \begin{bmatrix} \alpha + H\gamma \\ \beta \end{bmatrix}, \qquad v_2 = \begin{bmatrix} \gamma \\ -\alpha + K\beta \end{bmatrix},$$

where H and K are arbitrary (when $\alpha \neq 0$, they may be chosen to obtain vectors with nonzero entries, which are automatically independent).

Finally, the two possible commutators are

$$C_1 = [A_1, B] = \begin{bmatrix} \beta & -\alpha \\ 0 & -\beta \end{bmatrix}, \qquad C_2 = [A_2, B] = \begin{bmatrix} 0 & \gamma \\ -\beta & 0 \end{bmatrix}.$$

Let $A = A_1$, being $[A, B] = C_1$ singular if and only if $\beta = 0$. If C_1 is singular, then v_1 is a right eigenvector of B (α or γ have to be nonzero) and $u_1 v_1 = 0$. On the other hand, consider $v = \begin{bmatrix} x & y \end{bmatrix}^\top$. If it satisfies $uv = 0$ for some $u \in \mathcal{L}_A$, then $u_1 v = 0$, hence $y = 0$ (and $x \neq 0$). So, if $v \in \mathcal{R}_B$, it follows that $\beta x = 0$, thus $\beta = 0$ and C_1 is singular.

By choosing $A = A_2$, it follows that $[A, B] = C_2$ is not invertible if and only if $\beta\gamma = 0$. If C_2 is singular, either $\beta = 0$ or $\gamma = 0$, thus either $u_1 v_1 = 0$ or $u_2 v_2 = 0$. Conversely, if $uv = 0$ for $u \in \mathcal{L}_A$ and a generic $v \in \mathcal{R}_B$ as before, then either $u_1 v = 0$ or $u_2 v = 0$. In the first case, as we showed, $\beta = 0$; analogously, in the second case, $\gamma = 0$. Concluding, in both cases $\beta\gamma = 0$ and so C_2 is singular. □

To conclude this section, a result is given on the number of linearly independent matrices in the set $\{A^i S B^j\}_{0 \leq i < m, 0 \leq j < n}$ when condition (4) of Theorem 3.1 is not satisfied.

The general case demands an extremely complicated notation: only the case of cyclic and diagonalizable matrices A and B will be considered in this paper.

Theorem 3.8. *Let $S \in \mathbb{F}^{m \times n}$ and suppose that $A \in \mathbb{F}^{m \times m}$ and $B \in \mathbb{F}^{n \times n}$ are cyclic and diagonalizable. In particular, be $U \in \mathbb{E}^{m \times m}$ and $V \in \mathbb{E}^{n \times n}$ two invertible matrices, in some extension field $\mathbb{E}$ of $\mathbb{F}$, such that UAU^{-1} and $V^{-1}BV$ are diagonal.*

Then, the dimension of $\mathcal{V}_{A,B;S}$, is equal to the number of nonzero entries of USV.

Proof. Let α_h, $0 \leq h < m$ and β_k, $0 \leq k < n$, be the left eigenvalues of A associated with the rows of U and, respectively, the right eigenvalues of B associated with the columns of V.

Since A and B are cyclic and diagonalizable, they have no repeated eigenvalues, whence $\mathcal{V}^m_{\alpha_0,\ldots,\alpha_{m-1}}$ and $\mathcal{V}^n_{\beta_0,\ldots,\beta_{n-1}}$ are invertible Vandermonde matrices.

By Proposition 3.5, we have that

$$(V^\top \otimes U) R_{A,B;S} = \operatorname{diag}(USV)(\mathcal{V}^n_{\beta_0,\ldots,\beta_{n-1}} \otimes \mathcal{V}^m_{\alpha_0,\ldots,\alpha_{m-1}}),$$

where both Kronecker products are invertible. So, $\operatorname{rank} R_{A,B;S} = \operatorname{rank} \operatorname{diag}(USV)$, which is equal to the number of nonzero entries of USV.

The proof is concluded, since by definition (14), the (column) rank of $R_{A,B;S}$ is equal to the dimension of the space spanned by $\{A^i S B^j\}$. □

4. The irreducible case

For the remainder of the paper we will asssume that $\mathbb{F} = \mathbb{F}_q$ represents the finite field of order q.

The main result of this section will provide a necessary and sufficient condition for matrices A, B having irreducible characteristic polynomial which guarantees that condition (3) of Theorem 3.1 holds true:

Theorem 4.1. *Let $\mathbb{F}$ be a finite field and suppose that $A \in \mathbb{F}^{m\times m}$ and $B \in \mathbb{F}^{n\times n}$ have irreducible characteristic polynomials. Then,*

$$\mathcal{V}_{A,B;S} = \mathbb{F}^{m\times n}, \forall S \in \mathbb{F}^{m\times n} \setminus \{0\} \Leftrightarrow \gcd(m,n) = 1.$$

Proof. Define the $\mathbb{F}$-linear map

$$\begin{aligned} \psi : \quad \mathbb{F}^{m\times n} &\to \mathbb{F}^{m\times n} \\ Z = [z_{i,j}] &\mapsto \sum_{\substack{0\le i<m \\ 0\le j<n}} z_{i,j} A^i S B^j \end{aligned} \tag{15}$$

and note that $\mathcal{V}_{A,B;S}$ is the image of ψ. Therefore, we need to prove that $\ker\psi = \{0\}, \forall S \neq 0 \Leftrightarrow \gcd(m,n) = 1$. By (2) we obtain that

$$\mathbb{v}\left(\psi(Z)\right) = \mathbb{v}\left(\sum_{\substack{0\le i<m \\ 0\le j<n}} z_{i,j} A^i S B^j\right) = \sum_{\substack{0\le i<m \\ 0\le j<n}} z_{i,j} (B^j)^\top \otimes A^i \, \mathbb{v}(S).$$

Hence, by injectivity of $\mathbb{v}$, it follows that ψ is injective (for any choice of $S \neq 0$) if and only if the kernel of matrix $M = \sum_{0\le i<m, 0\le j<n} z_{i,j} (B^j)^\top \otimes A^i$ is trivial, i.e., M has no zero eigenvalues whenever $Z \neq 0$.

Observe first that, by the assumptions on A and B, the matrix rings $\mathbb{F}[A]$ and $\mathbb{F}[B]$ are fields. Moreover, all eigenvalue $\alpha \in \mathcal{E}_A$ and $\beta \in \mathcal{E}_B$ have $\mathbb{F}$-linearly independent powers up to degree $m-1$ and, respectively, $n-1$, being $\mathbb{F}(\alpha) \cong \mathbb{F}[A]$ and $\mathbb{F}(\beta) \cong \mathbb{F}[B]$, which are Galois extensions of $\mathbb{F}$ of degree m and, respectively, n.

By a classical result on Kronecker products (see, e.g., [3, Theorem 1, p. 411] for $\mathbb{F} = \mathbb{C}$, whose generalization to finite fields is straightorward) the set of eigenvalues of M is

$$\mathcal{E}_M = \left\{ \sum_{\substack{0\le i<m \\ 0\le j<n}} z_{i,j} \alpha^i \beta^j : \alpha \in \mathcal{E}_A, \beta \in \mathcal{E}_B \right\}, \tag{16}$$

where all eigenvalues are considered as elements in some common field extension.

So, $\ker\psi = \{0\}$ if and only if each sum in (16) is nonzero. In other words, for any two $\alpha \in \mathcal{E}_A$ and $\beta \in \mathcal{E}_B$, the products $\{\alpha^i\beta^j\}_{i<m,j<n}$ are $\mathbb{F}$-linearly independent. By [8, Proposition 5.1 and Theorem 5.5], this condition is equivalent to

$$\mathbb{F}(\alpha) \cap \mathbb{F}(\beta) = \mathbb{F}.$$

Since the intersection of $\mathbb{F}(\alpha)$ and $\mathbb{F}(\beta)$ is the field extension of $\mathbb{F}$ of degree $\gcd(m,n)$ (see [9, Theorem 2.6]), the proof is concluded. □

5. The cardinality of subsets of $\mathbb{F}[A]S\mathbb{F}[B]$

In this section we will explicitly compute the cardinality of the set $\mathbb{F}[A]S\mathbb{F}[B]$ whose relevance in Cryptography is discussed in [2, 10]. Define the space of polynomials

$$\mathcal{P}^k[s] = \{p(s) \in \mathbb{F}[s] : \deg p < k\},\ k = 0, 1, \dots$$

being, for instance, $\mathcal{P}^0 = \{0\}$ and $\mathcal{P}^1 = \mathbb{F}$.

Note that, given a square matrix M with $d = \deg \mu_M$,

$$\mathcal{P}^0[M] \subset \mathcal{P}^1[M] \subset \cdots \subset \mathcal{P}^{d-1}[M] \subset \mathcal{P}^d[M] = \mathcal{P}^k[M],\ \forall k \geq d.$$

The main objective of this section consists in calculating the cardinality of the sets

$$\mathcal{M}^{h,k}_{A,B;S} = \mathcal{P}^h[A]S\mathcal{P}^k[B] \subseteq \mathbb{F}^{m\times n},\ h,k \in \mathbb{N}_0.$$

Theorem 5.1. *Let $A \in \mathbb{F}^{m\times m}$, $B \in \mathbb{F}^{n\times n}$, and $S \in \mathbb{F}^{m\times n}$ such that $\mathcal{V}_{A,B;S} = \mathbb{F}^{m\times n}$. Then, for any $0 \leq h \leq m$ and $0 \leq k \leq n$,*

$$\left|\mathcal{M}^{h,k}_{A,B;S}\right| = \frac{(q^h-1)(q^k-1)}{q-1} + 1.$$

In order to demonstrate this statement, some specific notation and one preparatory lemma are needed.

First, for every $h \leq m$, let

$$\mathbb{F}^{h;m} = \{x \in \mathbb{F}^m : x_i = 0, \forall i = h, \dots, m-1\},$$

being therefore $\mathbb{F}^h \cong \mathbb{F}^{h;m} \subseteq \mathbb{F}^m$. Define, for every $h \leq m$ and $k \leq n$, the bilinear map

$$\begin{aligned} \varphi^{h,k} : \mathbb{F}^{h;m} \times \mathbb{F}^{k;n} &\to \mathbb{F}^{m \times n} \\ (x, y) &\mapsto x y^\top \end{aligned} \tag{17}$$

and, for the sake of simplicity, denote its image by

$$\varphi^{h,k} = \varphi^{h,k}(\mathbb{F}^{h;m} \times \mathbb{F}^{k;n}). \tag{18}$$

Lemma 5.2. *Let A, B, and S as in Theorem 5.1. Then* $\left|\mathcal{M}^{h,k}_{A,B;S}\right| = |\varphi^{h,k}|$.

Proof. It is easy to check that the map ψ defined in (15) induces a well defined restriction

$$\begin{aligned} \psi^{h,k} : \varphi^{h,k} &\to \mathcal{M}^{h,k}_{A,B;S} \\ M &\mapsto \psi(M) \end{aligned}$$

which is surjective. In fact, for every $M \in \mathcal{M}^{h,k}_{A,B;S}$, there exists $(x, y) \in \mathbb{F}^{h;m} \times \mathbb{F}^{k;n} \subseteq \mathbb{F}^m \times \mathbb{F}^n$ such that

$$M = \left(\sum_{0 \leq i < h} x_i A^i\right) S \left(\sum_{0 \leq j < k} y_j B^j\right) = \sum_{\substack{0 \leq i < m \\ 0 \leq j < n}} x_i y_j A^i S B^j = \psi(x y^\top) \in \psi^{h,k}(\varphi^{h,k}).$$

Whenever the conditions of Theorem 5.1 are satisfied, ψ is injective and therefore $\psi^{h,k}$ is a bijection between $\varphi^{h,k}$ and $\mathcal{M}^{h,k}_{A,B;S}$. □

Observe that this lemma shows that the cardinality of $\mathcal{M}^{h,k}_{A,B;S}$ is independent of the choice of A, B, and S when condition (3) is met.

The problem is now reduced to the computation of the cardinality of $\varphi^{h,k}$, defined in (18).

Proof (of Theorem 5.1). Consider again the map $\varphi^{h,k}$, defined in (17), and observe that

$$\mathbb{F}^{h;m} \times \mathbb{F}^{k;n} = (\varphi^{h,k})^{-1}(\varphi^{h,k}) = \bigcup_{Z \in \varphi^{h,k}} (\varphi^{h,k})^{-1}(Z).$$

Consequently, since the inverse images are disjoint,

$$q^h q^k = |\mathbb{F}^{h;m} \times \mathbb{F}^{k;n}| = \left| \bigcup_{Z \in \varphi^{h,k}} (\varphi^{h,k})^{-1}(Z) \right| = \sum_{Z \in \varphi^{h,k}} |(\varphi^{h,k})^{-1}(Z)|.$$

To compute the value of the summation, we have to consider two situations.

- When $Z = 0$, $\varphi^{h,k}(x, y) = xy^\top = 0$ if and only if all the products of each component of x and each component of y are zero if and only if $x = 0$ and $y = 0$ (1 case), $x = 0$ and $y \neq 0$ ($q^k - 1$ cases), or $x \neq 0$ and $y = 0$ ($q^h - 1$ cases). Therefore, $|(\varphi^{h,k})^{-1}(0)| = q^h + q^k - 1$.
- If $Z \neq 0$, observe that, by the bilinearity of $\varphi^{h,k}$, $\varphi^{h,k}(x, y) = \varphi^{h,k}(\alpha x, \alpha^{-1} y)$ for every $\alpha \in \mathbb{F} \setminus \{0\}$.

  On the other hand, if $\varphi^{h,k}(x, y) = \varphi^{h,k}(\tilde{x}, \tilde{y})$ then $\tilde{x} = \alpha x$ and $\tilde{y} = \alpha^{-1} y$ for some $\alpha \neq 0$. Indeed, considering only the indexes i and j such that $x_i y_j = \tilde{x}_i \tilde{y}_j \neq 0$, we get that

$$\frac{x_i}{\tilde{x}_i} = \frac{\tilde{y}_j}{y_j}.$$

  By the independency of the indices, it follows that $\alpha = \frac{x_i}{\tilde{x}_i} = \frac{\tilde{y}_j}{y_j}$ for every i, j. So, we conclude that $|(\varphi^{h,k})^{-1}(Z)| = |\mathbb{F} \setminus \{0\}| = q - 1$.

Putting all together,

$$\begin{aligned} q^h q^k &= |(\varphi^{h,k})^{-1}(0)| + \sum_{Z \in \varphi^{h,k} \setminus \{0\}} \left|(\varphi^{h,k})^{-1}(Z)\right| \\ &= q^h + q^k - 1 + \sum_{Z \in \varphi^{h,k} \setminus \{0\}} (q - 1) = q^h + q^k - 1 + (|\varphi^{h,k}| - 1)(q - 1), \end{aligned}$$

whence

$$|\varphi^{h,k}| = \frac{q^h q^k - q^h - q^k + 1}{q - 1} + 1 = \frac{(q^h - 1)(q^k - 1)}{q - 1} + 1.$$

Finally, the claim follows by Lemma 5.2. □

Acknowledgements

The authors are deeply grateful to the anonymous reviewers for their useful remarks and, in particular, for suggesting the use of Lemma 3.2 to simplify the statement and the proof of Theorem 3.1.

References

[1] G. Micheli, Cryptanalysis of a non-commutative key exchange protocol, Adv. Math. Commun. 9 (2015) 247–253. doi:10.3934/amc.2015.9.247.

[2] M.-C. Chang, On a matrix product question in cryptography, Linear Algebra Appl. 439 (7) (2013) 1742–1748. doi:10.1016/j.laa.2013.05.013.

[3] P. Lancaster, M. Tismenetsky, The theory of matrices – with applications, 2nd Edition, Computer science and applied mathematics, Academic Press, Orlando, 1985.

[4] M. L. J. Hautus, Controllability and observability conditions of linear autonomous systems, Nederl. Akad. Wetensch. Proc. Ser. A 72 (1969) 443–448.

[5] D. Shemesh, Common eigenvectors of two matrices, Linear Algebra Appl. 62 (1984) 11–18. doi:10.1016/0024-3795(84)90085-5.

[6] T. J. Laffey, Simultaneous reduction of sets of matrices under similarity, Linear Algebra Appl. 84 (0) (1986) 123–138. doi:10.1016/0024-3795(86)90311-3.

[7] H. Aslaksen, A. B. Sletsjøe, Generators of matrix algebras in dimension 2 and 3, Linear Algebra Appl. 430 (1) (2009) 1–6. doi:10.1016/j.laa.2006.05.022.

[8] P. M. Cohn, Algebra, Volume 3, 2nd Edition, John Wiley & Sons, Chichester, 1991.

[9] R. Lidl, H. Niederreiter, Finite Fields, 2nd Edition, Vol. 20 of Encyclopedia of Mathematics and its Applications, Cambridge University Press, Cambridge, 1997.

[10] G. Maze, C. Monico, J. Rosenthal, Public key cryptography based on semigroup actions, Adv. Math. Commun. 1 (4) (2007) 489–507. doi:10.3934/amc.2007.1.489.