BLOCK MATRICES IN LINEAR ALGEBRA

STEPHAN RAMON GARCIA AND ROGER A. HORN

Abstract. Linear algebra is best done with block matrices. As evidence in support of this thesis, we present numerous examples suitable for classroom presentation.

1. Introduction

This paper is addressed to instructors of a first course in linear algebra, who need not be specialists in the field. We aim to convince the reader that linear algebra is best done with block matrices. In particular, flexible thinking about the process of matrix multiplication can reveal concise proofs of important theorems and expose new results. Viewing linear algebra from a block-matrix perspective gives an instructor access to useful techniques, exercises, and examples.

Many of the techniques, proofs, and examples presented here are familiar to specialists in linear algebra or operator theory. We think that everyone who teaches undergraduate linear algebra should be aware of them. A popular current textbook says that block matrices "appear in most modern applications of linear algebra because the notation highlights essential structures in matrix analysis. . . " [5, p. 119].

The use of block matrices in linear algebra instruction aligns mathematics pedagogy better with topics in advanced courses in pure mathematics, computer science, data science, statistics, and other fields. For example, block-matrix techniques are standard fare in modern algorithms [3]. Textbooks such as [2–7] make use of block matrices.

We take the reader on a tour of block-matrix methods and applications. In Section 2, we use right-column partitions to explain several standard first-course results. In Section 3, we use left-column partitions to introduce the full-rank factorization, prove the invariance of the number of elements in a basis, and establish the equality of row and column rank. Instructors of a first linear algebra course will be familiar with these topics, but perhaps not with a block matrix / column partition approach to them. Section 4 concerns block-column matrices. Applications include justification of a matrix-inversion algorithm and a proof of the uniqueness of the reduced row echelon form. Block-row and block-column matrices are used in Section 5 to obtain inequalities for the rank of sums and products of matrices, along with algebraic characterizations of matrices that share the same column space or null space. The preceding material culminates in Section 6, in which we consider block matrices of several types and prove that the geometric multiplicity of an eigenvalue is at most its algebraic multiplicity. We also obtain a variety of determinantal results that are suitable for presentation in class. We conclude in Section 7 with Kronecker products and several applications.

Key words and phrases. Matrix, matrix multiplication, block matrix, Kronecker product, rank, eigenvalues.

Notation: We frame our discussion for complex matrices. However, all of our numerical examples involve only real matrices, which may be preferred by some first-course instructors. We use $\mathsf{M}_{m\times n}$ to denote the set of all $m \times n$ complex matrices; M_n denotes the set of all $n \times n$ complex matrices. Boldface letters, such as $\mathbf{a}, \mathbf{b}, \mathbf{c}$, denote column vectors; $\mathbf{e}_1, \mathbf{e}_2, \ldots, \mathbf{e}_n$ is the standard basis of $\mathbb{C}^n$. We regard elements of $\mathbb{C}^m$ as column vectors; that is, $m \times 1$ matrices. If $A \in \mathsf{M}_{m\times n}$, then each column of A belongs to $\mathsf{M}_{m\times 1}$. The transpose of a matrix A is denoted by A^{T}. The null space and column space of a matrix A are denoted by $\operatorname{null} A$ and $\operatorname{col} A$, respectively. The trace and determinant of a square matrix A are denoted by $\operatorname{tr} A$ and $\det A$, respectively.

2. Right-column partitions

If

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} 4 & 5 & 2 \\ 6 & 7 & 1 \end{bmatrix}, \tag{1}$$

then the entries of AB are dot products of rows of A with columns of B:

$$AB = \begin{bmatrix} 1\cdot 4 + 2 \cdot 6 & 1 \cdot 5 + 2 \cdot 7 & 1 \cdot 2 + 2 \cdot 1 \\ 3 \cdot 4 + 4 \cdot 4 & 3 \cdot 5 + 4 \cdot 7 & 3 \cdot 2 + 4 \cdot 1 \end{bmatrix} = \begin{bmatrix} 16 & 19 & 4 \\ 36 & 43 & 10 \end{bmatrix}. \tag{2}$$

But there are other ways to organize these computations. We examine right-column partitions in this section. If $A \in \mathsf{M}_{m\times r}$ and $B = [\mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_n] \in \mathsf{M}_{r\times n}$, then the jth column of AB is $A\mathbf{b}_j$. That is,

$$AB = [A\mathbf{b}_1\ A\mathbf{b}_2\ \ldots\ A\mathbf{b}_n]. \tag{3}$$

An intentional approach to column partitions can facilitate proofs of important results from elementary linear algebra.

Example 4. If A and B are the matrices from (1), then $B = [\mathbf{b}_1\ \mathbf{b}_2\ \mathbf{b}_3]$, in which

$$\mathbf{b}_1 = \begin{bmatrix} 4 \\ 6 \end{bmatrix}, \qquad \mathbf{b}_2 = \begin{bmatrix} 5 \\ 7 \end{bmatrix}, \quad \text{and} \quad \mathbf{b}_3 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Partitioned matrix multiplication yields the expected answer (2):

$$\begin{aligned} [A\mathbf{b}_1\ A\mathbf{b}_2\ A\mathbf{b}_3] &= \left[\begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 4 \\ 6 \end{bmatrix}\ \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 5 \\ 7 \end{bmatrix}\ \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix}\begin{bmatrix} 2 \\ 1 \end{bmatrix} \right] \\ &= \left[\begin{bmatrix} 16 \\ 36 \end{bmatrix}\ \begin{bmatrix} 19 \\ 43 \end{bmatrix}\ \begin{bmatrix} 4 \\ 10 \end{bmatrix} \right] = \begin{bmatrix} 16 & 19 & 4 \\ 36 & 43 & 10 \end{bmatrix} \\ &= AB. \end{aligned}$$

Example 5. Matrix-vector equations can be bundled together. For example, suppose that $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k$ are eigenvectors of $A \in \mathsf{M}_n$ for the eigenvalue λ and let $X = [\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_k] \in \mathsf{M}_{n\times k}$. Then

$$AX = [A\mathbf{x}_1\ A\mathbf{x}_2\ \ldots\ A\mathbf{x}_k] = [\lambda\mathbf{x}_1\ \lambda\mathbf{x}_2\ \ldots\ \lambda\mathbf{x}_k] = \lambda X.$$

This observation can be used to prove that the geometric multiplicity of an eigenvalue is at most its algebraic multiplicity; see Example 36.

The following example provides a short proof of an important implication in "the invertible matrix theorem," which is in the core of a first course in linear algebra.

Example 6 (Universal consistency yields right inverse)**.** If $A\mathbf{x} = \mathbf{b}$ is consistent for each $\mathbf{b} \in \mathbb{C}^n$, there are $\mathbf{b}_i \in \mathbb{C}^n$ such that $A\mathbf{b}_i = \mathbf{e}_i$ for $i = 1, 2, \ldots, n$. Then

$$A[\mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_n] = [A\mathbf{b}_1\ A\mathbf{b}_2\ \ldots\ A\mathbf{b}_n] = [\mathbf{e}_1\ \mathbf{e}_2\ \ldots\ \mathbf{e}_n] = I,$$

so $AB = I$ for $B = [\mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_n]$.

In the preceding example, we obtained a right inverse for a square matrix A. The fact that a right inverse for A is also a left inverse is nontrivial; it can fail for linear transformations if the underlying vector space is not finite dimensional [2, P.2.7]. Here is an explanation that is based on column partitions.

Example 7 (One-sided inverses are two-sided inverses)**.** If $A, B \in \mathsf{M}_n$ and $AB = I$, then $A(B\mathbf{x}) = \mathbf{x}$ for all $\mathbf{x} \in \mathbb{C}^n$ and hence $\operatorname{col} A = \mathbb{C}^n$. The Dimension Theorem [2, Cor. 2.5.4] ensures that $\operatorname{null} A = \{\mathbf{0}\}$. Partition $I - BA = [\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_n]$ according to its columns. Then

$$\begin{aligned}[][A\mathbf{x}_1\ A\mathbf{x}_2\ \ldots\ A\mathbf{x}_n] = A[\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_n] &= A(I - BA)\\ &= A - (AB)A = A - IA = 0,\end{aligned}$$

so each $\mathbf{x}_i = \mathbf{0}$ since $\operatorname{null} A = \{\mathbf{0}\}$. Thus, $I - BA = 0$ and hence $BA = I$.

Although it cannot be recommended as a practical numerical algorithm, Cramer's rule is an important concept. Why does it work?

Example 8 (Cramer's rule)**.** Let $A = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_n] \in \mathsf{M}_n$ be invertible, let $\mathbf{b} \in \mathbb{C}^n$, and let $A_i \in \mathsf{M}_n$ be the matrix obtained by replacing the ith column of A with $\mathbf{b}$. Then there is a unique $\mathbf{x} = [x_1\ x_2\ \ldots\ x_n]^{\mathsf{T}} \in \mathbb{C}^n$ such that $A\mathbf{x} = \mathbf{b}$. Cofactor expansion along the ith row of A reveals that (ith columns underlined)

$$\begin{aligned}x_i &= \det[\mathbf{e}_1\ \ldots\ \mathbf{e}_{i-1}\ \underline{\mathbf{x}}\ \mathbf{e}_{i+1}\ \ldots\ \mathbf{e}_n]\\ &= \det[A^{-1}\mathbf{a}_1\ \ \ldots\ \ A^{-1}\mathbf{a}_{i-1}\ \ \underline{A^{-1}\mathbf{b}}\ \ A^{-1}\mathbf{a}_{i+1}\ \ \ldots\ \ A^{-1}\mathbf{a}_n]\\ &= \det\big(A^{-1}[\mathbf{a}_1\ \ldots\ \mathbf{a}_{i-1}\ \underline{\mathbf{b}}\ \mathbf{a}_{i+1}\ \ldots\ \mathbf{a}_n]\big) = \det(A^{-1}A_i)\\ &= \frac{\det A_i}{\det A}.\end{aligned}$$

3. Left-column partitions

We have gotten some mileage out of partitioning the matrix on the right-hand side of a product. If we partition the matrix on the left-hand side of a product, other opportunities emerge. If $A = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_n] \in \mathsf{M}_{m\times n}$ and $\mathbf{x} = [x_1\ x_2\ \ldots\ x_n]^{\mathsf{T}} \in \mathbb{C}^n$, then

$$A\mathbf{x} = x_1\mathbf{a}_1 + x_2\mathbf{a}_2 + \cdots + x_n\mathbf{a}_n. \tag{9}$$

That is, $A\mathbf{x}$ is a linear combination of the columns of A.

The next example illustrates that relationships between geometric objects, such as vectors and subspaces, can often be framed algebraically.

Example 10 (Geometry and matrix algebra)**.** Let $A \in \mathsf{M}_{m\times n}$ and $B \in \mathsf{M}_{m\times k}$. We claim that

$$\operatorname{col} B \subseteq \operatorname{col} A \qquad \Longleftrightarrow \qquad \text{there exists an } X \in \mathsf{M}_{n\times k} \text{ such that } AX = B;$$

moreover, if the columns of A are linearly independent, then X is unique. If each column of $B = [\mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_k] \in \mathsf{M}_{m\times k}$ is a linear combination of the columns of $A \in \mathsf{M}_{m\times n}$, then (9) ensures that there are $\mathbf{x}_i \in \mathbb{C}^n$ such that $\mathbf{b}_i = A\mathbf{x}_i$ for each i;

if the columns of A are linearly independent, then the $\mathbf{x}_i$ are uniquely determined. Let $X = [\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_k] \in \mathsf{M}_{n\times k}$. Then

$$B = [\mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_k] = [A\mathbf{x}_1\ A\mathbf{x}_2\ \ldots\ A\mathbf{x}_k] = A[\mathbf{x}_1\ \mathbf{x}_2\ \ldots\ \mathbf{x}_k] = AX.$$

Conversely, if $AX = B$, then (9) indicates that each column of B lies in $\operatorname{col} A$.

The following example uses Example 10 to show that any two bases for the same subspace of $\mathbb{C}^n$ have the same number of elements [1], [2, P.3.38]. It relies on the fact that $\operatorname{tr} XY = \operatorname{tr} YX$ if both products are defined; see [2, (0.3.5)].

Example 11 (Number of elements in a basis)**.** If $\mathbf{a}_1, \mathbf{a}_2, \ldots, \mathbf{a}_r$ and $\mathbf{b}_1, \mathbf{b}_2, \ldots, \mathbf{b}_s$ are bases for the same subspace of $\mathbb{C}^n$, we claim that $r = s$. If

$$A = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_r] \in \mathsf{M}_{n\times r} \quad \text{and} \quad B = [\mathbf{b}_1\ \mathbf{b}_2\ \ldots\ \mathbf{b}_s] \in \mathsf{M}_{n\times s},$$

then $\operatorname{col} A = \operatorname{col} B$. Example 10 ensures that $B = AX$ and $A = BY$, in which $X \in \mathsf{M}_{r\times s}$ and $Y \in \mathsf{M}_{s\times r}$. Thus,

$$A(I_r - XY) = A - AXY = A - BY = A - A = 0.$$

Since A has linearly independent columns, each column of $I_r - XY$ is zero; that is, $XY = I_r$. A similar argument shows that $YX = I_s$ and hence

$$r = \operatorname{tr} I_r = \operatorname{tr} YX = \operatorname{tr} XY = \operatorname{tr} I_s = s.$$

Another consequence of the principle in Example 10 is a second explanation of the equality of left and right inverses.

Example 12 (One-sided inverses are two-sided inverses)**.** Suppose that $A, B \in \mathsf{M}_n$ and $AB = I$. If $B\mathbf{x} = \mathbf{0}$, then $\mathbf{x} = I\mathbf{x} = A(B\mathbf{x}) = A\mathbf{0} = \mathbf{0}$. This shows that $\operatorname{null} B = \{\mathbf{0}\}$. The Dimension Theorem ensures that $\operatorname{col} B = \mathbb{C}^n$, so there is an $X \in \mathsf{M}_n$ such that $I = BX$ (this is where we use Example 10). Then $BA = BAI = BABX = BIX = BX = I$.

A fundamental result from elementary linear algebra is the equality of $\operatorname{rank} A$ and $\operatorname{rank} A^{\mathsf{T}}$; that is, "column rank equals row rank." The identity (9) permits us to give a simple explanation.

Example 13 (Equality of row and column rank)**.** For $A \in \mathsf{M}_{m\times n}$, we claim that $\operatorname{rank} A = \operatorname{rank} A^{\mathsf{T}}$. We may assume that $k = \operatorname{rank} A \geq 1$. Let the columns of $B \in \mathsf{M}_{m\times k}$ be a basis for $\operatorname{col} A$. Example 10 ensures that there is an $X \in \mathsf{M}_{k\times n}$ such that $A = BX$. Thus, $A^{\mathsf{T}} = X^{\mathsf{T}} B^{\mathsf{T}}$, so $\operatorname{col} A^{\mathsf{T}} \subseteq \operatorname{col} X^{\mathsf{T}}$. Then

$$\operatorname{rank} A^{\mathsf{T}} = \dim \operatorname{col} A^{\mathsf{T}} \leq \dim \operatorname{col} X^{\mathsf{T}} \leq k = \operatorname{rank} A.$$

Now apply the same reasoning to A^{T} and obtain $\operatorname{rank} A = \operatorname{rank} A^{\mathsf{T}}$.

We finish this section with a matrix factorization that plays a role in many block-matrix arguments.

Example 14 (Full-rank factorization)**.** Let $A = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_n] \in \mathsf{M}_{m\times n}$ be nonzero, let $r = \operatorname{rank} A$, and let the columns of $X \in \mathsf{M}_{m\times r}$ be a basis for $\operatorname{col} A$. We claim that there is a unique $Y \in \mathsf{M}_{r\times n}$ such that $A = XY$; moreover, $\operatorname{rank} Y = \operatorname{rank} X = r$. Since the r columns of X are a basis for $\operatorname{col} A$, we have $\operatorname{rank} X = r$ and $\operatorname{col} A = \operatorname{col} X$. Example 10 ensures that there is a $Y \in \mathsf{M}_{r\times n}$ such that $A = XY$. Moreover, Y is unique because each column of A is a unique linear combination of the columns of X. Finally, invoke Example 13 to compute

$$r = \operatorname{rank} A^{\mathsf{T}} = \dim \operatorname{col}(Y^{\mathsf{T}} X^{\mathsf{T}}) \leq \dim \operatorname{col} Y^{\mathsf{T}} \leq r.$$

Therefore, $\operatorname{rank} Y = \operatorname{rank} Y^{\mathsf{T}} = \dim \operatorname{col} Y^{\mathsf{T}} = r$.

In the preceding example, the matrix X is never unique. One way to construct a basis for $\operatorname{col} A$ is related to the reduced row echelon form (RREF) of A. Let $A_0 = \mathbf{0} \in \mathbb{C}^n$; for each $j = 1, 2, \ldots, n$ let $A_j = [\mathbf{a}_1\ \mathbf{a}_2\ \ldots\ \mathbf{a}_j]$. For $j \in \{1, 2, \ldots, n\}$, we say that $\mathbf{a}_j$ is a *basic column* of A if $\mathbf{a}_j \notin \operatorname{col} A_{j-1}$ (that is, if $\operatorname{rank} A_j > \operatorname{rank} A_{j-1}$). The basic columns of A comprise a basis for $\operatorname{col} A$ and correspond to the pivot columns of the RREF of A; see [6, Problem 3.9.8].

4. Block columns

Let $A \in \mathsf{M}_{m\times r}$ and $B \in \mathsf{M}_{r\times n}$. Write

$$B = [B_1\ B_2],$$

in which $B_1 \in \mathsf{M}_{r\times k}$ and $B_2 \in \mathsf{M}_{r\times(n-k)}$; that is, group the first k columns of B to create B_1 and group the remaining $n-k$ columns of B to create B_2. Then,

$$AB = A[B_1\ B_2] = [AB_1\ AB_2]; \tag{15}$$

this is the block version of (3). It can be generalized to involve multiple blocks B_i. We consider two pedagogically-oriented applications of the block-column approach (15) to matrix multiplication: a justification of the "side-by-side" matrix inversion algorithm and a proof of the uniqueness of the reduced row echelon form of a matrix. First, we consider some examples that illustrate (15).

Example 16. Let A and B be as in (1) and write $B = [B_1\ B_2]$, in which

$$B_1 = \begin{bmatrix} 4 & 5 \\ 6 & 7 \end{bmatrix} \qquad \text{and} \qquad B_2 = \begin{bmatrix} 2 \\ 1 \end{bmatrix}.$$

Then

$$AB = [AB_1\ AB_2] = \left[\begin{bmatrix} 16 & 19 \\ 36 & 43 \end{bmatrix} \begin{bmatrix} 4 \\ 10 \end{bmatrix} \right],$$

as computed in (2).

Example 17 (Extending to a basis)**.** If the list $\mathbf{x}_1, \mathbf{x}_2, \ldots, \mathbf{x}_k \in \mathbb{C}^n$ is linearly independent, then it can be extended to a basis of $\mathbb{C}^n$. Equivalently, if $X \in \mathsf{M}_{n\times k}$ has linearly independent columns, then there is a $Y \in \mathsf{M}_{n\times(n-k)}$ such that $[X\ Y] \in \mathsf{M}_n$ is invertible. This observation has lots of applications; see Example 36.

Example 18 (Inversion algorithm)**.** Let $A \in \mathsf{M}_n$ be invertible and let R be a product of elementary matrices that encode a sequence of row operations that row reduces A to I. Then $RA = I$; that is, $R = A^{-1}$. Then (15) ensures that

$$R[A\ I] = [RA\ R] = [I\ A^{-1}].$$

Thus, if one can row reduce the block matrix $[A\ I]$ to $[I\ X]$, then $X = A^{-1}$.

Our second application of block columns is the uniqueness of the RREF. The RREF underpins almost everything in a typical first linear algebra course. It is used to parametrize solution sets of systems of linear equations and to compute the rank of a small matrix (for practical computations other procedures are preferred [3]).

Example 19 (Uniqueness of RREF)**.** We claim that each $A \in \mathsf{M}_{m\times n}$ has a unique reduced row echelon form E. If $A = 0$, then $E = 0$ so we assume that $A \neq 0$ and proceed by induction on the number of columns of A. In the base case $n = 1$, the RREF $E = \mathbf{e}_1$ is uniquely determined. Suppose that $n \geq 2$, partition $A = [A' \ \mathbf{a}]$, and suppose that $R \in \mathsf{M}_m$ encodes a sequence of row operations (which need not be unique) that reduce A to its RREF, which we partition as $E = [E' \ \mathbf{y}]$. Then $RA = [RA' \ R\mathbf{a}] = [E' \ \mathbf{y}]$. The induction hypothesis ensures that $RA' = E'$ is the unique RREF of A'. Let $r = \operatorname{rank} A'$. There are two cases to consider: either $\mathbf{a} \in \operatorname{col} A'$ or $\mathbf{a} \notin \operatorname{col} A'$. If $\mathbf{a} \notin \operatorname{col} A'$, then it is a basic column of A and $\mathbf{y} = \mathbf{e}_{r+1}$ is uniquely determined. If $\mathbf{a} \in \operatorname{col} A'$, then it is a unique linear combination of the basic columns of A'; that is, $\mathbf{a} = A'\mathbf{x}$, in which $\mathbf{x}$ is uniquely determined by the condition that it has a zero entry in each position corresponding to a nonbasic column of A'. Then $\mathbf{y} = R\mathbf{a} = RA'\mathbf{x} = E'\mathbf{x}$, in which both E' and $\mathbf{x}$ are uniquely determined.

5. Block rows and columns

What we have done for columns we can also do for rows. The following examples illustrate a few results derived from block-matrix multiplication. Chief among these are several important rank inequalities and characterizations of matrices with the same column space or null space.

Example 20. A numerical example illustrates the general principle. Write

$$A = \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 3 & 4 \\ 0 & 5 & 0 \end{array}\right] = [X \ Z] \quad \text{and} \quad B = \left[\begin{array}{cc} 3 & 0 \\ 1 & 4 \\ \hline 0 & 1 \end{array}\right] = \begin{bmatrix} Y \\ W \end{bmatrix}.$$

Then

$$\begin{aligned} AB &= [X \ Z] \begin{bmatrix} Y \\ W \end{bmatrix} = XY + ZW \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 3 \\ 0 & 5 \end{bmatrix} \begin{bmatrix} 3 & 0 \\ 1 & 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \\ 0 \end{bmatrix} [0 \ 1] \\ &= \begin{bmatrix} 3 & 0 \\ 3 & 12 \\ 5 & 20 \end{bmatrix} + \begin{bmatrix} 0 & 2 \\ 0 & 4 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 3 & 16 \\ 5 & 20 \end{bmatrix}. \end{aligned} \tag{21}$$

A computation verifies that this evaluation of AB agrees with the standard method.

The *rank* of $A \in \mathsf{M}_{m\times n}$ is the dimension of $\operatorname{col} A$. Bundled row and column partitions permit us to derive inequalities for $\operatorname{rank}(A+B)$ and $\operatorname{rank} AB$ without fiddling with bases, linear combinations, and spans. Block-matrix notation simplifies and streamlines our work.

Example 22 (Rank is subadditive)**.** For $A, B \in \mathsf{M}_{m\times n}$, we claim that

$$\operatorname{rank}(A + B) \leq \operatorname{rank} A + \operatorname{rank} B. \tag{23}$$

We may assume that $r = \operatorname{rank} A \geq 1$ and $s = \operatorname{rank} B \geq 1$ since there is nothing to prove if $r = 0$ or $s = 0$. Let $A = XY$ and $B = ZW$ be full-rank factorizations; see Example 14. Since $[X \ Z] \in \mathsf{M}_{m\times(r+s)}$, we have

$$\operatorname{rank}(A + B) = \dim \operatorname{col}(A + B) = \dim \operatorname{col}(XY + ZW)$$

$$= \dim \operatorname{col} \left([X\ Z] \begin{bmatrix} Y \\ W \end{bmatrix} \right) \leq \dim \operatorname{col}[X\ Z]$$
$$\leq r + s = \operatorname{rank} A + \operatorname{rank} B.$$

The preceding result could be proved by a counting argument: produce bases for $\operatorname{col} A$ and $\operatorname{col} B$ and observe that $\operatorname{col}(A+B) \subseteq \operatorname{col} A + \operatorname{col} B$. However, Example 22 has a natural advantage. Instead of dealing with the notational overhead of columns and bases, we let a block matrix do the work. This approach produces other applications too. For example, it is difficult to see a counting argument that reproduces the following result.

Example 24 (Sylvester's rank inequality)**.** For $A \in \mathsf{M}_{m\times k}$ and $B \in \mathsf{M}_{k\times n}$, we claim that

$$\operatorname{rank} A + \operatorname{rank} B - k \leq \operatorname{rank} AB.$$

Let $r = \operatorname{rank} AB$. If $r \geq 1$, then let $AB = XY$ be a full-rank factorization (Example 14), in which $X \in \mathsf{M}_{m\times r}$ and $Y \in \mathsf{M}_{r\times n}$. Define

$$C = \begin{cases} A & \text{if } r = 0, \\ [A\ X] \in \mathsf{M}_{m\times(k+r)} & \text{if } r \geq 1, \end{cases} \quad \text{and} \quad D = \begin{cases} B & \text{if } r = 0, \\ \begin{bmatrix} B \\ -Y \end{bmatrix} \in \mathsf{M}_{(k+r)\times n} & \text{if } r \geq 1. \end{cases}$$

Then $CD = 0$, so $\operatorname{col} D \subseteq \operatorname{null} C$ and

$$\operatorname{rank} A + \operatorname{rank} B \leq \operatorname{rank} C + \operatorname{rank} D \leq \operatorname{rank} C + \operatorname{nullity} C$$
$$= k + r = k + \operatorname{rank} AB$$

The following two examples reinforce an important point. Relationships between geometric objects (subspaces here) can be revealed by matrix arithmetic.

Example 25 (Matrices with the same column space)**.** Let $A, B \in \mathsf{M}_{m\times n}$. We claim that

$$\operatorname{col} A = \operatorname{col} B \iff \text{there is an invertible } S \in \mathsf{M}_n \text{ such that } A = BS.$$

The implication ($\Leftarrow$) is straightforward; we focus on ($\Rightarrow$). Suppose that $\operatorname{col} A = \operatorname{col} B$ and let the columns of $X \in \mathsf{M}_{m\times r}$ be a basis for $\operatorname{col} A$. Example 14 ensures that there are matrices $Y, Z \in \mathsf{M}_{r\times n}$ such that

$$A = XY, \qquad B = XZ, \quad \text{and} \quad \operatorname{rank} Y = \operatorname{rank} Z = r.$$

Let $U, V \in \mathsf{M}_{(m-r)\times n}$ be such that

$$R = \begin{bmatrix} Y \\ U \end{bmatrix} \in \mathsf{M}_m \quad \text{and} \quad T = \begin{bmatrix} Z \\ V \end{bmatrix} \in \mathsf{M}_m$$

are invertible. Then

$$A = XY = [X\ 0] \begin{bmatrix} Y \\ U \end{bmatrix} = [X\ 0]R$$

and

$$B = XZ = [X\ 0] \begin{bmatrix} Z \\ V \end{bmatrix} = [X\ 0]T,$$

so

$$A = [X\ 0]R = [X\ 0]T(T^{-1}R) = BS,$$

in which $S = T^{-1}R$ is invertible.

In a first linear algebra course, row reduction is often used to solve systems of linear equations. Students are taught that A and B have the same null space if $A = EB$, in which E is an elementary matrix. Since a matrix is invertible if and only if it is the product of elementary matrices, it follows that A and B have the same null space if they are row equivalent. What about the converse?

Example 26 (Matrices with the same null space)**.** Let $A, B \in \mathsf{M}_{m\times n}$. Then Example 25 ensures that

$$\begin{aligned} \operatorname{null} A = \operatorname{null} B &\iff (\operatorname{col} A^*)^\perp = (\operatorname{col} B^*)^\perp \\ &\iff \operatorname{col} A^* = \operatorname{col} B^* \\ &\iff A^* = B^*S \text{ for some invertible } S \in \mathsf{M}_m \\ &\iff A = RB \text{ for some invertible } R \in \mathsf{M}_m \end{aligned}$$

Thus, if a sequence of elementary row operations is performed on B to obtain a new matrix $A = RB$, then the linear systems $A\mathbf{x} = \mathbf{0}$ and $B\mathbf{x} = \mathbf{0}$ have the same solutions. The latter are easily described if R is chosen so that A is in row echelon form.

6. Block matrices

Having seen the advantages of block row and column partitions, we are now ready to consider both simultaneously. Let

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \quad \text{and} \quad B = \begin{bmatrix} B_{11} & B_{12} \\ B_{21} & B_{22} \end{bmatrix},$$

in which the sizes of the submatrices involved are appropriate for the following matrix multiplications to be defined:

$$AB = \begin{bmatrix} A_{11}B_{11} + A_{12}B_{21} & A_{11}B_{12} + A_{12}B_{22} \\ A_{21}B_{11} + A_{22}B_{21} & A_{21}B_{12} + A_{22}B_{22} \end{bmatrix}. \tag{27}$$

In particular, the diagonal blocks of A and B are square and the dimensions of the off-diagonal blocks are determined by context. Multiplication of larger block matrices is conducted in an analogous manner.

Example 28. Here is a numerical example of block matrix multiplication. We use horizontal and vertical bars to highlight our partitions, although we refrain from doing so in later examples. If

$$A = \left[\begin{array}{cc|c} 1 & 0 & 2 \\ 0 & 3 & 4 \\ \hline 0 & 5 & 0 \end{array}\right] \quad \text{and} \quad B = \left[\begin{array}{c|c} 3 & 0 \\ 1 & 4 \\ \hline 0 & 1 \end{array}\right],$$

then (27) ensures that

$$\begin{aligned} AB &= \begin{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 3 \\ 1 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix}[0] & \begin{bmatrix} 1 & 0 \\ 0 & 3 \end{bmatrix}\begin{bmatrix} 0 \\ 4 \end{bmatrix} + \begin{bmatrix} 2 \\ 4 \end{bmatrix}[1] \\ [0\ 5]\begin{bmatrix} 3 \\ 1 \end{bmatrix} + [0][0] & [0\ 5]\begin{bmatrix} 0 \\ 4 \end{bmatrix} + [0][1] \end{bmatrix} \\ &= \begin{bmatrix} \begin{bmatrix} 3 \\ 3 \end{bmatrix} & \begin{bmatrix} 2 \\ 16 \end{bmatrix} \\ [5] & [20] \end{bmatrix} = \begin{bmatrix} 3 & 2 \\ 3 & 16 \\ 5 & 20 \end{bmatrix}. \end{aligned}$$

This agrees with (21) and with the usual computation of the matrix product.

We are now ready for a symbolic example. Although there are more general formulas for the inverse of a 2×2 block matrix [2, P.3.28], the following special case is sufficient for our purposes.

Example 29 (Inverse of a block triangular matrix)**.** We claim that if $Y \in \mathsf{M}_n$ and $Z \in \mathsf{M}_m$ are invertible, then

$$\begin{bmatrix} Y & X \\ 0 & Z \end{bmatrix}^{-1} = \begin{bmatrix} Y^{-1} & -Y^{-1}XZ^{-1} \\ 0 & Z^{-1} \end{bmatrix}. \tag{30}$$

How can such a result be discovered? Perform row reduction with block matrices, being careful to take into account the noncommutativity of matrix multiplication:

(1) $\begin{bmatrix} Y^{-1} & 0 \\ 0 & I \end{bmatrix} \begin{bmatrix} Y & X \\ 0 & Z \end{bmatrix} = \begin{bmatrix} I & Y^{-1}X \\ 0 & Z \end{bmatrix}$ multiply first row by Y^{-1},

(2) $\begin{bmatrix} I & 0 \\ 0 & Z^{-1} \end{bmatrix} \begin{bmatrix} I & Y^{-1}X \\ 0 & Z \end{bmatrix} = \begin{bmatrix} I & Y^{-1}X \\ 0 & I \end{bmatrix}$ multiply second row by Z^{-1},

(3) $\begin{bmatrix} I & -Y^{-1}X \\ 0 & I \end{bmatrix} \begin{bmatrix} I & Y^{-1}X \\ 0 & I \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & I \end{bmatrix}$ add $-Y^{-1}X$ times the second row to first row.

The formula (30) for the inverse of a 2×2 block upper triangular matrix can be used to prove that the inverse of an upper triangular matrix is upper triangular.

Example 31 (Inverse of an upper triangular matrix)**.** We claim that if $A = [a_{ij}] \in \mathsf{M}_n$ is upper triangular and has nonzero diagonal entries, then A^{-1} is upper triangular. We proceed by induction on n. The base case $n = 1$ is clear. For the induction step, let $n \geq 2$ and suppose that every upper triangular matrix of size less than n with nonzero diagonal entries has an inverse that is upper triangular. Let $A \in \mathsf{M}_n$ be upper triangular and partition it as

$$A = \begin{bmatrix} B & \star \\ 0 & a_{nn} \end{bmatrix},$$

in which $B \in \mathsf{M}_{n-1}$ is upper triangular and $\star$ indicates an $(n-1) \times 1$ submatrix whose entries are unimportant. It follows from (30) that

$$A^{-1} = \begin{bmatrix} B^{-1} & \star \\ 0 & a_{nn}^{-1} \end{bmatrix}.$$

The induction hypothesis ensures that B^{-1} is upper triangular and hence so is A^{-1}. This completes the induction.

Determinants are a staple of many introductory linear algebra courses. Numerical recipes are often given for 2×2 and 3×3 matrices. Various techniques are occasionally introduced to evaluate larger determinants. Since the development of eigenvalues and eigenvectors is often based upon determinants via the characteristic polynomial (although this is not how modern numerical algorithms approach the subject [3]), techniques to compute determinants of larger matrices should be a welcome addition to the curriculum. This makes carefully-crafted problems involving

4×4 or 5×5 matrices accessible to manual computation. Many of the following examples can be modified by the instructor to provide a host of interesting determinant and eigenvalue problems.

To begin, we make an observation: if $A \in \mathsf{M}_n$, then

$$\det \begin{bmatrix} A & 0 \\ 0 & I_m \end{bmatrix} = \det A = \det \begin{bmatrix} I_m & 0 \\ 0 & A \end{bmatrix}. \tag{32}$$

There are several ways to establish these identities, each suited to a different approach to determinants. For example, one could establish the first equality (32) by row reduction. The same row operations used to compute the RREF of A are used to compute the RREF of the first block matrix. Thus, the first two determinants are equal. One could also induct on m. In the base case $m = 0$, the given block matrices are simply A itself. The inductive step follows from Laplace (cofactor) expansion along either the last or the first row of the the given block matrix. The identities (32) also follow readily from the combinatorial definition of the determinant.

If $A \in \mathsf{M}_n$, then its characteristic polynomial $p_A(z) = \det(zI - A)$ is a monic polynomial of degree n and its zeros are the eigenvalues of A. The following example indicates how to compute the characteristic polynomial of a block triangular matrix.

Example 33. Let A and D be square. Then

$$\begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = \begin{bmatrix} I & 0 \\ 0 & D \end{bmatrix} \begin{bmatrix} I & B \\ 0 & I \end{bmatrix} \begin{bmatrix} A & 0 \\ 0 & I \end{bmatrix}, \tag{34}$$

so (32) and the multiplicativity of the determinant ensure that

$$\det \begin{bmatrix} A & B \\ 0 & D \end{bmatrix} = (\det A)(\det D). \tag{35}$$

If M denotes the block matrix (34), then

$$p_M(z) = \det \begin{bmatrix} zI - A & -B \\ 0 & zI - D \end{bmatrix} = \det(zI - A)\det(zI - D) = p_A(z)p_D(z).$$

This is an important property of block-triangular matrices: the characteristic polynomial of the block matrix is the product of the characteristic polynomials of the diagonal blocks. This has many consequences; see Examples 36 and 37.

If $A \in \mathsf{M}_n$, then the *geometric multiplicity* of an eigenvalue λ is $\dim \operatorname{null}(A - \lambda I)$. The *algebraic multiplicity* of λ is its multiplicity as a root of p_A. A fundamental result is that the geometric multiplicity of an eigenvalue cannot exceed its algebraic multiplicity. Strict inequality can occur even for small matrices that arise in a first linear algebra course. For example, the elementary matrices

$$A = \begin{bmatrix} 1 & c \\ 0 & 1 \end{bmatrix}, \qquad c \neq 0$$

have this property. Fortunately, block-matrix multiplication provides a way to explain what is going on.

Example 36 (Geometric multiplicity $\leq$ algebraic multiplicity)**.** Let $A \in \mathsf{M}_n$ and let λ be an eigenvalue of A. Suppose that the columns of $X \in \mathsf{M}_{n \times k}$ form a basis for the corresponding eigenspace; see Example 5. Choose $Y \in \mathsf{M}_{n\times(n-k)}$ such that $S = [X\ Y] \in \mathsf{M}_n$ is invertible; see Example 17. Then $AX = \lambda X$ and

$$\begin{bmatrix} I_k & 0 \\ 0 & I_{n-k} \end{bmatrix} = I_n = S^{-1}S = [S^{-1}X\ \ S^{-1}Y], \quad \text{so} \quad S^{-1}X = \begin{bmatrix} I_k \\ 0 \end{bmatrix}.$$

Thus,

$$\begin{aligned} S^{-1}AS &= S^{-1}A[X\ Y] = S^{-1}[AX\ AY] = S^{-1}[\lambda X\ AY] \\ &= [\lambda S^{-1}X\ \ S^{-1}AY] = \begin{bmatrix} \lambda I_k & \star \\ 0 & C \end{bmatrix}, \end{aligned}$$

in which $\star$ denotes a $k \times (n-k)$ submatrix whose entries are of no interest. Since similar matrices have the same characteristic polynomial, Example 33 ensures that $p_A(z) = p_{S^{-1}AS}(z) = (z-\lambda)^k p_C(z)$. Consequently, $k = \operatorname{nullity}(A - \lambda I)$ is at most the multiplicity of λ as a zero of $p_A(z)$.

Students should be warned repeatedly that matrix multiplication is noncommutative. That is, if $A \in \mathsf{M}_{m\times n}$ and $B \in \mathsf{M}_{n\times m}$, then AB need not equal BA, even if both products are defined. Students may be pleased to learn that $AB \in \mathsf{M}_m$ and $BA \in \mathsf{M}_n$ are remarkably alike, despite potentially being of different sizes. This fact has an elegant explanation using block matrices.

Example 37 (AB versus BA)**.** If $A \in \mathsf{M}_{m\times n}$ and $B \in \mathsf{M}_{n\times m}$, then

$$\begin{bmatrix} AB & A \\ 0 & 0_n \end{bmatrix} \qquad \text{and} \qquad \begin{bmatrix} 0_m & A \\ 0 & BA \end{bmatrix} \tag{38}$$

are similar since

$$\begin{bmatrix} I_m & 0 \\ B & I_n \end{bmatrix} \begin{bmatrix} AB & A \\ 0 & 0_n \end{bmatrix} = \begin{bmatrix} 0_m & A \\ 0 & BA \end{bmatrix} \begin{bmatrix} I_m & 0 \\ B & I_n \end{bmatrix},$$

in which the intertwining matrix is invertible. Since similar matrices have the same characteristic polynomial, Example 33 ensures that

$$z^n p_{AB}(z) = z^m p_{BA}(z). \tag{39}$$

Thus, the nonzero eigenvalues of AB and BA are the same, with the same multiplicities. In fact, one can show that the Jordan canonical forms of AB and BA differ only in their treatment of the eigenvalue zero [2, Thm. 11.9.1].

The preceding facts about AB and BA are more than just curiosities. Example 37 can be used to compute the eigenvalues of certain large, structured matrices. Suppose that $A \in \mathsf{M}_n$ has rank $r < n$. If $A = XY$, in which $X, Y^{\mathsf{T}} \in \mathsf{M}_{n\times r}$ is a full-rank factorization (Example 14), then the eigenvalues of A are the eigenvalues of the $r \times r$ matrix YX, along with $n - r$ zero eigenvalues. Consider the following example.

Example 40. What are the eigenvalues of

$$A = \begin{bmatrix} 2 & 3 & 4 & \cdots & n+1 \\ 3 & 4 & 5 & \cdots & n+2 \\ 4 & 5 & 6 & \cdots & n+3 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ n+1 & n+2 & n+3 & \cdots & 2n \end{bmatrix}?$$

The column space of A is spanned by

$$\mathbf{e} = [1\ 1\ \ldots\ 1]^{\mathsf{T}} \qquad \text{and} \qquad \mathbf{r} = [1\ 2\ \ldots\ n]^{\mathsf{T}}$$

since the jth column of A is $\mathbf{r} + j\mathbf{e}$. The list $\mathbf{r}, \mathbf{e}$ is linearly independent, so it is a basis for $\operatorname{col} A$. Let $X = [\mathbf{r}\ \mathbf{e}]$ and observe that the jth column of A is $X[1\ j]^{\mathsf{T}}$. This yields a full-rank factorization (Example 14) $A = XY$, in which $Y = [\mathbf{e}\ \mathbf{r}]^{\mathsf{T}}$. Example 37 says that the eigenvalues of $A = XY$ are $n-2$ zeros and the eigenvalues of the 2×2 matrix

$$YX = \begin{bmatrix} \mathbf{e}^{\mathsf{T}} \\ \mathbf{r}^{\mathsf{T}} \end{bmatrix} [\mathbf{r}\ \mathbf{e}] = \begin{bmatrix} \mathbf{e}^{\mathsf{T}}\mathbf{r} & \mathbf{e}^{\mathsf{T}}\mathbf{e} \\ \mathbf{r}^{\mathsf{T}}\mathbf{r} & \mathbf{r}^{\mathsf{T}}\mathbf{e} \end{bmatrix} = \begin{bmatrix} \frac{1}{2}n(n+1) & n \\ \frac{1}{6}n(n+1)(2n+1) & \frac{1}{2}n(n+1) \end{bmatrix},$$

which are

$$n(n+1)\left(\frac{1}{2} \pm \sqrt{\frac{2n+1}{6(n+1)}}\right).$$

Block-matrix computations can do much more than provide bonus problems and alternative proofs of results in a first linear algebra course. Here are a few examples.

Example 41 (Sylvester's determinant identity)**.** If $X \in \mathsf{M}_{m\times n}$ and $Y \in \mathsf{M}_{n\times m}$, then

$$\det(I_m + XY) = \det(I_n + YX). \tag{42}$$

This remarkable identity of Sylvester relates the determinants of an $m \times m$ matrix and an $n \times n$ matrix. It follows from (35):

$$\begin{aligned} \det(I + XY) &= \det \begin{bmatrix} I + XY & 0 \\ Y & I \end{bmatrix} = \det\left(\begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \begin{bmatrix} I & -X \\ Y & I \end{bmatrix} \right) \\ &= \det \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \det \begin{bmatrix} I & -X \\ Y & I \end{bmatrix} = \det \begin{bmatrix} I & -X \\ Y & I \end{bmatrix} \det \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \\ &= \det\left(\begin{bmatrix} I & -X \\ Y & I \end{bmatrix} \begin{bmatrix} I & X \\ 0 & I \end{bmatrix} \right) = \det \begin{bmatrix} I & 0 \\ Y & I + YX \end{bmatrix} \\ &= \det(I + YX). \end{aligned}$$

Another explanation can be based on the fact that XY and YX have the same nonzero eigenvalues, with the same multiplicities (see Example 37). With the exception of the eigenvalue 1, the matrices $I_m + XY$ and $I_n + YX$ have the same eigenvalues with the same multiplicities. Since the determinant of a matrix is the product of its eigenvalues, (42) follows.

The following elegant identity permits the evaluation of the determinant of a rank-one perturbation of a matrix whose determinant is known. In particular, (44) is a rare example of the determinant working well with matrix addition.

Example 43 (Determinant of a rank-one update)**.** If $A \in \mathsf{M}_n$ is invertible and $\mathbf{u}, \mathbf{v} \in \mathbb{C}^n$, then Sylvester's identity (42) with $X = A^{-1}\mathbf{u}$ and $Y = \mathbf{v}^{\mathsf{T}}$ yields

$$\det(A + \mathbf{u}\mathbf{v}^{\mathsf{T}}) = (\det A)\det(I + A^{-1}\mathbf{u}\mathbf{v}^{\mathsf{T}}) = (\det A)(1 + \underbrace{\mathbf{v}^{\mathsf{T}}A^{-1}\mathbf{u}}_{\text{a scalar}}). \tag{44}$$

The identity (44) can be used to create large matrices whose determinants can be computed in a straightforward manner. For example,

$$\begin{bmatrix} 2 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 1 \\ 1 & 1 & 2 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 1 & 1 & 1 & 2 \end{bmatrix} = \begin{bmatrix} 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{bmatrix} + \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 & 1 \end{bmatrix} = A + \mathbf{u}\mathbf{v}^{\mathsf{T}} \tag{45}$$

in which $A = \operatorname{diag}(1, -1, 1, -1, 1)$ and $\mathbf{u} = \mathbf{v} = [1\ 1\ \ldots\ 1]^{\mathsf{T}}$. Since $A = A^{-1}$ and $\det A = 1$, an application of (44) reveals that the determinant of the matrix in (45) is 2.

If $a \neq 0$, then the right-hand side of

$$\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc \tag{46}$$

equals $a(d - ca^{-1}b)$, a formula that generalizes to 2×2 block matrices.

Example 47 (Schur complement)**.** Let

$$M = \begin{bmatrix} A & B \\ C & D \end{bmatrix}, \tag{48}$$

in which A and D are square and A is invertible. Take determinants in

$$\begin{aligned} \begin{bmatrix} I & 0 \\ -CA^{-1} & I \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} &= \begin{bmatrix} A & B \\ 0 & D - CA^{-1}B \end{bmatrix} \begin{bmatrix} I & -A^{-1}B \\ 0 & I \end{bmatrix} \\ &= \begin{bmatrix} A & 0 \\ 0 & D - CA^{-1}B \end{bmatrix} \end{aligned} \tag{49}$$

and use (35) to obtain

$$\det \begin{bmatrix} A & B \\ C & D \end{bmatrix} = (\det A) \det(D - CA^{-1}B); \tag{50}$$

there is an analogous formula if D is invertible. Schur's formula (50) reduces the computation of a large determinant to the computation of two smaller ones. Since left- or right-multiplication by invertible matrices leaves the rank of a matrix invariant, we derive the elegant formula

$$\operatorname{rank} M = \operatorname{rank} A + \operatorname{rank}(M/A).$$

in which $M/A = D - CA^{-1}B$ is the *Schur complement of A in M*.

Example 51. If A and C commute in (48), then A, B, C, D are square matrices of the same size and (50) reduces to $\det(AD - CB)$, which bears a striking resemblance to (46). For example,

$$\begin{aligned} \det \left[\begin{array}{cc|cc} 1 & 1 & 1 & 2 \\ 1 & 1 & 3 & 4 \\ \hline 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 1 \end{array} \right] &= \det \left(\begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \right) \\ &= \det \left(\begin{bmatrix} 2 & 2 \\ 2 & 2 \end{bmatrix} - \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \right) = \det \begin{bmatrix} 1 & 0 \\ -1 & -2 \end{bmatrix} = -2. \end{aligned}$$

Example 52. Partition

$$M = \left[\begin{array}{ccc|cc} 2 & 0 & 0 & 1 & 1 \\ 0 & 2 & 0 & 1 & 1 \\ 0 & 0 & 2 & 1 & 1 \\ \hline 1 & 1 & 1 & 4 & 1 \\ 1 & 1 & 1 & 1 & 4 \end{array}\right]$$

as in (48). Then (50) ensures that

$$\begin{aligned} \det M &= \det \begin{bmatrix} 2 & 0 & 0 \\ 0 & 2 & 0 \\ 0 & 0 & 2 \end{bmatrix} \det \left(\begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} - \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \end{bmatrix} \begin{bmatrix} \frac{1}{2} & 0 & 0 \\ 0 & \frac{1}{2} & 0 \\ 0 & 0 & \frac{1}{2} \end{bmatrix} \begin{bmatrix} 1 & 1 \\ 1 & 1 \\ 1 & 1 \end{bmatrix} \right) \\ &= 8 \det \left(\begin{bmatrix} 4 & 1 \\ 1 & 4 \end{bmatrix} - \begin{bmatrix} \frac{3}{2} & \frac{3}{2} \\ \frac{3}{2} & \frac{3}{2} \end{bmatrix} \right) = 8 \det \begin{bmatrix} \frac{5}{2} & -\frac{1}{2} \\ -\frac{1}{2} & \frac{5}{2} \end{bmatrix} \\ &= 8 \cdot 6 = 48. \end{aligned}$$

From a pedagogical perspective, such techniques are desirable since they permit the consideration of problems involving matrices larger than 3×3.

7. Kronecker products

We conclude with a discussion of Kronecker products. It illustrates again that block-matrix arithmetic can be a useful pedagogical tool.

If $A = [a_{ij}] \in \mathsf{M}_{m\times n}$ and $B \in \mathsf{M}_{p\times q}$, then the *Kronecker product* of A and B is the block matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & a_{12}B & \cdots & a_{1n}B \\ a_{21}B & a_{22}B & \cdots & a_{2n}B \\ \vdots & \vdots & \ddots & \vdots \\ a_{m1}B & a_{m2}B & \cdots & a_{mn}B \end{bmatrix} \in \mathsf{M}_{mp\times nq}.$$

Example 53. If

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \end{bmatrix} \quad \text{and} \quad B = [5\ 6],$$

then

$$A \otimes B = \begin{bmatrix} B & 2B \\ 3B & 4B \end{bmatrix} = \begin{bmatrix} 5 & 6 & 10 & 12 \\ 15 & 18 & 20 & 24 \end{bmatrix}$$

and

$$B \otimes A = [5A\ 6A] = \begin{bmatrix} 5 & 10 & 6 & 12 \\ 15 & 20 & 18 & 24 \end{bmatrix}.$$

The Kronecker product interacts with ordinary matrix multiplication and addition as follows (A, B, C, D are matrices c is a scalar):

(i) $(A \otimes B)(C \otimes D) = AC \otimes BD$;

(ii) $c(A \otimes B) = (cA) \otimes B = A \otimes (cB)$;

(iii) $(A + B) \otimes C = A \otimes C + B \otimes C$;

(iv) $A \otimes (B + C) = A \otimes B + A \otimes C$;

(v) $A \otimes (B \otimes C) = (A \otimes B) \otimes C$.

If A and B are square matrices of the same size then the eigenvalues of AB need not be products of eigenvalues of A and B. However, for square matrices A and B of any size, all of the eigenvalues of $A \otimes B$ are products of eigenvalues of A and B, and all possible products (by algebraic multiplicity) occur; see [2, P.10.39]. This fact (and a related version for sums of eigenvalues) can be used by instructors who wish to construct matrices with prescribed eigenvalues and multiplicities.

Example 54. If $A\mathbf{x} = \lambda\mathbf{x}$ and $B\mathbf{y} = \mu\mathbf{y}$, then

$$(A \otimes B)(\mathbf{x} \otimes \mathbf{y}) = (A\mathbf{x}) \otimes (B\mathbf{y}) = (\lambda\mathbf{x}) \otimes (\mu\mathbf{y}) = \lambda\mu(\mathbf{x} \otimes \mathbf{y})$$

and

$$\begin{aligned}[(A \otimes I) + (I \otimes B)](\mathbf{x} \otimes \mathbf{y}) &= (A \otimes I)(\mathbf{x} \otimes \mathbf{y}) + (I \otimes B)(\mathbf{x} \otimes \mathbf{y}) \\ &= A\mathbf{x} \otimes \mathbf{y} + \mathbf{x} \otimes B\mathbf{y} \\ &= \lambda\mathbf{x} \otimes \mathbf{y} + \mu\mathbf{x} \otimes \mathbf{y} \\ &= (\lambda + \mu)(\mathbf{x} \otimes \mathbf{y}).\end{aligned}$$

That is, if λ and μ are eigenvalues of A and B, respectively, then $\lambda\mu$ is an eigenvalue of $A \otimes B$ and $\lambda + \mu$ is an eigenvalue of $A \otimes I + I \otimes B$, respectively.

Example 55. The eigenvalues of

$$\begin{bmatrix} 3 & 4 & 6 & 8 \\ 2 & 1 & 4 & 2 \\ 12 & 16 & 9 & 12 \\ 8 & 4 & 6 & 3 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} \otimes \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}$$

are -5, -5, 1, and 25; these are $5 \times (-1)$, $(-1) \times 5$, $(-1) \times (-1)$, and 5×5. The eigenvalues of each factor are -1 and 5.

Example 56. The eigenvalues of

$$\begin{bmatrix} 4 & 4 & 2 & 0 \\ 2 & 2 & 0 & 2 \\ 4 & 0 & 6 & 4 \\ 0 & 4 & 2 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 2 \\ 4 & 3 \end{bmatrix} \otimes I_2 + I_2 \otimes \begin{bmatrix} 3 & 4 \\ 2 & 1 \end{bmatrix}$$

are -2, 4, 4, and 10; these are $(-1) + (-1)$, $(-1) + 5$, $5 + (-1)$, and $5 + 5$.

We conclude with a proof of a seminal result in abstract algebra: the algebraic numbers form a field. That such a result should have a simple proof using block matrices indicates the usefulness of the method.

An algebraic number is a complex number that is a zero of a monic polynomial with rational coefficients. Let

$$f(z) = z^n + c_{n-1}z^{n-1} + c_{n-2}z^{n-2} + \cdots + c_1 z + c_0, \qquad n \geq 1.$$

The *companion matrix* of f is $C_f = [-c_0]$ if $n = 1$ and is

$$C_f = \begin{bmatrix} 0 & 0 & \ldots & 0 & -c_0 \\ 1 & 0 & \ldots & 0 & -c_1 \\ 0 & 1 & \ldots & 0 & -c_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \ldots & 1 & -c_{n-1} \end{bmatrix} \quad \text{if } n \geq 2.$$

Induction and cofactor expansion along the top row of $zI - C_f$ shows that f is the characteristic polynomial of C_f. Consequently, a complex number is algebraic if and only if it is an eigenvalue of a matrix with rational entries.

Example 57 (The algebraic numbers form a field)**.** Let α, β be algebraic numbers and suppose that $p(\alpha) = q(\beta) = 0$, in which p and q are monic polynomials with rational coefficients. Then α, β are eigenvalues of the rational matrices C_p and C_q, respectively, $\alpha\beta$ is an eigenvalue of the rational matrix $C_p \otimes C_q$, and $\alpha + \beta$ is an eigenvalue of the rational matrix $C_p \otimes I + I \otimes C_q$. If $\alpha \neq 0$ and p has degree k, then there is a rational number c such that $cz^k p(z^{-1}) = f(z^{-1})$, in which f is a rational monic polynomial and $f(\alpha^{-1}) = 0$.

Acknowledgments: This work was partially supported by National Science Foundation Grants DMS-1800123 and DMS-1265973, and the David L. Hirsch III and Susan H. Hirsch Research Initiation Grant.

References

[1] Stephan Ramon Garcia. Linearly independent spanning sets. *Amer. Math. Monthly*, 124(8):722, 2017.

[2] Stephan Ramon Garcia and Roger A. Horn. *A second course in linear algebra*. Cambridge Mathematical Textbooks. Cambridge University Press, New York, 2017.

[3] Gene H. Golub and Charles F. Van Loan. *Matrix computations*. Johns Hopkins Studies in the Mathematical Sciences. Johns Hopkins University Press, Baltimore, MD, fourth edition, 2013.

[4] Roger A. Horn and Charles R. Johnson. *Matrix analysis*. Cambridge University Press, Cambridge, second edition, 2013.

[5] David C. Lay, Steven R. Lay, and Judi J. McDonald. *Linear Algebra and Its Applications*. Pearson, fifth edition, 2015.

[6] Carl Meyer. *Matrix analysis and applied linear algebra*. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000. With 1 CD-ROM (Windows, Macintosh and UNIX) and a solutions manual (iv+171 pp.).

[7] Fuzhen Zhang. *Matrix theory*. Universitext. Springer, New York, second edition, 2011. Basic results and techniques.

Department of Mathematics, Pomona College, 610 N. College Ave., Claremont, CA 91711

E-mail address: `stephan.garcia@pomona.edu`
URL: `http://pages.pomona.edu/~sg064747`

E-mail address: `rhorn@math.utah.edu`