Elemente der Mathematik 79 (2024), 30–33
DOI 10.4171/EM/496

Short note

A ring-theoretic approach to the double-sidedness of the matrix inverse

J. M. Almira and J. Ángel Cid

Abstract. We present an inductive proof of the double-sidedness of the matrix inverse based on a property that holds true for associative rings with unity.

1 Introduction

Throughout the article, R will denote an associative ring (not necessarily commutative) with unity 1. We say that $a \in R$ is right-invertible if there exists $b \in R$ such that $a \cdot b = 1$, and such b is called a right-inverse of a.

As usual, $a \in R$ is invertible if there exists $b \in R$ such that $a \cdot b = 1 = b \cdot a$ and $a^{-1} := b$ is the inverse of a. Finally, $a \in R$ is a left-divisor of zero if there exists $x \in R$, $x \neq 0$, such that $a \cdot x = 0$.

Although not explicitly stated in this way, a careful reading of the interesting note [10] shows the following quite unexpected relation between the *uniqueness of the right-inverse* and the *existence of the inverse* in a ring.

Main Lemma. *If $a \in R$ is right-invertible, with right-inverse $b \in R$, then the following claims are equivalent.*

(1) *The right-inverse of a is unique.*

(2) *a is not a left-divisor of zero.*

(3) *a is invertible.*

Proof. $\boxed{(1) \Rightarrow (2)}$ If $a \cdot x = 0$ for some $x \in R$, then $b + x$ is also a right-inverse of a since

$$a \cdot (b + x) = a \cdot b + a \cdot x = 1 + 0 = 1.$$

So, from (1), it follows that $b = b + x$, that is, $x = 0$.

$\boxed{(2) \Rightarrow (3)}$ Let $x = 1 - b \cdot a \in R$. Then

$$a \cdot x = a \cdot (1 - b \cdot a) = a - a \cdot (b \cdot a) = a - (a \cdot b) \cdot a = a - 1 \cdot a = 0,$$

and (2) implies $x = 0$, that is, $b \cdot a = 1$.

$\boxed{(3) \Rightarrow (1)}$ Clearly, the existence of a^{-1} implies that a is left-cancelable, and in particular, if we assume that $a \cdot b = 1 = a \cdot b'$, then

$$a^{-1} \cdot (a \cdot b) = a^{-1} \cdot (a \cdot b') \implies (a^{-1} \cdot a) \cdot b = (a^{-1} \cdot a) \cdot b' \implies b = b'. \qquad \blacksquare$$

The standard example of a right-invertible element in a ring which is not left-invertible, see for instance [1, Section 4], also shows that in general the right-inverse is not unique. If all the right-invertible elements in R are in fact invertible, then R is called *Dedekind-finite*. Several interesting examples of Dedekind-finite rings can be found in [11].

It is well known that $M_n(\mathbb{K})$, the set of $n \times n$ square matrices over an arbitrary field $\mathbb{K}$, is a Dedekind-finite ring. Many elementary proofs of this fact have been published; see for instance [1, 3–5, 7] and the references therein. Now, it is clear from the Main Lemma that, in order to prove the double-sidedness of the inverse of a matrix A, it is enough to show that

$$\exists B(A \cdot B = I_n) \quad \text{and} \quad \forall X(A \cdot X = \mathcal{O}_n \implies X = \mathcal{O}_n), \tag{1}$$

where I_n and $\mathcal{O}_n$ denote the n-order identity and the zero matrix, respectively. To show (1), it would be enough to demonstrate

$$\exists B(A \cdot B = I_n) \quad \text{and} \quad (A \cdot x = 0_n \implies x = 0_n)$$

for all vectors $x = (x_1, x_2, \ldots, x_n)^t \in \mathbb{K}^n$, where $0_n := (0, 0, \ldots, 0)^t \in \mathbb{K}^n$.

Equivalently, we must demonstrate that A defines an injective map. Obviously, A is surjective since $A \cdot B = I_n$, and taking into account that (see, e.g., [2, Theorem 3.4])

$$n = \dim \operatorname{range}(A) + \dim \ker(A) = n + \dim \ker(A),$$

it follows that $\ker(A) = \{0_n\}$, which is what we wanted to prove. Thus, we can summarize this result just saying that double-sidedness of the matrix inverse holds true because a finite-dimensional vector space cannot properly contain any subspace of the same dimension.

Remark. The Main Lemma can be generalized a little bit since we can substitute 1 by a more general element $d \in R$. In particular, the following holds: given $a \in R$, we define its center $Z(a) = \{x \in R : a \cdot x = x \cdot a\}$.

Proposition. *If $a, b \in R$ are such that $a \cdot b \in Z(a)$ and a is not a left-divisor of zero, then*

$$a \cdot b = b \cdot a.$$

Proof. Set $d = a \cdot b \in Z(a)$. We have that

$$a \cdot (d - b \cdot a) = a \cdot d - a \cdot (b \cdot a) = a \cdot d - (a \cdot b) \cdot a = a \cdot d - d \cdot a = 0,$$

which implies that $d = b \cdot a$ since a is not a left-divisor of zero. $\blacksquare$

Of course, the above is applicable when the elements of R are matrices. There exist, indeed, many other criteria for commutativity of matrices. For example, if A, B are simultaneously diagonalizable, they commute. This is so because there exist diagonal matrices

D_1, D_2 and an invertible matrix P such that $A = P^{-1} \cdot D_1 \cdot P$ and $B = P^{-1} \cdot D_2 \cdot P$. Hence

$$\begin{aligned} A \cdot B &= P^{-1} \cdot D_1 \cdot P \cdot P^{-1} \cdot D_2 \cdot P \\ &= P^{-1} \cdot D_1 \cdot D_2 \cdot P = P^{-1} \cdot D_2 \cdot D_1 \cdot P \\ &= P^{-1} \cdot D_2 \cdot P \cdot P^{-1} \cdot D_1 \cdot P = B \cdot A. \end{aligned}$$

Finally, we notice that the following somewhat similar result to the Main Lemma holds for groups (see [8, Corollary 1.41]).

Proposition. *If a set G with an associative operation has a unique left-neutral element and each element of G has a right-inverse, then G is a group.*

Moreover, the uniqueness of the left-neutral is a key assumption to prove this result, and in fact, if we omit it from the hypotheses, then G is not necessarily a group. Another more standard related result, without the uniqueness hypotheses, which requires the existence of both left-neutral and left-inverses, can be found in [6, 9].

2 A simple proof by induction

An elementary proof of property (1) that avoids the concept of dimension can be constructed by induction on the size n of the matrices A and B leading in this way to a simple proof of the double-sidedness of the matrix inverse.

Theorem. *Let $A, B \in M_n(\mathbb{K})$. If $A \cdot B = I_n$, then $B \cdot A = I_n$.*

Induction seems a natural strategy in order to prove (1) since the initial case $n = 1$ is obviously true and the structure of the matrix product allows to decompose it as a product of smaller size matrix blocks. A different inductive proof can be found in [1].

(i) If $n = 1$, then A, B are elements of $\mathbb{K}$, and $A \cdot B = 1$ means $A \neq 0$ so that $A \cdot X = 0$ implies $X = 0$.

(ii) Assume now that (1) holds true for matrices of some fixed size $n \geq 1$, and let $A, B \in M_{n+1}(\mathbb{K})$ such that $A \cdot B = I_{n+1}$ and $A \cdot X = \mathcal{O}_{n+1}$ for some matrix X. Since $A \cdot B = I_{n+1}$ implies $A \neq \mathcal{O}_{n+1}$, we have that A has a column which is not null. Making an elementary operation on the columns of A (that is, interchanging two columns if needed), we get a matrix with its first column not identically null, and then some elementary operations on its rows transform A into a matrix A^* with its first column equal to $(1, 0, \ldots, 0)^t \in \mathbb{K}^{n+1}$. Note that we can write $A^* = E \cdot A \cdot F$, where E is the product of some elementary matrices and F is the transpose of an elementary matrix, and hence both matrices E and F are invertible; see [12]. (Recall that the elementary matrices are the ones you obtain after making an elementary operation on the rows of the identity matrix.) Then, for $B^* = F^{-1} \cdot B \cdot E^{-1}$ and $X^* = F^{-1} \cdot X$, it is easy to check that

$$I_{n+1} = A^* \cdot B^* = \left[\begin{array}{c|c} 1 & v^t \\ \hline 0_n & \tilde{A} \end{array}\right] \cdot \left[\begin{array}{c|c} ? & ? \\ \hline ? & \tilde{B} \end{array}\right] = \left[\begin{array}{c|c} ? & ? \\ \hline ? & \tilde{A} \cdot \tilde{B} \end{array}\right]$$

and

$$\mathcal{O}_{n+1} = A^* \cdot X^* = \left[\begin{array}{c|c} 1 & v^t \\ \hline 0_n & \tilde{A} \end{array}\right] \cdot \left[\begin{array}{c|c} \delta & z^t \\ \hline x & \tilde{X} \end{array}\right] = \left[\begin{array}{c|c} \delta + v^t \cdot x & z^t + v^t \cdot \tilde{X} \\ \hline \tilde{A} \cdot x & \tilde{A} \cdot \tilde{X} \end{array}\right].$$

From the first equality, we have that $\tilde{A} \cdot \tilde{B} = I_n$, while the second implies $\tilde{A} \cdot \tilde{X} = \mathcal{O}_n$. Then $\tilde{X} = \mathcal{O}_n$ (by induction, applied to $\tilde{A}$), and also, $\tilde{A} \cdot x = 0_n$ implies that $x = 0_n$ (again, by induction, applied to $\tilde{A}$). Finally, taking these equalities into account, we get $0 = \delta + v^t x = \delta$ and $0_n^t = z^t + v^t \cdot \tilde{X} = z^t$. Thus, $X^* = \mathcal{O}_{n+1}$, and then also $X = F \cdot X^* = \mathcal{O}_{n+1}$. So the proof by induction is done.

Acknowledgments. We are grateful to the anonymous referee for his/her carefully reading of the manuscript and his/her constructive comments that improved its presentation. We also owe him/her the last proposition in the introduction.

References

[1] J. M. Almira, An elementary inductive proof that $AB = I$ implies $BA = I$ for matrices. *Rev. Acad. Canar. Cienc.* **29** (2017), 75–80

[2] S. Axler, *Linear algebra done right*. Undergrad. Texts Math., Springer, New York, 1996

[3] C. M. Bang, A condition for two matrices to be inverses of each other. *Amer. Math. Monthly* **81** (1974), 764–767

[4] R. A. Beauregard, A short proof of the two-sidedness of matrix inverses. *Math. Mag.* **80** (2007) 135–136

[5] J. A. Cid, An excursion through the double sidedness of the matrix inverse. *College Math. J.* **52** (2021), 54–56

[6] J. B. Fraleigh, *A first course in abstract algebra*. 7th edn., Addison-Wesley Publishing, Reading, 2003

[7] P. Hill, Classroom Notes: On the matrix equation $AB = I$. *Amer. Math. Monthly* **74** (1967), 848–849

[8] N. Hindman and D. Strauss, *Algebra in the Stone–Čech compactification*. 2nd revised and extended edn., De Gruyter Textbook, Walter de Gruyter, Berlin, 2012

[9] M. Kilp, U. Knauer and A. V. Mikhalev, *Monoids, acts and categories*. De Gruyter Exp. Math. 29, Walter de Gruyter, Berlin, 2000

[10] W. Rudin, Unique right inverses are two-sided. *Amer. Math. Monthly* **92** (1985), 489–490

[11] Y. Sharifi, Dedekind-finite rings. Blog entry available at https://ysharifi.wordpress.com/2010/09/17/dedekind-finite-rings/

[12] G. Strang, *Linear algebra and its applications*. 4th edn., Cengage Learning, 2006

J. M. Almira
Departamento de Ingeniería y Tecnología de Computadores
Área de Matemática Aplicada, Facultad de Informática
Universidad de Murcia, Campus Espinardo, 30100 Murcia, Spain
jmalmira@um.es

J. Ángel Cid
CITMAga, 15782 Santiago de Compostela, Spain;
Universidade de Vigo, Departamento de Matemáticas
Campus de Ourense, 32004, Spain
angelcid@uvigo.es