



## Gems of Exposition in Elementary Linear Algebra

David Carlson; Charles R. Johnson; David Lay; A. Duane Porter

*The College Mathematics Journal*, Volume 23, Issue 4 (Sep., 1992), 299-303.

Stable URL:

<http://links.jstor.org/sici?sici=0746-8342%28199209%2923%3A4%3C299%3AGOEIEL%3E2.0.CO%3B2-Z>

---

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at <http://www.jstor.org/about/terms.html>. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

*The College Mathematics Journal* is published by Mathematical Association of America. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at <http://www.jstor.org/journals/maa.html>.

---

*The College Mathematics Journal*  
©1992 Mathematical Association of America

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact [jstor-info@umich.edu](mailto:jstor-info@umich.edu).

©2002 JSTOR

## **Gems of Exposition in Elementary Linear Algebra\***

*David Carlson, San Diego State University, San Diego, CA 92182 carlson@sdsu.edu*

*Charles R. Johnson\*\*, College of William and Mary, Williamsburg, VA 23185*

*crjohn@wmvm1.bitnet*

*David Lay, University of Maryland, College Park, MD 20742 lay@math.umd.edu*

*A. Duane Porter, University of Wyoming, Laramie, WY 82071 adporter@corral.wyo.edu*

**The Gems Collection.** Do you have a favorite topic of elementary linear algebra that you teach in an especially nice way to your students? Or, an unusual proof or problem that is particularly good for communicating a fundamental idea? The Linear Algebra Curriculum Study Group would be pleased to have your contribution to a National Science Foundation sponsored project to collect and broadly disseminate such “gems of exposition in elementary linear algebra.” We share several sample “gems” in the next section.

The purpose of the gems collection is to provide classroom teachers of linear algebra with a variety of potentially useful items. Eventually these items may find their way into textbooks. Of course, the ultimate goal is to help the approximately 140,000 students who take linear algebra annually to learn and retain fundamental ideas.

The kinds of items we imagine for the gems collection include insightful proofs of important facts; novel short problems; open-ended problems that promote the synthesis of two or more ideas; challenge problems that might stimulate able students; longer expository units; machine-oriented computational exercises; and self-contained descriptions of applications. However, we do not wish to exclude any type of item that has potential for communicating fundamental linear algebra ideas to students. Other kinds of items, such as pedagogical devices or banks of true/false questions, might well be appropriate. We do not seek items that are routine or readily available in several current textbooks. However, particularly nice “lost” items that may be found in very few, or only out-of-print, textbooks would be appropriate. Proofs that are only “slick” may not illuminate a student who does not appreciate the starting point, but conceptual proofs or brief proofs that communicate an underlying idea can convey a lasting understanding of why something is true. We naturally seek the latter, especially proofs that relate different ideas and make use of recurring tools. The focus of the gems collection is upon material for the first course in linear algebra, but items associated with important ideas that don’t often fit into, or are noticeably beyond, the first course are also welcome. Better approaches could change our view of material that is appropriate for the first course, and second courses are a matter of great and increasing importance.

---

\*Work of the Linear Algebra Curriculum Study Group, including the Gems Project, is supported by National Science Foundation grant USE-91-53284.

\*\*This note was prepared for the Study Group by Johnson, but gems may be submitted to any of the four Study Group members listed. We will publish submitter’s names with their gems.

## Sample Gems

1. *The associativity of matrix multiplication.* Among the often-used algebraic properties that matrix multiplication *does* satisfy, associativity,  $A(BC) = (AB)C$ , is perhaps the most subtle for students, and the property, therefore, merits explanation in the elementary course. If matrices are viewed as representing linear transformations, then associativity may be viewed as a simple conceptual consequence of the associativity of the composition of functions, but this is probably not helpful for many students, especially as the discussion of associativity typically comes quite early in the course. On the other hand, the most common proof via manipulation of sums with its flurry of subscripts may not be much more helpful. A disguised version of the functional composition view may be both conceptual and concrete enough to be informative to beginning students. It is also based upon an appreciation of some of the fundamental views of the mechanics of matrix multiplication.

First, note two important facts about matrix multiplication.

(a) If  $B$  is a matrix and  $\mathbf{x}$  is a vector for which  $B\mathbf{x}$  is defined, then the vector  $B\mathbf{x}$  is a linear combination of the *columns* of  $B$ , with coefficients the entries of  $\mathbf{x}$ . That is,

$$B\mathbf{x} = x_1\mathbf{b}_1 + x_2\mathbf{b}_2 + \cdots + x_n\mathbf{b}_n,$$

in which the  $n$  columns of  $B$  are

$$B = [\mathbf{b}_1 \quad \mathbf{b}_2 \quad \cdots \quad \mathbf{b}_n],$$

and

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}.$$

b) If  $A$  and  $B$  are matrices for which  $AB$  is defined, then the  $i$ th column of  $AB$  is  $A$  times the  $i$ th column of  $B$ , i.e.

$$AB = [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_n].$$

Both of these facts, among others, should be stressed shortly after the definition of matrix multiplication. Now, let  $\mathbf{x}$  be an arbitrary column vector. By (a) and the fact that matrix multiplication acts linearly, we have

$$\begin{aligned} A(B\mathbf{x}) &= A(x_1\mathbf{b}_1 + \cdots + x_n\mathbf{b}_n) = x_1A\mathbf{b}_1 + \cdots + x_nA\mathbf{b}_n \\ &= [A\mathbf{b}_1 \quad A\mathbf{b}_2 \quad \cdots \quad A\mathbf{b}_n]\mathbf{x} = (AB)\mathbf{x}, \end{aligned}$$

the last step being observation (b). Now, if we let  $\mathbf{x}$  be an arbitrary column of  $C = [\mathbf{c}_1 \quad \cdots \quad \mathbf{c}_p]$ , we have

$$A(B\mathbf{c}_i) = (AB)\mathbf{c}_i$$

or again using (b)

$$A(BC) = (AB)C,$$

which is associativity.

2. *Cramer's rule.* Though it should not be advocated as a practical means for solving systems of linear equations  $A\mathbf{x} = \mathbf{b}$ , with  $A$   $n$ -by- $n$ ,  $\mathbf{x}$   $n$ -by-1, and  $\mathbf{b}$   $n$ -by-1, Cramer's rule is an historically fundamental idea of linear algebra that is useful for the analytic representation of the sensitivity of a component of the solution  $\mathbf{x}$  to the data  $A$ ,  $\mathbf{b}$ . It is frequently taught in the elementary course, either without proof or with an algebra-intensive computational proof that is more drudgerous than illuminating. Actually it can be viewed as a corollary to the multiplicativity of the determinant function, while again reinforcing important mechanics of matrix multiplication. To see this, adopt the notation  $A \stackrel{i}{\leftarrow} \mathbf{b}$  for the  $n$ -by- $n$  matrix that results from replacing the  $i$ th column of  $A$  with the column vector  $\mathbf{b}$  and rewrite the system

$$A\mathbf{x} = \mathbf{b}$$

as

$$A \left[ I \stackrel{i}{\leftarrow} \mathbf{x} \right] = A \stackrel{i}{\leftarrow} \mathbf{b}.$$

That the latter is an expanded version of the former again simply makes use of observation (b) from the last section. Taking determinants of both sides and using the multiplicativity of the determinant function then gives

$$(\det A) \left( \det \left[ I \stackrel{i}{\leftarrow} \mathbf{x} \right] \right) = \det \left[ A \stackrel{i}{\leftarrow} \mathbf{b} \right].$$

But  $\det[I \stackrel{i}{\leftarrow} \mathbf{x}] = x_i$  by Laplace expansion, so that if  $\det A \neq 0$ , we have

$$x_i = \frac{\det[A \stackrel{i}{\leftarrow} \mathbf{b}]}{\det A},$$

which is Cramer's rule. This proof also gives practice in the important skill of exploiting the structure of sparse matrices, in this case  $I \stackrel{i}{\leftarrow} \mathbf{x}$ .

3. *Relationships between  $AB$  and  $BA$ .* Let  $A$  be an  $m$ -by- $n$  matrix and  $B$  be an  $n$ -by- $m$  matrix with  $m \leq n$ . In the elementary course we stress that  $AB$  need not equal  $BA$ . In fact, they need not even be the same dimensions if  $m \neq n$ ! However, it is important to realize that  $AB$  and  $BA$  are not independent and, in fact, have much in common. In particular the two matrices have the same eigenvalues, counting multiplicities, with  $BA$  having an additional  $n - m$  eigenvalues equal to 0. This fact is seldom mentioned in the first course and may not be known to some instructors. With the following proof, perhaps it could be included in the first course. First notice that the  $(m + n)$ -by- $(m + n)$  partitioned matrices

$$\begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \quad \text{and} \quad \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}$$

are similar to each other via the partitioned block calculation:

$$\begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix} \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix} = \begin{bmatrix} AB & ABA \\ B & BA \end{bmatrix} = \begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix} \begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}.$$

Since

$$\begin{bmatrix} I_m & A \\ 0 & I_n \end{bmatrix}$$

is nonsingular, it provides the similarity. Because

$$\begin{bmatrix} AB & 0 \\ B & 0 \end{bmatrix}$$

is block triangular, its eigenvalues are those of the two diagonal blocks,  $AB$  and the  $n$ -by- $n$  0 matrix. Similarly, the eigenvalues of

$$\begin{bmatrix} 0 & 0 \\ B & BA \end{bmatrix}$$

are the eigenvalues of  $BA$ , together with  $m$  0's. Because the two partitioned matrices are similar and similar matrices have the same eigenvalues,  $AB$  and  $BA$  must have the same nonzero eigenvalues (counting multiplicities) and the additional  $n - m$  eigenvalues of  $BA$  must all be 0, verifying the assertion made earlier. With only a little more effort it may be seen that the Jordan structure associated with each nonzero eigenvalue is the same in  $AB$  and  $BA$ . The Jordan structure of the eigenvalue 0 can be different, even when  $m = n$ , though there is a close relation.

Aside from simplicity, the above proof reinforces the use of similarity and the important tool of manipulating partitioned matrices.

4. *Challenge problem: minimum rank.* Let  $A$ ,  $m_1$ -by- $n_1$ ,  $B$ ,  $m_1$ -by- $n_2$ , and  $C$ ,  $m_2$ -by- $n_1$ , be given real matrices (any field will do) and let  $X$ ,  $m_2$ -by- $n_2$ , be free to be chosen. Discover and prove a formula for the minimum value of

$$\text{rank} \begin{bmatrix} A & B \\ C & X \end{bmatrix},$$

in which the minimum is taken over all choices of  $X$ . Your formula should be in terms of

$$\text{rank}[A, B], \text{rank} \begin{bmatrix} A \\ C \end{bmatrix}, \text{ and } \text{rank}[A].$$

Besides stressing the concept of rank, which is very important for work beyond the first course, this problem, though challenging, can be done in several ways by entirely elementary means. In addition, it provides entre to the unsolved problem in which the unspecified entries are scattered arbitrarily, rather than lying in the block  $X$ .

5. *Challenge problem: linear transformations on matrices.* Examples of natural linear transformations on vector spaces other than  $R^n$  are rather dear, and linear transformations on spaces of matrices are generally considered complicated to

analyze at an elementary level. However, linear transformations on matrices are very important to understanding matrix equations, and a surprisingly lot of their structure can be understood in an entirely elementary way. Here is a sample problem that is not difficult after a small hint and can be generalized in several ways.

Let  $A$  and  $B$  be given 2-by-2 matrices each with distinct eigenvalues and define a linear transformation  $L$  from the 2-by-2 matrices into themselves via  $L(X) = AXB$ . Determine the eigenvalues and eigenvectors of  $L$  in terms of those of  $A$  and  $B$ . (Hint: if  $\mathbf{y}$  is a right eigenvector of  $A$ ,  $A\mathbf{y} = a\mathbf{y}$ , and  $\mathbf{z}^\top$  is a left eigenvector of  $B$ ,  $\mathbf{z}^\top B = b\mathbf{z}^\top$ , calculate  $L(\mathbf{y}\mathbf{z}^\top)$ .)

*Answer:* The eigenvalues of  $L$  are all possible products of pairs eigenvalues, one from  $A$  and the other from  $B$ , and the eigenvectors of  $L$  (2-by-2 matrices) are outer products of eigenvectors, one from  $A$  and the other from  $B$ . The assumption of diagonalizability, which assures a complete set of linearly independent eigenvectors, may be relaxed, though this provides a bit of difficulty at an elementary level.

**Notes on Sources.** Item 1 is due to David Lay and is the way he explains associativity to his students. Items 2 and 3 may both be found in [1], but they have many antecedents, both published and unpublished. For example, [7] and especially [6] contain proofs of Cramer's rule in the same spirit. Understanding of the relationship between  $AB$  and  $BA$  has an intriguing and relatively elaborate history that we do not attempt to recount here. There is a variety of elegant elementary proofs of various parts of the story, though we do not know the origins of the approach described here. A sample of other approaches may be found in [3], [4] and [5], and some history may be gleaned from their references (especially those in [4]). Several further approaches have been suggested to the Gems Project, and we welcome more. Item 4 is adapted from the work of H. Woerdeman (see [8] and references therein), which is much more general; also see [2] for an overview and some open problems. Item 5 is from an elementary treatment of matrix equations by C. Johnson and otherwise unpublished.

## References

1. R. Horn and C. R. Johnson, *Matrix Analysis*, Cambridge University Press, New York, 1985.
2. C. R. Johnson, Matrix completion problems: A survey, in *Matrix Theory and Applications, Proceedings of Symposia in Applied Mathematics 40*, American Mathematical Society, Providence, 1990.
3. W. V. Parker\*, The matrices  $AB$  and  $BA$ , *American Mathematical Monthly* 60 (1953) 316.
4. J. Schmid\*, A remark on characteristic polynomials, *American Mathematical Monthly* 77 (1970) 998–999.
5. W. M. Scott\*, On characteristic roots of matrix products, *American Mathematical Monthly* 48 (1941) 201–203.
6. S. M. Robinson\*, A short proof of Cramer's rule, *Mathematics Magazine* 43 (1970) 94–95.
7. D. E. Whitford and M. S. Klamkin\*, On an elementary derivation of Cramer's rule, *American Mathematical Monthly* 60 (1953) 186–187.
8. H. Woerdeman, Minimal rank completions for block matrices, *Linear Algebra and Its Applications* 121 (1989) 105–122.

---

\*Reprinted in *Selected Papers on Algebra*, S. Montgomery, E. W. Ralston, et al., eds., MAA, Washington, DC, 1977.