Expansion of a determinant by cofactors, revisited

S. K. Berberian (9-19-2011)

Let $\mathrm{A} = (a_{ij})$ be an $n \times n$ matrix over a field F. Write $|\mathrm{A}|$ for the determinant of A, and, for each pair of indices (i,j), let A^{ij} be the (i,j)'th *minor* of A, that is, the $(n-1)\times(n-1)$ matrix obtained from A by deleting its i'th row and j'th column. The *cofactor* of a_{ij} in A is the element A_{ij} of F defined by the formula

$$\mathrm{A}_{ij} = (-1)^{i+j}\big|\mathrm{A}^{ij}\big|.$$

The objective of this note is to give a simple conceptual proof of the formula

$$(1) \qquad |\mathrm{A}| = \sum_{j=1}^{n} a_{ij}\mathrm{A}_{ij} \qquad \text{for } i = 1, \dots, n$$

(expansion by the i'th row).

In what follows, it is useful to visualize a matrix $\mathrm{A} = (a_{ij}) \in \mathbf{M}_n(\mathrm{F})$ (the algebra of $n \times n$ matrices over F) as a presentation of its ordered n'ple of column vectors

$$\gamma_1 = \begin{pmatrix} a_{11} \\ a_{21} \\ a_{31} \\ \vdots \\ a_{n1} \end{pmatrix}, \ \gamma_2 = \begin{pmatrix} a_{12} \\ a_{22} \\ a_{32} \\ \vdots \\ a_{n2} \end{pmatrix}, \ \dots, \ \gamma_n = \begin{pmatrix} a_{1n} \\ a_{2n} \\ a_{3n} \\ \vdots \\ a_{nn} \end{pmatrix},$$

with

$$e_1 = \begin{pmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ e_2 = \begin{pmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \ \dots, \ e_n = \begin{pmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{pmatrix}$$

as canonical basis of the vector space V of $n \times 1$ matrices (cf. [2], p. 204, Th. 7.3.9). We regard the mapping $\mathrm{A} \mapsto |\mathrm{A}|$ as the unique mapping $f : \mathrm{V}^n \to \mathrm{F}$ that is *multilinear* (for each i, a linear function of γ_i while the other coordinates are held fixed), *alternate* (equal to 0 whenever $\gamma_i = \gamma_j$ for some pair i,j with $i \neq j$), and *normalized* ($f(e_1, \dots, e_n) = 1$), briefly, an alternate multilinear form on V^n, normalized with respect to the canonical basis of V (cf. [2], p. 197, Lemma 7.1.13). We assume shown that $|\mathrm{A}'| = |\mathrm{A}|$, where A' is the transpose of A (cf. [2], Th. 7.3.5).

The heart of the matter is to prove the formula for expansion by, say, the first row:

$$(2) \qquad |\mathrm{A}| = \sum_{j=1}^{n} a_{1j}\mathrm{A}_{1j}$$

in other words,

$$(2') \qquad f(\gamma_1, \dots, \gamma_n) = a_{11}\mathrm{A}_{11} + a_{12}\mathrm{A}_{12} + \cdots + a_{1n}\mathrm{A}_{1n}\,.$$

The strategy of the proof is to define a function $g : \mathrm{V}^n \to \mathrm{F}$ via the right side of $(2')$,

$$(2'') \qquad g(\gamma_1, \dots, \gamma_n) = a_{11}\mathrm{A}_{11} + a_{12}\mathrm{A}_{12} + \cdots + a_{1n}\mathrm{A}_{1n}\,,$$

and to show that $g = f$ by verifying that g is an alternate multilinear form such that $g(e_1, \ldots, e_n) = 1$.

(a) *Normalization*: When $\mathrm{A} = \mathrm{I}$ (the $n \times n$ identity matrix), formula $(2'')$ reduces to $g(e_1, \ldots, e_n) = 1 \cdot \mathrm{A}_{11} = \left|\mathrm{A}^{11}\right| = 1$ (A^{11} is the $(n-1) \times (n-1)$ identity matrix).

(b) *Multilinearity*: We are to show that $g(\gamma_1, \gamma_2, \ldots, \gamma_n)$ is a linear function of γ_j for each j, while the other γ's are held fixed. In fact, each term on the right side of $(2'')$ is a linear function of γ_j for each j, whence linearity for their sum. It will suffice to verify this for the first two terms and for γ_1 and γ_2 to exhibit the pattern of the proof. At work will be the properties of the determinants $\left|\mathrm{A}^{ij}\right|$ and the distributive law in F.

Linearity of $a_{11}\mathrm{A}_{11}$ as a function of γ_1: The coefficient a_{11} is a linear function of γ_1 (of which it is the first coordinate), whereas A_{11} is independent of γ_1.

Linearity of $a_{12}\mathrm{A}_{12}$ as a function of γ_1: A_{12} is a linear function of the first column of A^{12} (a subcolumn of γ_1), whereas a_{12} is independent of γ_1; the same argument applies to the terms $a_{13}\mathrm{A}_{13}, \ldots, a_{1n}\mathrm{A}_{1n}$. Thus $g(\gamma_1, \gamma_2, \ldots, \gamma_n)$ is a linear function of γ_1.

Linearity of $a_{11}\mathrm{A}_{11}$ as a function of γ_2: A_{11} is a linear function of the first column of A^{11} (a subcolumn of γ_2), whereas the coefficient a_{11} is independent of γ_2; the same argument applies to the terms $a_{13}\mathrm{A}_{13}, \ldots, a_{1n}\mathrm{A}_{1n}$.

Linearity of $a_{12}\mathrm{A}_{12}$ as a function of γ_2: the coefficient a_{12} is a linear function of γ_2 (of which it is the first coordinate), whereas A_{12} is independent of γ_2. Thus $g(\gamma_1, \gamma_2, \ldots, \gamma_n)$ is a linear function of γ_2.

The arguments for $\gamma_3, \ldots, \gamma_n$ are straightforward variations on the foregoing themes.

(c) *Alternateness*: Assuming j, k are indices such that $j \neq k$ and $\gamma_j = \gamma_k$, it remains only to show that $g(\gamma_1, \ldots, \gamma_n) = 0$. We will show that the formula $(2'')$ reduces to the two terms $a_{1j}\mathrm{A}_{1j}$ and $a_{1k}\mathrm{A}_{1k}$ and that their sum is 0. We can suppose that $j < k$. If h is an index such that $h \neq j$ and $h \neq k$, then $\left|\mathrm{A}^{1h}\right| = 0$ because deleting row 1 and column h from A leaves intact the equality of (what is left of) columns j and k, thus

$$g(\gamma_1, \ldots, \gamma_n) = a_{1j}\mathrm{A}_{1j} + a_{1k}\mathrm{A}_{1k} = a_{1j}(\mathrm{A}_{1j} + \mathrm{A}_{1k})$$

(because $a_{1j} = a_{1k}$ on account of $\gamma_j = \gamma_k$); we need only show that

$$\mathrm{A}_{1j} + \mathrm{A}_{1k} = 0\,. \tag{3}$$

If $1 < j < k < n$, the matrices in question are

$$\mathrm{A}^{1k} = \begin{pmatrix} a_{21} & a_{22} & \ldots & a_{2j} & a_{2,j+1} & \ldots & a_{2,k-2} & a_{2,k-1} & a_{2,k+1} & \ldots & a_{2n} \\ a_{31} & a_{32} & \ldots & a_{3j} & a_{3,j+1} & \ldots & a_{3,k-2} & a_{3,k-1} & a_{3,k+1} & \ldots & a_{3n} \\ \ldots \\ a_{n1} & a_{n2} & \ldots & a_{nj} & a_{n,j+1} & \ldots & a_{n,k-2} & a_{n,k-1} & a_{n,k+1} & \ldots & a_{nn} \end{pmatrix},$$

$$\mathrm{A}^{1j} = \begin{pmatrix} a_{21} & a_{22} & \ldots & a_{2,j-1} & a_{2,j+1} & \ldots & a_{2,k-1} & a_{2k} & a_{2,k+1} & \ldots & a_{2n} \\ a_{31} & a_{32} & \ldots & a_{3,j-1} & a_{3,j+1} & \ldots & a_{3,k-1} & a_{3k} & a_{3,k+1} & \ldots & a_{3n} \\ \ldots \\ a_{n1} & a_{n2} & \ldots & a_{n,j-1} & a_{n,j+1} & \ldots & a_{n,k-1} & a_{nk} & a_{n,k+1} & \ldots & a_{nn} \end{pmatrix}.$$

The j'th column of A^{1k} (whose column indices are j) is, by assumption, equal to the $(k-1)$'th column of A^{1j} (whose column indices are k); if the column of A^{1j} with column indices k is moved to the left past

$$(k-1) - (j+1) + 1 = k - j - 1$$

columns, one arrives at the matrix A^{1k}; it follows that

$$\left|\mathrm{A}^{1k}\right| = (-1)^{k-j-1}\left|\mathrm{A}^{1j}\right|,$$

thus

$$\begin{aligned}
\mathrm{A}_{1k} &= (-1)^{1+k}\left|\mathrm{A}^{1k}\right| \\
&= (-1)^{1+k}(-1)^{k-j-1}\left|\mathrm{A}^{1j}\right| \\
&= (-1)^{2k-j}\left|\mathrm{A}^{1j}\right| \\
&= (-1)^{-j}\left|\mathrm{A}^{1j}\right| \\
&= (-1)^{-j}\cdot(-1)^{-(1+j)}\mathrm{A}_{1j} = -\mathrm{A}_{1j}
\end{aligned}$$

whence $\mathrm{A}_{1j}+\mathrm{A}_{1k} = 0$. In the remaining cases that $j = 1$ or $k = n$ or both ($j = 1 < k < n$; $j = 1 < k = n$; $1 < j < k = n$), the matrices and arguments are modified in the obvious way.

To summarize, g has been shown to have the properties that characterize f; in other words, $g = f$. Thus g, acting on the column vectors of a matrix A, produces $|\mathrm{A}|$; in other words, the formula that defines g (expansion by cofactors by the first row) produces $|\mathrm{A}|$.

In particular, g acting on the column vectors of A' produces $|\mathrm{A}'| = |\mathrm{A}|$.

Let $\mathrm{B} = \mathrm{A}' = (b_{ij})$, where $b_{ij} = a_{ji}$. Note that $\mathrm{B}^{ij} = \left(\mathrm{A}^{ji}\right)'$; for, deleting row i and column j of A' produces the same matrix as transposing the result of deleting column i and row j of A. Whence

$$\mathrm{B}_{ij} = (-1)^{i+j}\left|\mathrm{B}^{ij}\right| = (-1)^{i+j}\left|\left(\mathrm{A}^{ji}\right)'\right| = (-1)^{j+i}\left|\mathrm{A}^{ji}\right| = \mathrm{A}_{ji}\,.$$

It follows that $b_{ij}\mathrm{B}_{ij} = a_{ji}\mathrm{A}_{ji}$ and so

$$\sum_{j=1}^{n} b_{ij}\mathrm{B}_{ij} = \sum_{j=1}^{n} a_{ji}\mathrm{A}_{ji} \tag{4}$$

for every i, thus the formula for expanding $|\mathrm{B}|$ by the i'th row of B is equal to the formula for expanding $|\mathrm{A}|$ by the i'th column of A. Putting $i = 1$, we know that the left side is the expansion of $|\mathrm{B}| = |\mathrm{A}|$ by the 1st row of B, *whence the formula for expanding* $|\mathrm{A}|$ *by the first column of* A.

For later use, summing instead on i we have

$$\sum_{i=1}^{n} b_{ij}\mathrm{B}_{ij} = \sum_{i=1}^{n} a_{ji}\mathrm{A}_{ji} \tag{4$'$}$$

for every j, relating the formula for expansion of $|\mathrm{A}|$ by the j'th row of A and the formula for expansion of $|\mathrm{B}|$ by the j'th column of B.

Expansion of $|\mathrm{A}|$ *by the* k*'th column of* A, $k > 1$: Let C be the result of interchanging columns 1 and k of A, so that $\mathrm{C} = (c_{ij})$, where $c_{i1} = a_{ik}$ for all i, $c_{ik} = a_{i1}$ for all i, and $c_{ij} = a_{ij}$ for all other pairs (i, j). The idea is to expand $|\mathrm{C}| = -|\mathrm{A}|$ by column 1 of C, then change the sign of the result. But we also want the resulting formula to agree with the right side of (4) (with the indices j, i replaced by i, k). We know that the expansion of $|\mathrm{C}|$ by column 1 of C is

$$|\mathrm{C}| = \sum_{i=1}^{n} c_{i1}\mathrm{C}_{i1} = \sum_{i=1}^{n} a_{ik}\mathrm{C}_{i1}\,.$$

We wish to show that

$$|\mathrm{A}| = \sum_{i=1}^{n} a_{ik}\mathrm{A}_{ik}\,;$$

since $|\mathrm{A}| = -|\mathrm{C}|$, it will suffice to show that $\mathrm{C}_{i1} = -\mathrm{A}_{ik}$ for all i. Let us inspect C^{i1}: it is the result of deleting column 1 of C (formerly column k of A) and row i of C (equal to row i of A with a_{ik} and a_{i1} interchanged), thus

$$\mathrm{C}^{i1} = \begin{pmatrix} a_{12} & a_{13} & \cdots & a_{1,k-1} & a_{11} & a_{1,k+1} & \cdots & a_{1n} \\ a_{22} & a_{23} & \cdots & a_{2,k-1} & a_{21} & a_{2,k+1} & \cdots & a_{2n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{i-1,2} & a_{i-1,3} & \cdots & a_{i-1,k-1} & a_{i-1,1} & a_{i-1,k+1} & \cdots & a_{i-1,n} \\ a_{i+1,2} & a_{i+1,3} & \cdots & a_{i+1,k-1} & a_{i+1,1} & a_{i+1,k+1} & \cdots & a_{i+1,n} \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ a_{n2} & a_{n3} & \cdots & a_{n,k-1} & a_{n1} & a_{n,k+1} & \cdots & a_{nn} \end{pmatrix}.$$

In this matrix, if the $(k-1)$'th column (with column indices 1) is moved to the left end, past $k-2$ columns, one obtains A^{ik}, therefore

$$\left|\mathrm{C}^{i1}\right| = (-1)^{k-2}\left|\mathrm{A}^{ik}\right| = (-1)^{k}\left|\mathrm{A}^{ik}\right|,$$

whence

$$\mathrm{C}_{i1} = (-1)^{i+1}\left|\mathrm{C}^{i1}\right| = (-1)^{i+1}(-1)^{k}(-1)^{-(i+k)}\mathrm{A}_{ik} = -\mathrm{A}_{ik}\,.$$

Expansion of $|\mathrm{A}|$ *by the* j*'th row of* A, $j > 1$: The left side of (4′) is the expansion of $|\mathrm{B}| = |\mathrm{A}|$ by the j'th column of B (by the result just proved), thus the right side establishes the expansion of $|\mathrm{A}|$ by the j'th row of A.

In brief, if one has proved the existence and uniqueness of a function $\mathrm{A} \mapsto |\mathrm{A}|$ of $\mathbf{M}_n(\mathrm{F})$ into F that is an alternate multilinear function of the column vectors of A satisfying $|\mathrm{I}| = 1$, and established that $|\mathrm{A}'| = |\mathrm{A}|$, then expansion by cofactors is a straightforward elementary corollary, albeit with a lot of steps.

Inspection of textbooks available to me revealed a surprising diversity of proofs. Most (including [2]) proceed by reassociating the $n!$ terms in the development of $|\mathrm{A}|$. The proof in [3] is accomplished in four sentences, but it comes at a high cost in prerequisites.

The proof in [4] is closest in spirit to the one given above, but proceeds in the reverse order: the formula for g provides a determinant function for 2×2 matrices, is extended to $n\times n$ matrices by induction, and is followed by efficient proofs of uniqueness and further properties (such as $|\mathrm{AB}| = |\mathrm{A}|\,|\mathrm{B}|$ and $\mathrm{A}\cdot \mathrm{adj}\ \mathrm{A} = |\mathrm{A}|\,\mathrm{I}$). Thus expansion by cofactors is built into the very definition of determinant. (I am indebted to Robert Burckel for the observation that this was the method employed in the book of Artin [1].)

References

[1] Artin, Emil, *Galois theory*, Notre Dame Mathematical Lectures, No. 2, 2nd edn., Notre Dame, Indiana, 1944, 82pp. [MR 5,225c]

[2] Berberian, S. K., *Linear algebra*, Oxford University Press, Oxford/New York/Tokyo, 1992. [Zbl **0755**.15001]

[3] Bourbaki, N., *Algebra. Vol.I* (Chs. I–III), reprint of the 1974 edition, Springer, New York, 1989. [MR **90d**:00002].

[4] Hoffman, K. and Kunze, R., *Linear algebra*, Prentice-Hall, 1961; 2nd edn., 1971. [MR **23** #A3146]; MR **43** #1998]