Determinants of n×n Matrices (optional)

By Arkady Etkin

Recall that $\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \det\begin{pmatrix} a_{11}i + a_{12}j + a_{13}k \\ a_{21}i + a_{22}j + a_{23}k \\ a_{31}i + a_{32}j + a_{33}k \end{pmatrix} = a_{11}a_{22}a_{33}\det\begin{pmatrix} i \\ j \\ k \end{pmatrix} + a_{11}a_{23}a_{32}\det\begin{pmatrix} i \\ k \\ j \end{pmatrix} +$

$$+ a_{12}a_{21}a_{33}\det\begin{pmatrix} j \\ i \\ k \end{pmatrix} + a_{12}a_{23}a_{31}\det\begin{pmatrix} j \\ k \\ i \end{pmatrix} + a_{13}a_{21}a_{32}\det\begin{pmatrix} k \\ i \\ j \end{pmatrix} + a_{13}a_{22}a_{31}\det\begin{pmatrix} k \\ j \\ i \end{pmatrix}, \qquad (1)$$

where each of the determinants in the sum above is either 1 or –1.

For example, $\det\begin{pmatrix} k \\ j \\ i \end{pmatrix} = (-1)\det\begin{pmatrix} k \\ i \\ j \end{pmatrix} = (-1)^2\det\begin{pmatrix} i \\ k \\ j \end{pmatrix} = (-1)^3\det\begin{pmatrix} i \\ j \\ k \end{pmatrix} = (-1)^3$

Observe that each scalar product in front of the determinants in (1) has the form $a_{1\alpha}a_{2\beta}a_{3\gamma}$ where (α, β, γ) is a rearrangement or *permutation* of (1,2,3). Indeed, if we replace the unit directional vectors i, j, k with their generalized names e_1, e_2, e_3 in (1), then we may notice that every term in the sum is of the form $a_{1\alpha}a_{2\beta}a_{3\gamma}\det\begin{pmatrix} e_\alpha \\ e_\beta \\ e_\gamma \end{pmatrix}$:

$$a_{11}a_{22}a_{33}\det\begin{pmatrix} e_1 \\ e_2 \\ e_3 \end{pmatrix} + a_{11}a_{23}a_{32}\det\begin{pmatrix} e_1 \\ e_3 \\ e_2 \end{pmatrix} + a_{12}a_{21}a_{33}\det\begin{pmatrix} e_2 \\ e_1 \\ e_3 \end{pmatrix} + a_{12}a_{23}a_{31}\det\begin{pmatrix} e_2 \\ e_3 \\ e_1 \end{pmatrix} +$$

$$+ a_{13}a_{21}a_{32}\det\begin{pmatrix} e_3 \\ e_1 \\ e_2 \end{pmatrix} + a_{13}a_{22}a_{31}\det\begin{pmatrix} e_3 \\ e_2 \\ e_1 \end{pmatrix} \qquad (2)$$

Since the summands in (2) follow a clear pattern, our immediate goal is to represent (2) in a more compact form. To achieve this goal, note that the rearrangements of (1,2,3) may be viewed as one-to-one and onto functions on the set {1,2,3}. For example, the permutation (2,3,1) may be represented by the function σ:{1,2,3}→{1,2,3}, where σ(1)=2, σ(2)=3, and σ(3)=1.

With this notation, $a_{12}a_{23}a_{31}\det\begin{pmatrix} e_2 \\ e_3 \\ e_1 \end{pmatrix}$ becomes $a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}\det\begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ e_{\sigma(3)} \end{pmatrix}$.

The permutation ι, known as the identity permutation, is the trivial rearrangement of (1,2,3). In other words, ι(1)=1, ι(2)=2, and ι(3)=3.
Since every summand is obtained by choosing one and only one term in each row and each column of the determinant matrix, we see that

$$\det\begin{pmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{pmatrix} = \sum_{\sigma\in S_3} a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)} \det\begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ e_{\sigma(3)} \end{pmatrix} \qquad (3)$$

Here, the sum is taken over all permutations of (1,2,3). The set of all such permutations is denoted by S_3.

In order to make equation (3) the definition of the 3 by 3 determinant, it is necessary to eliminate all reference to determinants on its right side. To do this, recall that each $\det\begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ e_{\sigma(3)} \end{pmatrix}$ will be 1 or –1. It will be 1 if the number of pairwise row interchanges is even and it will be –1 if this number is odd. For any σ in S_3, let N(σ) denote the minimal number of ***transpositions***, or pairwise interchanges of components in the vector (σ(1), σ(2), σ(3)) that are necessary to make it into (1,2,3). For example, if (σ(1), σ(2), σ(3)) = (3,2,1), then N(σ) = N(3,2,1) = 1+N(1,2,3) = 1+0 = 1. In particular, $\det\begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ e_{\sigma(3)} \end{pmatrix} = (-1)^1$ for the permutation σ in our example. The reader should take a moment to realize that for a generic σ, $\det\begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ e_{\sigma(3)} \end{pmatrix} = (-1)^{N(\sigma)}$.

We may now restate the definition of a 3 by 3 determinant in our new language.

Definition: Given a 3 by 3 matrix A with entries (a_{ij}), we define det(A) by

$$\sum_{\sigma\in S_3} a_{1\sigma(1)}a_{2\sigma(2)}a_{3\sigma(3)}(-1)^{N(\sigma)}$$

Note that in the definition above, we are chiefly concerned with the parity of N(σ) (i.e. whether N(σ) is even or not). Its exact value is of no interest to us. For instance, I could have brazenly (and incorrectly) stated that N(3,2,1) = 1 + N(2,3,1) = 2 + N(2,1,3) = 3+N(1,2,3) = 3, because $(-1)^1 = (-1)^3$. In fact, if I can find a sequence of transpositions in the vector (σ(1), σ(2), σ(3)) that converts it into (1,2,3), then I immediately know that $(-1)^{N(\sigma)}$ is equal to negative one raised to the ***order***, or the number of transpositions of this sequence. This is so, because, given any two such sequences of transpositions of order s and t respectively, we see that the sequence of s transpositions will convert $\det\begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ e_{\sigma(3)} \end{pmatrix}$ into $(-1)^s$, whereas the second sequence converts this determinant to $(-1)^t$ (why?). But this implies that $(-1)^s = \det\begin{pmatrix} e_{\sigma(1)} \\ e_{\sigma(2)} \\ e_{\sigma(3)} \end{pmatrix} = (-1)^t$. Thus, the number of transpositions needed to rearrange the unit normal vectors inside the determinant to their proper order is unique up to parity. That is, we can perform this rearrangement in few or in many steps, but if the number of steps taken by one procedure is odd, then the number of steps in all other procedures is odd. Conversely, if the number of steps is even, then the number of steps in all other procedures is also even.

Comprehension check: What is the parity of N(3,1,2)? In other words, is N(3,1,2) even or odd? What is the value of $\det\begin{pmatrix} e_3 \\ e_1 \\ e_2 \end{pmatrix}$?

The idea of volume may be generalized to higher dimensions. The determinant $\begin{vmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \end{vmatrix}$, for example, may be thought as the volume of a 4-D box with dimensions 1 by 1 by 1 by 1 (What would you say is its volume?). By essentially following the same patterns of reasoning that lead to the definition of 2×2 and 3×3 determinants, you will conclude that the n-D volume of a 'parallelepiped' spanned by vectors $\mathbf{a}_1,\ldots,\mathbf{a}_n \in R^n$ is the absolute value of $\det\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$, where determinant is any function with the following properties[a]:

(i) $$\det\begin{pmatrix} 1 & \cdots & 0 \\ \vdots & & \vdots \\ 0 & \ldots & 1 \end{pmatrix} = \det\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = 1$$

(ii) $$\det\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = 0 \text{ if } \mathbf{a}_i = \mathbf{a}_j \text{ for any } i \neq j$$

[a] The fact that we are able to give a definite formula for this function using properties (i)-(iii) shows that this function is unique.

(iii) $$\det\begin{pmatrix} a_1 \\ \vdots \\ a_i+\lambda b \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_1 \\ \vdots \\ a_i \\ \vdots \\ a_n \end{pmatrix} + \lambda \det\begin{pmatrix} a_1 \\ \vdots \\ b \\ \vdots \\ a_n \end{pmatrix}$$ For any $\mathbf{b} \in R^n$ and any scalar $\lambda \in R$

Properties (i)-(iii) imply that, just like in the 3×3 case, the $n\times n$ determinant is the sum of all n-product scalars (exactly one from each row and each column) multiplied by determinants of the corresponding permutations of unit directional vectors $\mathbf{e}_1,\ldots,\mathbf{e}_n$. That is,

$$\det\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{n1} & \cdots & a_{nn} \end{pmatrix} = \det\begin{pmatrix} a_{11}e_1+\cdots+a_{1n}e_n \\ \vdots \\ a_{n1}e_1+\cdots+a_{nn}e_n \end{pmatrix} = \sum_{\sigma\in S_n} a_{1\sigma(1)}\ldots a_{n\sigma(n)} \det\begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix}$$

This leads us to the following definition.

Definition: Let A be an $n\times n$ matrix (a_{ij}), then $\det(A) = \sum_{\sigma\in S_n} a_{1\sigma(1)}a_{2\sigma(2)}\ldots a_{n\sigma(n)}(-1)^{N(\sigma)}$, where the sum is taken over all permutations of (1,2,…,n).

For those of you who know a bit of combinatorial analysis, there is a total of n! permutations of the numbers (1,2,…,n). Thus, there are 6=3! permutations of the vector (1,2,3) and therefore 6 summands in the definition of 3×3 determinant.

The above definition is computationally cumbersome, but it has significant theoretical consequences. Before we can begin to explore a few of these consequences, we'll need to develop quite a bit of theory.

It is common to represent a permutation σ ∈ S_n with a 2 by n array $\begin{pmatrix} 1 & \cdots & n \\ \sigma(1) & \cdots & \sigma(n) \end{pmatrix}$. For example, the permutation in S_4 that rearranges (1,2,3,4) into (4,2,1,3) can be denoted by $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}$. For any two permutations μ, σ ∈ S_n, the function μ σ, defined by μ σ(i) = μ(σ(i)) is a permutation, because the combined action of μ and σ is itself a reshuffling of the vector (1,2,…n):

$$(1,2,\ldots n) \to \sigma \to (\sigma(1),\sigma(2),\ldots\sigma(n)) \to \mu \to (\mu(\sigma(1)),\mu(\sigma(2)),\ldots\mu(\sigma(n))) =$$
$$= (\mu\sigma(1),\mu\sigma(2),\ldots\mu\sigma(n))$$

Example: If μ = $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 1 & 4 & 2 \end{pmatrix}$ and σ = $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}$, then you should verify that μ σ = $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 2 & 1 & 3 & 4 \end{pmatrix}$

Observe that for σ ∈ S_n, there must be a permutation μ∈ S_n that brings (σ(1),…, σ(n)) back to its natural order (1,…,n). In algebraic notation, this means that μ σ = σ μ = ι, where ι is the identity permutation (i.e. the permutation that does not shuffle (1,…,n).We refer to this μ by its special name σ^{-1} because μ is the unique inverse function of σ.[b]

Remark: In the discussion above, we regarded μ as the inverse of σ. However it is more correct to say that σ and μ are inverses of one another. We are free to say that σ is the inverse of μ and refer to σ by the name of μ^{-1}.

Example: If σ = $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 2 & 1 & 3 \end{pmatrix}$, compute σ^{-1}.

Solution: σ^{-1} must undo the shuffling of (1,2,3,4) that results from the action of σ. That is, σ^{-1} applied to (σ(1), σ(2), σ(3), σ(4)), will give back the vector (1,2,3,4). Hence (σ^{-1}(σ(1)), σ^{-1}(σ(2)), σ^{-1}(σ(3)), σ^{-1}(σ(4))) = (σ^{-1}(4), σ^{-1}(2), σ^{-1}(1), σ^{-1}(3)) = (1,2,3,4). It follows that σ^{-1}(4) = 1, σ^{-1}(2) = 2, σ^{-1}(1) = 3, σ^{-1}(3) = 4. Thus $\sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 2 & 4 & 1 \end{pmatrix}$

Comprehension Check: Let ξ = $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}$, compute ξ^{-1}

Since any permutation of (1,…,n) can be uniquely identified as the inverse of another permutation, we can restate the definition of the $n \times n$ determinant in the following form.

$$\det(A) = \det\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix} = \sum_{\sigma \in S_n} (-1)^{N(\sigma^{-1})} a_{1\sigma^{-1}(1)} \ldots a_{n\sigma^{-1}(n)} \qquad (4)$$

This form will prove useful in relating the determinant of an $n \times n$ matrix A to the determinant of the transpose of A.

[b] More generally, if f is a function with domain A and range B (i.e. f: A→B), then the function g: B→A that satisfies f(g(x))= x for all x in B and g(f(y))=y for all y in A is called the inverse of f. This inverse function is unique. To see this, suppose that h: B→A is another inverse function. Then for any x in B, h(x) = h(f(g(x)) = (h∘f)(g(x)) = g(x). Here '∘' denotes function composition.

Definition: Given an $m \times n$ matrix A = $\begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix}$, we define the $n \times m$ matrix

$A^T = \begin{pmatrix} {a_{11}}^T & \cdots & {a_{1m}}^T \\ \vdots & & \vdots \\ {a_{n1}}^T & \cdots & {a_{nm}}^T \end{pmatrix}$, called the *transpose* of A, by $(a_{ij}^T) = (a_{ji})$. That is, A^T is obtained from A by making the first column of A the first row of A^T, the second column of A—the second row of A^T and so on. In other words,

$$A^T = \begin{pmatrix} {a_{11}}^T & \cdots & {a_{1m}}^T \\ \vdots & & \vdots \\ {a_{n1}}^T & \cdots & {a_{nm}}^T \end{pmatrix} = \begin{pmatrix} a_{11} & \cdots & a_{m1} \\ \vdots & & \vdots \\ a_{1n} & \cdots & a_{nm} \end{pmatrix}$$

Example: Let A = $\begin{pmatrix} 1 & 4 & 7 & 10 \\ 2 & 5 & 8 & 11 \\ 3 & 6 & 9 & 12 \end{pmatrix}$ then $A^T = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \\ 10 & 11 & 12 \end{pmatrix}$

Comprehension Check: Compute the transpose of the following matrices:

(a) $\begin{pmatrix} 1 & -3 & 4 \\ 2 & 2 & -10 \\ 4 & 0 & 0 \end{pmatrix}$

(b) (1 –7 4 5)

(c) $\begin{pmatrix} 0 \\ -3 \\ 4 \\ 6 \end{pmatrix}$

(d) (13)

If A is an $n \times n$ matrix then so is A^T. What might be the relationship between det(A) and det(A^T)?

$$\det(A^T) = \det\begin{pmatrix} a_{11}{}^T & \cdots & a_{1n}{}^T \\ \vdots & & \vdots \\ a_{n1}{}^T & \cdots & a_{nn}{}^T \end{pmatrix} = \sum_{\sigma\in S_n} (-1)^{N(\sigma)} a_{1\sigma(1)}{}^T \ldots a_{n\sigma(n)}{}^T \qquad (5)$$

Because $a_{1\sigma(i)}{}^T = a_{\sigma(i)1}$ for any $1 \le i \le n$, equation (5) can be written as

$$\det(A^T) = \sum_{\sigma\in S_n} (-1)^{N(\sigma)} a_{\sigma(1)1} \ldots a_{\sigma(n)n} = \sum_{\sigma\in S_n} (-1)^{N(\sigma)} a_{\sigma(1)\sigma^{-1}(\sigma(1))} \ldots a_{\sigma(n)\sigma^{-1}(\sigma(n))} =$$
$$= \sum_{\sigma\in S_n} (-1)^{N(\sigma)} a_{\alpha_1\sigma^{-1}(\alpha_1)} \ldots a_{\alpha_n\sigma^{-1}(\alpha_n)} \qquad (6)$$

By rearranging the terms of the scalar product $a_{\alpha_1\sigma^{-1}(\alpha_1)} \ldots a_{\alpha_n\sigma^{-1}(\alpha_n)}$ so that the row indices are in order, the last equality in (6) becomes

$$\sum_{\sigma\in S_n} (-1)^{N(\sigma)} a_{1\sigma^{-1}(1)} \ldots a_{1\sigma^{-1}(n)} \qquad (7)$$

If we can prove that $(-1)^{N(\sigma)} = (-1)^{N(\sigma^{-1})}$, then we will be able to identify formula (7) with formula (4). This will imply that $\det(A^T) = \det(A)$. The following proposition will establish this and more.

Proposition: Let μ, σ ∈ S_n be any two permutations. Then $(-1)^{N(\mu\sigma)} = (-1)^{N(\mu)+N(\sigma)}$

Proof: $$(-1)^{N(\mu\sigma)} = \det\begin{pmatrix} e_{\mu\sigma(1)} \\ \vdots \\ e_{\mu\sigma(n)} \end{pmatrix} = \det\begin{pmatrix} e_{\mu(\sigma(1))} \\ \vdots \\ e_{\mu(\sigma(n))} \end{pmatrix} \qquad (i)$$

Denote $e_{\sigma(1)}$ by f_1, $e_{\sigma(2)}$ by f_2,…, $e_{\sigma(n)}$ by f_n. Then the rightmost expression in (i) becomes $$\det\begin{pmatrix} f_{\mu(1)} \\ \vdots \\ f_{\mu(n)} \end{pmatrix} = (-1)^{N(\mu)} \det\begin{pmatrix} f_1 \\ \vdots \\ f_n \end{pmatrix} = (-1)^{N(\mu)} \det\begin{pmatrix} e_{\sigma(1)} \\ \vdots \\ e_{\sigma(n)} \end{pmatrix} =$$

$$= (-1)^{N(\mu)}(-1)^{N(\sigma)} \det\begin{pmatrix} e_1 \\ \vdots \\ e_n \end{pmatrix} = (-1)^{N(\mu)}(-1)^{N(\sigma)} = (-1)^{N(\mu)+N(\sigma)} \qquad \blacktriangledown$$

Corollary: Let σ ∈ S_n, then $(-1)^{N(\sigma)} = (-1)^{N(\sigma^{-1})}$.

Proof: $1 = (-1)^0 = (-1)^{N(\iota)} = (-1)^{N(\sigma\sigma^{-1})} = (-1)^{N(\sigma)+N(\sigma^{-1})} = (-1)^{N(\sigma)}(-1)^{N(\sigma^{-1})}$

In particular, $1 = (-1)^{N(\sigma)}(-1)^{N(\sigma^{-1})}$. Upon multiplying the last equation by $(-1)^{N(\sigma)}$ on both sides, we obtain $(-1)^{N(\sigma)} = (-1)^{2N(\sigma)}(-1)^{N(\sigma^{-1})} = (-1)^{N(\sigma^{-1})}$ ▼

Theorem: Let A be an $n \times n$ matrix. Then $\det(A) = \det(A^T)$.

Proof: We summarize the argument for convenience.

$$\det(A^T) = \sum_{\sigma\in S_n}(-1)^{N(\sigma)} a_{\sigma(1)1}...a_{\sigma(n)n} = \sum_{\sigma\in S_n}(-1)^{N(\sigma)} a_{\sigma(1)\sigma^{-1}(\sigma(1))}...a_{\sigma(n)\sigma^{-1}(\sigma(n))} =$$

$$= \sum_{\sigma\in S_n}(-1)^{N(\sigma)} a_{1\sigma^{-1}(1)}...a_{1\sigma^{-1}(n)} = \sum_{\sigma\in S_n}(-1)^{N(\sigma^{-1})} a_{1\sigma^{-1}(1)}...a_{1\sigma^{-1}(n)} = \det(A) \qquad \blacktriangledown$$

The above theorem has a useful corollary that unlocks further properties of the determinant.

Corollary: Let A be an $n \times n$ matrix with column vectors $A_1, \cdots, A_n$. Then $\det(A) = \det(A_1 \cdots A_n)$ has the following properties:

(i) $\det(A) = \det(A_1 \cdots A_n) = 0$ if $\mathbf{A}_i = \mathbf{A}_j$ for any i≠j

(ii) $\det(A_1 \cdots A_i + \lambda B \cdots A_n) = \det(A_1 \cdots A_i \cdots A_n) + \lambda \det(A_1 \cdots B \cdots A_n)$ For any column vector **B** ∈ R^n and any scalar $\lambda \in R$

(iii) Interchanging any two columns in the determinant results in a change of sign.

Proof: The column vectors $A_1, \cdots, A_n$ of A are the row vectors of A^T. Thus, the column properties (i)-(iii) of det(A) are the row properties of det(A^T). For instance, property (i) is true because, if $\mathbf{A}_i = \mathbf{A}_j$ for some i≠j, $\det(A_1 \cdots A_n) = \det(A) = \det(A^T) = \det\begin{pmatrix} A_1^T \\ \vdots \\ A_n^T \end{pmatrix} = 0$ ▼

Example:

(a) $\begin{vmatrix} 1 & 2 & 3 \\ 1 & 2 & 3 \\ 1 & 2 & 3 \end{vmatrix} = 2\begin{vmatrix} 1 & 1 & 3 \\ 1 & 1 & 3 \\ 1 & 1 & 3 \end{vmatrix} = 0$

(b) $\begin{vmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 4+3 \\ 2 & 5 & 5+3 \\ 3 & 6 & 6+3 \end{vmatrix} = \begin{vmatrix} 1 & 4 & 4 \\ 2 & 5 & 5 \\ 3 & 6 & 6 \end{vmatrix} + \begin{vmatrix} 1 & 4 & 3 \\ 2 & 5 & 3 \\ 3 & 6 & 3 \end{vmatrix} = 0 + \begin{vmatrix} 1 & 1+3 & 3 \\ 2 & 2+3 & 3 \\ 3 & 3+3 & 3 \end{vmatrix} = \begin{vmatrix} 1 & 1 & 3 \\ 2 & 2 & 3 \\ 3 & 3 & 3 \end{vmatrix} + \begin{vmatrix} 1 & 3 & 3 \\ 2 & 3 & 3 \\ 3 & 3 & 3 \end{vmatrix} = 0$

We can use the row and the column properties to define the determinant recursively. But first, we'll need another definition.

Definition: Let A be an $n \times n$ matrix. Then the *ijth minor* of A, $\overline{A_{ij}}$, is the $(n-1)\times(n-1)$ matrix obtained by deleting the ith row and jth column from A.

Example: Let A be the 4×4 matrix $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 9 & 10 & 11 & 12 \\ 13 & 14 & 15 & 16 \end{pmatrix}$. Compute

(a) $\overline{A_{11}}$

(b) $\overline{A_{32}}$

(c) $\overline{A_{24}}$

Solution:

(a) $\overline{A_{11}} = \begin{pmatrix} 6 & 7 & 8 \\ 10 & 11 & 12 \\ 14 & 15 & 16 \end{pmatrix}$

(b) $\overline{A_{32}} = \begin{pmatrix} 1 & 3 & 4 \\ 5 & 7 & 8 \\ 13 & 15 & 16 \end{pmatrix}$

(c) $\overline{A_{24}} = \begin{pmatrix} 1 & 2 & 3 \\ 9 & 10 & 11 \\ 13 & 14 & 15 \end{pmatrix}$

Lemma: Let A be an $n \times n$ matrix with row vectors $a_{11}\mathbf{e}_1, \mathbf{a}_2, \ldots, \mathbf{a}_n$. Then $\det(A) = a_{11}\det(\overline{A_{11}})$.

Proof: Recall that

$$\det(A) = \sum_{\sigma \in S_n} a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)} (-1)^{N(\sigma)}$$

However, in this particular case, the first row of A has only one nonzero entry.

$$\det(A) = \det\begin{pmatrix} a_{11}e_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix} = \det\begin{pmatrix} a_{11} & 0 & \cdots & 0 \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{pmatrix}$$

This means that for any σ∈ S_n such that σ(1) ≠ 1, $a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)} (-1)^{N(\sigma)} = 0$ (why?) Hence,

$$\det(A) = \sum_{\sigma \in S_n; \sigma(1)=1} a_{1\sigma(1)} a_{2\sigma(2)} \ldots a_{n\sigma(n)} (-1)^{N(\sigma)} = \sum_{\sigma \in S_n; \sigma(1)=1} a_{11} a_{2\sigma(2)} \ldots a_{n\sigma(n)} (-1)^{N(\sigma)}.$$

Note that each term in the sum above has a_{11} as a factor. Also observe that for $2 \le i, j \le n$, $a_{ij} = \overline{a_{i-1\,j-1}}$, where $\overline{a_{i-1\,j-1}}$ denotes the $i-1, j-1$ entry in the matrix $\overline{A_{11}}$. Since every σ∈ S_n satisfying σ(1) =1 permutes the numbers 2, 3,…, n among themselves, each such σ can be identified with a unique permutation μ∈ S_{n-1} by resetting 2→1, 3→2,…, n→n-1.[c] With these ideas, we may write

$$\det(A) = \sum_{\sigma \in S_n; \sigma(1)=1} a_{11} a_{2\sigma(2)} \ldots a_{n\sigma(n)} (-1)^{N(\sigma)} = a_{11} \sum_{\mu \in S_{n-1}} \overline{a_{1\mu(1)}}\, \overline{a_{2\mu(2)}} \ldots \overline{a_{n-1\mu(n-1)}} (-1)^{N(\mu)}.$$

[c] For example, if σ∈ S_4 transforms the vector (1,2,3,4) into (1,4,3,2), then, by deleting 1 and setting each of the remaining numbers one unit back, we get (3,2,1). This 3-tuple vector is the work of permutation μ ∈ S_3, which is identical to σ in everything but its name.

And since the right-hand-side of the last equation is just the entry in the first row and first column of A multiplied by the determinant of the minor $\overline{A_{11}}$, the desired result follows. ▼

The next example illustrates an important step in the theorem that follows.

Example: Let A be the 4×4 matrix $\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & 0 & 11 & 0 \\ 13 & 14 & 15 & 16 \end{pmatrix}$. Then $\det\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & 0 & 11 & 0 \\ 13 & 14 & 15 & 16 \end{pmatrix} =$

$$(-1)\det\begin{pmatrix} 1 & 2 & 3 & 4 \\ 0 & 0 & 11 & 0 \\ 5 & 6 & 7 & 8 \\ 13 & 14 & 15 & 16 \end{pmatrix} = (-1)^2\det\begin{pmatrix} 0 & 0 & 11 & 0 \\ 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 13 & 14 & 15 & 16 \end{pmatrix} = (-1)^{2+1}\det\begin{pmatrix} 0 & 11 & 0 & 0 \\ 1 & 3 & 2 & 4 \\ 5 & 7 & 6 & 8 \\ 13 & 15 & 14 & 16 \end{pmatrix} =$$

$$= (-1)^{2+2}\det\begin{pmatrix} 11 & 0 & 0 & 0 \\ 3 & 1 & 2 & 4 \\ 7 & 15 & 6 & 8 \\ 15 & 13 & 14 & 16 \end{pmatrix} = (-1)^{3+3}\det\begin{pmatrix} 11 & 0 & 0 & 0 \\ 3 & 1 & 2 & 4 \\ 7 & 15 & 6 & 8 \\ 15 & 13 & 14 & 16 \end{pmatrix}$$

By the lemma above, the determinant of the last matrix is

$$(-1)^{3+3}11\det\begin{pmatrix} 1 & 2 & 4 \\ 15 & 6 & 8 \\ 13 & 14 & 16 \end{pmatrix}.$$

Notice that $\overline{A_{33}} = \begin{pmatrix} 1 & 2 & 4 \\ 15 & 6 & 8 \\ 13 & 14 & 16 \end{pmatrix}$. Hence,

$$\det\begin{pmatrix} 1 & 2 & 3 & 4 \\ 5 & 6 & 7 & 8 \\ 0 & 0 & 11 & 0 \\ 13 & 14 & 15 & 16 \end{pmatrix} = (-1)^{3+3}11\det\begin{pmatrix} 1 & 2 & 4 \\ 15 & 6 & 8 \\ 13 & 14 & 16 \end{pmatrix} = (-1)^{3+3}11\det\left(\overline{A_{33}}\right)$$

The above example suggests that if row i of an $n\times n$ matrix A has a single nonzero entry a_{ij}, then $\det(A) = (-1)^{i+j}a_{ij}\det(\overline{A_{ij}})$. This is demonstrated within the proof of the theorem below.

Theorem: Let A be an $n\times n$ matrix. Then, for $1\le i\le n$, $\det(A)=\sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(\overline{A_{ij}})$.

Proof: $$\det(A)=\det\begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n}\\ \vdots & & & \vdots\\ a_{i1} & a_{i2} & \cdots & a_{in}\\ \vdots & & & \vdots\\ a_{n1} & a_{n2} & \cdots & a_{nn}\end{pmatrix}=\det\begin{pmatrix} a_1\\ \vdots\\ a_{i1}e_1+\cdots+a_{in}e_n\\ \vdots\\ a_n\end{pmatrix}=\sum_{j=1}^{n}\det\begin{pmatrix} a_1\\ \vdots\\ a_{ij}e_j\\ \vdots\\ a_n\end{pmatrix} \qquad \text{(i)}$$

Where $\det\begin{pmatrix} a_1\\ \vdots\\ a_{ij}e_j\\ \vdots\\ a_n\end{pmatrix}=\det\begin{pmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n}\\ \vdots & & & & \vdots\\ 0 & \cdots & a_{ij} & \cdots & 0\\ \vdots & & & & \vdots\\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn}\end{pmatrix}$.

By shifting the ith row up $i-1$ times and the jth column to the left $j-1$ times, we get that

$$\det\begin{pmatrix} a_1\\ \vdots\\ a_{ij}e_j\\ \vdots\\ a_n\end{pmatrix}=(-1)^{i-1}(-1)^{j-1}\det\begin{pmatrix} a_{ij} & 0 & \cdots & 0 & 0 & \cdots & 0\\ a_{1j} & a_{11} & \cdots & a_{1j-1} & a_{1j+1} & \cdots & a_{1n}\\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots\\ a_{i-1j} & a_{i-11} & \cdots & a_{i-1j-1} & a_{i-1j+1} & \cdots & a_{i-1n}\\ a_{i+1j} & a_{i+11} & \cdots & a_{i+1j-1} & a_{i+1j+1} & \cdots & a_{i+1n}\\ \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots\\ a_{nj} & a_{n1} & \cdots & a_{nj-1} & a_{nj+1} & \cdots & a_{nn}\end{pmatrix} \qquad \text{(ii)}$$

Notice that the determinant in the right side of equation (ii) satisfies the hypothesis of the lemma. Note also that the matrix that results from deleting the first row and the first column is $\overline{A_{ij}}$. Hence, (ii) is equivalent to $(-1)^{i+j-2}a_{ij}\det(\overline{A_{ij}})=(-1)^{i+j}a_{ij}\det(\overline{A_{ij}})$. Summing over all j, we get the desired result. ▼