

Supplemental Materials for EE203001 Students

II. Determinant Functions

Chung-Chin Lu

Department of Electrical Engineering
National Tsing Hua University

May 22, 2003

1 Three Axioms for a Determinant Function

Let d be a scalar-valued function on the space $M_{n \times n}$ of all $n \times n$ matrices. Let A_1, A_2, \dots, A_n be the n rows of an $n \times n$ matrix A . We will denote the value $d(A)$ of A under the function d as $d(A) = d(A_1, A_2, \dots, A_n)$, indicating that d is a scalar-valued function of n n -dimensional row vectors A_1, A_2, \dots, A_n . The scalar-valued function d is called a determinant function of order n if it satisfies the following three axioms:

Axiom 1. *Homogeneity in each row.* If matrix B is obtained from matrix A by multiplying one row, says the i th row, of A by a scalar α , then $d(B) = \alpha d(A)$, i.e.,

$$d(A_1, \dots, \alpha A_i, \dots, A_n) = \alpha d(A_1, \dots, A_i, \dots, A_n).$$

Axiom 2. *Invariance under row addition.* If matrix B is obtained from matrix A by adding one row, says the k th row, of A to another row, says the i th row, of A , then $d(B) = d(A)$, i.e.,

$$d(A_1, \dots, A_i + A_k, \dots, A_k, \dots, A_n) = d(A_1, \dots, A_i, \dots, A_k, \dots, A_n).$$

Axiom 3. *The determinant of the identity matrix is one.*

$$d(I_{n \times n}) = d(e_1, e_2, \dots, e_n) = 1.$$

The primary purpose of this supplement is to show that there exists such a determinant function d of order n and it is unique. Before being able to do so, we will derive important properties of such a determinant function d if exists.

Theorem 1 If some row of matrix A is the zero vector, then $d(A) = 0$.

Proof. Without loss of generality, assume the i th row A_i of A is the zero vector. Then

$$\begin{aligned} d(A) &= d(\dots, A_{i-1}, 0, A_{i+1}, \dots) = d(\dots, A_{i-1}, (-1)0, A_{i+1}, \dots) \\ &= (-1)d(\dots, A_{i-1}, 0, A_{i+1}, \dots) = (-1)d(A), \end{aligned}$$

where Axiom 1 is applied to the third equality. Thus we must have $d(A) = 0$. \square

Theorem 2 If matrix B is obtained from matrix A by adding a scalar multiple of one row, says the k th row, of A to another row, says the i th row, of A , then $d(B) = d(A)$, i.e.,

$$d(A_1, \dots, A_i + \alpha A_k, \dots, A_k, \dots, A_n) = d(A_1, \dots, A_i, \dots, A_k, \dots, A_n)$$

for any scalar α .

Proof. It is trivial if $\alpha = 0$. Assume that $\alpha \neq 0$. Then we have

$$\begin{aligned} d(A_1, \dots, A_i, \dots, A_k, \dots, A_n) &= \frac{1}{\alpha} d(A_1, \dots, A_i, \dots, \alpha A_k, \dots, A_n) \\ &= \frac{1}{\alpha} d(A_1, \dots, A_i + \alpha A_k, \dots, \alpha A_k, \dots, A_n) \\ &= \left(\frac{\alpha}{\alpha}\right) d(A_1, \dots, A_i + \alpha A_k, \dots, A_k, \dots, A_n), \end{aligned}$$

where Axiom 1 is applied to the 1st and the 3rd equalities and Axiom 2 is applied to the 2nd equality. This completes the proof. \square

Theorem 3 If matrix B is obtained from matrix A by interchanging two rows of A , says the i th and the k th rows with $i \neq k$, then $d(B) = (-1)d(A)$, i.e.,

$$d(A_1, \dots, A_k, \dots, A_i, \dots, A_n) = (-1)d(A_1, \dots, A_i, \dots, A_k, \dots, A_n).$$

Proof. We compute

$$\begin{aligned} d(A_1, \dots, A_i, \dots, A_k, \dots, A_n) &= d(A_1, \dots, A_i + A_k, \dots, A_k, \dots, A_n) \\ &= d(A_1, \dots, A_i + A_k, \dots, A_k + (-1)(A_i + A_k), \dots, A_n) \\ &= d(A_1, \dots, A_i + A_k, \dots, (-1)A_i, \dots, A_n) \\ &= d(A_1, \dots, A_i + A_k + (-1)A_i, \dots, (-1)A_i, \dots, A_n) \\ &= d(A_1, \dots, A_k, \dots, (-1)A_i, \dots, A_n) \\ &= (-1)d(A_1, \dots, A_k, \dots, A_i, \dots, A_n), \end{aligned}$$

where Axiom 1 is applied to the last equality, Axiom 2 is applied to the 1st and 4th equality, and Theorem 2 is applied to the 2nd equality. \square

Theorem 4 If two rows of matrix A are equal, then $d(A) = 0$.

Proof. Since if we switch the two equal rows of A , the resulted matrix remains the same as A but the determinant value must change sign by the previous theorem, we must have $d(A) = 0$. \square

Note that Theorems 1–4 are consequences of Axioms 1 and 2 and are independent of Axiom 3.

2 Gauss-Jordan Process and the Uniqueness of Determinant Functions

In this section, we will show that a determinant function d of order n is unique if exists by considering the effect of performing elementary row operations on an $n \times n$ matrix. Recall that we perform three types of elementary row operations in Gauss-Jordan elimination process:

1. Interchanging two rows.
2. Multiplying a row by a nonzero scalar.
3. Adding to one row a scalar multiple of another.

By Theorem 3, a type 1 row operation will produce a sign change in the determinant of a square matrix. By Axiom 1, a type 2 row operation will leave the determinant of a square matrix unchanged and so does a type 3 row operation by Theorem 2.

Consider an upper triangular matrix

$$U = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ 0 & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}.$$

If u_{nn} is zero, then the last row of U is the zero vector and by Theorem 1, $d(U) = 0 = u_{11}u_{22} \cdots u_{nn}$. If $u_{nn} \neq 0$, then by applying at most $(n-1)$ type 3 row operations, we have

$$d(U) = d\left(\begin{bmatrix} u_{11} & u_{12} & \cdots & 0 \\ 0 & u_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}\right).$$

If $u_{n-1,n-1} = 0$, then again we have $d(U) = 0 = u_{11}u_{22} \cdots u_{nn}$. If not, then again by applying at most $(n-2)$ type 3 row operations, we have

$$d(U) = d\left(\begin{bmatrix} u_{11} & u_{12} & \cdots & 0 & 0 \\ 0 & u_{22} & \cdots & 0 & 0 \\ \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & \cdots & u_{n-1,n-1} & 0 \\ 0 & 0 & \cdots & 0 & u_{nn} \end{bmatrix}\right).$$

Continuing this process, we either have some $u_{ii} = 0$ such that $d(U) = 0 = u_{11}u_{22} \cdots u_{nn}$ or have all $u_{ii} \neq 0$ such that

$$d(U) = d\left(\begin{bmatrix} u_{11} & 0 & \cdots & 0 \\ 0 & u_{22} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & u_{nn} \end{bmatrix}\right) = (u_{11}u_{22} \cdots u_{nn}) d(I_{n \times n}). \quad (1)$$

In general, an $n \times n$ matrix A can be transformed to an upper triangular matrix U by applying a sequence of elementary row operations. If there are p row exchanges and q scalar multiplications by nonzero scalars c_1, c_2, \dots, c_q included in the sequence, then we have

$$d(U) = (-1)^p(c_1 c_2 \cdots c_q) d(A),$$

i.e.,

$$d(A) = (-1)^p(c_1 c_2 \cdots c_q)^{-1} d(U) = (-1)^p(c_1 c_2 \cdots c_q)^{-1}(u_{11} u_{22} \cdots u_{nn}) d(I_{n \times n}), \quad (2)$$

by (1). Since Theorems 2 and 3 depend only on Axioms 1 and 2 and are independent of Axiom 3, (2) is obtained from Axioms 1 and 2 and is independent of Axiom 3. Thus we conclude that if f is a scalar-valued function of $n \times n$ matrices satisfying Axioms 1 and 2, then we must have

$$f(A) = \alpha f(I_{n \times n}), \quad (3)$$

where the scalar α depends on the matrix A (and probably depends on the Gauss-Jordan process proceeded as you might conceive). The following lemma gives a further characterization of (3) if there exists a determinant function d of order n (which in turn shows that the scalar α in (3) depends only on the matrix A and is independent of the Gauss-Jordan process proceeded).

Lemma 5 *If f is another scalar-valued function on the space $M_{n \times n}$ of all $n \times n$ matrices satisfying Axioms 1 and 2, then we have*

$$f(A) = d(A)f(I_{n \times n})$$

for all $A \in M_{n \times n}$.

Proof. Define a scalar-valued function on the space $M_{n \times n}$ as

$$g(A) = f(A) - d(A)f(I_{n \times n}).$$

It is easy to see that g satisfies both Axioms 1 and 2, but not Axiom 3. In fact, we have

$$g(I_{n \times n}) = f(I_{n \times n}) - d(I_{n \times n})f(I_{n \times n}) = f(I_{n \times n}) - f(I_{n \times n}) = 0,$$

since $d(I_{n \times n}) = 1$ by Axiom 3. But from (3), we have

$$g(A) = \alpha g(I_{n \times n}) = 0$$

for all A . This completes the proof. \square

Now by (1) and Axiom 3, if d is a determinant function of order n , then

$$d(U) = u_{11} u_{22} \cdots u_{nn}$$

for any upper triangular matrix U with diagonal entries $u_{11}, u_{22}, \dots, u_{nn}$. Note that the value $d(U)$ is invariant for any determinant function d of order n . In fact, this is true for any $n \times n$ matrix A , as stated in the following uniqueness theorem.

Theorem 6 [Uniqueness Theorem] If f is another scalar-valued function on the space $M_{n \times n}$ of all $n \times n$ matrices satisfying all the three axioms, then we have

$$f(A) = d(A) \quad \forall A \in M_{n \times n}.$$

Proof. By the above lemma, we have

$$f(A) = d(A)f(I_{n \times n}).$$

By Axiom 3, we have $f(I_{n \times n}) = 1$ and thus $f(A) = d(A)$. \square

Although the above theorem states that a determinant function d is unique if exists, we do not settle down the existence question yet. But for a 1×1 matrix $A = [a_{11}]$, the mapping

$$d(A) = a_{11}$$

satisfies the three axioms and is the unique determinant function of order 1. It is clear that d is a linear function of the single row vector of A . And for a 2×2 matrix $A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$, it is easy to check that the following definition

$$d(A) = a_{11}a_{22} - a_{12}a_{21}$$

satisfies the three axioms and thus gives the unique determinant function of order 2. When we regard the determinant function d of order 2 as a scalar function of two 2-dimensional row vectors, we have the following additive property in each row:

$$\begin{aligned} d(A_1 + A'_1, A_2) &= (a_{11} + a'_{11})a_{22} - (a_{12} + a'_{12})a_{21} = (a_{11}a_{22} - a_{12}a_{21}) + (a'_{11}a_{22} - a'_{12}a_{21}) \\ &= d(A_1, A_2) + d(A'_1, A_2) \\ d(A_1, A_2 + A'_2) &= a_{11}(a_{22} + a'_{22}) - a_{12}(a_{21} + a'_{21}) = (a_{11}a_{22} - a_{12}a_{21}) + (a_{11}a'_{22} - a_{12}a'_{21}) \\ &= d(A_1, A_2) + d(A_1, A'_2). \end{aligned}$$

Together with the homogeneity axiom in each row, we have

$$\begin{aligned} d(\alpha A_1 + \beta A'_1, A_2) &= \alpha d(A_1, A_2) + \beta d(A'_1, A_2) \\ d(A_1, \alpha A_2 + \beta A'_2) &= \alpha d(A_1, A_2) + \beta d(A_1, A'_2), \end{aligned}$$

for any scalars α, β , which says that d is a linear function of one row when the other row is held fixed. We call d to be 2-linear. In the next section, we will show that if a determinant function d of order $(n - 1)$ exists and is $(n - 1)$ -linear, then a determinant function d^1 of order n also exists and is n -linear. Thus by induction, there is a unique determinant function d of order n for every n and d is n -linear.

¹While we have a little bit abused the notation, we are able to distinguish different determinant function d for different order n in the context.

3 Multilinearity and the Existence of a Determinant Function of Order n

Assume that d is a determinant function of order $n - 1$ and is $(n - 1)$ -linear. Let A_{ij} be the $(n - 1) \times (n - 1)$ matrix obtained from an $n \times n$ matrix A by deleting the i th row and the j th column of A . This submatrix A_{ij} of A is called the i, j minor of A . For each j , $1 \leq j \leq n$, consider the following scalar-valued function E_j as

$$E_j(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} d(A_{ij}), \quad \forall A \in M_{n \times n}.$$

Let B be the matrix obtained from matrix A by multiplying the k th row A_k of A by a scalar α . Then we have

$$\begin{aligned} E_j(B) &= E_j(A_1, \dots, \alpha A_k, \dots, A_n) = \sum_{i=1}^n (-1)^{i+j} b_{ij} d(B_{ij}) \\ &= (-1)^{k+j} (\alpha a_{kj}) d(A_{kj}) + \sum_{\substack{i=1 \\ i \neq k}}^n (-1)^{i+j} a_{ij} (\alpha d(A_{ij})) \\ &= \alpha \sum_{i=1}^n (-1)^{i+j} a_{ij} d(A_{ij}) = \alpha E_j(A_1, \dots, A_k, \dots, A_n) = \alpha E_j(A), \end{aligned}$$

where (1) the k, j minor B_{kj} of B is the same as the k, j minor A_{kj} of A and (2) for all $1 \leq i \leq n$, $i \neq k$, the i, j minor B_{ij} of B is obtained by multiplying the k th row or $(k - 1)$ th row ($i > k$ or $i < k$) of the i, j minor A_{ij} of A by scalar α . Thus the function E_j satisfies Axiom 1.

Now let B be the matrix obtained from matrix A by adding the l th row to the k th row of A . Then we have

$$\begin{aligned} E_j(B) &= E_j(A_1, \dots, A_k + A_l, \dots, A_l, \dots, A_n) = \sum_{i=1}^n (-1)^{i+j} b_{ij} d(B_{ij}) \\ &= (-1)^{k+j} (a_{kj} + a_{lj}) d(A_{kj}) + (-1)^{l+j} a_{lj} (d(A_{lj}) + d(\hat{A}_{lj})) + \sum_{\substack{i=1 \\ i \neq k, l}}^n (-1)^{i+j} a_{ij} d(A_{ij}) \\ &= (-1)^{k+j} a_{lj} d(A_{kj}) + (-1)^{l+j+|k-l|-1} a_{lj} d(A_{kj}) + \sum_{i=1}^n (-1)^{i+j} a_{ij} d(A_{ij}) \\ &= E_j(A_1, \dots, A_k, \dots, A_l, \dots, A_n) = E_j(A), \end{aligned}$$

where (1) the k, j minor B_{kj} of B is the same as the k, j minor A_{kj} of A , (2) the l, j minor B_{lj} of B is obtained by adding the shortened $(n - 1)$ -dimensional row vector $A_l^{(j)}$ (obtained from the l th row vector A_l of A by removing its j th component) to the shortened $(n - 1)$ -dimensional row vector $A_k^{(j)}$ of the l, j minor A_{lj} of A and then from the $(n - 1)$ -linearity

of d , $d(B_{lj}) = d(A_{lj}) + d(\hat{A}_{lj})$ with \hat{A}_{lj} being the matrix obtained from A_{lj} by replacing the shortened row vector $A_k^{(j)}$ with the shortened row vector $A_l^{(j)}$, and (3) for all $1 \leq i \leq n$, $i \neq k, l$, the i, j minor B_{ij} of B is obtained by adding the shortened row vector $A_i^{(j)}$ of the i, j minor A_{ij} of A to the shortened row $A_k^{(j)}$ of A_{ij} . Note that matrix \hat{A}_{lj} can be obtained from the k, j minor A_{kj} of A by exchanging the shortened row vector $A_l^{(j)}$ of A_{kj} with the adjacent row vector $|k-l|-1$ times and then $d(\hat{A}_{lj}) = (-1)^{|k-l|-1}d(A_{kj})$. Thus the function E_j satisfies Axiom 2.

For $A = I_{n \times n}$, we have

$$E_j(I_{n \times n}) = (-1)^{j+j}d((I_{n \times n})_{jj}) = d(I_{(n-1) \times (n-1)}) = 1,$$

which says that E_j satisfies Axiom 3. Then by Theorem 6, E_j is the unique determinant function of order n , also denoted as d , and

$$d(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} d(A_{ij}), \quad \forall A \in M_{n \times n}, \quad (4)$$

for any j , $1 \leq j \leq n$. Now, for each k , $1 \leq k \leq n$, we have from in (4)

$$\begin{aligned} & d(A_1, \dots, A_k + A'_k, \dots, A_n) \\ &= (-1)^{k+j} (a_{kj} + a'_{kj}) d(\dots, A_{k-1}^{(j)}, A_{k+1}^{(j)}, \dots) + \sum_{\substack{i=1 \\ i \neq k}}^n (-1)^{i+j} a_{ij} d(\dots, A_k^{(j)} + A_k'^{(j)}, \dots) \\ &= \left((-1)^{k+j} a_{kj} d(\dots, A_{k-1}^{(j)}, A_{k+1}^{(j)}, \dots) + \sum_{\substack{i=1 \\ i \neq k}}^n (-1)^{i+j} a_{ij} d(\dots, A_k^{(j)}, \dots) \right) \\ & \quad + \left((-1)^{k+j} a'_{kj} d(\dots, A_{k-1}^{(j)}, A_{k+1}^{(j)}, \dots) + \sum_{\substack{i=1 \\ i \neq k}}^n (-1)^{i+j} a_{ij} d(\dots, A_k'^{(j)}, \dots) \right) \\ &= d(A_1, \dots, A_k, \dots, A_n) + d(A_1, \dots, A'_k, \dots, A_n), \end{aligned}$$

which shows that the determinant function d of order n is n -linear.

Now we have proved that if a determinant function d of order $(n-1)$ exists and is $(n-1)$ -linear, then a determinant function d of order n exists and is n -linear. Since a determinant function d of order 2 exists and is 2-linear, we conclude by induction that for every n , there exists a unique determinant function of order n and it is n -linear, as stated in the following theorem.

Theorem 7 [Existence Theorem] For every positive integer n , there is a unique determinant function of order n . Furthermore, this function is n -linear.

This unique function is frequently denoted as \det . The expansion formula in (4) is called the expansion of the determinant $\det(A)$ of A by the j th column minors, rewritten here as

$$\det(A) = \sum_{i=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}). \quad (5)$$

We next show that the determinant $\det(A)$ of matrix A can be expanded by minors in a row.

For each i , $1 \leq i \leq n$, consider the i th row A_i of an $n \times n$ matrix A ,

$$A_i = a_{i1}I_1 + a_{i2}I_2 + \cdots + a_{in}I_n,$$

where $I_1 = (1, 0, \dots, 0), \dots, I_n = (0, \dots, 0, 1)$ are standard row vectors. Then we have

$$\begin{aligned} \det(A) &= d(A_1, \dots, A_{i-1}, \sum_{j=1}^n a_{ij}I_j, A_{i+1}, \dots, A_n) \\ &= \sum_{j=1}^n a_{ij}d(A_1, \dots, A_{i-1}, I_j, A_{i+1}, \dots, A_n) \\ &= \sum_{j=1}^n a_{ij} \det(A'_{ij}) \end{aligned} \quad (6)$$

by the n -linearity of d , where

$$A'_{ij} = \begin{bmatrix} A_1 \\ \vdots \\ A_{i-1} \\ I_j \\ A_{i+1} \\ \vdots \\ A_n \end{bmatrix} = \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & a_{i-1,j} & \cdots & a_{i-1,n} \\ 0 & \cdots & 1 & \cdots & 1 \\ a_{i+1,1} & \cdots & a_{i+1,j} & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix}.$$

By adding a multiple $-a_{kj}I_j$ of the i th row I_j of A'_{ij} to the k th row A_k of A'_{ij} for each k , $1 \leq k \leq n$ and $k \neq i$, we obtain the matrix

$$A''_{ij} = \begin{bmatrix} a_{11} & \cdots & 0 & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{i-1,1} & \cdots & 0 & \cdots & a_{i-1,n} \\ 0 & \cdots & 1 & \cdots & 1 \\ a_{i+1,1} & \cdots & 0 & \cdots & a_{i+1,n} \\ \vdots & \ddots & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & 0 & \cdots & a_{nn} \end{bmatrix}$$

and by Theorem 2,

$$\det(A''_{ij}) = \det(A'_{ij}). \quad (7)$$

With the expansion of the determinant $d(A''_{ij})$ by the j th column minors in (5), we have

$$\det(A''_{ij}) = (-1)^{i+j} \det(A_{ij}) \quad (8)$$

since the i, j minor of A''_{ij} is equal to the i, j minor A_{ij} of A . By (6)–(8), we have

$$\det(A) = \sum_{j=1}^n (-1)^{i+j} a_{ij} \det(A_{ij}), \quad (9)$$

which is called the expansion of the determinant $\det(A)$ of A by the i th row minors.

The next theorem is an application of the multilinearity of the determinant function \det .

Theorem 8 *If the rows of an $n \times n$ matrix A are linearly dependent, then $\det(A) = 0$.*

Proof. Suppose there exist scalars c_1, c_2, \dots, c_n , not all zeros and says $c_k \neq 0$, such that $\sum_{i=1}^n c_i A_i = 0$. Then we have $A_k = \sum_{i=1, i \neq k}^n t_i A_i$, where $t_i = -c_i/c_k$. Thus

$$\begin{aligned} \det(A) &= d(\dots, A_{k-1}, \sum_{\substack{i=1 \\ i \neq k}}^n t_i A_i, A_{k+1}, \dots) \\ &= \sum_{\substack{i=1 \\ i \neq k}}^n t_i d(\dots, A_{k-1}, A_i, A_{k+1}, \dots). \end{aligned}$$

But for each $i \neq k$, row A_i is equal to at least one of the rows $A_1, \dots, A_{k-1}, A_{k+1}, \dots, A_n$ and hence

$$d(\dots, A_{k-1}, A_i, A_{k+1}, \dots) = 0$$

by Theorem 4. Thus $\det(A) = 0$. □

The next theorem is an application of the expansion formulas in (5) and (9).

Theorem 9 *An $n \times n$ matrix A and its transpose A^t have the same determinants, i.e., $\det(A) = \det(A^t)$.*

Proof. The proof is by induction on n , the size of the matrix A . For $n = 1$, we have $A = A^t = [a_{11}]$ and $\det(A) = \det(A^t) = a_{11}$ trivially. Assume that the theorem is true for $(n - 1)$. Let $A = [a_{ij}]$ be an $n \times n$ matrix and $B = A^t = [b_{ij}]$. Then $b_{ij} = a_{ji}$. By expanding $\det(B)$ by the i th row minors, we have

$$\det(B) = \sum_{j=1}^n (-1)^{i+j} b_{ij} \det(B_{ij}).$$

The i, j minor B_{ij} of B is equal to the transpose of the j, i minor A_{ji} of A , i.e., $B_{ij} = A_{ji}^t$. And by the induction step, $\det(A_{ji}) = \det(A_{ji}^t)$. Hence we have

$$\det(B) = \sum_{j=1}^n (-1)^{j+i} a_{ji} \det(A_{ji}) = \det(A)$$

from the expansion of $\det(A)$ by the i th column minors. We conclude that $\det(A) = \det(A^t)$. □

Since the row vectors $(A^t)_i$ of A^t is just the transpose $(a_i)^t$ of the column vectors a_i of A , we have

$$\det(A) = \det(A^t) = d((A^t)_1, \dots, (A^t)_i, \dots, (A^t)_n) = d((a_1)^t, \dots, (a_i)^t, \dots, (a_n)^t) = d(a_1, \dots, a_i, \dots, a_n)$$

which implies that we can regard the determinant function \det as a scalar function of n n -dimensional column vectors a_1, a_2, \dots, a_n . Thus by Theorem 9, all the row properties of the determinant function in Theorems 1–4 and 7–8 have the corresponding versions for columns as stated in the following corollary.

Corollary 10 *The determinant function $\det(A) = d(a_1, \dots, a_i, \dots, a_n)$ of an $n \times n$ matrix $A = [a_1 a_2 \dots a_n]$, where a_1, a_2, \dots, a_n are n column vectors of A , has the following column properties:*

1. *It is n -linear with respect to column vectors, i.e., for each i , $1 \leq i \leq n$,*

$$d(a_1, \dots, \alpha a_i + \beta a'_i, \dots, a_n) = \alpha d(a_1, \dots, a_i, \dots, a_n) + \beta d(a_1, \dots, a'_i, \dots, a_n),$$

for any scalars α, β .

2. *It is invariant under the addition of a multiple of a column of A to another column of A , i.e., for any scalar α ,*

$$d(a_1, \dots, a_i + \alpha a_k, \dots, a_k, \dots, a_n) = d(a_1, \dots, a_i, \dots, a_k, \dots, a_n).$$

3. *It changes its sign under the exchange of two columns of A , i.e.,*

$$d(a_1, \dots, a_k, \dots, a_i, \dots, a_n) = (-1)d(a_1, \dots, a_i, \dots, a_k, \dots, a_n).$$

And $\det(A) = 0$ if A has linearly dependent columns. \square

4 The Cofactor Matrix and Cramer's Rule

The scalar $(-1)^{i+j} \det(A_{ij})$ corresponding to the i, j minor A_{ij} of an $n \times n$ matrix A is called the cofactor of the i, j entry a_{ij} of A , denoted as $\text{cof } a_{ij}$, i.e.,

$$\text{cof } a_{ij} = (-1)^{i+j} \det(A_{ij}).$$

The $n \times n$ matrix

$$\text{cof } A = [\text{cof } a_{ij}]_{i,j=1}^n$$

is called the cofactor matrix of A .

Consider the product $A(\text{cof } A)^t$ of A and the transpose of the cofactor matrix of A . The i, j entry of the product is

$$(A(\text{cof } A)^t)_{ij} = \sum_{k=1}^n a_{ik} \text{cof } a_{jk} = \sum_{k=1}^n (-1)^{j+k} a_{ik} \det(A_{jk}) = d(\dots, A_{j-1}, A_i, A_{j+1}, \dots)$$

by the expansion formula along the j th row and is

$$(A(\text{cof } A)^t)_{ij} = \begin{cases} \det(A), & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}$$

by Theorem 4. Thus we have

$$A(\text{cof } A)^t = \det(A)I_{n \times n}.$$

Next Consider the product $(\text{cof } A)^t A$. The i, j entry of this product is

$$((\text{cof } A)^t A)_{ij} = \sum_{k=1}^n \text{cof } a_{ki} a_{kj} = \sum_{k=1}^n (-1)^{k+i} a_{kj} \det(A_{ki}) = d(\dots, a_{i-1}, a_j, a_{i+1}, \dots)$$

by the expansion formula along the i th column and is

$$((\text{cof } A)^t A)_{ij} = \begin{cases} \det(A), & \text{if } i = j, \\ 0, & \text{if } i \neq j, \end{cases}$$

by Corollary 10. Thus we have

$$(\text{cof } A)^t A = \det(A)I_{n \times n}.$$

We conclude in the following theorem.

Theorem 11 For any $n \times n$ matrix A with $n \geq 2$, we have

$$A(\text{cof } A)^t = (\text{cof } A)^t A = \det(A)I_{n \times n}.$$

In particular, if $\det(A) \neq 0$, then A is invertible and its inverse is

$$A^{-1} = \frac{1}{\det(A)}(\text{cof } A)^t.$$

□

As an application of the above theorem, we consider a system of n linear equations in n unknowns x_1, x_2, \dots, x_n ,

$$Ax = \begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n1} & a_{n2} & \cdots & a_{nn} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} = b$$

If $\det(A) \neq 0$, then A is invertible and the above system of linear equations has a unique solution

$$x = A^{-1}b = \frac{1}{\det(A)}(\text{cof } A)^t b,$$

where

$$x_i = \frac{1}{\det(A)} \sum_{j=1}^n \text{cof } a_{ji} b_j = \frac{1}{\det(A)} \sum_{j=1}^n (-1)^{j+i} b_j \det(A_{ji}) = \frac{d(a_1, \dots, a_{i-1}, b, a_{i+1}, \dots, a_n)}{d(a_1, \dots, a_{i-1}, a_i, a_{i+1}, \dots, a_n)} \quad (10)$$

for each i , $1 \leq i \leq n$. Equation (10) is called Cramer's rule.

5 The Product Formula for Determinants

Consider the product AB of two $n \times n$ matrices A and B . It is easy to see that the i th row $(AB)_i$ of AB is just the i th row A_i of A times B , i.e.,

$$(AB)_i = A_i B, \quad \forall i.$$

Thus we have

$$\det(AB) = d(A_1 B, A_2 B, \dots, A_n B)$$

And by keeping B fixed, we can define a scalar-valued function $f(A) = \det(AB)$ of matrix A and

$$f(A_1, A_2, \dots, A_n) = d(A_1 B, A_2 B, \dots, A_n B).$$

Since for each i ,

$$\begin{aligned} f(A_1, \dots, \alpha A_i, \dots, A_n) &= d(A_1 B, \dots, \alpha A_i B, \dots, A_n B) \\ &= \alpha d(A_1 B, \dots, A_i B, \dots, A_n B) \\ &= \alpha f(A_1, \dots, A_i, \dots, A_n), \end{aligned}$$

and for $i \neq k$,

$$\begin{aligned} f(A_1, \dots, A_i + A_k, \dots, A_k, \dots, A_n) &= d(A_1 B, \dots, (A_i + A_k) B, \dots, A_k B, \dots, A_n B) \\ &= d(A_1 B, \dots, A_i B, \dots, A_k B, \dots, A_n B) \\ &= f(A_1, \dots, A_i, \dots, A_k, \dots, A_n), \end{aligned}$$

function f satisfies Axioms 1 and 2 for a determinant function and by Lemma 5,

$$f(A) = \det(A) f(I_{n \times n}).$$

Since

$$f(I_{n \times n}) = \det(I_{n \times n} B) = \det(B),$$

we have the following theorem.

Theorem 12 [Product Formula] For any two square matrices A and B of the same size, we have

$$\det(AB) = \det(A) \det(B).$$

□

As an application of the above theorem, consider an invertible $n \times n$ matrix A , i.e.,

$$AA^{-1} = I_{n \times n}.$$

By the product formula, we have

$$\det(A) \det(A^{-1}) = \det(I_{n \times n}) = 1.$$

Together with Theorem 11, we have the following theorem.

Theorem 13 A square matrix A is invertible if and only if $\det(A) \neq 0$. In this case, we have

$$\det(A^{-1}) = \det(A)^{-1}.$$

□

Another application of the product formula is to prove the converse of Theorem 8 and the last sentence of Corollary 10.

Theorem 14 If an $n \times n$ matrix A has n linearly independent rows (columns), then $\det(A) \neq 0$.

Proof. Firstly assume that the n columns a_1, a_2, \dots, a_n of A are linearly independent in the linear space V of all n -dimensional column vectors. Since V is n -dimensional, $\{a_1, a_2, \dots, a_n\}$ is a basis of V and every n -dimensional column vector in V can be expressed uniquely as a linear combination of a_1, a_2, \dots, a_n . In particular, for standard unit column vectors e_i , we have

$$\begin{aligned} e_1 &= \beta_{11}a_1 + \beta_{21}a_2 + \dots + \beta_{n1}a_n \\ e_2 &= \beta_{12}a_1 + \beta_{22}a_2 + \dots + \beta_{n2}a_n \\ &\vdots \quad \vdots \quad \vdots \\ e_n &= \beta_{1n}a_1 + \beta_{2n}a_2 + \dots + \beta_{nn}a_n, \end{aligned}$$

i.e.,

$$[e_1 e_2 \dots e_n] = [a_1 a_2 \dots a_n] \begin{bmatrix} \beta_{11} & \beta_{12} & \dots & \beta_{1n} \\ \beta_{21} & \beta_{22} & \dots & \beta_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \beta_{n1} & \beta_{n2} & \dots & \beta_{nn} \end{bmatrix},$$

i.e.,

$$I_{n \times n} = AB.$$

By the product formula, we have $\det(A) \det(B) = \det(I_{n \times n}) = 1$ and then $\det(A) \neq 0$. Secondly assume that A has n linearly independent row vectors. Then A^t has n linearly independent column vectors. As proved in above, $\det(A^t) \neq 0$. Thus we have $\det(A) \neq 0$ since $\det(A) = \det(A^t)$. □

Corollary 15 An $n \times n$ matrix A has n linearly independent rows (columns) if and only if $\det(A) \neq 0$. □

Combined with Theorem 13 and Corollary 15, we have the following corollary.

Corollary 16 An $n \times n$ matrix A is invertible if and only if A has n linearly independent rows (columns). □

Note that Corollary 16 can be proved directly without the concept of determinants.