Chapter 4

Determinants

4.1 Definition Using Expansion by Minors

Every square matrix A has a number associated to it and called its *determinant*, denoted by $\det(A)$.

One of the most important properties of a determinant is that it gives us a criterion to decide whether the matrix is invertible:

A matrix A is invertible iff $\det(A) \neq 0$.

It is possible to define determinants in terms of a fairly complicated formula involving $n!$ terms (assuming A is a $n \times n$ matrix) but this way to proceed makes it more difficult to prove properties of determinants.

Consequently, we follow a more algorithmic approach due to Mike Artin.

We will view the determinant as *a function of the rows of an $n \times n$ matrix*.

Formally, this means that

$$\det\colon (\mathbb{R}^n)^n \to \mathbb{R}.$$

We will define the determinant *recursively* using a process called *expansion by minors*.

Then, we will derive properties of the determinant and prove that there is a unique function satisfying these properties.

As a consequence, we will have an axiomatic definition of the determinant.

For a 1×1 matrix $A = (a)$, we have

$$\det(A) = \det(a) = a.$$

For a 2×2 matrix,

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

it will turn out that

$$\det(A) = ad - bc.$$

The determinant has a geometric interpretation as a signed area, in higher dimension as a signed volume.

In order to describe the recursive process to define a determinant we need the notion of a minor.

Definition 4.1. Given any $n \times n$ matrix with $n \geq 2$, for any two indices i, j with $1 \leq i, j \leq n$, let A_{ij} be the $(n-1) \times (n-1)$ matrix obtained by deleting row i and colummn j from A and called a *minor*:

$$
A_{ij} = \begin{bmatrix}
 & & & & \times & & \\
 & & & & \times & & \\
\times & \times & \times & \times & \times & \times & \times \\
 & & & & \times & & \\
 & & & & \times & & \\
 & & & & \times & & \\
 & & & & \times & &
\end{bmatrix}
$$

For example, if

$$
A = \begin{bmatrix}
2 & -1 & 0 & 0 & 0 \\
-1 & 2 & -1 & 0 & 0 \\
0 & -1 & 2 & -1 & 0 \\
0 & 0 & -1 & 2 & -1 \\
0 & 0 & 0 & -1 & 2
\end{bmatrix}
$$

then

$$
A_{2\,3} = \begin{bmatrix}
2 & -1 & 0 & 0 \\
0 & -1 & -1 & 0 \\
0 & 0 & 2 & -1 \\
0 & 0 & -1 & 2
\end{bmatrix}.
$$

We can now proceed with the definition of determinants.

Definition 4.2. Given any $n \times n$ matrix $A = (a_{ij})$, if $n = 1$, then

$$\det(A) = a_{11},$$

else

$$\det(A) = a_{11}\det(A_{11}) + \cdots + (-1)^{i+1} a_{i1}\det(A_{i1}) + \cdots + (-1)^{n+1} a_{n1}\det(A_{n1}), \quad (*)$$

the expansion by minors on the first column.

When $n = 2$, we have

$$\det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = a_{11}\det[a_{22}] - a_{21}\det[a_{12}] = a_{11}a_{22} - a_{21}a_{12},$$

which confirms the formula claimed earlier.

When $n = 3$, we get

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}\det\begin{bmatrix} a_{22} & a_{23} \\ a_{32} & a_{33} \end{bmatrix} - a_{21}\det\begin{bmatrix} a_{12} & a_{13} \\ a_{32} & a_{33} \end{bmatrix} + a_{31}\det\begin{bmatrix} a_{12} & a_{13} \\ a_{22} & a_{23} \end{bmatrix},$$

and using the formula for a 2×2 determinant, we get

$$\begin{aligned} \det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} &= a_{11}(a_{22}a_{33} - a_{32}a_{23}) - a_{21}(a_{12}a_{33} - a_{32}a_{13}) \\ &\quad + a_{31}(a_{12}a_{23} - a_{22}a_{13}). \end{aligned}$$

As we can see, the formula is already quite complicated!

Given a $n \times n$-matrix $A = (a_{i\,j})$, its determinant $\det(A)$ is also denoted by

$$\det(A) = \begin{vmatrix} a_{1\,1} & a_{1\,2} & \dots & a_{1\,n} \\ a_{2\,1} & a_{2\,2} & \dots & a_{2\,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n\,1} & a_{n\,2} & \dots & a_{n\,n} \end{vmatrix}.$$

We now derive some important and useful properties of the determinant.

Recall that we view the determinant $\det(A)$ as a function of the rows of the matrix A, so we can write

$$\det(A) = \det(A_1, \dots, A_n),$$

where $A_1, \dots, A_n$ are the rows of A.

Proposition 4.1. *The determinant function* $\det\colon (\mathbb{R}^n)^n \to \mathbb{R}$ *satisfies the following properties:*

(1) $\det(I) = 1$, *where* I *is the identity matrix.*

(2) The determinant is linear in each of its rows; this means that

$$\begin{aligned}\det(A_1, \ldots, A_{i-1}, B + C, A_{i+1}, \ldots, A_n) = \\ \det(A_1, \ldots, A_{i-1}, B, A_{i+1}, \ldots, A_n) \\ + \det(A_1, \ldots, A_{i-1}, C, A_{i+1}, \ldots, A_n)\end{aligned}$$

and

$$\begin{aligned}\det(A_1, \ldots, A_{i-1}, \lambda A_i, A_{i+1}, \ldots, A_n) = \\ \lambda \det(A_1, \ldots, A_{i-1}, A_i, A_{i+1}, \ldots, A_n).\end{aligned}$$

(3) If two adjacent rows of A *are equal, then* $\det(A) = 0$. *This means that*

$$\det(A_1, \ldots, A_i, A_i, \ldots, A_n) = 0.$$

Property (2) says that det is a *multilinear map*, and property (3) says that det is an *alternating map*.

Proposition 4.2. *The determinant function* $\det\colon (\mathbb{R}^n)^n \to \mathbb{R}$ *satisfies the following properties:*

(4) If two adjacent rows are interchanged, then the determinant is multiplied by -1*; thus,*

$$\det(A_1, \ldots, A_{i+1}, A_i, \ldots, A_n) \\ = -\det(A_1, \ldots, A_i, A_{i+1}, \ldots, A_n).$$

(5) If two rows are identical then the determinant is zero; that is,

$$\det(A_1, \ldots, A_i, \ldots, A_i, \ldots, A_n) = 0.$$

(6) If any two distinct rows of A *are are interchanged, then the determinant is multiplied by* -1*; thus,*

$$\det(A_1, \ldots, A_j, \ldots, A_i, \ldots, A_n) \\ = -\det(A_1, \ldots, A_i, \ldots, A_j, \ldots, A_n).$$

(7) If a multiple of a row is added to another row, the determinant is unchanged; that is,

$$\det(A_1, \ldots, A_i + \lambda A_j, \ldots, A_n) \\ = \det(A_1, \ldots, A_i, \ldots, A_n).$$

(8) If any row of A *is zero, then* $\det(A) = 0$*.*

Using property (6), it is easy to show that the expansion by minors formula $(*)$ can be adapted to any column. Indeed, we have

$$\begin{aligned}\det(A) = (-1)^{j+1}a_{1j}\det(A_{1j}) + \cdots + (-1)^{j+i}a_{ij}\det(A_{ij}) \\ + \cdots + (-1)^{j+n}a_{nj}\det(A_{nj}). \quad (**)\end{aligned}$$

The beauty of this approach is that properties (6) and (7) describe the effect of the elementary operations $P(i,j)$ and $E_{i,j,\lambda}$ on the determinant:

Indeed, (6) says that

$$\det(P(i,j)A) = -\det(A), \tag{a}$$

and (7) says that

$$\det(E_{i,j;\lambda}A) = \det(A). \tag{b}$$

Furthermore, linearity (propery (2)) says that

$$\det(E_{i,\lambda}A) = \lambda\det(A). \tag{c}$$

Substituting the identity I for A in the above equations, since $\det(I) = 1$, we find the determinants of the elementary matrices:

(1) For any permutation matrix $P(i, j)$ $(i \neq j)$, we have

$$\det(P(i, j)) = -1.$$

(2) For any row operation $E_{i,j;\lambda}$ (adding λ times row j to row i), we have

$$\det(E_{i,j;\lambda}) = 1.$$

(3) For any row operation $E_{i,\lambda}$ (multiplying row i by λ), we have

$$\det(E_{i,\lambda}) = \lambda.$$

The above properties together with the equations (a), (b), (c) yield the following important proposition:

Proposition 4.3. *For every $n \times n$ matrix A and every elementary matrix E, we have*

$$\det(EA) = \det(E)\det(A).$$

We can now use Proposition 4.3 and the reduction to row echelon form to compute $\det(A)$.

Indeed, recall that we showed (just before Proposition 2.15)) that every square matrix A can be reduced by elementary operations to a matrix A' which is either the identity or else whose last row is zero,

$$A' = E_k \cdots E_1 A.$$

If $A' = I$, then $\det(A') = 1$ by (1), else if A' has a zero row, then $\det(A') = 0$ by (8).

Furthermore, by induction using Proposition 4.3 (see the proof of Proposition 4.7), we get

$$\det(A') = \det(E_k \cdots E_1 A) = \det(E_k) \cdots \det(E_1) \det(A).$$

Since all the determinants, $\det(E_k)$ of the elementary matrices E_i are known, we see that the formula

$$\det(A') = \det(E_k) \cdots \det(E_1) \det(A)$$

determines $\det(A)$. As a consequence, we have the following characterization of a determinant:

Theorem 4.4. *(Axiomatic Characterization of the Determinant) The determinant* det *is the unique function $f\colon (\mathbb{R}^n)^n \to \mathbb{R}$ satisfying properties (1), (2), and (3) of Proposition 4.1.*

Instead of evaluating a determinant using expansion by minors on the columns, we can use expansion by minors on the rows.

Indeed, define the function D given

$$D(A) = (-1)^{i+1} a_{i\,1} D(A_{i\,1}) + \cdots + (-1)^{i+n} a_{i\,n} D(A_{i\,n}), \quad (\dagger)$$

with $D([a]) = a$.

Then, it is fairly easy to show that the properties of Proposition 4.1 hold for D, and thus, by Theorem 4.4, the function D also defines the determinant, that is,

$$D(A) = \det(A).$$

Proposition 4.5. *For any square matrix A, we have $\det(A) = \det(A^{\top})$.*

We also obtain the important characterization of invertibility of a matrix in terms of its determinant.

Proposition 4.6. *A square matrix A is invertible iff* $\det(A) \neq 0$.

We can now prove one of the most useful properties of determinants.

Proposition 4.7. *Given any two $n \times n$ matrices A and B, we have*

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

In order to give an explicit formula for the determinant, we need to discuss some properties of permutation matrices.

4.2 Permutations and Permutation Matrices

Let $[n] = \{1, 2 \ldots, n\}$, where $n \in \mathbb{N}$, and $n > 0$.

Definition 4.3. A *permutation on n elements* is a bijection $\pi\colon [n] \to [n]$. When $n = 1$, the only function from $[1]$ to $[1]$ is the constant map: $1 \mapsto 1$. Thus, we will assume that $n \geq 2$. A *transposition* is a permutation $\tau\colon [n] \to [n]$ such that, for some $i < j$ (with $1 \leq i < j \leq n$), $\tau(i) = j$, $\tau(j) = i$, and $\tau(k) = k$, for all $k \in [n] - \{i, j\}$. In other words, a transposition exchanges two distinct elements $i, j \in [n]$.

If τ is a transposition, clearly, $\tau \circ \tau = \mathrm{id}$. We have already encountered transpositions before, but represented by the matrices $P(i, j)$.

We will also use the terminology product of permutations (or transpositions), as a synonym for composition of permutations.

Clearly, the composition of two permutations is a permutation and every permutation has an inverse which is also a permutation.

Therefore, the set of permutations on $[n]$ is a *group* often denoted $\mathfrak{S}_n$. It is easy to show by induction that the group $\mathfrak{S}_n$ has $n!$ elements.

There are various ways of denoting permutations. One way is to use a functional notation such as

$$\begin{pmatrix} 1 & 2 & \cdots & i & \cdots & n \\ \pi(1) & \pi(2) & \cdots & \pi(i) & \cdots & \pi(n) \end{pmatrix}.$$

For example the permutation $\pi\colon [4] \to [4]$ given by

$$\begin{aligned} \pi(1) &= 3 \\ \pi(2) &= 4 \\ \pi(3) &= 2 \\ \pi(4) &= 1 \end{aligned}$$

is represented by

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}.$$

The above representation has the advantage of being compact, but a matrix representation is also useful and has the advantage that composition of permutations corresponds to matrix multiplication.

A permutation can be viewed as an operation permuting the rows of a matrix. For example, the permutation

$$\begin{pmatrix} 1 & 2 & 3 & 4 \\ 3 & 4 & 2 & 1 \end{pmatrix}$$

corresponds to the matrix

$$P_\pi = \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

Observe that the matrix P_π has a single 1 on every row and every column, all other entries being zero, and that if we multiply any 4×4 matrix A by P_π on the left, then the rows of $P_\pi A$ are permuted according to the permutation π;

that is, the $\pi(i)$th row of $P_\pi A$ is the ith row of A.

For example,

$$
\begin{aligned}
P_\pi A &= \begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} \\
&= \begin{bmatrix} a_{41} & a_{42} & a_{43} & a_{44} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix}.
\end{aligned}
$$

Equivalently, the ith row of $P_\pi A$ is the $\pi^{-1}(i)$th row of A.

In order for the matrix P_π to move the ith row of A to the $\pi(i)$th row, the $\pi(i)$th row of P_π must have a 1 in column i and zeros everywhere else;

this means that the ith column of P_π contains the basis vector $e_{\pi(i)}$, the vector that has a 1 in position $\pi(i)$ and zeros everywhere else.

This is the general situation and it leads to the following definition.

Definition 4.4. Given any permutation $\pi\colon [n] \to [n]$, the *permutation matrix* $P_\pi = (p_{ij})$ representing π is the matrix given by

$$p_{ij} = \begin{cases} 1 & \text{if } i = \pi(j) \\ 0 & \text{if } i \neq \pi(j); \end{cases}$$

equivalently, the jth column of P_π is the basis vector $e_{\pi(j)}$.

A *permutation matrix* P is any matrix of the form P_π (where P is an $n \times n$ matrix, and $\pi\colon [n] \to [n]$ is a permutation, for some $n \geq 1$).

Remark: There is a confusing point about the notation for permutation matrices.

A permutation matrix P acts on a matrix A by multiplication on the left by permuting the rows of A.

As we said before, this means that the $\pi(i)$th row of $P_\pi A$ is the ith row of A, or equivalently that the ith row of $P_\pi A$ is the $\pi^{-1}(i)$th row of A.

But then, observe that the row index of the entries of the ith row of PA is $\pi^{-1}(i)$, and not $\pi(i)$! See the following example:

$$\begin{bmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{41} & a_{42} & a_{43} & a_{44} \end{bmatrix} = \begin{bmatrix} a_{41} & a_{42} & a_{43} & a_{44} \\ a_{31} & a_{32} & a_{33} & a_{34} \\ a_{11} & a_{12} & a_{13} & a_{14} \\ a_{21} & a_{22} & a_{23} & a_{24} \end{bmatrix},$$

where

$$\begin{aligned} \pi^{-1}(1) &= 4 \\ \pi^{-1}(2) &= 3 \\ \pi^{-1}(3) &= 1 \\ \pi^{-1}(4) &= 2. \end{aligned}$$

Proposition 4.8. *The following properties hold:*

(1) Given any two permutations $\pi_1, \pi_2 \colon [n] \to [n]$, the permutation matrix $P_{\pi_2 \circ \pi_1}$ representing the composition of π_1 and π_2 is equal to the product $P_{\pi_2} P_{\pi_1}$ of the permutation matrices P_{π_1} and P_{π_2} representing π_1 and π_2; that is,

$$P_{\pi_2 \circ \pi_1} = P_{\pi_2} P_{\pi_1}.$$

(2) The matrix $P_{\pi_1^{-1}}$ representing the inverse of the permutation π_1 is the inverse $P_{\pi_1}^{-1}$ of the matrix P_{π_1} representing the permutation π_1; that is,

$$P_{\pi_1^{-1}} = P_{\pi_1}^{-1}.$$

Furthermore,

$$P_{\pi_1}^{-1} = (P_{\pi_1})^\top.$$

The following proposition shows the importance of transpositions.

Proposition 4.9. *For every $n \geq 2$, every permutation $\pi \colon [n] \to [n]$ can be written as a nonempty composition of transpositions.*

Remark: When $\pi = \mathrm{id}_n$ is the identity permutation, we can agree that the composition of 0 transpositions is the identity.

Proposition 4.9 shows that the transpositions generate the group of permutations $\mathfrak{S}_n$.

Since we already know that the determinant of a transposition matrix is -1, Proposition 4.9 implies that for every permutation matrix P, we have

$$\det(P) = \pm 1.$$

We can say more. Indeed if a given permutation π is factored into two different products of transpositions $\tau_p \circ \cdots \circ \tau_1$ and $\tau'_q \circ \cdots \circ \tau'_1$, because

$$det(P_\pi) = \det(P_{\tau_p}) \cdots \det(P_{\tau_1}) = \det(P_{\tau'_q}) \cdots \det(P_{\tau'_1}),$$

and $\det(P_{\tau_i}) = \det(P_{\tau'_j}) = -1$, *the natural numbers p and q have the same parity*.

Consequently, for every permutation σ of $[n]$, the *parity* of the number p of transpositions involved in any decomposition of σ as $\sigma = \tau_p \circ \cdots \circ \tau_1$ is an invariant (only depends on σ).

Definition 4.5. For every permutation σ of $[n]$, the parity $\epsilon(\sigma)$ of the number of transpositions involved in any decomposition of σ is called the *signature* of σ. We have $\epsilon(\sigma) = \det(P_\sigma)$.

Remark: When $\pi = \mathrm{id}_n$ is the identity permutation, since we agreed that the composition of 0 transpositions is the identity, it it still correct that $(-1)^0 = \epsilon(\mathrm{id}) = +1$.

It is also easy to see that $\epsilon(\pi' \circ \pi) = \epsilon(\pi')\epsilon(\pi)$.

In particular, since $\pi^{-1} \circ \pi = \mathrm{id}_n$, we get $\epsilon(\pi^{-1}) = \epsilon(\pi)$.

We are now ready to give an explicit formula for a determinant.

Given an $n \times n$ matrix A (with $n \geq 2$), we can view its first row A_1 as the sum of the n rows

$$[a_{11}\, 0 \cdots 0],\ [0\, a_{12}\, 0 \cdots 0],\ \ldots, [0 \cdots 0\, a_{1n}],$$

and we can expand A by linearity as

$$\begin{aligned}\det(A) = \det \begin{bmatrix} a_{11} & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} + \det \begin{bmatrix} 0 & a_{12} & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \\ \vdots & \vdots & \cdots & \vdots \end{bmatrix} \\ + \cdots + \det \begin{bmatrix} \cdots & \cdots & 0 & a_{1n} \\ \cdots & \cdots & \vdots & \vdots \\ \vdots & \cdots & \vdots & \vdots \\ \cdots & \cdots & \vdots & \vdots \end{bmatrix}.\end{aligned}$$

We can repeat this process on the second row, the third row, etc.

At the end, we obtain a sum of determinants of matrices of the form

$$M = \begin{bmatrix} & & a_{1?} & & \\ & a_{2?} & & & \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ & & & & a_{n?} \end{bmatrix}$$

having a single entry left in each row, all the others being zero.

Observe that all the determinants involving matrices having a zero column will be zero.

Actually, the only determinants that survive are those that have a single entry a_{ij} in each row and each column.

Such matrices are very similar to permutation matrices. In fact, they must be of the form $M_\pi = (m_{ij})$ for some permutation π of $[n]$, with

$$m_{ij} = \begin{cases} a_{ij} & \text{if } i = \pi(j) \\ 0 & \text{if } i \neq \pi(j). \end{cases}$$

Consequently, by multilinearity of determinants, we have

$$\begin{aligned}\det(A) &= \sum_{\pi \in \mathfrak{S}_n} a_{\pi(1)1} \cdots a_{\pi(n)n} \det(P_\pi) \\ &= \sum_{\pi \in \mathfrak{S}_n} \epsilon(\pi) a_{\pi(1)1} \cdots a_{\pi(n)n}.\end{aligned}$$

We summarize the above derivation as the following proposition which gives the *complete expansion* of the determinant.

Proposition 4.10. *For any $n \times n$ matrix $A = (a_{ij})$, we have*

$$\det(A) = \sum_{\pi \in \mathfrak{S}_n} \epsilon(\pi) a_{\pi(1)1} \cdots a_{\pi(n)n}.$$

Note that since $\det(A) = \det(A^\top)$, we also have

$$\det(A) = \sum_{\pi \in \mathfrak{S}_n} \epsilon(\pi) a_{1\pi(1)} \cdots a_{n\pi(n)}.$$

These formulae are more of theoretical than of practical importance. However, these formulae do show that the determinant is a polynomial function in the n^2 variables a_{ij}, and this has some importance consequences.

Remark: There is a geometric interpretation of determinants which we find quite illuminating. Given n linearly independent vectors $(u_1, \ldots, u_n)$ in $\mathbb{R}^n$, the set

$$P_n = \{\lambda_1 u_1 + \cdots + \lambda_n u_n \mid 0 \leq \lambda_i \leq 1,\ 1 \leq i \leq n\}$$

is called a *parallelotope*. If $n = 2$, then P_2 is a *parallelogram* and if $n = 3$, then P_3 is a *parallelepiped*, a skew box having u_1, u_2, u_3 as three of its corner sides.

Then, it turns out that $\det(u_1, \ldots, u_n)$ is the *signed volume* of the parallelotope P_n (where volume means n-dimensional volume).

The sign of this volume accounts for the orientation of P_n in $\mathbb{R}^n$.

As we saw, the determinant of a matrix is a multilinear alternating map of its rows.

This fact, combined with the fact that the determinant of a matrix is also a multilinear alternating map of its columns is often useful for finding short-cuts in computing determinants.

We illustrate this point on the following example which shows up in polynomial interpolation.

Example 4.1. Consider the so-called *Vandermonde determinant*

$$V(x_1, \ldots, x_n) = \begin{vmatrix} 1 & 1 & \ldots & 1 \\ x_1 & x_2 & \ldots & x_n \\ x_1^2 & x_2^2 & \ldots & x_n^2 \\ \vdots & \vdots & \ddots & \vdots \\ x_1^{n-1} & x_2^{n-1} & \ldots & x_n^{n-1} \end{vmatrix}.$$

We claim that

$$V(x_1, \ldots, x_n) = \prod_{1 \leq i < j \leq n} (x_j - x_i),$$

with $V(x_1, \ldots, x_n) = 1$, when $n = 1$.

4.3 Inverse Matrices and Determinants

In the next two sections, K is a commutative ring and when needed, a field.

Definition 4.6. Let K be a commutative ring. Given a matrix $A \in \mathrm{M}_n(K)$, let $\widetilde{A} = (b_{i\,j})$ be the matrix defined such that

$$b_{i\,j} = (-1)^{i+j} \det(A_{j\,i}),$$

the cofactor of $a_{j\,i}$. The matrix $\widetilde{A}$ is called the *adjugate* of A, and each matrix $A_{j\,i}$ is called a *minor* of the matrix A.

Note the reversal of the indices in

$$b_{i\,j} = (-1)^{i+j} \det(A_{j\,i}).$$

Thus, $\widetilde{A}$ is the transpose of the matrix of cofactors of elements of A.

Proposition 4.11. *Let K be a commutative ring. For every matrix $A \in \mathrm{M}_n(K)$, we have*

$$A\widetilde{A} = \widetilde{A}A = \det(A)I_n.$$

As a consequence, A is invertible iff $\det(A)$ is invertible, and if so, $A^{-1} = (\det(A))^{-1}\widetilde{A}$.

When K is a field, an element $a \in K$ is invertible iff $a \neq 0$.

In this case, the second part of the proposition can be stated as A is invertible iff $\det(A) \neq 0$.

Note in passing that this method of computing the inverse of a matrix is usually not practical.

We now consider some applications of determinants to linear independence and to solving systems of linear equations.

4.4 Systems of Linear Equations and Determinants

We now characterize when a system of linear equations of the form $Ax = b$ has a unique solution.

Proposition 4.12. *Given an $n \times n$-matrix A over a field K, the following properties hold:*

(1) For every column vector b, there is a unique column vector x such that $Ax = b$ iff the only solution to $Ax = 0$ is the trivial vector $x = 0$, iff $\det(A) \neq 0$.

(2) If $\det(A) \neq 0$, the unique solution of $Ax = b$ is given by the expressions

$$x_j = \frac{\det(A^1, \ldots, A^{j-1}, b, A^{j+1}, \ldots, A^n)}{\det(A^1, \ldots, A^{j-1}, A^j, A^{j+1}, \ldots, A^n)},$$

known as Cramer's rules.

(3) The system of linear equations $Ax = 0$ has a nonzero solution iff $\det(A) = 0$.

As pleasing as Cramer's rules are, it is usually impractical to solve systems of linear equations using the above expressions.

4.5 The Cayley–Hamilton Theorem

We conclude this chapter with an interesting and important application of Proposition 4.11, the *Cayley–Hamilton theorem*.

The results of this section apply to matrices over any commutative ring K.

First, we need the concept of the characteristic polynomial of a matrix.

Definition 4.7. If K is any commutative ring, for every $n \times n$ matrix $A \in \mathrm{M}_n(K)$, the *characteristic polynomial* $P_A(X)$ of A is the determinant

$$P_A(X) = \det(XI - A).$$

The characteristic polynomial $P_A(X)$ is a polynomial in $K[X]$, the ring of polynomials in the indeterminate X with coefficients in the ring K.

For example, when $n = 2$, if

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix},$$

then

$$P_A(X) = \begin{vmatrix} X - a & -b \\ -c & X - d \end{vmatrix} = X^2 - (a + d)X + ad - bc.$$

We can substitute the matrix A for the variable X in the polynomial $P_A(X)$, obtaining a *matrix* P_A. If we write

$$P_A(X) = X^n + c_1 X^{n-1} + \cdots + c_n,$$

then

$$P_A = A^n + c_1 A^{n-1} + \cdots + c_n I.$$

We have the following remarkable theorem.

Theorem 4.13. *(Cayley–Hamilton) If K is any commutative ring, for every $n \times n$ matrix $A \in \mathrm{M}_n(K)$, if we let*

$$P_A(X) = X^n + c_1 X^{n-1} + \cdots + c_n$$

be the characteristic polynomial of A, then

$$P_A = A^n + c_1 A^{n-1} + \cdots + c_n I = 0.$$

As a concrete example, when $n = 2$, the matrix

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

satisfies the equation

$$A^2 - (a + d)A + (ad - bc)I = 0.$$

4.6 Further Readings

Thorough expositions of the material covered in Chapter 1, 2, 3, and Chapter 4 can be found in Strang [27, 26], Lax [20], Meyer [22], Artin [1], Lang [18], Mac Lane and Birkhoff [21], Hoffman and Kunze [16], Dummit and Foote [10], Bourbaki [3, 4], Van Der Waerden [29], Serre [24], Horn and Johnson [15], and Bertin [2].