The Definition of a Determinant

The standard definition of a determinant requires some knowledge of permutations — which is where this section begins.

Let n denote a positive integer and set $X_n = \{1, 2, \cdots, n\}$. A permutation of n elements is a one-to-one correspondence

$$\sigma : X_n \longrightarrow X_n .$$

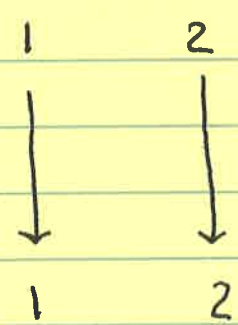
Recall that a one-to-one correspondence is a function which is both one-to-one and onto. S_n denotes the set of all permutations of n elements.

For our purposes, it is useful to depict elements of S_n - using arrow diagrams. For instance S_1 consists of a single element, σ_1, as depicted below.

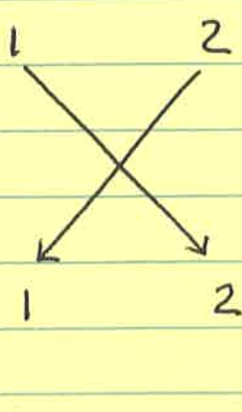
σ_1:

$S_2 = \{\sigma_1, \sigma_2\}$, as depicted below.

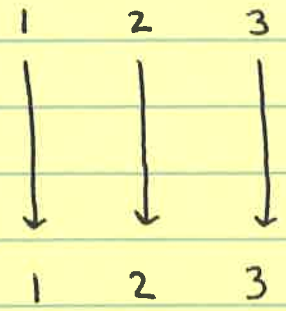
σ_1

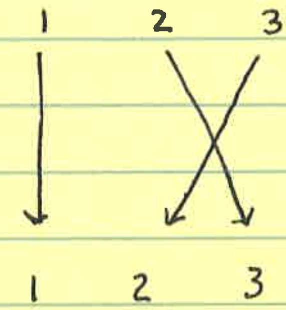


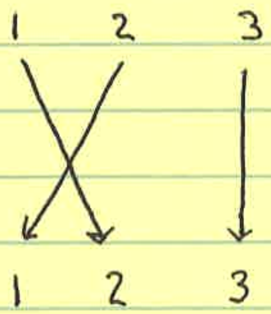
σ_2

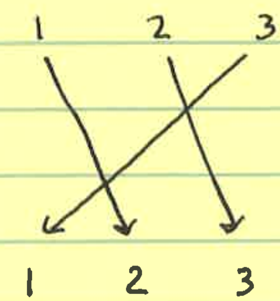


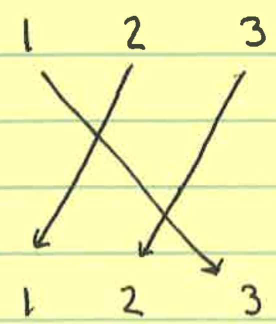
$S_3 = \{\sigma_1, \sigma_2, \sigma_3, \sigma_4, \sigma_5, \sigma_6\}$, as depicted below.

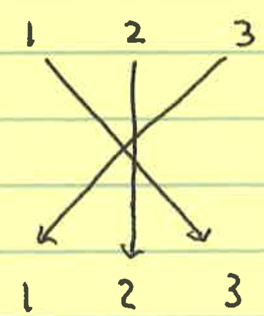
σ_1: 

σ_2: 

σ_3: 

σ_4: 

σ_5: 

σ_6: 

In general, S_n has $n!$ (n factorial) elements.

Permutations come in two basic flavors - technically called parities - called even and odd. Observe that the arrow diagrams which we have drawn have the property that at most two arrows intersect at any one point. (This can always be arranged by moving an arrows endpoints slightly.) The parity of a permutation, when this is the case, is determined by the number of crossings of arrows. If the number of crossings is even, then the permutation is even. Similarly, if the number of crossings is odd, then the permutation is odd.

For instance, in S_2, σ_1 is even while σ_2 is odd. In S_3, $\sigma_1, \sigma_4, \sigma_5$ are even while σ_2, σ_3 and σ_6 are odd.

If $\sigma \in S_n$, then the expression $(-1)^\sigma$ is defined as follows:

$$(-1)^\sigma = \begin{cases} 1 & \text{if } \sigma \text{ is even} \\ -1 & \text{if } \sigma \text{ is odd} \end{cases}.$$

$(-1)^\sigma$ is frequently called the sign of the permutation σ.

Now, suppose $A = [a_{ij}]$ is a $n \times n$ matrix. Then the determinant of A, denoted by $\det(A)$, is defined to be

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

Let's check this definition against computations of 2×2 and 3×3 determinants used in Calculus. In Calculus,

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

Using the definition

$$\det \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = (-1)^{\sigma_1} a_{1\sigma_1(1)} a_{2\sigma_1(2)} + (-1)^{\sigma_2} a_{1\sigma_2(1)} a_{2\sigma_2(2)}$$

$$= (1)\, a_{11}a_{22} + (-1)\, a_{12}a_{21}$$

Note that the two values agree.

For 3×3 determinants, in Calculus

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

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

$$+ \det\begin{bmatrix} a_{21} & a_{22} \\ a_{31} & a_{32} \end{bmatrix} a_{13}$$

$$= a_{11}(a_{22}a_{33} - a_{23}a_{32})$$

$$- a_{12}(a_{21}a_{33} - a_{23}a_{31})$$

$$+ a_{13}(a_{21}a_{32} - a_{22}a_{31})$$

$$= a_{11}a_{22}a_{33} - a_{11}a_{23}a_{32}$$

$$- a_{12}a_{21}a_{33} + a_{12}a_{23}a_{31}$$

$$+ a_{13}a_{21}a_{32} - a_{13}a_{22}a_{31}.$$

According to the definition

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

$$= (-1)^{\sigma_1} a_{1\sigma_1(1)} a_{2\sigma_1(2)} a_{3\sigma_1(3)} + (-1)^{\sigma_2} a_{1\sigma_2(1)} a_{2\sigma_2(2)} a_{3\sigma_2(3)}$$

$$+ (-1)^{\sigma_3} a_{1\sigma_3(1)} a_{2\sigma_3(2)} a_{3\sigma_3(3)} + (-1)^{\sigma_4} a_{1\sigma_4(1)} a_{2\sigma_4(2)} a_{3\sigma_4(3)}$$

$$+ (-1)^{\sigma_5} a_{1\sigma_5(1)} a_{2\sigma_5(2)} a_{3\sigma_5(3)} + (-1)^{\sigma_6} a_{1\sigma_6(1)} a_{2\sigma_6(2)} a_{3\sigma_6(3)}$$

$$= (1) a_{11} a_{22} a_{33} + (-1) a_{11} a_{23} a_{32}$$

$$+ (-1) a_{12} a_{21} a_{33} + (1) a_{12} a_{23} a_{31}$$

$$+ (1) a_{13} a_{21} a_{32} + (-1) a_{13} a_{22} a_{31} .$$

Once again, note that the two values

agree.

Immediate Consequences of the Definition

This section is devoted to proving three useful observations, which follow rather quickly from the definition of the determinant.

Proposition #33: Suppose $A = [a_{ij}]$ is a $n \times n$ matrix having a row consisting solely of zeros. Then $\det(A) = 0$.

Proof: Suppose the k-th row of A consists solely of zeros. Then $a_{kj} = 0$, for $j = 1, \cdots, n$. Then

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

Since $a_{k\sigma(k)} = 0$ for every $\sigma \in S_n$, we have

$$\det(A) = \sum_{\sigma \in S_n} 0 = 0.$$

Suppose $A = [a_{ij}]$ is a $n \times n$ matrix. We say A is upper triangular provided $a_{ij} = 0$ whenever $i > j$. Put another way, every nonzero entry in A lies on, or above, the main diagonal. Similarly, we say A is lower triangular provided $a_{ij} = 0$ whenever $j > i$.

<u>Proposition #34</u>: Suppose $A = [a_{ij}]$ is an $n \times n$ matrix. If A is either upper triangular or lower triangular, then $\det(A) = a_{11} a_{22} \cdots a_{nn}$.

<u>Proof</u>: We'll suppose A is upper triangular. The lower triangular case is similar.

The argument boils down to counting. Let $\sigma \in S_n$ and consider the expression

$$a_{1\sigma(1)} \cdots a_{n\sigma(n)}.$$

Since $a_{ij} = 0$ whenever $i > j$, in order for $a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ to be nonzero, we must have that

$$\sigma(i) \geq i$$

for $i = 1, \cdots, n$.

Suppose $\sigma(i) \geq i$ for $i = 1, \cdots, n$ and there is some $k \in \{1, \cdots, n-1\}$ such that $\sigma(k) = \ell > k$. Since $\sigma(i) \geq i$ for $i = 1, \cdots, n$, and σ is one-to-one, the $n - \ell + 1$ numbers $\sigma(\ell), \sigma(\ell+1), \cdots, \sigma(n)$ must all be strictly greater than ℓ. But there are only $n - \ell$ elements of $\{1, \cdots, n\}$ which are strictly greater than ℓ. Since $n - \ell + 1 > n - \ell$ - we have a contradiction.

We are left with only one possible $\sigma \in S_n$ such that $\sigma(i) \geq i$ for $i = 1, \cdots, n$. This being

$$\sigma(i) = i$$

for $i = 1, \cdots, n$.

Thus, $\det(A) = (-1)^{\sigma} a_{1\sigma(1)} \cdots a_{n\sigma(n)}$ where $\sigma(i) = i$ for $i = 1, \cdots, n$. So

$$\det(A) = (1)\, a_{11} \cdots a_{nn}.$$

<u>Proposition #35</u>: Suppose $A = [a_{ij}]$ is an $n \times n$ matrix. Then

$$\det(A) = \det(A^t).$$

<u>Proof</u>: Let $A^t = [\alpha_{ij}]$. Then

$\alpha_{ij} = a_{ji}$ for all $i, j = 1, \cdots, n$.

It follows that

$$\det(A^t) = \sum_{\sigma \in S_n} (-1)^\sigma \alpha_{1\sigma(1)} \cdots \alpha_{n\sigma(n)}$$

$$= \sum_{\sigma \in S_n} (-1)^\sigma a_{\sigma(1)1} \cdots a_{\sigma(n)n}.$$

Since $\sigma \in S_n$, the numbers $\sigma(1), \cdots, \sigma(n)$ are just the numbers $1, \cdots, n$ in some permuted order. Thus, we can rearrange the product $a_{\sigma(1)1} \cdots a_{\sigma(n)n}$ so that

$$a_{\sigma(1)1} \cdots a_{\sigma(n)n} = a_{\sigma(i_1)i_1} \cdots a_{\sigma(i_n)i_n}$$

-where $\sigma(i_1) = 1, \cdots, \sigma(i_n) = n$. Now, since σ is a one-to-one correspondence,

σ has an inverse function, σ^{-1}.

Note that $\sigma(i_j) = j$ implies that $\sigma^{-1}(j) = \sigma^{-1}(\sigma(i_j)) = i_j$.

It follows that

$$a_{\sigma(i_1)i_1} \cdots a_{\sigma(i_n)i_n} = a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)}.$$

Therefore

$$a_{\sigma(1)1} \cdots a_{\sigma(n)n} = a_{1\sigma^{-1}(1)} \cdots a_{n\sigma^{-1}(n)}.$$

Since this is valid for every $\sigma \in S_n$, we obtain

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

We now claim that $(-1)^{\sigma} = (-1)^{\sigma^{-1}}$.

The easy way to see this is to draw the arrow diagram for σ and then observe that the arrow diagram for σ^{-1} is obtained from the arrow diagram for σ by turning the arrow diagram for σ upside down and reversing the orientations on these arrows. This clearly implies σ and σ^{-1} have the same parity — so $(-1)^{\sigma} = (-1)^{\sigma^{-1}}$.

Thus

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

Or, if we set $\tau = \sigma^{-1}$,

$$\det(A^t) = \sum_{\tau^{-1} \in S_n} (-1)^{\tau} a_{1\tau(1)} \cdots a_{n\tau(n)} .$$

Now, the point is that as τ^{-1} varies over S_n, then τ also varies over S_n. Formally, this can be expressed as follows: If $\tau \in S_n$, then τ has a unique inverse, τ^{-1}. This implies that there is a well-defined function

$$F : S_n \longrightarrow S_n$$

defined by

$$F(\tau) = \tau^{-1}.$$

Its easy to see that F is onto — if $\lambda \in S_n$, then $F(\lambda^{-1}) = (\lambda^{-1})^{-1} = \lambda$.

Now, since S_n is a finite set, having exactly $n!$ elements, F being onto implies F is also one-to-one (the Pigeon Hole Principle). But this implies F is a one-to-one correspondence. Thus $F^{-1}: S_n \to S_n$ exists and F^{-1} is also a one-to-one correspondence. As $F^{-1}(\tau^{-1}) = \tau$ – we see that as τ^{-1} varies over S_n, so must τ vary over S_n.

This last argument implies that

$$\sum_{\tau^{-1} \in S_n} (-1)^{\tau} a_{1\tau(1)} \cdots a_{n\tau(n)}$$

$$= \sum_{\tau \in S_n} (-1)^{\tau} a_{1\tau(1)} \cdots a_{n\tau(n)} -$$

since, in either sum, τ will represent each element of S_n exactly one time.

As $\det(A) = \sum_{\tau \in S_n} (-1)^{\tau} a_{1\tau(1)} \cdots a_{n\tau(n)}$,

we have established that

$$\det(A^t) = \det(A).$$

HW #42: Compute the following determinants.

1) $$\det \begin{bmatrix} t-1 & 2 & 4 \\ 0 & t-3 & 5 \\ 0 & 0 & t-6 \end{bmatrix}$$

2) $$\det \begin{bmatrix} a_{11} & a_{12} & 0 \\ a_{21} & a_{22} & 0 \\ a_{31} & a_{32} & 0 \end{bmatrix}$$

3) $\det(I_n)$.