Math 204-2, Fall 2008
Notes on Determinants

1. Introduction

The goal of these notes is to give a proof of the following important theorem regarding determinants.

Theorem 1.1. *Let A be an $n \times n$ matrix. Then* $\det(A)$ *may be computed by cofactor expansion along any row or column. In other words, for any numbers i_0 and j_0 between* 1 *and n, we have*

$$\det(A) = \sum_{j=1}^{n} a_{i_0 j} A_{i_0 j} = \sum_{i=1}^{n} a_{i j_0} A_{i j_0},$$

where $A_{ij} = (-1)^{i+j} \det(M_{ij})$ is the (i,j)th cofactor of A.

We'll prove this theorem by giving a more explicit, non-inductive description of the determinant, and then we'll show that each sum above is equal to this explicit formula. In order to describe this formula, we need to discuss *permutations*.

2. Permutations

A permutation is just a way of rearranging the numbers $\{1, \ldots, n\}$. More formally, we can think of a permutation as a function that assigns to each number $1, \ldots, n$ a new number, in a 1-1 fashion.

Definition 2.1. *A* permutation *of the set $\{1, \ldots, n\}$ is a 1-1 function*

$$\sigma : \{1, \ldots, n\} \longrightarrow \{1, \ldots, n\}.$$

We denote the set of all permutations of $\{1, \ldots, n\}$ by Σ_n.

Here 1-1 means that no two numbers are sent to the same place under the function σ, i.e. if $i \neq j$ then $\sigma(i) \neq \sigma(j)$. Since the domain and range of σ each contain n elements, the fact that σ is 1-1 implies that it must also be *onto*, that is, every number j can be written as $\sigma(i)$ for some i. We call a function that is both 1-1 and onto a *bijection*.

We will need to break up permutations into two types, the *even* permutations and the *odd* permutations. In order to do this, we need to consider *inversions.*

Definition 2.2. *If $\sigma \in \Sigma_n$ is a permutation, then an* inversion *in σ is a pair (i, j) with $1 \leqslant i < j \leqslant n$ such that $\sigma(i) > \sigma(j)$.*

So inversions are numbers whose order is reversed by the permutation σ.

Example 2.3. *Let $\sigma \in \Sigma_5$ be the permutation defined by*

$$\sigma(1) = 3, \ \ \sigma(2) = 1, \ \ \sigma(3) = 4, \ \ \sigma(4) = 2, \ \ \sigma(5) = 5.$$

Then the inversions in σ are the pairs $(1, 2)$, $(1, 4)$, $(3, 4)$.

Definition 2.4. *A permutation $\sigma \in \Sigma_n$ is called* even *if σ contains an even number of inversions, and σ is called* odd *if it contains an odd number of inversions.*

For example, the permutation in Example 2.3 is *odd* because it contains 3 inversions.

In our description of the determinant, the various signs appearing will depend on whether certain permutations are odd or even. For this reason, we make the following definition.

Definition 2.5. *The* sign *of a permutation σ is defined by*

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

Hence the permutation in Example 2.3 has sign -1.

3. Explicit definition of the determinant

We can now give a new description of the determinant, which we will prove agrees with the old (inductive) definition.

This definition is motivated by noticing that when one expands out the determinant of a 3×3 matrix A, the terms one gets are products of 3 terms of A, and each such product consists of a single entry from each row and column of A. For example, one of the terms is $a_{11}a_{22}a_{33}$.

Products of this form can be described using permutations. Let $A = (a_{ij})$ be an $n \times n$ matrix. Given any permutation $\sigma \in \Sigma_n$, we can form the product

$$a_{1\sigma(1)}a_{2\sigma(2)} \ldots a_{n\sigma(n)}.$$

In this product, the kth entry comes from row k and column $\sigma(k)$. Clearly such a product contains exactly one entry from each row, and since σ is a bijection, it also contains exactly one entry from each column.

Definition 3.1. *For any $n \times n$ matrix A, we define $D(A)$ to be the sum*

$$\sum_{\sigma \in \Sigma_n} sgn(\sigma) a_{1\sigma(1)} a_{2\sigma(2)} \dots a_{n\sigma(n)}.$$

To compare this number $D(A)$ to our inductive definition of the determinant, we will need to consider the numbers $D(M_{ij})$ where M_{ij} is an $(n-1) \times (n-1)$ minor of A. Since the rows and columns of the minor matrices of an $n \times n$ matrix A are most naturally numbered using subsets of $\{1, \dots, n\}$ rather than the set $\{1, \dots, n-1\}$ itself, we generalize the function D a bit.

Let A be a matrix whose rows and columns are numbered using *any* increasing lists of numbers. We will write $R(A)$ for the set of numbers corresponding to the rows of A, and $C(A)$ for the set of numbers corresponding to the columns of A. For example, if A is a 4×4 matrix whose rows and columns are numbered in the usual fashion, then the minor matrix M_{23} contains rows $1, 3$ and 4 of A and columns $1, 2$ and 4 of A, so we would set $R(M_{23}) = \{1, 3, 4\}$ and $C(M_{23}) = \{1, 2, 4\}$. Note that this numbering of rows and columns is additional information that we tack on to our matrix, and we'll call a matrix together with such a numbering a *labeled* matrix. If A is any labeled matrix, then we will write its entries in the form a_{ij}, where $i \in R(A)$ and $j \in C(A)$.

If A is a square, labeled matrix, then $R(A)$ and $C(A)$ have the same size, and we can define $\Sigma_{R(A),C(A)}$ to be the set of all bijections $\sigma : R(A) \to C(A)$. For example, if $R(A) = \{1, 3, 4\}$ and $C(A) = \{1, 2, 4\}$, then the function $\sigma(1) = 2$, $\sigma(3) = 1$, $\sigma(4) = 4$ is an element of $\Sigma_{R(A),C(A)}$.

If σ is an element of $\Sigma_{R(A),C(A)}$, then we define an inversion in σ to be a pair of numbers $i, j \in R(A)$ such that $i < j$ but $\sigma(i) > \sigma(j)$. We then set $\operatorname{sgn}(\sigma) = 1$ if σ contains an even number of inversions, and we set $\sigma = -1$ if σ contains an odd number of inversions. Notice that when $R(A) = C(A) = \{1, \dots, n\}$, this is exactly our original definition of $\operatorname{sgn}(\sigma)$.

Definition 3.2. *Let A be a square, labeled matrix, and say $R(A) = \{r_1 < r_2 < \ldots < r_n\}$. Then we define $D(A)$ to be the sum*

$$\sum_{\sigma \in \Sigma_{R(A),C(A)}} sgn(\sigma) a_{r_1\sigma(r_1)} \cdots a_{r_n\sigma(r_n)}.$$

Note that this definition agrees with the previous one when $R(A) = C(A) = \{1, \ldots, n\}$.

We also need to think about how cofactor expansion works for labeled matrices.

Definition 3.3. *Let A be a square, labeled matrix, with $R(A) = \{r_1 < r_2 < \cdots, r_n\}$ and $C(A) = \{c_1 < c_2 < \cdots < c_n\}$. Then we define the determinant of A inductively via cofactor expansion along the first row of A:*

$$\det(A) = \sum_{k=1}^{n} a_{r_1 c_k} (-1)^{1+k} \det(M_{r_1 c_k}),$$

where $M_{r_l c_k}$ denotes the labeled *matrix obtained from A by deleting row r_l and column c_k; this new matrix is labeled by $R(M_{r_l c_k}) = \{r_1 < r_2 < \cdots r_{l-1} < r_{l+1} < \cdots r_n\}$ and $C(M_{r_l c_k}) = \{c_1 < c_2 < \cdots c_{k-1} < c_{k+1} < \cdots c_n\}$.*

Similarly, we define the cofactor expansion of A along row r_i by

$$\sum_{k=1}^{n} a_{r_i c_k} (-1)^{i+k} \det(M_{r_i c_k}),$$

and along column c_j by

$$\sum_{k=1}^{n} a_{r_k c_j} (-1)^{k+j} \det(M_{r_k c_j}).$$

Notice that these definitions are really nothing new at all; they give exactly the same numbers as if we simply relabeled A in the normal way. This is just useful in keeping track of the labeling of the minors.

The main result of these notes is:

Theorem 3.4. *Let A be an $n \times n$ matrix (labeled in the usual way). Then for any i and j between 1 and n we have*

$$\sum_{k=1}^{n} a_{ik} A_{ik} = D(A) = \sum_{k=1}^{n} a_{kj} A_{kj}.$$

In other words, computing the determinant by cofactor expansion along any row or column always results in the same number, namely $D(A)$.

Proof. We will prove the theorem by induction on n, and in fact we will prove it for *all* labeled matrices. When $n = 1$, A is simply a single number $a \in \mathbb{R}$, and all of the above numbers are simply a. We now assume that for all *labeled* $(n-1) \times (n-1)$ matrices A,

$$\sum_{k=1}^{n} a_{r_i c_k} (-1)^{i+k} \det(M_{r_i c_k}) = D(A)$$

$$= \sum_{k=1}^{n} a_{r_k c_j} (-1)^{k+j} \det(M_{r_k c_j}),$$

and we consider an $n \times n$ matrix A labeled by $R(A) = \{r_1 < r_2 < \cdots, r_n\}$ and $C(A) = \{c_1 < c_2 < \cdots < c_n\}$.

Let us consider cofactor expansion of A along row r_i. This is given by the sum

$$(1) \qquad \sum_{j=1}^{n} (-1)^{i+j} a_{r_i c_j} \det M_{r_i c_j}.$$

Now, each minor matrix $M_{r_i c_j}$ is a labeled $n \times n$ matrix, so we know by our induction hypothesis that

$$\det M_{r_i c_j} =$$

$$(2) \qquad \sum_{\sigma \in \Sigma_{R(A)_i, C(A)_j}} \operatorname{sgn}(\sigma) a_{r_1 \sigma(r_1)} \cdots a_{r_{i-1} \sigma(r_{i-1})} a_{r_{i+1} \sigma(r_{i+1})} \cdots a_{r_n \sigma(r_n)}.$$

where the new label sets $R(A)_i$ and $C(A)_j$ are obtained by removing r_i and c_j.

Combining (1) and (2), we find that cofactor expansion along row r_i gives the sum

(3)

$$\sum_{j=1}^{n} (-1)^{i+j} a_{r_i c_j} \left(\sum_{\sigma \in \Sigma_{R(A)_i C(A)_j}} \operatorname{sgn}(\sigma) a_{r_1 \sigma(r_1)} \cdots a_{r_{i-1} \sigma(r_{i-1})} a_{r_{i+1} \sigma(r_{i+1})} \cdots a_{r_n, \sigma(r_n)} \right)$$

$$= \sum_{j=1}^{n} \sum_{\sigma \in \Sigma_{R(A)_i C(A)_j}} (-1)^{i+j} \operatorname{sgn}(\sigma) a_{r_i c_j} a_{r_1 \sigma(r_1)} \cdots a_{r_{i-1} \sigma(r_{i-1})} a_{r_{i+1} \sigma(r_{i+1})}.$$

We need to show that this sum agrees with

$$D(A) = \sum_{\sigma \in \Sigma_{R(A),C(A)}} \operatorname{sgn}(\sigma) a_{r_1\sigma(r_1)} \cdots a_{r_n\sigma(r_n)}. \tag{4}$$

First, note that each $\sigma \in \Sigma_{R(A),C(A)}$ sends r_i to c_{j_0} for some j_0, and hence restricts to a function $\sigma' \in \Sigma_{R(A)_i, C(A)_{j_0}}$. I claim that the term corresponding to σ in the sum (4) agrees with the term ($j = j_0$, $\sigma = \sigma'$) in the sum (3), i.e. that

$$\operatorname{sgn}(\sigma) a_{r_i c_{j_0}} a_{r_1\sigma(r_1)} \cdots a_{r_{i-1}\sigma(r_{i-1})} a_{r_{i+1}\sigma(r_{i+1})}$$
$$= (-1)^{i+j_0} \operatorname{sgn}(\sigma') a_{r_i c_{j_0}} a_{r_1\sigma'(r_1)} \cdots a_{r_{i-1}\sigma'(r_{i-1})} a_{r_{i+1}\sigma'(r_{i+1})}.$$

Note that in the latter expression, $\sigma'(r_l) = \sigma(r_l)$ for each $l \neq i$ (by definition of σ'), so these two expressions contain exactly the same terms a_{kl} and we just need to check that the signs agree, i.e. that

$$\operatorname{sgn}(\sigma) = (-1)^{i+j_0} \operatorname{sgn}(\sigma').$$

We need to count the number of inversions in σ, i.e. the number of pairs $r_l < r_k$ such that $\sigma(r_l) > \sigma(r_k)$. Denote this number by $\operatorname{inv}(\sigma)$. Note that every inversion in σ' is an inversion in σ, and the remaining inversions in σ are those involving r_i. Hence we have

(5)
$$\operatorname{inv}(\sigma) = \operatorname{inv}(\sigma') + \#\{r_k \in R(A) | r_k < r_i \text{ and } \sigma(r_k) > \sigma(r_i) = c_{j_0}\}$$
$$+ \#\{r_l \in R(A) | r_i < r_l \text{ and } \sigma(r_i) > \sigma(r_l)\}.$$

We will write

$$x = \#\{r_k \in R(A) \,|\, r_k < r_i \text{ and } \sigma(r_k) > \sigma(r_i) = c_{j_0}\}. \tag{6}$$

We will now compute the number

$$y = \#\{r_l \in R(A) \,|\, r_i < r_l \text{ and } c_{j_0} = \sigma(r_i) > \sigma(r_l)\}$$

in terms of i, j_0, and x.

Let σ^{-1} denote the inverse function to σ, that is, $\sigma^{-1}(c_l)$ is the number r_k such that $\sigma(r_k) = c_l$. Sending r_l to $\sigma(r_l)$ provides a bijective function from

$$\{r_l \in R(A) \,|\, r_i < r_l \text{ and } \sigma(r_i) > \sigma(r_l)\}$$

to the set

$$\{c_l \in C(A) \,|\, c_{j_0} = \sigma(r_i) > c_l \text{ and } r_i < \sigma^{-1}(c_l)\},$$

so the latter set has y elements as well.

There are a total of $j_0 - 1$ elements $c_l \in C(A)$ with $c_{j_0} > c_l$, and we have just seen that y of them satisfy $r_i < \sigma^{-1}(c_l)$. Hence $j_0 - 1 - y$ of them satisfy $\sigma^{-1}(c_l) < r_i$, or in other words

$$j_0 - 1 - y = \#\{c_l \in C(A) \mid c_l < c_{j_0} \text{ and } \sigma^{-1}(c_l) < r_i\}.$$

There are a total of $i - 1$ elements $c_l \in C(A)$ with $\sigma^{-1}(c_l) < r_i$, and we have just seen that $j_0 - 1 - y$ of them satisfy $c_l < c_{j_0}$. Hence the other $i - 1 - (j_0 - 1 - y) = i - j_0 + y$ of them satisfy $c_l > c_{j_0}$, or in other words

$$j_0 - i + y = \#\{c_l \in C(A) \mid c_l > c_{j_0} = \sigma(r_i) \text{ and } \sigma^{-1}(c_l) < r_i\}.$$

Finally, sending r_l to $\sigma(r_l)$ provides a bijective function from the set

$$\{r_l \in R(A) \mid \sigma(r_l) > \sigma(r_i) = c_{j_0} \text{ and } r_l < r_i\}$$

to the previous set. This last set has size x (see (6)), so we find that $j_0 - i + y = x$.

Looking back at equation (5), we see that the total number of inversions in σ is given by

$$\begin{aligned}\operatorname{inv}(\sigma) &= \operatorname{inv}(\sigma') + x + y = \operatorname{inv}(\sigma') + (j_0 - i + y) + y \\ &= \operatorname{inv}(\sigma') + j_0 + i - 2i + 2y\end{aligned}$$

so $\operatorname{inv}(\sigma) - \operatorname{inv}(\sigma') = j_0 + i + 2(y - i)$. So if $j_0 + i$ is even, then $\operatorname{inv}(\sigma)$ and $\operatorname{inv}(\sigma')$ have the same parity (i.e. both are odd or both are even), meaning that

$$\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma') = (-1)^{i+j_0}\operatorname{sgn}(\sigma'),$$

as desired. On the other hand, if $i + j_0$ is odd, then $\operatorname{inv}(\sigma)$ and $\operatorname{inv}(\sigma')$ have opposite parity, so again

$$\operatorname{sgn}(\sigma) = -\operatorname{sgn}(\sigma') = (-1)^{i+j_0}\operatorname{sgn}(\sigma'),$$

as desired.

This completes the proof, at least for cofactor expansion along rows. A symmetric argument works for cofactor expansion along columns. □