LECTURE 8: DETERMINANTS AND PERMUTATIONS

MA1111: LINEAR ALGEBRA I, MICHAELMAS 2016

1. Determinants

In the last lecture, we saw some applications of invertible matrices. We would now like to describe how to detect whether a matrix is invertible. Last time, we proved that if a matrix is invertible, then its RREF is I_n. In the first tutorial, we showed that if $ad - bc \neq 0$, then the matrix $\left(\begin{smallmatrix} a & b \\ c & d \end{smallmatrix}\right)$ reduces to I_2 and is hence invertible. Following the steps of that proof, it isn't hard to see that in fact this is an if and only if statement.

It will turn out for every square matrix of any size that there is a number associated to it, called the **determinant**, which vanishes if and only if the matrix isn't invertible. In the case of 2×2 matrices, this happens to be the number $ad-bc$. For general n, there is a unique function $\det A$ of matrices of size $n \times n$ satisfying a few simple properties. We will first think of the determinant as a function of n variables, $\det(r_1, \ldots, r_n)$, where r_i is the i-th row of A, thought of as a vector. The characterizing properties of det are the following:

(1) det is a multilinear function.
(2) det is **alternating**, which means that if any two rows of A are equal, say $r_i = r_j$, then the determinant is 0.
(3) The value of det on I_n is 1.

Here is a simple consequence of the above properties: if we swap two rows in A, say rows i and j, then we have, by the alternating property and by multilinearity,

$$\begin{aligned} 0 &= \det(r_1, \ldots, r_i + r_j, \ldots, r_i + r_j, \ldots, r_n) \\ &= \det(r_1, \ldots, r_i, \ldots, r_i, \ldots, r_n) + \det(r_1, \ldots, r_i, \ldots, r_j, \ldots, r_n) \\ &+ \det(r_1, \ldots, r_j, \ldots, r_i, \ldots, r_n) + \det(r_1, \ldots, r_j, \ldots, r_j, \ldots, r_n) \\ &= \det(r_1, \ldots, r_i, \ldots, r_j, \ldots, r_n) + \det(r_1, \ldots, r_j, \ldots, r_i, \ldots, r_n), \end{aligned}$$

so that **swapping two rows of a matrix multiplies the determinant by -1.** Moreover, if we add a multiple c of row j to row i, we find, again using multilinearity and the alternating property, that

$$\begin{aligned} &\det(r_1, \ldots, r_i + c r_j, \ldots, r_j, \ldots, r_n) \\ &= \det(r_1, \ldots, r_i, \ldots, r_j, \ldots, r_n) + c\det(r_1, \ldots, r_j, \ldots, r_j, \ldots, r_n) \\ &= \det(r_1, \ldots, r_i, \ldots, r_j, \ldots, r_n), \end{aligned}$$

Date: October 14, 2016.

so that **adding a multiple of one row to another leaves a determinant unchanged.** Finally, if we multiply a row by a constant, then multilinearity again shows that **the determinant is multiplied by the same constant.**

Example. *We use the properties above to find that (with $e_1 = (1, 0)$ and $e_2 = (0, 1)$)*

$$\det\begin{pmatrix} a & b \\ c & d \end{pmatrix} = \det((a, b), (c, d)) = \det(ae_1 + be_2, ce_1 + de_2)$$
$$ac\det(e_1, e_1) + ad\det(e_1, e_2) + bc\det(e_2, e_1) + bd\det(e_2, e_2)$$
$$= ad\det(e_1, e_2) - bc\det(e_1, e_2) = ad\det I_2 - bc\det I_2 = ad - bc.$$

Example. *We compute (with $e_1 = (1, 0, 0)$, $e_2 = (0, 1, 0)$, and $e_3 = (0, 0, 1)$)*

$$\det\begin{pmatrix} 1 & 2 & 3 \\ 1 & 0 & 2 \\ 0 & 1 & 0 \end{pmatrix} = \det(e_1 + 2e_2 + 3e_3, e_1 + 2e_3, e_2)$$
$$= \det(e_1, e_1 + 2e_3, e_2) + 2\det(e_2, e_1 + 2e_3, e_2) + 3\det(e_3, e_1 + 2e_3, e_2)$$
$$= \det(e_1, e_1, e_2) + 2\det(e_1, e_3, e_2) + 2\det(e_2, e_1, e_2) + 4\det(e_2, e_3, e_2) + 3\det(e_3, e_1, e_2) + 6\det(e_3, e_3, e_2)$$
$$= 2\det(e_1, e_3, e_2) + 3\det(e_3, e_1, e_2) = -2\det(e_1, e_2, e_3) + 3\det(e_1, e_2, e_3) = -2\det I_3 + 3\det I_3 = 1.$$

After doing a few such numerical examples, you discover that there are some patterns which seem to emerge. We will explain this by giving another definition of the determinant. Firstly, however, we need to describe a new mathematical object.

2. Permutations

A **permutation** of a set with n elements is simply a rearrangement of the elements. We will usually describe these by taking as a set of size n the set of the first n natural numbers $1, 2, \ldots, n$. For example, we may rearrange the numbers $1, 2, 3$ according to a permutation π, giving $3, 1, 2$. One convenient way of keeping track of this action is to use the two row notation:

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

In this notation, the numbers in the second row are the results of applying the permutation to the elements of the first row. There is one very special type of permutation we shall need, called a **transpostion**. We will denote by (ij) the permutation which only switches i and j and leaves all other elements unchanged. Note that we must know by context what n is, since (ij) can denote a permutation of any number n of elements. For example,

$$(23) = \begin{pmatrix} 1 & 2 & 3 \\ 1 & 3 & 2 \end{pmatrix}.$$

More generally, a k-**cycle** is a permutation $(a_1, a_2, \ldots, a_k)$ which sends a_1 to a_2, a_2 to $a_3, \ldots$ a_{k-1} to a_k and finally a_k to a_1, and leaves all elements not listed fixed. For example, we have

$$(152) = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 1 & 3 & 4 & 2 \end{pmatrix}.$$

We can define products of permutations by writing them next to one another and by applying the actions of each working from right to left (the reason being that these are really compositions of functions). This will give a new permutation.

Example. *If we consider two permutations*

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 2 & 4 & 1 & 3 & 5 \end{pmatrix}, \qquad \sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 4 & 3 & 2 & 1 \end{pmatrix},$$

then

$$\pi\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 5 & 3 & 1 & 4 & 2 \end{pmatrix}, \qquad \sigma\pi = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 \\ 4 & 2 & 5 & 3 & 1 \end{pmatrix}.$$

Note that multiplication of permutations is **not** *commutative.*

What we are interested in is decomposing permutations into cycles. This can be done using the following simple procedure.

Algorithm. *To write a permutation π as a product of disjoint cycles (i.e., a product of cycles with no common elements between any two of them), pick the first number among $1, 2, \ldots n$ which isn't fixed (going to itself) by π. This is the first element of the first cycle. To find the rest of the first cycle, keep applying π to successive elements until you get back to the first element you started with, in which case you have "closed off" the first cycle. Now repeat this process on the set of all remaining numbers from $1, \ldots n$ until every element is either fixed by π or is in one of the cycles you have already written down.*

Example. *For the permutations in the last example, we have*

$$\pi = (1243),$$

and

$$\sigma = (15)(24).$$

We also have

$$\pi\sigma = (1523),$$

and

$$\sigma\pi = (1435).$$

The key application of these cycle decompositions is the following result, whose proof would require too much time for the application we have in mind in this class.

Theorem. *If π is is any permutation, then there is a certain unique number ± 1, called* $\operatorname{sign}(\pi)$ *or the sign of π, associated to π. If π can be written as a product of m transpositions (this in general isn't unique, but the claim is that its true for* **any** *representation as a product of transpositions that you find), then* $\operatorname{sign}(\pi) = (-1)^m$.

We also say that π is **even** if $\operatorname{sign}(\pi) = +1$ and that π is **odd** if $\operatorname{sign}(\pi) = -1$. The point is that every permutation can be written as a product of transpositions. This can be found by first finding the cycle decomposition of the preceding algorithm and then using the following elementary decomposition of any cycle into transpositions:

$$(a_1, \ldots a_k) = (a_1 a_k) \cdots (a_1 a_3)(a_1 a_2).$$

This directly shows that any k cycle has sign $(-1)^{k+1}$.

Note that the parity of an k-cycle as a permutation is **opposite** the parity of k as an integer.

Thus, if we use the algorithm above, then the parity of any permutation which is a product of cycles of lengths $k_1, \ldots, k_\ell$ can be read off as the product $(-1)^{k_1+1} \cdots (-1)^{k_\ell+1}$.

Example. *Assuming the notation of the last example, we can use the cycle decompositions directly to read off the signs of all the permutations involved:*

$$\operatorname{sign}(\pi) = (-1)^{4+1} = -1,$$

and

$$\operatorname{sign}(\sigma) = (-1)^3(-1)^3 = +1,$$
$$\operatorname{sign}(\pi\sigma) = (-1)^5 = -1,$$

and

$$\operatorname{sign}(\sigma\pi) = (-1)^5 = -1.$$

3. Leibniz form of the determinant

Using the reasoning in the above examples for determinants, we can write down a general formula for determinants using permutations. As we extrapolate from the examples above, we can see that if we want to use the defining properties of a determinant to compute it, we first write each row vector in terms of the **standard basis vectors** $e_1, \ldots, e_n$, where e_i is the i-th row of I_n, and then use multilinearity to expand. We will then get a sum of products of coordinates of the row vectors, namely matrix entries, with one term in each product coming from each row, times values of determinants on some ordering of the e_i's. Now, by the alternating property, whenever one of these e_j functions appears twice, we will get a zero in that term. Otherwise, we are exactly in the case that the entries in the corresponding term are just a permutation of $e_1, \ldots, e_n$. Finally, using the property that switching two rows just multiplies the determinant by a factor of -1, we take each of these remaining terms and perform a series of transpositions (swapping two entries), to reduce it to the sign of the corresponding permutation times the determinant of I_n, which is of course 1. All of this is summarized in the following result, where if $x = 1, \ldots, n$ and π is a permutation, then $\pi(x)$ is the result of applying π to x.

Theorem. *The determinant function defined by the properties above can be computed for any matrix A as*

$$\det A = \sum_{\pi} \operatorname{sign}(\pi) A_{1\pi(1)} A_{2\pi(2)} \cdots A_{n\pi(n)},$$

where in the sum, π runs over all permutations of $1, 2, \ldots n$.