Solution Set 6D

Alex Levin
Math23a
Prof. Boller

December 3, 2004

(8) Recall the definition of a transpose of a matrix as referred to in homework problem #5.8 and prove the following:

Theorem 1. *In A is an $n \times n$ matrix, then* $\det(A^t) = \det(A)$.

Proof. Let

$$A = \begin{pmatrix} a_{1,1} & a_{1,2} & \dots & a_{1,n} \\ a_{2,1} & a_{2,2} & \dots & a_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{n,1} & a_{n,2} & \dots & a_{n,n} \end{pmatrix} \tag{1}$$

and

$$A^t = \begin{pmatrix} a'_{1,1} & a'_{1,2} & \dots & a'_{1,n} \\ a'_{2,1} & a'_{2,2} & \dots & a'_{2,n} \\ \vdots & \vdots & \ddots & \vdots \\ a'_{n,1} & a'_{n,2} & \dots & a'_{n,n} \end{pmatrix} \tag{2}$$

where, of course,

$$a'_{i,j} = a_{j,i}. \tag{3}$$

By the definition of the determinant, we obtain

$$\det(A^t) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a'_{\sigma(1),1} \dots a'_{\sigma(n),n}. \tag{4}$$

Notice that since each σ is a permutation and thus $\sigma(j)$ runs over all of $\{1, 2, \dots, n\}$ as j runs over $\{1, 2, \dots, n\}$, we can rewrite the term in (4) corresponding to a certain permutation σ as

$$\operatorname{sgn}(\sigma) a'_{\sigma(i_1),i_1} \dots a'_{\sigma(i_n),i_n} \tag{5}$$

where each i_r satisfies $\sigma(i_r) = r$. Again, we can find $i_1, \ldots, i_n$ because σ is a permutation; in fact, $i_r = \sigma^{-1}(r)$. So, summing over all permutation gives us

$$\det(A^t) = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a'_{1,\sigma^{-1}(1)} \ldots a'_{n,\sigma^{-1}(n)} \tag{6}$$

$$= \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma) a_{\sigma^{-1}(1),1} \ldots a_{\sigma^{-1}(n),n} \tag{7}$$

by (3).

For each permutation, there is a unique inverse. So, we can rewrite the sum in (7) as a sum over σ^{-1}. Now, we claim that $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$. More generally, $\operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) = \operatorname{sgn}(\sigma\tau)$. Indeed, if we write σ as k transpositions, and τ as m transpositions, then $\sigma\tau$ will be expressible as $m + k$ transpositions. Then, noting (from class) that the number of transpositions that can be used to express any given permutation is always either even or odd (this makes the signum well-defined), we obtain

$$\operatorname{sgn}(\sigma\tau) = (-1)^{k+m} \tag{8}$$

$$= (-1)^k(-1)^m \tag{9}$$

$$= \operatorname{sgn}(\sigma)\operatorname{sgn}(\tau) \tag{10}$$

So, since $\sigma\sigma^{-1}$ gives the identity permutation, which clearly has signum 0, the above property tells us that we must have $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.

Given all of this, (7) gives us

$$\det(A^t) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a_{\pi(1),1} \ldots a_{\pi(n),n} \tag{11}$$

$$= \det(A) \tag{12}$$

as desired. Note that we have just relabeled σ^{-1} as π.

□

This problem went extremely well for most people. The proofs were nice and clean, and the most common mistake was just forgetting to state (or to prove) that $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\sigma^{-1})$.

The flipside to this is the fact that the proofs sounded very similar; in many of the papers I read, the line "If $\sigma(i) = j$, then $i = \sigma^{-1}(j)$" appeared verbatim. In the future, please cite your sources.

Finally, a few people tried proving this result by expansion by minors. This is a mess. Not only are the minors all different (so you need to use some sort of induction), but also, you end up using the fact that to compute the determinant, one can go down either row or column in the expansion. This, however, is a consequence of $\det(A^t) = \det(A)$.

(9) Show that $A : V \to V$ is invertible if and only if $\det(A) \neq 0$.

Let V be n-dimensional. Notice that A is invertible if and only if its columns are linearly independent. Indeed, since the matrix A essentially represents $(Ae_1, \ldots, Ae_n)$, where $e_1, \ldots, e_n$ is the standard basis, once we have n linearly independent columns, we know that $Ae_1, \ldots, Ae_n$ are all linearly independent, hence $\dim \operatorname{im}(A) = n$, and so A is surjective and thus invertible.

For the reverse implication, suppose that A is invertible, and that there is some linear relation on its columns. In other words, suppose $c_1 Ae_1 + \ldots + c_n Ae_n = 0$. By linearity, this is just $A(c_1 e_1 + \ldots + c_n e_n)$. Then, since A is invertible, it has trivial kernel, which means that $c_1 e_1 + \ldots + c_n e_n = 0$. Linear independence of $e_1, \ldots, e_n$ implies that $c_1 = \ldots = c_n = 0$, so that the columns of A are linearly independent.

The determinant, $\det(A)$, is just an appropriately normalized n-linear, alternating form evaluated on the columns of A. We know that the output of such a form is non-zero if and only if the input consists of linearly independent vectors. By the preceding paragraphs, then, $\det(A) \neq 0 \iff A$ is invertible.

Most people proved the backwards direction (A is invertible $\implies \det(A) \neq 0$) by saying that if A^{-1} exists, then $AA^{-1} = I$, so $\det(AA^{-1}) = \det(A)\det(A^{-1}) = \det(I) = 1$, and noting that this cannot happen if $\det(A) = 0$. This is fine, except we never really proved the identity $\det(AB) = \det(A)\det(B)$.

What was actually very surprising is that a vast majority of students, in showing that $\det(A) \neq 0$ implies A is invertible, demonstrated that the columns of A are linearly independent, and then worked very hard to prove that A is injective (or surjective) and hence invertible. But once we have n linearly independent vectors in the image of A (i.e. the columns of the matrix), we know A is surjective since $\dim \operatorname{im}(A) = n$; additional work is unnecessary.