Mathematics MATH 223, Fall 2007
Linear Algebra
On determinants

These notes were prepared for a different class. For this course we expect you to know the properties contained in the THEOREM — what they say, and how to use them. You are not responsible for the proofs. You are also responsible for knowing the cofactor expansions, as stated in Proposition 2 (and explained in the paragraph before that statement). I hope that at least some of you will find the fully rigourous explanations interesting and/or helpful.

Suppose that $A = (a_{j,k})_{1\le j\le n, 1\le k\le n}$ is an $n \times n$ matrix over any field F. We assume $n \ge 2$. We define the *determinant* $det(A)$ as follows:

$$det(A) = \sum_{\sigma\in n} sgn(\sigma) \prod_{j=1}^{n} a_{j,\sigma(j)}.$$

Here S_n is the collection of all permutations of the set $\{1, \ldots, n\}$, where a *permutation* of a set Z is a function $\sigma : Z \longrightarrow Z$ which is both one-to-one and onto. (Note, in our case $Z = \{1, \ldots, n\}$ is finite, so any function from Z to itself that is one-to-one is necessarily onto, and vice versa. This is not always the case if Z is infinite.) There are $n!$ elements of S_n and they form a group under function composition. If $\sigma, \pi \in S_n$, we write $\sigma\pi$ for the composition so $\sigma\pi(k) = \sigma(\pi(k))$ for each $k \in Z$. $\sigma\pi$ will be a permutation also.

We will write $\begin{pmatrix} k & \ell \end{pmatrix}$ for the permutation τ such that $\tau(k) = \ell$, $\tau(\ell) = k$ and $\tau(m) = m$ for every $m \ne k, \ell$. (Here k, ℓ and m stand for distinct positive integers $\le n$.) Such a permutation is called a *transposition.*

If $\sigma \in S_n$, $sgn(\sigma)$ will be either ± 1, depending on σ. It is not hard to see that any element $\sigma \in S_n$ can be expressed as a product $\tau_1 \cdots \tau_k$, where each τ_k is a transposition. This expression for σ is not unique, but we have the following:
PROPOSITION 1. If $\sigma = \tau_1 \cdots \tau_k = \rho_1 \cdots \rho_m$ where each τ and ρ is a transposition, then k and m are either both even or both odd.

Assuming this proposition, we call the permutation σ of $\{1, \ldots, n\}$ *even* if $\sigma = \tau_1 \cdots \tau_k$ with k even and each τ a transposition; otherwise σ is (uh) *odd.* We set $sgn(\sigma) = 1$ if σ is even, $sgn(\sigma) = -1$ if σ is odd.

In case $n = 3$, the six permutations are the identity (which is even), the three transpositions $\begin{pmatrix} 1 & 2 \end{pmatrix}$, $\begin{pmatrix} 1 & 3 \end{pmatrix}$ and $\begin{pmatrix} 2 & 3 \end{pmatrix}$, which are odd, and the two 3-cycles $\begin{pmatrix} 1 & 2 & 3 \end{pmatrix} = \begin{pmatrix} 1 & 3 \end{pmatrix}\begin{pmatrix} 1 & 2 \end{pmatrix}$ and $\begin{pmatrix} 1 & 3 & 2 \end{pmatrix} = \begin{pmatrix} 1 & 2 \end{pmatrix}\begin{pmatrix} 1 & 3 \end{pmatrix}$, which are even. The determinant is then
$+1a_{1,1}a_{2,2}a_{3,3} - 1a_{1,2}a_{2,1}a_{3,3} - 1a_{1,3}a_{2,2}a_{3,1} - 1a_{1,1}a_{2,3}a_{3,2} + 1a_{1,2}a_{2,3}a_{3,1} + 1a_{1,3}a_{2,1}a_{3,2}$. (The last term, for instance, corresponds to the 3-cycle $\sigma = \begin{pmatrix} 1 & 3 & 2 \end{pmatrix}$ such that $\sigma(1) = 3$, $\sigma(2) = 1$ and $\sigma(3) = 2$.)

We will prove the proposition shortly, and then the following properties, which include the serious ones we need for determinants.
THEOREM.

1. $det(I) = 1$. In fact, if A is either upper or lower triangular, then $det(A) = \prod_{j=1}^{n} a_{j,j}$.
2. If B is obtained from A by switching two rows, then $det(B) = -det(A)$.
3. If A has two identical rows, or a row of all zeroes, then $det(A) = 0$.
4. If B is obtained from A by replacing the jth row R_j of A by αR_j, where α is a scalar, then $det(B) = \alpha det(A)$.
5. If A, A' and A'' have all rows identical except their jth rows, which are R_j, R_j' and R_j'' respectively, and $R_j'' = R_j + R_j'$, then $det(A'') = det(A) + det(A')$.
6. If R_j and R_k are distinct rows of A, and B is obtained from A by replacing R_j by $R_j + \alpha R_k$ for some scalar α, then $det(A) = det(B)$.
7. $det(A) = 0$ if and only if A is not invertible.
8. For any $n \times n$ A and B, $det(AB) = det(A)det(B)$.
9. If A is invertible, $det(A^{-1}) = (det(A))^{-1}$.
10. For any invertible P, $det(A) = det(P^{-1}AP)$.
11. $det(A) = det(A^T)$.

There are also the cofactor expansions, which we will explain later; these are useful for computing determinants.

Note that 2), 4) and 6) describe what happens to the determinant if we perform a single row operation on A. These properties, together with $det(I) = 1$, can be taken as defining the determinant. (Except that, if F has characteristic 2, we need to add that if A has a row of zeroes, its determinant is 0.) Properties 4) and 5) are referred to a the *multilinearity properties* of the determinant; they say that, although the determinant is not a linear function of the matrix, it is linear when regarded as a function of any single row (holding the other rows constant).

Proof of Proposition 1: There are several ways to do this; we will use a “brute force” method. We need two lemmas, the first of which is a matter of simple calculation, and left to the reader:

LEMMA 1. Suppose that k, ℓ, m, p are distinct. (If $n = 2$, we will not need the last two; if $n = 3$ we won’t use 4.)

1. $\begin{pmatrix} k & \ell \end{pmatrix}\begin{pmatrix} k & \ell \end{pmatrix} = \iota$, the identity.
2. $\begin{pmatrix} k & \ell \end{pmatrix}\begin{pmatrix} k & m \end{pmatrix} = \begin{pmatrix} \ell & m \end{pmatrix}\begin{pmatrix} k & \ell \end{pmatrix}$.
3. $\begin{pmatrix} k & \ell \end{pmatrix}\begin{pmatrix} \ell & m \end{pmatrix} = \begin{pmatrix} \ell & m \end{pmatrix}\begin{pmatrix} k & m \end{pmatrix}$.

4. $\begin{pmatrix} k & \ell \end{pmatrix}\begin{pmatrix} m & p \end{pmatrix} = \begin{pmatrix} m & p \end{pmatrix}\begin{pmatrix} k & \ell \end{pmatrix}$.

The second lemma actually contains the work.

LEMMA 2. There does not exist a product $\tau_1 \cdots \tau_{2a+1}$ of an odd number of transpositions which is equal to the identity. (So a is a natural number.)

Proof: By induction on a. Clearly, the identity is not a single transposition. We assume (as our induction hypothesis) that there is no product of $2a-1$ transpositions which is the identity, and show that no product of $2a+1$ of them can be ι.

So suppose that $\pi = \tau_1 \cdots \tau_{2a+1}$ is a product of transpositions in S_n. Let k be any number that occurs in any of the τ_j's. Write $\pi = \tau_1 \cdots \tau_q \begin{pmatrix} k & \ell \end{pmatrix} \tau_{q+2} \cdots \tau_{2a+1}$, where k doesn't occur in any of the first q τ's. It is possible that $q = 0$, but if $q+1 = 2a+1$, then $\pi(\ell) = k$ and in that case we have that $\pi \neq \iota$.

If $q+1 \neq 2a+1$, there are four possibilities for τ_{q+2}; they are represented in the previous lemma. In the first case, if $\tau_{q+2} = \begin{pmatrix} k & \ell \end{pmatrix}$, then we can cancel τ_{q+1} and τ_{q+2} and have π written as a product of $2a-1$ transpositions, which prevents it from being the identity. In each of the other three cases, we replace $\tau_{q+1}\tau_{q+2}$ according to the lemma to get π written as a product of $2a+1$ transpositions, such that none of the first $q+1$ of them feature k. (k remains fixed throughout this proof.)

We repeat this process as often as necessary, moving the leftmost occurrence of k further to the right each time out. If we ever encounter a transposition $\begin{pmatrix} k & \ell' \end{pmatrix}$ next to itself, we cancel these occurrences, expressing π as a product of $2a-1$ transpositions. In this case we finish by the induction hypothesis. If this never happens, then we can move k to the end; there will be an expression of π as a product of $2a+1$ transpositions such that k occurs only in the rightmost one. If this is $\begin{pmatrix} k & \ell^* \end{pmatrix}$, then $\pi(\ell^*) = k \neq \ell^*$. This completes the proof of the lemma.

To finish the proof of the proposition, suppose for a contradiction that $\sigma = \tau_1 \cdots \tau_k = \rho_1 \cdots \rho_m$ where k and m have different parity. Then $\iota = \tau_1 \cdots \tau_k \rho_m \cdots \rho_1$ expresses the identity as a product of an odd number of transpositions, and we won't have that.

The reason this proposition is necessary is, of course, to ensure that the definition of the determinant *makes sense.*

We now turn to the proof of the main theorem.

The proof of 1) is easy straight from the definition. The product of the diagonal elements occurs as a summand (with positive sign) in any determinant because it corresponds to the identity permutation. In $\sigma \neq \iota$ is any other permutation, there is j with $\sigma(j) < j$ and then if A is upper triangular, $a_{j,\sigma(j)} = 0$. Thus the product corresponding to σ is zero; the determinant of an upper triangular matrix is the sum of the product of the diagonal elements with a bunch of zeroes. The case where is A is lower triangular is similar, and the special case when $A = I$ immediate.

2) is more delicate, but quite believable. The idea is that any product that

occurs in A also occurs in B, except with the opposite sign. Indeed, suppose that B is obtained from A by interchanging rows k and ℓ; let $\tau = \begin{pmatrix} k & \ell \end{pmatrix}$. Then for any permutation σ, in the product $b_{1,\sigma(1)} \cdots b_{n,\sigma(n)}$ we have that $b_{j,\sigma(j)} = a_{j,\sigma(j)} = a_{j,\sigma\tau(j)}$ for any $j \neq k, \ell$ and $b_{k,\sigma(k)} = a_{\ell,\sigma(k)} = a_{\ell,\sigma\tau(\ell)}$ and $b_{\ell,\sigma(\ell)} = a_{k,\sigma\tau(k)}$. Thus the product in the expansion for $det(B)$ corresponding to σ is the same as the product in the expansion of $det(A)$ corresponding to $\sigma\tau$. Obviously, σ and $\sigma\tau$ have opposite signs. This shows 2).

The first part of 3) is an immediate consequence of 2), except in characteristic 2. If A has two identical rows, we switch them and find $det(A) = -det(A)$, which say that, as long as the characteristic of F is not 2, that $det(A) = 0$. But in characteristic 2, we see that every product will occur twice (as in the last paragraph) if A has two identical rows, and therefore they all cancel out in that case, too. In case A has a row (or a column) of all zeroes, since the expansion of $det(A)$ will have at least one factor that is zero in each product, we see immediately from the definition that $det(A) = 0$.

For 4), for each σ, the product in $det(B)$ corresponding to σ is $b_{1,\sigma(1)} \cdots b_{j-1,\sigma(j-1)} b_{j,\sigma(j)} b_{j+1,\sigma(j+1)} \cdots b_{n,\sigma(n)} =$ $a_{1,\sigma(1)} \cdots a_{j-1,\sigma(j-1)} (\alpha a_{j,\sigma(j)}) a_{j+1,\sigma(j+1)} \cdots a_{n,\sigma(n)}$ so by the distributive law, the sum over all σ's for B is α times the sum over all σ's for A. That is, $det(B) = \alpha det(A)$.

For 5), let $a_{j,k}$ be the j,k-entry in A, $a'_{j,k}$ the j,k-entry in a' and $a''_{j,k} = a_{j,k} + a'_{j,k}$ the j,k entry of A''. Of course, for any $\ell \neq j$, the ℓ,k entry in all three is the same (we will call it $a_{\ell,k}$). Then for any σ, the product corresponding to σ in the development of $det(A'')$ is $a_{1,\sigma(1)} \cdots a_{j-1,\sigma(j-1)} a''_{j,\sigma(j)} a_{j+1,\sigma(j+1)} \cdots a_{n,\sigma(n)} =$ $a_{1,\sigma(1)} \cdots a_{j-1,\sigma(j-1)} (a_{j,\sigma(j)} + a'_{j,\sigma(j)}) a_{j+1,\sigma(j+1)} \cdots a_{n,\sigma(n)} =$ $a_{1,\sigma(1)} \cdots a_{j-1,\sigma(j-1)} a_{j,\sigma(j)} a_{j+1,\sigma(j+1)} \cdots a_{n,\sigma(n)}$ $+ a_{1,\sigma(1)} \cdots a_{j-1,\sigma(j-1)} a'_{j,\sigma(j)} a_{j+1,\sigma(j+1)} \cdots a_{n,\sigma(n)}$. Since this is true for every σ, $det(A'') = det(A) + det(A')$.

For 6), let C be the matrix with jth row αR_k, but otherwise looks like A. It follows from 4) that the determinant of C is α times that of a matrix with two identical rows and hence $det(C) = 0$. But 5) tells us that $det(B) = det(A) + det(C)$, so indeed $det(B) = det(A)$.

Together 2), 4) and 6) tell us how a determinant changes as we perform elementary row operations on a matrix. Note that these never change a determinant from a nonzero scalar to zero, or vice versa. But row-reducing a matrix doesn't change whether or not it's invertible. In row-reducing, we eventually reach the identity matrix (invertible, with nonzero deteminant) or a matrix with a row of zeroes (of determinant zero, and not invertible). That is, $det(A) = 0$ if and only if $det(R) = 0$ where R is the RREF form of A; also A is noninvertible if and only if R is. Now R is invertible if and only if it is I with nonzero determinant. This shows 7). [Actually, row-reducing and keeping track of how the determinant is affected — usually not at all — is often the most effort-effective way to find determinants; of course we needn't go all the way to the identity, but can stop once we have a triangular matrix.]

For 8), we again consider row-reducing the matrix. First consider the case when A is not invertible. Then AB cannot be invertible either — if $(AB)C = I$, then $A(BC) = I$. In this case, by the last bit, both $det(A)$ and $det(AB)$ are 0.

If A is invertible, then there are elementary matrices $E_1, \ldots, E_k$ such that $A = E_1 \cdots E_k$. Then $AB = E_1 \cdots E_k B$.

We note that $det(EC) = det(E)det(C)$ for any matrix C and any elementary matrix E. This is immediate from parts 2), 4) and 6) above, once we note that the determinant of an elementary matrix corresponding to a row-switch is -1 (as it can itself be obtained from the identity by a row-switch), that the determinant of a diagonal matrix with all 1's except α in one spot is α (which follows from part 1)), and the determinant of the elementary matrix for $R_j \mapsto R_j + \alpha R_k$ is 1 (which also follows from part 1)).

Now using this repeatedly $det(AB) = det(E_1)det(E_2 \cdots E_k B) = \ldots = det(E_1) \cdots det(E_k)det(B)$ and $det(A) = det(E_1)det(E_2 \cdots E_k) = \ldots = det(E_1) \cdots det(E_k)$. This does it.

9) is now immediate, since $1 = det(I) = det(A^{-1}A) = det(A^{-1})det(A)$. 10) is, too, but it's important enough to warrant separate mention.

11) is also easy if we go through elementary matrices; first, either A and A^T are both noninvertible and both have determinant 0, or neither is. In this case, $A = E_1 \cdots E_k$ where each E_j is elementary, and $A^T = E_k^T \cdots E_1^T$. It's direct that $det(E) = det(E^T)$ in case E is elementary.

10) has the corollary that, if $T : V \longrightarrow V$ is any operator on a finite-dimensional space, it makes sense to talk about $det(T)$. Just let it be $det([T]_B)$ for any basis B of V. If C is another basis, then $[T]_C = P^{-1}[T]_B P$ for some invertible P and therefore the value is independent of the choice of basis.

Now we let $A_{j,k}$ be the $(n-1) \times (n-1)$ submatrix of A obtained by deleting the jth row and kth column; we leave the order of elements otherwise the same. (So A is $n \times n$ and $1 \le j, k \le n$.) It is traditional to leave this rather imprecise definition in this form, and who am I to flout tradition (at least in this case)? Then the *j, k-cofactor* of A is $(-1)^{j+k}det(A_{j,k})$. One of the standard ways to define the determinant is to pick some row and/or column (usually either the first row or first column) and use the following

PROPOSITION 2. Let A be an $n \times n$ matrix and $1 \le j \le n$. Then $det(A) = \sum_{k=1}^{n} a_{j,k} C_{j,k}$, where $C_{j,k}$ is the j, k-cofactor. Also, $det(A) = \sum_{k=1} a_{k,j} C_{k,j}$.

The first formula is the cofactor expansion along the jth row, the second along the jth column.

Proof: We prove the row version, for the case $j = 1$, first. Let A_k be the matrix that looks just like A except that the first row has zeroes everywhere but the kth place, where it is $a_{1,k}$. Then by additivity along the first row, $det(A) = det(A_1) + \cdots + det(A_n)$. It is enough to show that $det(A_k) = (-1)^{1+k} a_{1,k} det(A_{1,k})$ for each k.

If $k = 1$, then A_1 has $a_{1,1}$ in the top left corner, all zeroes for the rest of the top row, and $A_{1,1}$ underneath all those zeroes. If σ is a permutation such that $\sigma(1) \neq 1$, then the product corresponding to σ is zero (as $a_{1,\sigma(1)} = 0$). The surviving permutations all fix 1, and each of these can be identified with a permutation of $\{1, \dots, n-1\}$. (Technically, if $\sigma \in S_n$ fixes 1, then we associate $\hat{\sigma} \in S_{n-1}$ where $\hat{\sigma}(\ell) = \sigma(\ell + 1)$; clearly σ and $\hat{\sigma}$ have the same sign.)

Each product in the expansion of $det(A_{1,1})$ becomes, upon multiplying by $a_{1,1}$, a product in the expansion of A_1 with the same sign, so indeed $det(A_1) = a_{1,1}det(A_{1,1})$. (This is a little loose, but I hope the idea is clear.)

Now it follows from parts 2) and 11) in our main theorem that switching two columns in a matrix switches the sign of the determinant. If we do this to A_k $k-1$ times we produce a matrix with $a_{1,k}$ in the top left corner, and $A_{1,k}$ under the line of zeroes following $a_{1,k}$. (We move the column headed by $a_{1,k}$ one place to the left each time.) By the last paragraphs, this matrix has determinant $a_{1,k}det(A_{1,k})$ and so A_k has determinant $(-1)^{k-1}a_{1,k}det(A_{1,k})$ as required. (Of course $(-1)^{k-1} = (-1)^{k+1}$.)

For the general case for rows, notice first that $j-1$ row switches (of adjacent rows) put the jth row on top, leaving the order of the rest alone. Call the resulting matrix B. $det(B) = (-1)^{j-1}det(A)$. By the last case, $det(B) = \sum_{k=1}^{n}(-1)^{1+k}b_{1,k}det(B_{1,k})$. But $b_{1,k} = a_{j,k}$ and $B_{1,k} = A_{j,k}$, so we finish by multiplying by $(-1)^{j-1}$.

Finally, the column formula follows by taking transposes of everything.

Another result involving the determinant requires the notion of the *classical adjoint* $adj(A)$ of a square matrix A. It is defined as the matrix whose j,k-entry is $(-1)^{j+k}det(A_{k,j})$, the k,j cofactor. (Note the reversal of indices.)

PROPOSITION 3. *$Aadj(A) = adj(A)A = det(A)I$. Thus, if $det(A) \neq 0$, $A^{-1} = \frac{1}{det(A)}adj(A)$.*

To prove that the product $Aadj(A) = det(A)I$, we first consider the cofactor expansion for $det(A)$ along the jth row. It is $a_{j,1}(-1)^{j+1}det(A_{j,1}) + \cdots + a_{j,n}(-1)^{j+n}det(A_{j,n})$. The elements $a_{j,1}, \dots, a_{j,n}$ are of course the entries of the jth row of A; but the elements $(-1)^{j+1}det(A_{j,1}), \dots, (-1)^{j+n}det(A_{j,n})$ are the entries of the jth column of $adj(A)$. Thus this cofactor expansion gives the j,j-entry of $Aadj(A)$. All the diagonal entries of $Aadj(A)$ are $det(A)$.

Now suppose $j \neq k$ and consider the matrix M which looks like A except that both its jth row and its kth row are the jth row of A. As M has two identical rows, $det(M) = 0$; also $M_{k,\ell} = A_{k,\ell}$ for each ℓ. Expanding $det(M)$ along the kth row gives $0 = det(M) = m_{k,1}(-1)^{k+1}det(M_{k,1}) + \cdots + m_{k,n}(-1)^{k+n}det(M_{k,n}) = a_{j,1}(-1)^{k+1}det(A_{k,1}) + \cdots + a_{j,n}(-1)^{k+n}det(A_{k,n})$. This last sum is the j,k-entry of $Aadj(A)$, which is thus 0. This finishes the proof that $Aadj(A) = det(A)I$.

The proof that $adj(A)A = det(A)I$ is similar, using column expansions, and anyway only necessary if $det(A) = 0$. The inverse formula in case $det(A) \neq 0$ follows instantly.

This formula is not very efficient for actual calculations, but it does have a useful corollary. Note that each cofactor $(-1)^{j+k}det(A_{j,k})$ is ± a sum of products of elements from A, so if the entries of A are integers, then so are

those of $adj(A)$. Thus,
COROLLARY. Suppose that A is an $n \times n$ matrix with integer entries. Then $det(A) = \pm 1$ implies that A^{-1} has integer entries.

The converse of this is also true; why?

I also include the slick (Bourbaki) proof of proposition 1. I don't know which one I like better. Fix any distinct real numbers $x_1, \ldots, x_n$. For $\sigma \in S_n$, we let

$$sgn(\sigma) = \frac{\prod_{1 \leq j < k \leq n}(x_{\sigma(j)} - x_{\sigma(k)})}{\prod_{1 \leq j < k \leq n}(x_j - x_k)}.$$

Note that the factors in the numerator are the same as those in the denominator, except possibly in the opposite order; $sgn(\sigma) = \pm 1$.

Then we verify that the map $sgn : S_n \longrightarrow \{1, -1\}$ is a group homomorphism, where the operation on $\{1, -1\}$ is ordinary multiplication. That is, $sgn(\pi\sigma) = sgn(\pi)sgn(\sigma)$. This is because

$$sgn(\pi\sigma) = \frac{\prod_{1 \leq j < k \leq n}(x_{\pi\sigma(j)} - x_{\pi\sigma(k)})}{\prod_{1 \leq j < k \leq n}(x_j - x_k)} =$$

$$\frac{\prod_{1 \leq j < k \leq n}(x_{\sigma(j)} - x_{\sigma(k)})}{\prod_{1 \leq j < k \leq n}(x_j - x_k)} \cdot \frac{\prod_{1 \leq j < k \leq n}(x_{\pi\sigma(j)} - x_{\pi\sigma(k)})}{\prod_{1 \leq j < k \leq n}(x_{\sigma(j)} - x_{\sigma(k)})} =$$

$$sgn(\sigma)\frac{\prod_{1 \leq j < k \leq n}(x_{\pi\sigma(j)} - x_{\pi\sigma(k)})}{\prod_{1 \leq j < k \leq n}(x_{\sigma(j)} - x_{\sigma(k)})}.$$

Now this last quotient is $sgn(\pi)$ because, for any $j \neq k$, suppose that $\sigma(j) = \ell$ and $\sigma(k) = m$. In case $\ell < m$, we have the factor $\frac{x_{\pi(\ell)} - x_{\pi(m)}}{x_\ell - x_m} = \frac{x_{\pi\sigma(j)} - x_{\pi\sigma(k)}}{x_{\sigma(j)} - x_{\sigma(k)}}$ in the product defining $sgn(\pi)$; if $\ell > m$, we have the factor $\frac{x_{\pi(m)} - x_{\pi(\ell)}}{x_m - x_\ell} = \frac{x_{\pi\sigma(j)} - x_{\pi\sigma(k)}}{x_{\sigma(j)} - x_{\sigma(k)}}$. (Note the reversal of – um – signs.)

Next notice that $sgn \begin{pmatrix} 1 & 2 \end{pmatrix} = -1$, since in the numerator the factors $x_1 - x_j$ and $x_2 - x_j$ get switched by the transposition for $j > 2$, $x_j - x_k$ gets unmoved if both j and k are bigger than 2, but $x_1 - x_2$ gets replaced by $x_2 - x_1$.

Now if $a \neq 1, 2$ we have that $\begin{pmatrix} 1 & a \end{pmatrix} = \begin{pmatrix} 2 & a \end{pmatrix}\begin{pmatrix} 1 & 2 \end{pmatrix}\begin{pmatrix} 2 & a \end{pmatrix}$, so $sgn \begin{pmatrix} 1 & a \end{pmatrix} = -1$, too. If $b \neq 1, a$, then $\begin{pmatrix} a & b \end{pmatrix} = \begin{pmatrix} 1 & b \end{pmatrix}\begin{pmatrix} 1 & a \end{pmatrix}\begin{pmatrix} 1 & b \end{pmatrix}$, so $sgn \begin{pmatrix} a & b \end{pmatrix} = -1$. That is, every transposition has sign -1. From this it follows that the odd permutations have sign -1, and the even ones have sign +1.