Introduction You are aware that the matrix $\begin{pmatrix} a & b \\ c & d \end{pmatrix}$ is invertible if and only if $ad - bc \neq 0$. We write

$$\begin{vmatrix} a & b \\ c & d \end{vmatrix} = \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} := ad - bc.$$

You also know (by having done the arithmetic?) that

$$\det \begin{pmatrix} \alpha a & \alpha b \\ c & d \end{pmatrix} = \alpha \det \begin{pmatrix} a & b \\ c & d \end{pmatrix}, \quad \det \begin{pmatrix} c & d \\ a & b \end{pmatrix} = -\det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and that} \quad \det I = \det \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} = 1.$$

It is also true (and you should do the arithmetic) that

$$\det \begin{pmatrix} a_1 + a_2 & b_1 + b_2 \\ c & d \end{pmatrix} = \det \begin{pmatrix} a_1 & b_1 \\ c & d \end{pmatrix} + \det \begin{pmatrix} a_2 & b_2 \\ c & d \end{pmatrix}$$

and (by switching rows twice) that

$$\det \begin{pmatrix} a & b \\ \alpha c & \alpha d \end{pmatrix} = \alpha \det \begin{pmatrix} a & b \\ c & d \end{pmatrix} \quad \text{and} \quad \det \begin{pmatrix} a & b \\ c_1 + c_2 & d_1 + d_2 \end{pmatrix} = \det \begin{pmatrix} a & b \\ c_1 & d_1 \end{pmatrix} + \det \begin{pmatrix} a & b \\ c_2 & d_2 \end{pmatrix}.$$

In words, the determinant of a 2×2 matrix is a function of its rows that is a linear function of each each row (the other row being held constant), that changes sign when the rows are switched, and that takes the value 1 when the matrix is I_2. Linearity in each row is called "multilinearity" ("bilinearity" in 2 dimensions). The sign–change property is called "alternating." Thus the determinant of a 2×2 matrix is an "alternating multilinear function" of the 2 rows of a matrix and the determinant of I_2 is 1.

Let us turn this around. Suppose the we are given a function $F(h_1, h_2)$ defined for pairs of *row* vectors $h_1 \in (\mathbb{R}^2)^T$ and $h_2 \in (\mathbb{R}^2)^T$ and we are given that $F(h_1, h_2)$ is alternating and bilinear, and $F(e_1^T, e_2^T) = 1$. We can think of this as a function $F(V)$ defined on (the rows of) 2×2 matrices $V \in \mathbb{R}^{2\times 2}$. If we let $V = I_2$, $F(V) = F(I) = F(e_1^T, e_2^T) = 1$.

Digression: This creates a notation problem! So far we have thought of an $m \times n$ matrix V as a row of m columns, each column being in $\mathbb{R}^n$: $V = (v_1 \cdots v_n)$. We can also think of V as a column of rows: $V = (h_1, \ldots, h_m)$. For the sake of "classical" determinants it is customary to start out thinking of the determinant as a function of the rows of a matrix. We will stick to the notation $V = (v_{ij})$, so in terms of the entries v_{ij} the rows become $h_i = (v_{i1} \cdots v_{in})$, $1 \leq i \leq n$. We could also write $e_i^T V$ for h_i and we will eventually do that. For now the "v" for "vertical" and "h" for "horizontal" notation will be easier to take visually.

These properties: alternation, multilinearity, $F(I) = 1$ "determine" everything about determinants of $n \times n$ matrices (determinants are defined only for square matrices). Determinants are important, mostly for theoretical reasons, but they have properties that are useful for calculations. Determinants also arise in testing whether a symmetric matrix is positive definite, negative definite or neither. We now return to the 2×2 context. We were studying an alternating bilinear function F such that $F(I) = 1$.

Let us evaluate four special cases: $F(e_1^T, e_2^T) = 1$ is given, $F(e_2^T, e_1^T) = -F(e_1^T, e_2^T) = -1$. When we switch e_1^T and e_1^T, $F(e_1^T, e_1^T) = -F(e_1^T, e_1^T)$ so $F(e_1^T, e_1^T) = 0$. For similar reasons, $F(e_2^T, e_2^T) = 0$.

With $h_1 = (\, v_{11} \quad v_{21} \,)$ and $h_2 = (\, v_{21} \quad v_{22} \,)$ we have $h_1 = v_{11}e_1^T + v_{12}e_2^T$ and $h_2 = v_{12}e_1^T + v_{22}e_2^T$ so

$$\begin{aligned} F(h_1, h_2) &= F(v_{11}e_1^T + v_{12}e_2^T, v_{21}e_1^T + v_{22}e_2^T) \\ &= v_{11}F(e_1^T, v_{21}e_1^T + v_{22}e_2^T) + v_{12}F(e_2^T, v_{21}e_1^T + v_{22}e_2^T) \\ &= v_{11}v_{21}F(e_1^T, e_1^T) + v_{11}v_{22}F(e_1^T, e_2^T) \\ &\quad + v_{12}v_{21}F(e_2^T, e_1^T) + v_{12}v_{22}F(e_2^T, e_2^T) \\ &= v_{11}v_{21}0 + v_{11}v_{22} \cdot 1 \\ &\quad + v_{12}v_{21}(-1) + v_{12}v_{22}0 \\ &= v_{11}v_{22} - v_{12}v_{21}. \end{aligned} \tag{1}$$

Thus *every* bilinear alternating function on $\mathbb{R}^{2\times 2}$ that is one when evaluated at I_2 *must* be the determinant you know about. The calculation we just did can be carried out in exactly the same way for a trilinear alternating function $F(V)$ on $\mathbb{R}^{3\times 3}$ that satisfies $F(I_3) = 1$. Instead of the four terms on the right after the middle equal sign above, we would get 27 terms involving $F(e_i^T, e_j^T, e_k^T)$, each subscript being one of 1, 2 and 3. By now we know that if i j and k are *not* all different, $F(e_i^T, e_j^T, e_k^T) = 0$. To pick *all–different* i j and k we choose i first – three possibilities. We pick j next – two possibilities because i has been chosen. Once i and j are chosen, there is no choice for k: it is *the* remaining one. This gives just 6 ways of choosing different i j and k. These six choices are 123, 132, 231, 213, 312 and 321. Each is called a *permutation* of the "letters" 1 2 and 3, that is, an "arrangement" of them. Permutations are relevant to determinants because they appear as the subscripts in the terms $F(e_i^T, e_j^T, e_k^T)$ in the three–variable version of (1) and we have to figure out whether $F(e_i^T, e_j^T, e_k^T) = \pm 1$ (assuming i j and k are all different). So, in $F(e_3^T, e_1^T, e_2^T)$ we make row switches, chosen to put e_3^T, e_1^T, e_2^T in their "natural" order, e_1^T, e_2^T, e_3^T. Here is one way to do it:

$$F(e_3^T, e_1^T, e_2^T) = -F(e_1^T, e_3^T, e_2^T) = -(-F(e_1^T, e_2^T, e_3^T)) = -(-1) = 1.$$

That way took two switches. We could have done it this way:

$$F(e_3^T, e_1^T, e_2^T) = -F(e_3^T, e_2^T, e_1^T) = -(-F(e_2^T, e_3^T, e_1^T)) = -(-(-F(e_1^T, e_3^T, e_2^T))) = -(-(-(-F(e_1^T, e_2^T, e_3^T)))) = 1$$

which took 4 switches. It is a Theorem that no matter how the switching is done, the number of switches required to return a given permutation on n letters to natural order is always even or always odd. Thus if π is a permutation, and it takes k switches to return π to natural order, we define the "sign" of π to be $(-1)^k =: \mathrm{sgn}\,(\pi)$. This is an unambiguous definition because k is always even or always odd. Thus $\mathrm{sgn}\,(312) = 1$. In fact,

$$F(e_i, e_j, e_k) = \mathrm{sgn}\,(ijk).$$

There is a standard way to put a permutation back in natural order. Starting on the left, switch each number with the ones to its right that it exceeds. Example: In $643512 \to 463512 \to 436512 \to 435612 \to 435162 \to 435126$; it took 5 switches to move 6 past all the numbers it exceeds. If we look at each of these steps, we see that in *each one* it would take 2 switches to move 5 past 1 and 2. Without actually writing the steps down, we can count how many steps the standard method takes: for each number in the permutation, count how many numbers to its right are smaller than it is, add them up (keep a running total), and raise -1 to that power. It's also possible to just count "minus, plus, minus, ..." but that seldom works for me. In our example 643512 this procedure gives $5+3+2+2+0+0 = 12$ and $(-1)^{12} = 1$, so $\mathrm{sgn}\,(643512) = 1$. To actually put the numbers back in order in these numbers of steps, move the numbers into place, starting with the biggest one, then next biggest, and so on. This will be done for this example in more detail later.

It will be useful to write down a useful notational device for the case of 3×3 and bigger matrices, that of "summing over an index set." We begin by letting $S_3 := \{123,\ 132,\ 231,\ 213,\ 312,\ 321\}$. Since S_3 is a finite set, the order in which we add real numbers does not matter. We denote the "typical" element of S_3 by π. That is, $\pi = \pi_1\pi_2\pi_3$ stands for any of 123, 132, 231, 213, 312 and 321, which one being unspecified. Then, given numbers a_π, one for each π in S_3, we write $\sum_{\pi\in S_3} a_\pi$ to mean $a_{123} + a_{132} + a_{231} + a_{213} + a_{312} + a_{321}$, or $a_{231} + a_{321} + a_{123} + a_{312} + a_{213} + a_{132}$, or any other of the $720 = 6!$ ways of adding the six numbers a_π.

When we apply this to the trilinear alternating function $F(h_1, h_2, h_3)$ such that $F(e_1^T, e_2^T, e_3^T) = 1$, we first write $h_i = v_{i1}e_1^T + v_{i2}e_2^T + v_{i3}e_3^T$, $1 \le i \le 3$, then substitute these into $F(h_1, h_2, h_3)$ and use trilinearity and the fact that all but six of the resulting terms will disappear to get

$$F(V) = F(v_{ij}) = F(h_1, h_2, h_3) = \sum_{\pi\in S_3} v_{1\pi_1}v_{2\pi_2}v_{3\pi_3}F(e_{\pi_1}^T, e_{\pi_2}^T, e_{\pi_3}^T).$$

We have seen that $F(e_{\pi_1}^T, e_{\pi_2}^T, e_{\pi_3}^T) = \mathrm{sgn}(\pi)$, so we can rewrite the previous sum as

$$F(h_1, h_2, h_3) = \sum_{\pi\in S_3} v_{1\pi_1}v_{2\pi_2}v_{3\pi_3}\mathrm{sgn}(\pi) = \sum_{\pi\in S_3} \mathrm{sgn}(\pi)v_{1\pi_1}v_{2\pi_2}v_{3\pi_3}.$$

When written out in full detail, this would be

$$\begin{aligned}F(h_1,\,h_2,\,h_3) &= \sum_{\pi\in S_3} \operatorname{sgn}(\pi) v_{1\pi_1}v_{2\pi_2}v_{3\pi_3}\\ &= \operatorname{sgn}(123)v_{11}v_{22}v_{33} + \operatorname{sgn}(132)v_{11}v_{23}v_{32} + \operatorname{sgn}(231)v_{12}v_{23}v_{31}\\ &\quad + \operatorname{sgn}(213)v_{12}v_{21}v_{33} + \operatorname{sgn}(312)v_{13}v_{21}v_{32} + \operatorname{sgn}(321)v_{13}v_{22}v_{31}\\ &= v_{11}v_{22}v_{33} - v_{11}v_{23}v_{32} + v_{12}v_{23}v_{31} - v_{12}v_{21}v_{33} + v_{13}v_{21}v_{32} - v_{13}v_{22}v_{31},\end{aligned}$$

which you probably already know. Once again, we see that a trilinear alternating function $F(h_1,\,h_2,\,h_3)$ such that $F(e_1^T,\,e_2^T,\,e_3^T) = 1$ is uniquely determined. We have found a formula for F but we have not shown that our formula gives us a trilinear alternating function that assigns the value 1 to I_3. Showing that $F(I_3) = 1$ is easy from the formula $F(h_1,\,h_2,\,h_3) = v_{11}v_{22}v_{33} - v_{11}v_{23}v_{32} + v_{12}v_{23}v_{31} - v_{12}v_{21}v_{33} + v_{13}v_{21}v_{32} - v_{13}v_{22}v_{31}$ because all the off–diagonal terms are zero so we get $F(h_1,\,h_2,\,h_3) = v_{11}v_{22}v_{33} = 1$. It takes a little more eyestrain to verify, from the formula, that the effect of switching two rows is to change the sign of the original. For one thing, we need to remember that the *first* subscripts are row numbers! For example,

$$F(h_1,\,h_2,\,h_3) = v_{11}v_{22}v_{33} - v_{11}v_{23}v_{32} + v_{12}v_{23}v_{31} - v_{12}v_{21}v_{33} + v_{13}v_{21}v_{32} - v_{13}v_{22}v_{31} \tag{2}$$

and

$$F(h_3,\,h_2,\,h_1) = v_{13}v_{22}v_{31} - v_{13}v_{21}v_{32} + v_{12}v_{21}v_{33} - v_{12}v_{23}v_{31} + v_{11}v_{23}v_{32} - v_{11}v_{22}v_{33}.$$

In the second formula, the terms match as follows with terms in the first formula (check that there is a sign change in each case): $1 \to 6$, $2 \to 5$, $3 \to 4$, $4 \to 3$, $5 \to 2$, $6 \to 1$. Thus $F(h_3,\,h_2,\,h_1) = -F(h_1,\,h_2,\,h_3)$. You should check the other two cases! Switching will be used to prove trilinearity!

Checking trilinearity can be done by checking linearity in the first variable only! We can check linearity in any other variable by switching to the first one, applying linearity in the first variable, then switching back. The two switches cancel the negatives. Putting a constant with any variable, say the second, gives

$$\begin{aligned}F(h_1,\,ch_2,\,h_3) &= v_{11}cv_{22}v_{33} - v_{11}cv_{23}v_{32} + v_{12}cv_{23}v_{31} - v_{12}cv_{21}v_{33} + v_{13}cv_{21}v_{32} - v_{13}cv_{22}v_{31}\\ &= c(v_{13}v_{22}v_{31} - v_{13}v_{21}v_{32} + v_{12}v_{21}v_{33} - v_{12}v_{23}v_{31} + v_{11}v_{23}v_{32} - v_{11}v_{22}v_{33})\\ &= cF(h_1,\,h_2,\,h_3).\end{aligned}$$

Then $F(ch_1,\,h_2,\,h_3) = -F(h_2,\,ch_1,\,h_3) = -cF(h_2,\,h_1,\,h_3) = -c(-F(h_1,\,h_2,\,h_3)) = cF(h_1,\,h_2,\,h_3)$. The other case is done similarly.

To finish checking linearity we must show that for all row vectors h_1, $\tilde{h}_1$, h_2, h_3,

$$F(h_1 + \tilde{h}_1,\,h_2,\,h_3) = F(h_1,\,h_2,\,h_3) + F(\tilde{h}_1,\,h_2,\,h_3).$$

This will be left to you! The idea is to replace every v_{1j} in (2) (the first factor in each triple product there) by $v_{1j} + \tilde{v}_{1j}$, do the necessary arithmetic, and regroup.

The general case We let S_n denote the set of all $n!$ permutations $\pi = \pi_1\pi_2\cdots\pi_n$. In using the multilinearity of F the main idea is that all the subscripts in $F(e_{j_1}^T,\,e_{j_2}^T,\,\ldots,\,e_{j_n}^T)$ have to be different (otherwise $F(e_{j_1}^T,\,e_{j_2}^T,\,\ldots,\,e_{j_n}^T) = 0$). But then **(important:)** each of the numbers $1,\,2,\,\ldots,\,n$ appears once and only once in the list $j_1,\,j_2,\,\ldots,\,j_n$. The terms that survive are then those whose subscripts form a permutation $\pi \in S_n$. Here is the Important Theorem.

Alternating Multilinear Function Theorem: *The function $F : \mathbb{R}^{n\times n} \to \mathbb{R}$ given by the formula*

$$F(V) = F(v_{ij}) = F(h_1,\,h_2,\,\ldots,\,h_n) = F(e_1^TV,\,e_2^TV,\,\ldots,\,e_n^TV) = \sum_{\pi\in S_n} \operatorname{sgn}(\pi) v_{1\pi_1}v_{2\pi_2}\cdots v_{n\pi_n} \tag{2.5}$$

is an alternating multilinear function and $F(I_n) = 1$. Moreover, every alternating multilinear function $F : \mathbb{R}^{n\times n} \to \mathbb{R}$ such that $F(I_n) = 1$ is given by the same formula.

This is an important theorem whose proof, not given here, is somewhat similar to what we have done. Proving that the present formula gives an alternating function is the hard part in the general case, because we don't want to go thru the $n!$ terms in the sums! So we invent a way to "multiply" permutations, and we call the permutation that switches k and ℓ $\tau_{k\ell}$ (for *transposition,* a more formal version of "switch"). We show that $\text{sgn}(\pi\tau_{k\ell}) = \text{sgn}(\pi)\text{sgn}(\tau_{k\ell})$. Then we use the facts that $\text{sgn}(\tau_{k\ell}) = -1$ and that $\pi\tau_{k\ell}$ runs thru all of S_n as π runs thru S_n, in the formulas we get in the Theorem before and after we switch rows (see "Additional Material," at the end of this tour).

Definition of the determinant of a matrix *If* $A = (a_{ij})$ *is an* $n \times n$ *matrix, we define*

(3)
$$\det(A) := \sum_{\pi \in S_n} sgn(\pi) a_{1\pi_1} a_{2\pi_2} \cdots a_{n\pi_n} \quad (= F(e_1^T A,\, e_2^T A,\, \ldots,\, e_n^T A)).$$

We have defined $\det(A)$ in terms of the rows of A. One thing we need to do to get flexibility in applying determinants is to view the determinant as a function of the columns as well. Other properties are also needed. We turn to developing them next. The approach is to list one Fact after another, and apply them to develop the properties and Theorems that we need. Some of the Facts will be rather "technical." Each will be preceded by a "bullet:" •.

Properties of the determinants of matrices

• *If a matrix* A *has a zero row then* $\det(A) = 0$. Let us suppose that row k is zero. Then in the formula (3) each term contains a factor $a_{ik} = 0$ for some i and therefore every term in the defining sum is zero.

• *If a matrix* A *has only zero entries below the main diagonal (or only zero entries above the main diagonal) then* $\det(A)$ *is the product of the entries on the main diagonal.* The formula (3) for the determinant has one term $a_{11}a_{22}\cdots a_{nn}$, corresponding to the permutation π with $\pi_i = i$ for all i. Each of the other terms has the factor $a_{1\pi_1}a_{2\pi_2}\cdots a_{n\pi_n}$ in which $i \neq \pi_i$ for some i. For such π at least one factor $a_{i\pi_i}$ must lie above and another one below the main diagonal (unless $\pi = i$ for all i), so in all these terms the product of the entries along the main diagonal is zero if A is lower triangular or upper triangular.

Why is it true that at least one factor $a_{i\pi_i}$ must lie above and another one below the main diagonal (unless $\pi_k = k$ for all k)? Let's suppose i is such that $\pi_i \neq i$. Then for some $j \neq i$ we would also have to have $\pi_j \neq j$. Otherwise, $\pi_j = j$ for all $j \neq i$. Since i must appear somewhere as a second subscript (otherwise the second subscripts would not all be different), there has to exist ℓ such that $\pi_\ell = i$, and $\ell \neq i$ because $\pi_i \neq i$. But because $\ell \neq i$, $\pi_\ell = \ell$. Therefore $\ell = \pi_\ell = i \neq \ell$. This is a contradiction, so our asumption that $\pi_j = j$ for all $j \neq i$ is wrong and there is some $j \neq i$ such that $\pi_j \neq j$. Then $\pi_i > i$ or $i > \pi_i$, and similarly $\pi_j > j$ or $j > \pi_j$. If $\pi_i > i$ then $a_{i\pi_i}$ lies above the diagonal. Thus we want $\pi_i > i$ and $j > \pi_j$ or vice versa. We cannot have $\pi_k \geq k$ for all k or $\pi_k \leq k$ for all k, unless $\pi_k = k$ for all k. This is true because $\pi_1 + \cdots \pi_n = 1 + \cdots + n$.

• *If two rows of a matrix* A *are proportional then* $\det(A) = 0$. Let us suppose that $e_1^T A = c e_2^T A$, i.e. that row 1 is a multiple of row 2. This is a speecial case but we can easily convert any other case to this one. Then

$$\begin{aligned}\det(A) &= F(e_1^T A,\, e_2^T A,\, \ldots,\, e_n^T A) = F(c e_2^T A,\, e_2^T A,\, \ldots,\, e_n^T A)\\ &= cF(e_2^T A,\, e_2^T A,\, \ldots,\, e_n^T A) = -cF(e_2^T A,\, e_2^T A,\, \ldots,\, e_n^T A) \ \ \text{(switch rows 1 and 2)}\\ &= 0 \ \ \text{because the only real number equal to its own negative is } 0.\end{aligned}$$

In particular, if two rows of A are equal then $\det(A) = 0$. Can you say how to deal with the general case, using this one and row switches?

(3.5) • *If we start with a matrix* A *and create a matrix* B *by multiplying a row of* A *by a constant* K, *then* $\det(B) = K\det(A)$. Proof is trivial. Can you do it? If $K = 0$ then $\det(B) = 0$. If $K \neq 0$ then the matrix $D_i(K) := I_n - e_i e_i^T + K e_i e_i^T = I_n + (K-1)e_i e_i^T$ has ones on the diagonal except in row i (and column i), where it has K, and it has zeroes elsewhere. Then $B = D_i(K)A$. We have proved that

$$\det(D_i(K)A) = K\det(A) = \det(D_i(K))\det(A).$$

Can you show that $\det(D_i(K)) = K$? We did this Fact because $D_i(K)$ is an ERO matrix (Elementary Row Operation matrix). For later use, we notice that when $K \neq 0$, $1/K = \det(D_i(1/K)) = \det(D_i(K)^{-1}) = \det(D_i(K))^{-1}$.

• *If we start with a matrix A and create a matrix B by switching rows k and ℓ of A, where $k \neq \ell$, then* $\det(B) = -\det(A)$. Proof is trivial. Can you do it? We construct the matrix $P_{k\ell} := I_n - e_k e_k^T - e_\ell e_\ell^T + e_k e_\ell^T + e_\ell e_k^T$ (to understand this quickly sketch P_{24} in the 5×5 case). Then (and you should be able to show why) $B = P_{k\ell} A$, and we have proved (with your help) that

$$\det(P_{k\ell} A) = -\det(A) = \det(P_{k\ell}) \det(A).$$

This is another ERO matrix determinant formula. For later use, we notice that $P_{k\ell}^{-1} = P_{k\ell}$ so that $-1 = \det(P_{k\ell}) = \det(P_{k\ell}^{-1}) = \det(P_{k\ell})^{-1}$.

• *If we start with a matrix A and create a matrix B by adding to row k of A a multiple of row ℓ of A, where $k \neq \ell$, then* $\det(B) = \det(A)$. To verify this we first suppose that $k = 1$ and $\ell = 2$. Then

$$\begin{aligned}\det(B) &= F(e_1^T A + K e_2^T A,\, e_2^T A,\, \ldots,\, e_n^T A)\\ &= F(e_1^T A,\, e_2^T A,\, \ldots,\, e_n^T A) + F(K e_2^T A,\, e_2^T A,\, \ldots,\, e_n^T A) \quad \text{(by linearity)}\\ &= F(e_1^T A,\, e_2^T A,\, \ldots,\, e_n^T A) + 0 \quad \text{because the second term has two proportional rows.}\\ &= \det(A).\end{aligned}$$

You should complete the proof by switching rows k and 1 of B and also switch rows ℓ and 2 of B, apply the case we just did, then reverse the switches you did (if you wish). We can create B using an ERO matrix. We define $S_{k\ell}(K) := I + K e_k e_\ell^T$. Then $B = S_{k\ell}(K) A$, so we have proved that

$$\det(S_{k\ell}(K) A) = \det(A) = \det(S_{k\ell}(K)) \det(A).$$

In particular, $\det(S_{k\ell}(K)) = 1$. Can you show it directly? What does $S_{k\ell}(K)$ look like? For later use, we notice that $S_{k\ell}(K)^{-1} = S_{k\ell}(-K)$ so that $\det(S_{k\ell}(K)^{-1}) = \det(S_{k\ell}(-K)) = 1 = \det(S_{k\ell}(K))^{-1}$.

• *If we start with a matrix A and ERO matrices $E_1, \ldots, E_N$ and create the matrix $B := E_N \cdots E_1 A$, then*

$$\det(B) = \left(\prod_{i=1}^{N} \det(E_i) \right) \det(A).$$

In all of our determinant formulas involving EROs E we had $\det(EA) = \det(E)\det(A)$. Thus by associativity of matrix multiplication we pull out one "$\det(E)$" at a time and get

$$\begin{aligned}\det(B) = \det(E_N \cdots E_1 A) &= \det(E_N) \det(E_{N-1} \cdots E_1 A)\\ &= \det(E_N) \det(E_{N-1}) \det(E_{N-2} \cdots E_1 A)\\ &= \cdots =\\ &= \det(E_N) \det(E_{N-1}) \det(E_{N-2}) \cdots \det(E_2) \det(E_1 A)\\ &= \left(\prod_{i=1}^{N} \det(E_{N+1-i}) \right) \det(A).\end{aligned} \tag{4}$$

• **Toward the product formula** It is time for a **clever trick:** in (4), we choose $A = I$. Now (4) reads

$$\det(B) = \det(E_N \cdots E_1 I) = \det(E_N \cdots E_1) = \left(\prod_{i=1}^{N} \det(E_{N+1-i}) \right) \det(I) = \prod_{i=1}^{N} \det(E_{N+1-i}). \tag{5}$$

This is important enough to be called a Theorem.

(5.5) **Theorem** *The determinant of a product of finitely many ERO matrices $E_1, \ldots, E_N$ is equal to the product of their determinants:*

$$\det(E_1 \cdots E_N) = \prod_{i=1}^{N} \det(E_i).$$

Now that we have the formula (4), we will apply it to the formula $A = (E_N \cdots E_1)^{-1}B$, knowing as we now do that EROs are invertible, that products of invertible matrices are invertible and that the inverse of a product is the procuct of the inverses *in the reverse order*. First we rewrite our formula:

$$A = (E_N \cdots E_1)^{-1}B = \left(\prod_{i=1}^{N} E_i^{-1}\right) B.$$

We know now that the inverse of an ERO matrix is an ERO matrix too, so by (4) and (5)

$$\begin{aligned} \det(A) &= \left(\prod_{i=1}^{N} \det(E_i^{-1})\right)\det(B) = \det\left(\prod_{i=1}^{N} E_i^{-1}\right)\det(B) \\ &= \det\left(\prod_{i=1}^{N} E_{N+1-i}\right)^{-1}\det(B) \\ \text{and}& \\ \det(B) &= \left(\prod_{i=1}^{N} \det(E_{N+1-i})\right)\det(A) = \det\left(\prod_{i=1}^{N} E_{N+1-i}\right)\det(A). \end{aligned} \tag{6}$$

Gauss–Jordan elimination is done by multiplying on the left by EROs. We get $(\ A \mid I\ ) \to (\ R \mid M\ )$. Here, all of the matrices are $n \times n$, $M = \prod_{i=1}^{N} E_{N+1-i}$ and $B = R$. The formulas in (6) become (we recall $A = M^{-1}R$):

$$\det(A) = \det(M^{-1}R) = \det(M^{-1})\det(R) \quad \text{and} \quad \det(R) = \det(MA) = \det(M)\det(A). \tag{7}$$

Since R is $n \times n$ and upper triangular, $\det(R)$ is the product of the entries on its main diagonal. This product will be zero unless there is a leading one, on the main diagonal, in each row. This gives us a Fact:

• *If A is an $n \times n$ matrix and $(\ A \mid I\ ) \to (\ R \mid M\ )$ by Gauss–Jordan elimination then $\det(A) = 0$ if and only if $\det(R) = 0$, and $\det(A) \neq 0$ if and only if A is invertible and then $\det(R) = 1$, so that*

$$\det(A) = \det(M^{-1}) = \det\left(\prod_{i=1}^{N} E_i^{-1}\right) = \left(\prod_{i=1}^{N} \det(E_i^{-1})\right) \quad \textit{if } A \textit{ is invertible, and } \det(A) = 0 \textit{ otherwise.}$$

This is proved by examination of (7), because $\det(M)$, being the product of the non-zero determinants of EROs, is non-zero. We also made use of our knowledge of Gauss–Jordan elimination! The next Fact is an "old" one:

(7.5) • *If A is an $n \times n$ matrix then A is invertible if and only if* $\ker(A) = \{0\}$.

Proof: If A is invertible and $x \in \ker(A)$, so that $Ax = 0$, then $x = Ix = (A^{-1}A)x = A^{-1}(Ax) = A^{-1}0 = 0$. That is, $\ker(A) = \{0\}$. If $\ker(A) = \{0\}$ then Gauss–Jordan elimination gives $(\ A \mid I\ ) \to (\ R \mid M\ )$ and $\ker(R) = \ker(A) = \{0\}$ so $R = I$ and therefore $R = I = MR$ which means that $M = A^{-1}$, or, A is invertible.

Another (partly) "old" Fact:
(7.7) • *If A is an $n \times n$ matrix then $\det(A) = 0$ if and only if* $\ker(A) \neq \{0\}$.

Proof: If $\det(A) = 0$, then $R \neq I$, which means that there is some column of R without a leading one. We know that this column contributes a basis element for $\ker(A)$ (see the note "Two rref's are equal iffi their kernels are equal," especially the formulas $(*)$ therein), so $\ker(A) \neq \{0\}$. On the other hand, if $\ker(A) \neq \{0\}$, then $\ker(R) = \ker(A) \neq \{0\}$, so $R \neq I$, so some main–diagonal entry of R is zero, so $\det(R) = 0$, so by (7) $\det(A) = 0$.

Here is another application of (7) and the formulas (6). Though technical, it well deserves the title "Theorem."

(8) • **Theorem:** *If A is an $n \times n$ matrix and $(A \mid I) \to (R \mid M)$ by Gauss–Jordan elimination then A is invertible if and only if A can be expressed as the product of finitely many ERO matrices. In that case, $R = I$, $M = A^{-1}$ and $\det(A)$ is given by*

$$\det(A) = \det(M)^{-1} = \prod_{i=1}^{N} \det(E_i^{-1}). \tag{9}$$

Theorem (8) is just a combined version of (7) and the formulas (6). The next Fact is another version of Theorem (8) and either of its two preceding Facts. It deserves to be called a Theorem.

(10) • **Theorem:** *If A is an $n \times n$ matrix then A is invertible if and only if $\det(A) \neq 0$.*

Proof: If A is invertible, then $\det(A) = \det(M)^{-1} = \prod_{i=1}^{N} \det(E_i^{-1}) \neq 0$. If $\det(A) \neq 0$ then $\ker(A) = \{0\}$ so $\ker(R) = \ker(A) = \{0\}$ and thus $R = I$ so $MA = R = I$, hence A is invertible.

(11) • **Theorem:** *If A and B are $n \times n$ matrices, then AB is invertible if and only if A and B are both invertible.*

Proof: If A and B are both invertible, so is AB (and $(AB)^{-1} = B^{-1}A^{-1}$). This is an "old" Fact! If AB is invertible, then $I = AB(AB)^{-1} = A[B(AB)^{-1}]$ so $B(AB)^{-1} = A^{-1}$. You should "do" the similar argument that shows B is invertible.

(12) • **Theorem:** *If A and B are $n \times n$ matrices, then*

$$\det(AB) = \det(A)\det(B).$$

Proof: If A and B are both invertible, each of them can be expressed as the product of finitely many ERO matrices: $A = \prod_{i=1}^{N} E_i^{-1}$, $B = \prod_{i=1}^{M} \tilde{E}_i^{-1}$. By (9) $\det(A) = \prod_{i=1}^{N} \det(E_i^{-1})$. Again by (9) $\det(B) = \prod_{i=1}^{M} \det(\tilde{E}_i^{-1})$. Therefore

$$\begin{aligned}
\det(A)\det(B) &= \prod_{i=1}^{N} \det(E_i^{-1}) \prod_{i=1}^{M} \det(\tilde{E}_i^{-1}) \\
&= \det\left(\prod_{i=1}^{N} E_i^{-1}\right) \det\left(\prod_{i=1}^{M} \tilde{E}_i^{-1}\right) \quad \text{by (5.5)} \\
&= \det\left(\prod_{i=1}^{N} E_i^{-1} \prod_{i=1}^{M} \tilde{E}_i^{-1}\right) \quad \text{by (5.5)} \\
&= \det(AB).
\end{aligned}$$

If either one of A and B is not invertible, then $\det(A)\det(B) = 0$ since at least one of the determinants is zero, by (7.5) and (7.7). But then AB is not invertible either, by (11), so by (7.7), $\det(AB) = 0$. For these reasons, $\det(A)\det(B) = \det(AB)$ because both sides of this equation are zero. This completes the proof of Theorem (12).

• **A "method" for calculating determinants** Let us illustrate the method with an example:

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}.$$

Then by linearity in the first row, $\det(A) = 1 \cdot \det(e_1^T, h_2, h_3) + 2 \cdot \det(e_1^T, h_2, h_3) + 3 \cdot \det(e_1^T, h_2, h_3)$, where $h_1 = (1 \quad 2 \quad 3)$, $h_1 = (4 \quad 5 \quad 4)$, $h_1 = (3 \quad 2 \quad 1)$. Let's look at this in detail:

$$\det(A) = 1 \cdot \det\begin{pmatrix} 1 & 0 & 0 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix} + 2 \cdot \det\begin{pmatrix} 0 & 1 & 0 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix} + 3 \cdot \det\begin{pmatrix} 0 & 0 & 1 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}.$$

In each one of these determinants we can add multiples of the first row to the other rows without changing the value of the determinant, by (3.5). Then we can "zero out" the column below each 1, so

$$\det(A) = 1\cdot\det\begin{pmatrix}1&0&0\\0&5&4\\0&2&1\end{pmatrix} + 2\cdot\det\begin{pmatrix}0&1&0\\4&0&4\\3&0&1\end{pmatrix} + 3\cdot\det\begin{pmatrix}0&0&1\\4&5&0\\3&2&0\end{pmatrix}.$$

A nice thing about Theorem (12) is that now we can perform *column* operations on matrices by multiplying on the *right* by an ERO matrix – which then becomes an Elementary Column Operation matrix, or ECO. Our example now has three determinants, each with lots of zeroes. We can switch *columns* 1 and 2 in the middle determinant, which creates the matrix $\begin{pmatrix}1&0&0\\0&4&4\\0&3&1\end{pmatrix} = \begin{pmatrix}0&1&0\\4&0&4\\3&0&1\end{pmatrix}P_{12}$, where $P_{12} = \begin{pmatrix}0&1&0\\1&0&0\\0&0&1\end{pmatrix}$, and $\det(P_{12}) = -1$ because P_{12}, now acting as an ECO, is still an ERO that switches rows (when it multiplies on the left).

Thus by Theorem (12) $\det\begin{pmatrix}0&1&0\\4&0&4\\3&0&1\end{pmatrix} = -\det\begin{pmatrix}1&0&0\\0&4&4\\0&3&1\end{pmatrix}$. In the third matrix we will multiply on the right by $P_{23}P_{12}$, whose determinant is $+1$, and we get

$$\det(A) = 1\cdot\det\begin{pmatrix}1&0&0\\0&5&4\\0&2&1\end{pmatrix} - 2\cdot\det\begin{pmatrix}1&0&0\\0&4&4\\0&3&1\end{pmatrix} + 3\cdot\det\begin{pmatrix}1&0&0\\0&4&5\\0&3&2\end{pmatrix}.$$

We can do this to determinants of an size!

Now let's look at the defining formula for the determinant of a matrix B:

$$\det(B) = \sum_{\pi\in S_n}\operatorname{sgn}(\pi)b_{1\pi_1}b_{2\pi_2}\cdots b_{n\pi_n}.$$

Suppose that, in B $b_{11} = 1$ and $b_{i1} = 0 = b_{1j}$ if $i > 1$ and $j > 1$. This is what we have in our exemplary matrices. Then in every term that has $\pi_1 \neq 1$, $\operatorname{sgn}(\pi)b_{1\pi_1}b_{2\pi_2}\cdots b_{n\pi_n} = 0$. The only terms that survive are those that have $\pi_1 = 1$. The set of $\pi\in S_n$ that have $\pi_1 = 1$ is really the set of permutations of $2, 3, \ldots, n$. Therefore,

(13) • **Lemma:** *If B is an $n\times n$ matrix such that $b_{11} = 1$ and $b_{i1} = 0 = b_{1j}$ if $i > 1$ and $j > 1$, then*

$$\det(B) = \det\begin{pmatrix}b_{22}&\cdots&b_{2n}\\\vdots&\ddots&\vdots\\b_{n2}&\cdots&b_{nn}\end{pmatrix} = \det(B_{11}),$$

the determinant of the $(n-1)\times(n-1)$ matrix B_{11}, which is obtained by removing the first row and first column of B.

In our example this gives

$$\det(A) = 1\cdot\det\begin{pmatrix}5&4\\2&1\end{pmatrix} - 2\cdot\det\begin{pmatrix}4&4\\3&1\end{pmatrix} + 3\cdot\det\begin{pmatrix}4&5\\3&2\end{pmatrix} = 1\cdot(-3) - 2\cdot(-8) + 3\cdot(-7) = -8.$$

Now here is the "method" of expanding a determinant along a row. Any row works fine because we can use row and column switches (we have to keep track of how many) and row operations to put a matrix into the form of Lemma (13). To use the method, we define, (in an $n\times n$ matrix A), A_{ij} to be the $(n-1)\times(n-1)$ matrix obtained by removing row i and column j from A, where $1\le i\le n$ and $1\le j\le n$.

• *If A is an $n\times n$ matrix then for $1\le i\le n$,*

(14) $$\det(A) = \sum_{j=1}^{n}(-1)^{i+j}a_{ij}\det(A_{ij}).$$

In our example we had $i = 1$. It is sometimes easier to expand along some row other than the first, for example if a row has many zeroes in it, for then we can skip evaluating the corresponding determinants, $\det(A_{ij})$, which are called *subdeterminants.*

We still have to discuss the determinant of A^T. It will turn out to be equal to the determinant of A. Then we will be able to expand determinants along columns too. The formula is the same as the one in (14) except that the sum runs from $i = 1$ to n:

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}). \tag{15}$$

Another method of evaluating determinants – maybe the best one

• We actually "already know" this method. Given an an $n \times n$ matrix A we can carry out Gauss–Jordan elimination until we arrive at an upper triangular matrix. If any diagonal entry is zero, we can stop because the determunant is then known to be zero. Otherwise, the determinant is the product of the diagonal entries of the upper triangular matrix, *divided by* the product of the non-zero numbers we multiplied rows by, times -1 if we switched rows an odd number of times. This can be read off of (7) if we describe the "partial" Gauss–Jordan procees by $(A \mid I) \to (U \mid K)$, where U is upper triangular. This modified version of (7) becomes

$$\det(U) = \det(KA) = \det(K)\det(A), \quad \text{so that} \quad \det(A) = \det(U)/\det(K). \tag{16}$$

To calculate $\det(K)$ we keep a running product of the determinants of the EROs we use. Row switches can be counted to see if the number of them is odd. Adding a multiple of one row to another does not change anything. If we multiply a row by a non-zero number, we multiply our running product by that.

But we can do more! If we want to we can do column operations too, and keep track the same way, because of Theorem (12).

Let's use this row–or–column Gauss–Jordan elimination method on our example

$$A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}.$$

We can add row 1 to row 3 and then subtract row 3 from row 2 and use (14) with $i = 2$ because of the zeroes:

$$\det(A) = \det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix} = \det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 4 & 4 & 4 \end{pmatrix} = \det\begin{pmatrix} 1 & 2 & 3 \\ 0 & 1 & 0 \\ 4 & 4 & 4 \end{pmatrix} = (-1)^{2+2}\det\begin{pmatrix} 1 & 3 \\ 4 & 4 \end{pmatrix} = -8.$$

Notice that we did not need to "zero out" the rest of column 2. I did that in the example when we first did it to make the pattern stand out better. We can mix in column operations too: subtract column 1 from column 3, add row 3 to row 1, use (15) (column expansion) with $j = 3$:

$$\det(A) = \det\begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix} = \det\begin{pmatrix} 1 & 2 & 2 \\ 4 & 5 & 0 \\ 3 & 2 & -2 \end{pmatrix} = \det\begin{pmatrix} 4 & 4 & 0 \\ 4 & 5 & 0 \\ 3 & 2 & -2 \end{pmatrix} = (-1)^{3+3}(-2)\det\begin{pmatrix} 4 & 4 \\ 4 & 5 \end{pmatrix} = -8.$$

The determinant of the transpose

(16.5) • **Theorem:** *For every* $n \times n$ *matrix* A, $\det(A^T) = \det(A)$.

Proof: To show that $\det(A^T) = \det(A)$ we need to know a little more about permutations. Permutations $\pi \in S_n$ are actually functions defined on $\{1, 2, \ldots, n\}$ that take values in $\{1, 2, \ldots, n\}$:

$$\pi : \{1, 2, \ldots, n\} \to \{1, 2, \ldots, n\}.$$

For reasons I do not know we write π_i or πi instead of $\pi(i)$. Permutations are one–to–one and onto. "One–to–one" means that $\pi_i = \pi_j$ implies $i = j$. "Onto" means that for every $j \in \{1, 2, \ldots, n\}$ there exists $i \in \{1, 2, \ldots, n\}$ such that $\pi_i = j$. If π and τ are in S_n we can define their"product" $\tau\pi$ by applying π first and then τ: $(\tau\pi)_i := \tau_{\pi_i}$. This is just like $g \circ f(x) = g(f(x))$. The permutation $12\cdots n$ is usually called "the identity permutation," is denoted e and $e\pi = \pi e = \pi$ for every π; e acts like multiplying by 1.

Let's recall the formula (3) for the determinant, and apply it to A^T:

$$\det(A^T) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) a_{\pi_1 1} a_{\pi_2 2} \cdots a_{\pi_n n}$$

because to form the transpose we switch subscripts. We want to somehow rearrange the product $a_{\pi_1 1} a_{\pi_2 2} \cdots a_{\pi_n n}$ so that it looks like $a_{1\sigma_1 1} a_{2\sigma_2} \cdots a_{n\sigma_n}$ for some $\sigma \in S_n$. We construct a σ that "inverts" or "undoes" π. Every i, $1 \le i \le n$, is π_j for some j, i.e., $i = \pi_j$ for some j. And the j is unique (because π is one–to–one), so we can write $j = \sigma_i$ for the unique j such that $\pi_j = i$. This procedure defines σ as a function from $\{1, 2, \ldots, n\}$ to $\{1, 2, \ldots, n\}$. Then $\pi_{\sigma_i} = i$ for all i. That is, $\pi\sigma = e$, which implies that σ is one–to–one (if $\sigma_i = \sigma_k$ then $i = \pi_{\sigma_i} = \pi_{\sigma_k} = k$). On the other hand, σ_{π_j} is the uniq1ue k such that $\pi_k = \pi_j$, so $k = j$ (because π is one–to–one). That is, $\sigma_{\pi_j} = j$ for all j, so $\sigma\pi = e$, which implies that σ is onto (because, given j, we pick $i := \pi_j$ and then $\sigma_i = \sigma_{\pi_j} = j$). This procedure defines a permutation in S_n. And σ "undoes" π and vice versa. Then in $a_{\pi_1 1}$, $1 = \sigma_{\pi_1}$ so $a_{\pi_1 1} = a_{\pi_1 \sigma_{\pi_1}}$. When we find where to put $a_{\pi_1 1}$ in the rearranged product so that its first subscript is in order, the second subscript will be σ of it. For example, if $n = 6$ and $\pi = 643512$, then $\pi_1 = 6$ so $1 = \sigma_{\pi_1} = \sigma_6$. Then in the rearranged product, $a_{\pi_1 1} = a_{61} = a_{6\sigma_6}$, and so on. Here is the rest of σ: $\pi_2 = 4$ so $\sigma_4 = 2$; $\pi_3 = 3$ so $\sigma_3 = 3$; $\pi_4 = 5$ so $\sigma_5 = 4$; $\pi_5 = 1$ so $\sigma_1 = 5$; $\pi_6 = 2$ so $\sigma_2 = 6$. That is, $\sigma = 563241$. This means that in this example

$$\operatorname{sgn}(\pi) a_{\pi_1 1} a_{\pi_2 2} \cdots a_{\pi_6 6} = \operatorname{sgn}(\pi) a_{61} a_{42} a_{33} a_{54} a_{15} a_{26} = \operatorname{sgn}(\pi) a_{15} a_{26} a_{33} a_{42} a_{54} a_{61} = \operatorname{sgn}(\pi) a_{1\sigma_1} a_{2\sigma_2} \cdots a_{6\sigma_6}.$$

This equation is true in general, by the same argument, which is only hard because we have to concentrate on how the subscripts of π and σ work. It would be nice if $\operatorname{sgn}(\pi) = \operatorname{sgn}(\sigma)$. This is indeed true. How do we show it? Let's look at the example again, this time writing it like a graph:

$$\pi = \begin{pmatrix} 6 & 4 & 3 & 5 & 1 & 2 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix}. \tag{17}$$

When we go thru the standard switches with smaller numbers to the right to put the top row in order, we'll carry the bottom row along:

$$\begin{pmatrix} 6 & 4 & 3 & 5 & 1 & 2 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix} \to \begin{pmatrix} 4 & 3 & 5 & 1 & 2 & 6 \\ 2 & 3 & 4 & 5 & 6 & 1 \end{pmatrix} \quad \text{in 5 switches with smaller numbers,}$$

$$\begin{pmatrix} 4 & 3 & 5 & 1 & 2 & 6 \\ 2 & 3 & 4 & 5 & 6 & 1 \end{pmatrix} \to \begin{pmatrix} 4 & 3 & 1 & 2 & 5 & 6 \\ 2 & 3 & 5 & 6 & 4 & 1 \end{pmatrix} \quad \text{in 2 switches with smaller numbers,}$$

$$\begin{pmatrix} 4 & 3 & 1 & 2 & 5 & 6 \\ 2 & 3 & 5 & 6 & 4 & 1 \end{pmatrix} \to \begin{pmatrix} 3 & 1 & 2 & 4 & 5 & 6 \\ 3 & 5 & 6 & 2 & 4 & 1 \end{pmatrix} \quad \text{in 3 switches with smaller numbers,}$$

$$\begin{pmatrix} 3 & 1 & 2 & 4 & 5 & 6 \\ 3 & 5 & 6 & 2 & 4 & 1 \end{pmatrix} \to \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 3 & 2 & 4 & 1 \end{pmatrix} \quad \text{in 2 switches with smaller numbers,}$$

for a total of 12 switches, starting with 6, then 5, and so on. Now let's take the last array, and flip it over:

$$\begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 \\ 5 & 6 & 3 & 2 & 4 & 1 \end{pmatrix} \to \begin{pmatrix} 5 & 6 & 3 & 2 & 4 & 1 \\ 1 & 2 & 3 & 4 & 5 & 6 \end{pmatrix} = \sigma.$$

When we once again count switches to put the top row of σ back in order, we'll get $4 + 4 + 2 + 1 + 1 + 0 = 12$. Thus $\operatorname{sgn}(\sigma) = \operatorname{sgn}(\pi)$, so

$$\operatorname{sgn}(\pi) a_{\pi_1 1} a_{\pi_2 2} \cdots a_{\pi_6 6} = \operatorname{sgn}(\sigma) a_{1\sigma_1} a_{2\sigma_2} \cdots a_{6\sigma_6}.$$

There is an easier way to show the equality $\mathrm{sgn}(\sigma) = \mathrm{sgn}(\pi)$! First we build a matrix $P_\pi = (\pi_{ij})$ from (17) by letting the ij-th entry of P_π be 1 if $j = \pi_i$ and 0 otherwise (i.e. $\pi_{ij} = e_i^T P_\pi e_j = \delta_{\pi_i j}$):

$$P_\pi := \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} = \sum_{i=1}^{n} e_{\pi_i} e_i^T .$$

This is essentially the graph of π as a function from $\{1, 2, \ldots, n\}$ to $\{1, 2, \ldots, n\}$, except that we use ones instead of dots, and zeroes instead of blanks. Moreover, it's upside down because we read matrices from left to right and top to bottom. Let's notice that $P_\pi e_i = e_{\pi_i}$ – this is why we constructed P_π this way.

This construction can be done for permutations of any size!

Let us also notice that every P_π has orthonormal columns so every P_π is an orthogonal matrix, and therefore $P_\pi^T = P_\pi^{-1}$. Therefore $P_\pi^{-1} = P_\sigma$ because we know $P_\sigma P_\pi e_i = P_\sigma e_{\pi_i} = e_{\sigma_{\pi_i}} = e_i$ for all i.

Next we calculate $\det(P_\pi)$. We'll replace the dummy sum–variable in the formula by τ:

$$\det(P_\pi) = \sum_{\tau \in S_n} \mathrm{sgn}(\tau) \pi_{1\tau_1} \pi_{2\tau_2} \cdots \pi_{n\tau_n} = \sum_{\tau \in S_n} \mathrm{sgn}(\tau) \delta_{\pi_1 \tau_1} \delta_{\pi_2 \tau_2} \cdots \delta_{\pi_n \tau_n} .$$

In order for the product $\delta_{\pi_1 \tau_1} \delta_{\pi_2 \tau_2} \cdots \delta_{\pi_n \tau_n}$ to be non-zero it must be true that $\pi_i = \tau_i$ for every i, $1 \le i \le n$. There is thus only one non-zero term among all the $n!$ terms, so the sum reduces to

$$\det(P_\pi) = \mathrm{sgn}(\pi) \delta_{\pi_1 \pi_1} \delta_{\pi_2 \pi_2} \cdots \delta_{\pi_n \pi_n} = \mathrm{sgn}(\pi) 1 \cdot 1 \cdots 1 = \mathrm{sgn}(\pi).$$

We can now show that $\mathrm{sgn}(\pi) = \mathrm{sgn}(\sigma)$. We know that $P_\sigma P_\pi = I$, so

$$1 = \det(I) = \det(P_\sigma P_\pi) = \det(P_\sigma) \det(P_\pi) = \mathrm{sgn}(\pi) \mathrm{sgn}(\sigma).$$

Since $\mathrm{sgn}(\pi) = \pm 1$ and $\mathrm{sgn}(\pi)\mathrm{sgn}(\sigma) = 1$, $\mathrm{sgn}(\pi) = \mathrm{sgn}(\sigma)$.

We know that to every $\pi \in S_n$ there corresponds $\sigma \in S_n$ that undoes π, and that π undoes σ. We could also write π^{-1} for σ. Therefore, as π runs thru S_n, σ runs thru S_n as well. Hence for every π, if σ is the permutation that "undoes" π,

$$\mathrm{sgn}(\pi) a_{\pi_1 1} a_{\pi_2 2} \cdots a_{\pi_n n} = \mathrm{sgn}(\sigma) a_{1\sigma_1} a_{2\sigma_2} \cdots a_{n\sigma_n},$$

and thus

$$\det(A^T) = \sum_{\pi \in S_n} \mathrm{sgn}(\pi) a_{\pi_1 1} a_{\pi_2 2} \cdots a_{\pi_n n} = \sum_{\sigma \in S_n} \mathrm{sgn}(\sigma) a_{1\sigma_1} a_{2\sigma_2} \cdots a_{n\sigma_n} = \det(A).$$

Now formula (15) is proved! That was the formula

$$\det(A) = \sum_{i=1}^{n} (-1)^{i+j} a_{ij} \det(A_{ij}) \tag{18}$$

for expanding a determinant as a function of the *columns* of a matrix as well as a function of the rows. In other terms,

$$\det(A^T) = F(e_1^T A^T, \ldots, e_n^T A^T) = F(v_1^T, \ldots, v_n^T).$$

Then we can expand along the "rows" $v_1^T, \ldots, v_n^T$ which have the coordinates of the original columns.

Formula (18) looks a bit like the formula for a matrix product. If we define

$$\mathrm{adj}(A) := \Big((-1)^{i+j} \det(A_{ij}) \Big)^T = \Big((-1)^{i+j} \det(A_{ji}) \Big)$$

then the ij-th coordinate of $\text{adj}(A)A$ is $\sum_{k=1}^{n}(-1)^{i+k}\det(A_{ki})a_{kj} = \sum_{k=1}^{n}(-1)^{k+i}a_{kj}\det(A_{ki})$.

If $i = j$, the jj-th coordinate of $\text{adj}(A)A$ is $\sum_{k=1}^{n}(-1)^{j+k}\det(A_{kj})a_{kj} = \sum_{k=1}^{n}(-1)^{k+j}a_{kj}\det(A_{kj}) = \det(A)$, by formula (18) with dummy variable k in place of i.

But if $i \neq j$, the ij-th coordinate of $\text{adj}(A)A$ is $\sum_{k=1}^{n}(-1)^{k+i}a_{kj}\det(A_{ki})$. This is what we would get if we put column j of A in place of column i. You might have to look again at the text about formulas (14) and (15) to see that this is indeed true. But then the matrix has two copies of column j, one in column j and the other replacing column i. This natrix thus has determinant zero. Thus the off–diagonal entries of $\text{adj}(A)A$ are all zero, and the ones on the diagonal are all $\det(A)$ so

$$\text{adj}(A)A = \det(A)I. \tag{18.3}$$

The same thing happens when we examine $A\text{adj}(A)$, but we use formula (14), and we get

$$A\text{adj}(A) = \det(A)I. \tag{18.6}$$

This gives us a Theorem.

(19) • **Theorem:** *If A is an $n \times n$ matrix and $\det(A) \neq 0$ then*

$$A^{-1} = \frac{1}{\det(A)}\,adj(A).$$

Calculating A^{-1} this way is fine if $n = 2$ or $n = 3$. It is marginally OK if $n = 4$, but for $n > 4$ it's usually very inefficient.

(20) • **Theorem (Cramer's Rule):** *If A is an $n \times n$ matrix and $\det(A) \neq 0$ then the solution of the equation $Ax = b$ is*

$$A^{-1}b = \frac{adj(A)b}{\det(A)}.$$

The i-th coordinate of $A^{-1}b$ is thus the determinant of $A(b,\, i) := A + (b - Ae_i)e_i^T$ (the matrix obtained from A by replacing column i of A by b), divided by $\det(A)$:

$$e_i \bullet (A^{-1}b) = \frac{\det(A(b,\, i))}{\det(A)}.$$

Proof: The first equation follows from Theorem (19). The i-th coordinate of $A^{-1}b$ is then

$$\frac{1}{\det(A)}\sum_{k=1}^{n}(-1)^{k+i}\det(A_{ki})b_k = \frac{1}{\det(A)}\sum_{k=1}^{n}(-1)^{k+i}b_k\det(A_{ki}),$$

which (by the discussion preceding Theorem (19), in the case $i = j$) is exactly what is stated in Theorem (20) when we replace column i of A by b.

• **Example** Let's calculate $\text{adj}(A)$ for the matrix $A = \begin{pmatrix} 1 & 2 & 3 \\ 4 & 5 & 4 \\ 3 & 2 & 1 \end{pmatrix}$. We have to recall that A_{ij} is the matrix obtained from A by deleting its row i and column j, and we must multiply $\det(A_{ij})$ by $(-1)^{i+j}$. Then

$$\text{adj}(A)^T = \begin{pmatrix} \det\begin{pmatrix} 5 & 4 \\ 2 & 1 \end{pmatrix} & -\det\begin{pmatrix} 4 & 4 \\ 3 & 1 \end{pmatrix} & \det\begin{pmatrix} 4 & 5 \\ 3 & 2 \end{pmatrix} \\ -\det\begin{pmatrix} 2 & 3 \\ 2 & 1 \end{pmatrix} & \det\begin{pmatrix} 1 & 3 \\ 3 & 1 \end{pmatrix} & -\det\begin{pmatrix} 1 & 2 \\ 3 & 2 \end{pmatrix} \\ \det\begin{pmatrix} 2 & 3 \\ 5 & 4 \end{pmatrix} & -\det\begin{pmatrix} 1 & 3 \\ 4 & 4 \end{pmatrix} & \det\begin{pmatrix} 1 & 2 \\ 4 & 5 \end{pmatrix} \end{pmatrix} = \begin{pmatrix} -3 & 8 & -7 \\ -4 & 8 & -4 \\ -7 & 8 & -3 \end{pmatrix}$$

so that

$$\operatorname{adj}(A) = \begin{pmatrix} -3 & -4 & -7 \\ 8 & 8 & 8 \\ -7 & -4 & -3 \end{pmatrix}.$$

You should check the equations (18.3), (18.6) and the equation in Theorem (19) for this example.

- **An Example of a different sort** Some time ago we found that a matrix of the form $A = I_n - ab^T$ (where a and b are in $\mathbb{R}^n$ and neither is zero) is invertible if and only if $a \bullet b \neq 1$, i.e., $1 - a \bullet b \neq 0$. We should suspect that $1 - a \bullet b$ and $\det(I - ab^T)$ are related. They are the same. You should try to calculate the determinant when $a = (1, -2, 3, -4)$ and $b = (1, 3, 3, 1)$.

The method that works in all cases is to exploit multilinearity in gory detail, but take advantage of the notation for summing over an index set!

The i-th row of $I_n - ab^T$ is $e_i^T(I_n - ab^T) = e_i^T - a_i b^T$ (do you agree?). Thus

$$\det(I_n - ab^T) = F(e_1^T - a_1 b^T,\, e_2^T - a_2 b^T,\, \ldots,\, e_n^T - a_n b^T).$$

Multilinearity allows to use the idea of the Distributive Law that we would use to expand a product of the form $(x_1 - a_1)(x_2 - a_2)\cdots(x_n - a_n)$: choose one term from each factor $x_i - a_i$ (choose x_i or $-a_i$), multiply them together, and add all the products. There are 2^n different ways to choose one term from each factor, so we have to add 2^n products of n numbers. We need an index set with 2^n elements. A set that will do the job is the set of all possible strings of n zeroes and ones. A notation that works well is

$$B_n := \{\epsilon \in \mathbb{R}^n : \epsilon = (\epsilon_1,\, \epsilon_2,\, \ldots \epsilon_n), \quad \text{where each } \ \epsilon_i \in \{0,\, 1\}\}.$$

If $n = 3$, for example, $B_3 = \{(0,0,0), (0,0,1), (0,1,0), (0,1,1), (1,0,0), (1,0,1), (1,1,0), (1,1,1)\}$ and the B stands for "binary."

We can then write

$$(x_1 - a_1)(x_2 - a_2)\cdots(x_n - a_n) = \sum_{\epsilon \in B_n} \prod_{i=1}^{n} ((1 - \epsilon_i)x_i - \epsilon_i a_i).$$

I am sure this looks *very* complicated, but it does the job! Since $\epsilon_i = 0$ or $\epsilon_i = 1$, $(1 - \epsilon_i)x_i - \epsilon_i a_i = x_i$ if $\epsilon_i = 0$ and $(1 - \epsilon_i)x_i - \epsilon_i a_i = -a_i$ if $\epsilon_i = 1$. We will do the same for the determinant we are working with, and we will see that all but $n+1$ of the terms in that sum of 2^n terms are zero!

We return to the determinant formula above:

$$\begin{aligned} \det(I_n - ab^T) &= F(e_1^T - a_1 b^T,\, e_2^T - a_2 b^T,\, \ldots,\, e_n^T - a_n b^T) \\ &= \sum_{\epsilon \in B_n} F((1 - \epsilon_1)e_1^T - \epsilon_1 a_1 b^T,\, \ldots,\, (1 - \epsilon_n)e_n^T - \epsilon_n a_n b^T). \end{aligned}$$

Each term in the big sum is actually the determinant of the matrix with rows $(1 - \epsilon_i)e_i^T - \epsilon_i a_i b^T$, $1 \leq i \leq n$. Let's look at an arbitrary term in the big sum, and suppose that at least two of the $\epsilon_i = 1$. Then

$$F((1 - \epsilon_1)e_1^T - \epsilon_1 a_1 b^T,\, \ldots,\, (1 - \epsilon_n)e_n^T - \epsilon_n a_n b^T)$$

will contain at least two rows of the form $-a_i b^T$. But then the matrix will have two proportional rows, and will thus be zero.

Hence the only terms that survive are those that have no ones at all in ϵ, or exactly one of them. The term with no ones is $F(e_1^T,\, \ldots,\, e_n^T) = F(I) = 1$. For each i, $1 \leq i \leq n$, the term with $\epsilon_i = 1$ and all other $\epsilon_j = 0$ is (let's take $1 < i < n$)

$$F(e_1^T, \ldots, a_i b^T, \ldots, e_n^T) = \det \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ 0 & 1 & \cdots & 0 & 0 & 0 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 1 & 0 & 0 & \cdots & 0 & 0 \\ a_ib_1 & a_ib_2 & \cdots & a_ib_{i-1} & a_ib_i & a_ib_{i+1} & \cdots & a_ib_{n-1} & a_ib_n \\ 0 & 0 & \cdots & 0 & 0 & 1 & \cdots & 0 & 0 \\ \vdots & \vdots & \cdots & \vdots & \vdots & \vdots & \cdots & \vdots & \vdots \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 1 & 0 \\ 0 & 0 & \cdots & 0 & 0 & 0 & \cdots & 0 & 1 \end{pmatrix}.$$

We can use row operations to "zero" out everything in row i except the entry in column i. The resulting matrix is diagonal, and its determunant is $a_i b_i$. When we add the $n+1$ terms left, we get $\det(I_n - ab^T) = 1 - a \bullet b$. Can you say how to find $\det(I_n + ab^T)$?

Additional material

- **The formula** (2.5) **(the determinant formula) gives an alternating function**

We have to be careful in this argument to use nothing that we proved after we defined the determinant, because we used the alternating property. The purpose of this argument is to prove the alternating property! Having said that, we can, however, use anthing that did not use any property of determinants. We can use the definition of "multiplying" permutations, given in the proof of Theorem (16.5), for example.

Our formula (2.5) to work with is

$$F(V) = F(v_{ij}) = F(h_1, h_2, \ldots, h_n) = F(e_1^T V, e_2^T V, \ldots, e_n^T V) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) v_{1\pi_1} v_{2\pi_2} \cdots v_{n\pi_n}.$$

If we switch two rows of V, say rows k and ℓ, with $1 < k < \ell < n$, we get a modified matrix $U = (u_{ij})$ and we need to show that $F(U) = -F(V)$. We have

$$F(U) = \sum_{\pi \in S_n} \operatorname{sgn}(\pi) u_{1\pi_1} u_{2\pi_2} \cdots u_{n\pi_n},$$

and a typical term in the big sum is

$$\operatorname{sgn}(\pi) u_{1\pi_1} \cdots u_{k\pi_k} \cdots u_{\ell\pi_\ell} \cdots u_{n\pi_n} = \operatorname{sgn}(\pi) v_{1\pi_1} \cdots v_{\ell\pi_k} \cdots v_{k\pi_\ell} \cdots v_{n\pi_n}$$

because $u_{ij} = v_{ij}$ unless $i = k$ or $i = \ell$, and then we have $u_{kj} = v_{\ell j}$ and $u_{\ell j} = v_{kj}$. In the term on the right the first subscripts are out of order. The second subscripts form π. Let's put those first subscripts back in order:

$$\operatorname{sgn}(\pi) u_{1\pi_1} \cdots u_{k\pi_k} \cdots u_{\ell\pi_\ell} \cdots u_{n\pi_n} = \operatorname{sgn}(\pi) v_{1\pi_1} \cdots v_{k\pi_\ell} \cdots v_{\ell\pi_k} \cdots v_{n\pi_n}.$$

Now the second subscripts are not in the same order as π. We had

$$\pi_1 \quad \cdots \quad \pi_k \quad \cdots \quad \pi_\ell \quad \cdots \quad \pi_n \quad \text{and now we have} \quad \pi_1 \quad \cdots \quad \pi_\ell \quad \cdots \quad \pi_k \quad \cdots \quad \pi_n\,,$$

where all the π_i's not shown are in their original places in π. This new ordering is a new permutation, and we need to know how it is "mathematically" related to π. We bring in the permutation $\tau_{k\ell}$ that switches k and ℓ and leaves all other i's alone:

$$(\tau_{k\ell})_i := \begin{cases} i, & \text{if} \quad i \neq k \quad \text{and} \quad i \neq \ell; \\ \ell, & \text{if} \quad i = k; \\ k, & \text{if} \quad i = \ell. \end{cases}$$

In other words, $(\tau_{k\ell})_k = \ell$, $(\tau_{k\ell})_\ell = k$ and $(\tau_{k\ell})_i = i$ otherwise. What is $P_{\tau_{k\ell}}$ when $n = 6$, $k = 2$ and $\ell = 5$? To mathematically relate π and the new permutation we recall how permutations were "multiplied" by functional composition (this was discussed in the proof of Theorem (16.5)) and we calculate $\pi\tau_{k\ell}$:

$$[\pi\tau_{k\ell}]_i = \pi_{(\tau_{k\ell})_i} = \begin{cases} \pi_i, & \text{if} \quad i \neq k \quad \text{and} \quad i \neq \ell; \\ \pi_\ell, & \text{if} \quad i = k; \\ \pi_k, & \text{if} \quad i = \ell. \end{cases}$$

Thus the new permutation is $\pi\tau_{k\ell}$, and we can write

$$\text{sgn}(\pi) u_{1\pi_1} \cdots u_{k\pi_k} \cdots u_{\ell\pi_\ell} \cdots u_{n\pi_n} = \text{sgn}(\pi) v_{1[\pi\tau_{k\ell}]_1} \cdots v_{k[\pi\tau_{k\ell}]_k} \cdots v_{\ell[\pi\tau_{k\ell}]_\ell} \cdots v_{n[\pi\tau_{k\ell}]_n}.$$

In order to make the term on the right fit the pattern of the determinant formula we need to know how $\text{sgn}(\pi)$ and $\text{sgn}(\pi\tau_{k\ell})$ are related. We are forced to use our original definition of $\text{sgn}(\pi)$ and $\text{sgn}(\pi\tau_{k\ell})$! We recall that we find $\text{sgn}(\pi)$ by carrying out the scheme

$$\begin{array}{ccccccccc} \pi_1 & \cdots & \pi_k & \cdots & \pi_\ell & \cdots & \pi_n & & \\ c_1+ & \cdots & +c_k+ & \cdots & +c_\ell+ & \cdots & +c_n & = s(\pi), \end{array} \qquad \text{and then writing } \text{sgn}(\pi) := (-1)^{s(\pi)},$$

where c_i is the number of the numbers π_j to the right of π_i that are less than π_i.

If we can show that $\text{sgn}(\pi\tau_{k\ell}) = -\text{sgn}(\pi)$ we will be done with the proof (except for "cleanup"). We will do so in two steps. First we will show it if $\ell = k + 1$, then (fairly easily) show that the other cases are true too. We will still assume $1 < k < k + 1 < n$, but the conditions $1 < k$ and $k + 1 < n$ are not at all necessary; they just make the diagrams look better. We start with the scheme for finding $\text{sgn}(\pi)$:

$$\begin{array}{cccccccc} \pi_1 & \cdots & \pi_k & & \pi_{k+1} & \cdots & \pi_n & \\ c_1 + & \cdots & + c_k & + & c_{k+1} + & \cdots & + c_n & = s(\pi); \end{array} \qquad \text{sgn}(\pi) = (-1)^{s(\pi)}.$$

For $\text{sgn}(\pi\tau_{k\ell})$ we'll use c'_i for the new counts, and we have its scheme:

$$\begin{array}{cccccccc} \pi_1 & \cdots & \pi_{k+1} & & \pi_k & \cdots & \pi_n & \\ c'_1 + & \cdots & + c'_k & + & c'_{k+1} + & \cdots & + c'_n & = s(\pi\tau_{k\ell}); \end{array} \qquad \text{sgn}(\pi\tau_{k\ell}) = (-1)^{s(\pi\tau_{k\ell})}.$$

There are two easy things to notice: $c'_i = c_i$ if $i < k$ or if $i > k + 1$, by the way these counts are defined (do you agree?). We have to consider two subcases: $\pi_k > \pi_{k+1}$ and $\pi_k < \pi_{k+1}$. In the first case, π_{k+1} is one of the numbers to the right of π_k that are less than π_k. Thus c_k "includes" π_{k+1}. In the scheme for $\text{sgn}(\pi\tau_{k\ell})$, π_{k+1} is no longer to the right of π_k. Thus the count we make for π_k in its new location, which is now c'_{k+1}, is one less than it was before: $c'_{k+1} = c_k - 1$. The count we make for π_{k+1} in *its* new location, which is now c'_k, will be the same as it was before; all the nubers that were to its right and less than it are still there, and π_k, now to the right of π_{k+1}, is *not* less than π_{k+1}. Thus $c'_k = c_{k+1}$. We can write down what the scheme has become:

$$\begin{array}{cccccccc} \pi_1 & \cdots & \pi_{k+1} & & \pi_k & \cdots & \pi_n & \\ c_1 + & \cdots & + c_{k+1} & + & c_k - 1 + & \cdots & + c_n & = s(\pi\tau_{k\ell}) = s(\pi) - 1; \end{array} \qquad \text{sgn}(\pi\tau_{k\ell}) = (-1)^{s(\pi)-1} = -\text{sgn}(\pi).$$

This completes the subcase $\pi_k > \pi_{k+1}$. In the other subcase, $\pi_k < \pi_{k+1}$, which you should do, the count will increase by one instead of decreasing by one. Thus in either subcase, $\text{sgn}(\pi\tau_{k\ell}) = -\text{sgn}(\pi)$.

Next we consider the other cases, $k < \ell$ with $k + 1 < \ell$. In these cases we can switch π_k with the ones between it and π_ℓ and then switch with π_ℓ. This will take $\ell - k$ switches, and each will change the sign of $\text{sgn}(\pi)$ once. Then we switch π_ℓ, now just to the left of π_k, back in the other direction until it occupies the place where π_k was to begin with. This will take just $\ell - k - 1$ switches, because we do not have to switch with π_k. The total number of switches is therefore $2(\ell - k) - 1$, an odd number, so $\text{sgn}(\pi\tau_{k\ell}) = (-1)^{2(\ell-k)-1}\text{sgn}(\pi) = -\text{sgn}(\pi)$.

Now we can write

$$\begin{aligned} F(U) &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) u_{1\pi_1} u_{2\pi_2}\cdots u_{n\pi_n}, \\ &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) v_{1\pi_1}\cdots v_{\ell\pi_k}\cdots v_{k\pi_\ell}\cdots v_{n\pi_n} \\ &= \sum_{\pi\in S_n} \operatorname{sgn}(\pi) v_{1[\pi\tau_{k\ell}]_1}\cdots v_{k[\pi\tau_{k\ell}]_k}\cdots v_{\ell[\pi\tau_{k\ell}]_\ell}\cdots v_{n[\pi\tau_{k\ell}]_n} \\ &= \sum_{\pi\in S_n} -\operatorname{sgn}(\pi\tau_{k\ell}) v_{1[\pi\tau_{k\ell}]_1}\cdots v_{k[\pi\tau_{k\ell}]_k}\cdots v_{\ell[\pi\tau_{k\ell}]_\ell}\cdots v_{n[\pi\tau_{k\ell}]_n} \\ &= -\sum_{\pi\in S_n} \operatorname{sgn}(\pi) v_{1\pi_1}\cdots v_{n\pi_n} = -F(V). \end{aligned}$$

because $\pi\tau_{k\ell}$ runs thru all the permutations in S_n as π runs thru all the permutations in S_n (we can see that this is true because $\tau_{k\ell}\tau_{k\ell}=e$ [you should verify that $\tau_{k\ell}\tau_{k\ell}=e$]). This completes the proof!

- **How the formula (2.5) arises from an alternating multilinear function on the rows of an $n\times n$ matrix that is 1 when evaluated at I_n.**

These are gory details that once again use the notation for summing over an index set. We have $h_i = \sum_{j=1}^n v_{ij}e_j^T$ for $1\le i\le n$, so

$$F(h_1,\, h_2,\, \ldots,\, h_n) = F\left(\sum_{j_1=1}^n v_{ij_1}e_{j_1}^T,\ \sum_{j_2=1}^n v_{ij_2}e_{j_2}^T,\ \ldots,\ \sum_{j_n=1}^n v_{ij_n}e_{j_n}^T\right).$$

Now we use linearity in the first variable, getting

$$F(h_1,\, h_2,\, \ldots,\, h_n) = \sum_{j_1=1}^n v_{ij_1} F\left(e_{j_1}^T,\ \sum_{j_2=1}^n v_{ij_2}e_{j_2}^T,\ \ldots,\ \sum_{j_n=1}^n v_{ij_n}e_{j_n}^T\right).$$

When we do the same in each variable, we keep pulling a summation and the v_{ij_k} out, $k=1,\,2,\,\ldots,\,n$. When we are done we will have

$$F(h_1,\, h_2,\, \ldots,\, h_n) = \sum_{j_1=1}^n v_{ij_1}\sum_{j_2=1}^n v_{ij_2}\cdots\sum_{j_n=1}^n v_{ij_n} F\left(e_{j_1}^T,\, e_{j_2}^T,\, \ldots,\, e_{j_n}^T\right).$$

We can then put all the v's back inside the summations and get

$$F(h_1,\, h_2,\, \ldots,\, h_n) = \sum_{j_1=1}^n\sum_{j_2=1}^n\cdots\sum_{j_n=1}^n v_{ij_1}v_{ij_2}\cdots v_{ij_n} F\left(e_{j_1}^T,\, e_{j_2}^T,\, \ldots,\, e_{j_n}^T\right).$$

In each term we know that $F\left(e_{j_1}^T,\, e_{j_2}^T,\, \ldots,\, e_{j_n}^T\right)=0$ if $j_k=j_\ell$ for some $k\ne\ell$ because now we are *assuming* that $F(V)$ is alternating. Hence only those lists of j_k's that are all different leave open the possibility that $F\left(e_{j_1}^T,\, e_{j_2}^T,\, \ldots,\, e_{j_n}^T\right)\ne 0$. But then every $i\in\{1,\,2,\,\ldots,\,n\}$ is in the list, so $j_1j_2\cdots j_n$ is a permutation of $\{1,\,2,\,\ldots,\,n\}$. We can then replace the summation $\sum_{j_1=1}^n\sum_{j_2=1}^n\cdots\sum_{j_n=1}^n$ (which is a sum over n^n indices) by the summation $\sum_{\pi\in S_n}$ (a sum over "only" $n!$ indices), define $\operatorname{sgn}(\pi) := F\left(e_{\pi_1}^T,\, e_{\pi_2}^T,\, \ldots,\, e_{\pi_n}^T\right)$, and so obtain the formula in (2.5).

- **Some problems**
  - "Identify" the matrix with rows $e_{\pi_1}^T,\, e_{\pi_2}^T,\, \ldots,\, e_{\pi_n}^T$.
  - Show that, if π and $\tilde\pi$ are in S_n then $P_\pi P_{\tilde\pi} = P_{\pi\tilde\pi}$.

This (nearly) shows that there is a one–to–one correspondence between the "permutation matrices" P_π and the permutations $\pi \in S_n$.

- Find the "counts" for (643512), (645312) and (643152).
- Show that if A is invertible then A can be expesssed as the product of finitely many ERO matrices.
- Find $\det \begin{pmatrix} 2 & 3 & 3 & 1 \\ -2 & -5 & -6 & -2 \\ 3 & 9 & 10 & 3 \\ -4 & -12 & -12 & -3 \end{pmatrix}$. Hint: Subtract I_4 from the matrix.
- Find $\det \begin{pmatrix} 3 & 4 & 7 \\ 8 & 8 & 8 \\ 7 & 4 & 3 \end{pmatrix}$.
- Suppose A is $n \times n$ and invertible. Find $\det(\mathrm{adj}(A))$.
- Show that $\pi\tau_{k\ell}$ runs thru all the permutations in S_n as π runs thru all the permutations in S_n.