LECTURE 16

Determinants

We are now going to break with the text in two ways. First of all, since we have just spent the last couple weeks discussing various applications of matrices to linear algebra, to me it seems much more natural to move directly to the theory of determinants rather than to a discussion of inner products. Secondly, I will opt for a more utililitarian definition of determinant (by defining how it by how it's computed rather than the properties it is to have).

We saw a couple lectures ago that while it is always possible to see if an $n \times n$ matrix is invertible by checking that its reduced row echelon form coincides with the identity matrix, it is a pretty tedious check to carry out. Wouldn't it be nice if there was a simple function of the matrix entries whose value would immediately tell you whether or not a matrix is invertible. Well, it turns out that there is such a function, the *determinant function*; that's simple enough to define, but a bit harder to compute.

It will also turn out that these determinant functions have lots and lots of applications.

Let me begin with Curtis's definition. Recall that an $n \times n$ matrix corresponds to an ordered list of n elements of $\mathbb{F}^n$.

DEFINITION 16.1. *A **determinant function** D is a function that maps ordered lists of n elements of $\mathbb{F}^n$ to an element of $\mathbb{F}$ satisfying the following three requirements:*

(1) *If $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is the standard basis for $\mathbb{F}^n$, then*

$$D\left([\mathbf{e}_1, \ldots, \mathbf{e}_n]\right) = 1_{\mathbb{F}}$$

(which is equivalent to requiring the value of D on the identity matrix is $1_{\mathbb{F}}$).

(2) $D\left([a_1, \ldots, a_i + a_j, \ldots, a_j, \ldots, a_n]\right) = D\left([a_1, \ldots, a_i, \ldots, a_j, \ldots, a_n]\right)$

(3) $D\left([a_1, \ldots, a_{i-1}, \lambda a_i, a_{i+1}, \ldots, a_n]\right) = \lambda D\left([a_1, \ldots, a_n]\right)$

REMARK 16.2. Note that since any vector in $\mathbb{F}^n$ can be constructed from the standard basis vectors via repeated application of vector addition and scalar multiplication, that (1) together with the rules (2) and (3) provide a means for computing $D\left([a_1, \ldots, a_n]\right)$ for an arbitrary list of n vectors in $\mathbb{F}^n$. What is not so obvious from this defintion is that the results of such computations is independent of the way they are carried out. Of course, this would have to be the case if a determinant function is to actually be well-defined via this definition. But where's the proof? Secondly, even if you can prove that the above definition can be consistently satisfied, how do you know there is only one solution (i.e. that the determinant function thus defined is unique). For this reason, I prefer the following definition.

DEFINITION 16.3. *Let $\mathbf{M}$ be an $n \times n$ matrix with entries in $\mathbb{F}$. By definition the $(ij)^{th}$ -minor $\mathbf{M}_{ij}$ of $\mathbf{M}$ is the $(n-1) \times (n-1)$ matrix formed by deleting the i^{th} row and j^{th} column of $\mathbf{M}$ from $\mathbf{M}$. The determinant of $\mathbf{M}$ is then defined by the recursive formula*

(i) $\det([a]) = a$ *for any* 1×1 *matrix* a

(ii) *if* $n > 1$, *then*

$$\det(\mathbf{M}) = \sum_{i=1}^{n} (-1)^{i+j} M_{ij} \det(\mathbf{M}_{ij}) \qquad \textit{(for any fixed colunn index } j\textit{)}$$

or

$$\det(\mathbf{M}) = \sum_{j=1}^{n} (-1)^{i+j} M_{ij} \det(\mathbf{M}_{ij}) \qquad \textit{(for any fixed row index } i\textit{)}$$

Of course, there is still some ambiguity in the definition; how do we know that the the expansions on the right hand side are independent of which column or row we choose to keep fixed.

So here's a third, slightly more elaborate, definition of the determinant function.

Let $\mathfrak{S}_n$ denote the set of permutations of the numbers $1, \ldots, n$. So for example

$$\mathfrak{S}_3 = \{[1,2,3], [1,3,2], [2,1,3], [2,3,1], [3,1,2], [3,2,1]\}$$

We can separate elements $\mathfrak{S}_n$ into two classes, called *even* and *odd*, by the following algorithm. Fix a permutation $\sigma = [\sigma_1, \ldots, \sigma_n] \in \mathfrak{S}_n$. Anytime we have a situation where an integer σ_j occurs before an σ_j in σ but with $\sigma_i > \sigma_j$ we'll say σ has an inversion. For example,

$$\sigma = [2,4,3,1]$$

has an inversions amongst the entries $(1,2)$, $(1,3)$, $(1,4)$ (because $1 < 2$, $1 < 3$, $1 < 4$, but $2,3,4$ precede 1 in σ) and an inversion $(3,4)$ (because $3 < 4$ but 4 precedes 3 in σ). Thus σ has a total of 4 inversions. We say that σ is an *even permutation* if σ has an even number of inversions, and σ is an *odd permutation* if σ has an odd number of permutations.

Definition 16.4. *The* ***sign*** *of a permutation $\sigma \in \mathfrak{S}_n$ is*

$$\varepsilon(\sigma) := (-1)^{\#\ \textit{inversions in}\ \sigma}$$

We think of ε as a function that maps permutations to ± 1.

Here is an equivalent way of defining the function $\varepsilon : \mathfrak{S}_n \to \{\pm 1\}$. Let $\sigma = [\sigma_1, \ldots, \sigma_n]$ be a permutation and consider the product

$$\prod_{1 \leq i < j \leq n} \frac{x_{\sigma_i} - x_{\sigma_j}}{x_i - x_j} \tag{2}$$

Note that in denominator we have exactly one factor $x_k - x_\ell$ for each ordered pair $k < \ell$ of distinct integers $k, \ell \in \{1, \ldots, n\}$. In the numerator, however, we'll either have a factor $x_k - x_\ell$ or a factor $x_\ell - x_k$. The latter case occuring whenever we have a situation where $\ell = \sigma_i$, $k = \sigma_j$ but $\iota < j$. That is, whenever we have an inversion in σ we'll have a factor in the numerator that is -1 times a factor in the denominator. Thus,

$$\prod_{1 \leq i < j \leq n} \frac{x_{\sigma_i} - x_{\sigma_j}}{x_i - x_j} = (-1)^{\#\ \text{inversions in}\ \sigma} = \varepsilon(\sigma)$$

- We'll need the following fact latter on

Fact 16.5. *If σ differs from σ' by a single interchange, then $\varepsilon(\sigma') = -\varepsilon(\sigma)$.*

Proof. First we'll prove an easy special case, when $j = i + 1$. Suppose

$$\begin{aligned} \sigma &= [\sigma_1, \ldots, \sigma_{i-1}, \sigma_i, \sigma_{i+1}, \sigma_{i+2}, \ldots, \sigma_n] \\ \sigma' &= [\sigma_1, \ldots, \sigma_{i-1}, \sigma_{i+1}, \sigma_i, \sigma_{i+2}, \ldots, \sigma_n] \end{aligned}$$

Note that the except for the ordering of the specific pair σ_i, σ_{i+1}, all the relative ordering of the entries pairs of entries of σ and σ' are the same. However, when we look at the relative orderings of the i^{th} and $(i+1)^{th}$ entries, we have

$$\begin{array}{lllll} \sigma_i < \sigma_{i+1} & \Rightarrow & \sigma'_i > \sigma'_{i+1} & & (\text{since } \sigma'_i \equiv \sigma_{i+1} \text{ and } \sigma'_{i+1} \equiv \sigma_i) \\ \sigma_i > \sigma_{i+1} & \Rightarrow & \sigma'_i < \sigma'_{i+1} & & (\text{same reason}) \end{array}$$

Therefore, the number of inversions of σ will differ from the number of inversions of σ' by exactly ± 1. Hence,

$$\varepsilon(\sigma') = (-1)^{\left(\#\text{ inversions of }\sigma'\right)} = (-1)^{(\#\text{ inversions of }\sigma)\ \pm 1} = \varepsilon(\sigma)(-1)^{\pm 1} = -\varepsilon(\sigma) \quad .$$

Now let's consider the more general case where σ_i and σ_j are not necessarily neighboring entries in σ. Suppose $\sigma = [\sigma_1, \dots, \sigma_i, \dots, \sigma_j, \dots, \sigma_n]$ and $\sigma' = [\sigma_1, \dots, \sigma_j, \dots, \sigma_i, \dots, \sigma_n]$ (i.e, σ' differs from σ by a single interchange $\sigma_i \longleftrightarrow \sigma_j$), We can systematically convert σ to σ' via a sequence of nearest neighbor exchanges. Let s_i be the operation that interchanges the i^{th} entry of a permutation with the $(i+1)^{th}$ entry. Then

$$s_j s_{j-1} \cdots s_i(\sigma) = [\sigma_1, \dots, \sigma_{i-1}, \sigma_{i+1}, \dots, \sigma_j, \sigma_i, \dots, \sigma_n]$$

(each successive s_* bumps σ_i further along to the right). Next we apply the interchanges $s_{j-1}, s_{j-2}, \dots, s_i$ in succession to bump σ_j off to left until it occupies the place originally occupied by σ_i. Thus,

$$\sigma' = s_i s_{i+1} \cdots s_{j-1} s_j s_{j-1} \cdots s_i(\sigma)$$

Note that this is always an odd number of nearest neighbor interchanges (notice how s_j sits right at the center of composition flanked by $s_i \cdots s_{j-1}$ on one side and $s_{j-1} \cdots s_i$ on the other). By our earlier result, each of these nearest neighbor interchanges has the effect of flipping the sign of $\varepsilon(\sigma)$. Since we have an odd number of them, the total sign flip will be $(-1)^{\text{odd number}} = -1$. So our desired conclusion follows. $\square$

DEFINITION 16.6. *Let $\mathfrak{S}_n$ be the set of permutations of the numbers $1, 2, \dots, n$. The determinant of an $n \times n$ matrix $\mathbf{M}$ over a field $\mathbb{F}$ is the element of $\mathbb{F}$ determined by the following formula*

$$\det(\mathbf{M}) = \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) M_{1\sigma_1} M_{2\sigma_2} \cdots M_{n\sigma_n}. \tag{3}$$

PROPOSITION 16.7. *The determinant function* det *has the following properties.*

(i) *$det(\mathbf{I}_n) = 1$ if $\mathbf{I}_n$ is the $n \times n$ identity matrix.*
(ii) *If $\mathbf{M}'$ is a matrix obtained from $\mathbf{M}$ by replacing one row of $\mathbf{M}$ by its scalar multiple by $\lambda \in \mathbb{F}$, then*

$$\det(\mathbf{M}') = \lambda \det(\mathbf{M}) \quad .$$

(iii) *If $\mathbf{M}''$ is a matrix obtained from $\mathbf{M}$ by replacing one row of $\mathbf{M}$ by its vector sum with another row of $\mathbf{M}$ then*

$$\det(\mathbf{M}'') = \det(\mathbf{M}) \quad .$$

Proof.

(i) If $\mathbf{M} = \mathbf{I}_n$, then $M_{ij} = \begin{cases} 1 & \text{if } i = j \\ 0 & \text{if } i \neq j \end{cases}$. Thus, the products $M_{1\sigma_1} M_{2\sigma(2)} \cdots M_{n\sigma_n}$ on the right hand side of (**) will be non-zero only when $1 = \sigma_1$, $2 = \sigma_2$, $\dots$, $n = \sigma_n$. That is to say, only one term in the sum on the right hand side of (**) will be non-zero; it will correspond to $\sigma = [1, 2, \dots, n]$. But then since

$$\varepsilon([1, 2, \dots, n]) \equiv \frac{P(x_1, x_2, \dots, x_n)}{P(x_1, x_2, \dots, x_n)} = +1$$

it follows that

$$\det(\mathbf{I}_n) = \varepsilon([1, 2, \dots, n]) M_{11} M_{22} \cdots M_{nn} = (+1) \cdot 1 \cdot 1 \cdot \dots \cdot 1 = 1 \quad .$$

(ii) Note that each term on the right hand side has exactly one factor from each row of $\mathbf{M}$ (observe that each term in the sum is a product of entries in distinct rows, with each row contributing one factor). Thus, if we create $\mathbf{M}'$ by scalar multiplying the i^{th} row of $\mathbf{M}$ by $\lambda \in \mathbb{F}$, then exact one factor $M_{i\sigma_i}$ in each term of $\det(\mathbf{M})$ picks up a factor of λ. Thus,

$$\begin{aligned}
\det(\mathbf{M}') &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) M_{1\sigma_1} M_{2\sigma_2} \cdots (\lambda M_{i\sigma_i}) \cdots M_{n\sigma_n} \\
&= \lambda \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) M_{1\sigma_1} M_{2\sigma_2} \cdots M_{n\sigma_n} \\
&= \lambda \det(\mathbf{M})
\end{aligned}$$

(iii) Suppose $\mathbf{M}''$ is obtained form $\mathbf{M}$ by adding row j to row i. Then

$$\begin{aligned}
\det(\mathbf{M}'') &= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) M_{1\sigma_1} M_{2\sigma_2} \cdots (M_{i\sigma_i} + M_{j\sigma_i}) \cdots M_{j\sigma_j} \cdots M_{n\sigma} \\
&= \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) M_{1\sigma_1} M_{2\sigma_2} \cdots M_{i\sigma_i} \cdots M_{j\sigma_j} \cdots M + \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) M_{1\sigma_1} M_{2\sigma_2} \cdots M_{j\sigma_i} \cdots M_{j\sigma_j} \cdots M \\
&= \det(\mathbf{M}) + \sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) M_{1\sigma_1} M_{2\sigma_2} \cdots M_{j\sigma_i} \cdots M_{j\sigma_j} \cdots M_{n\sigma_n} \quad .
\end{aligned}$$

Now let's look at the second term on the right:

$$\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) M_{1\sigma_1} M_{2\sigma_2} \cdots M_{j\sigma_i} \cdots M_{j\sigma_j} \cdots M_{n\sigma_n}$$

Note that the factor $\varepsilon(\sigma)$ is flip sign under the interchanges $i \longleftrightarrow j$. That is to say, if

$$\sigma = [\sigma_1, \ldots, \sigma_i, \ldots, \sigma_j, \ldots, \sigma_n]$$

and

$$\sigma' = [\sigma_1, \ldots, \sigma_j, \ldots, \sigma_i, \ldots, \sigma_n]$$

then we'll have $\varepsilon(\sigma') = -\varepsilon(\sigma)$. Note also since we sum over **all** permutations, for every term

$$(*) \qquad \varepsilon(\sigma) M_{1\sigma_1} M_{2\sigma_2} \cdots M_{j\sigma_i} \cdots M_{j\sigma_j} \cdots M_{n\sigma_n}$$

there will be a corresponding term

$$\varepsilon(\sigma') M_{1\sigma_1} M_{2\sigma_2'} \cdots M_{j\sigma_i'} \cdots M_{j\sigma_j'} \cdots M_{n\sigma_n'} = -\varepsilon(\sigma) M_{1\sigma_1} M_{2\sigma_2} \cdots M_{j\sigma_j} \cdots M_{j\sigma_i} \cdots M_{n\sigma_n}$$

which, after reordering factors, is identical to (*) except opposite in sign. Hence, for each term (*) in the sum there will be another term that cancelling it. Thus,

$$\sum_{\sigma \in \mathfrak{S}_n} \varepsilon(\sigma) M_{1\sigma_1} M_{2\sigma_2} \cdots M_{j\sigma_i} \cdots M_{j\sigma_j} \cdots M_{n\sigma_n} = 0$$

We conclude

$$\det(\mathbf{M}) = \det(\mathbf{M}'') \quad .$$

□

We thus see that our determinant function satisfies the properties of Curtis's definition.

Corollary 16.8. *Let* $\mathbf{M}$ *be an* $n \times n$ *matrix.*

(i) If $\mathbf{M}'$ is a matrix obtained from $\mathbf{M}$ by interchanging two rows then $\det(\mathbf{M}') = -\det(\mathbf{M}) \quad .$
(ii) If $\mathbf{M}''$ is a matrix obtained from $\mathbf{M}$ by adding a scalar mulitple of one row to another row then $\det(\mathbf{M}'') = \det(\mathbf{M}) \quad .$

Proof.

(i) A row interchange can be implemented using the operations (ii) and (iii) of the preceding theorem. This I will demonstrate by example on a 2×2 matrix

$$\begin{pmatrix} a & b \\ c & d \end{pmatrix} \xrightarrow[R_2 \to R_2 + R_1]{R_1 \to -R_1} \begin{pmatrix} -a & -b \\ a+c & b+d \end{pmatrix} \xrightarrow{R_1 \to R_1 + R_2} \begin{pmatrix} c & d \\ a+c & b+d \end{pmatrix}$$

$$\xrightarrow{R_1 \to -R_1} \begin{pmatrix} -c & -d \\ a+c & b+d \end{pmatrix} \xrightarrow[R_1 \to -R_1]{R_2 \to R_2 + R_1} \begin{pmatrix} c & d \\ a & b \end{pmatrix}$$

By the preceding theorem the only time the determinant would change is when we multiplied a row by -1. Since we did this three times we have a total factor of $(-1)(-1)(-1) = -1$. If we substitute $R_1 \to R_i$ and $R_2 \to R_j$, then this same sequence of operations could be used to implement a row interchange between the i^{th} and j^{th} rows of a general $n \times n$ matrix. Again since there a total of three sign flips would occur in the determinants of the intermediary matrices, the statement to be proved follows.

(ii) This is proved similarly. If one applies the following sequence of operations

$$\begin{array}{lcll} R_i & \to & \frac{1}{\lambda} R_i & \left(\text{determinant changes by a factor } \frac{1}{\lambda}\right) \\ R_i & \to & R_i + R_j & (\text{determinant doesn't change}) \\ R_i & \to & \lambda R_i & (\text{determinant changes by a factor } \lambda) \end{array}$$

the overall effect will be the replacement of row i with its sum with λ times row j. According to the preceding theorem, the change in the determinant would be a factor of $\left(\frac{1}{\lambda}\right)(\lambda) = 1$; that is to say, the determinant would not change under this overall transformation. $\square$

PROPOSITION 16.9. *Let* $\mathbf{M}_{ij}$ *be the* $(ij)^{th}$ *minor of an* $n \times n$ *matrix* $\mathbf{M}$ *(as in Definition 5.3). Then if we define* $\det(\mathbf{M})$ *as in Definition 15.4 we have*

$$\det(\mathbf{M}) = \sum_{j=1}^{n} (-1)^{i+j} \det(\mathbf{M}_{ij}) \qquad \textit{for any fixed row index } i \quad .$$

Proof. Let's first concentrate on the case where $i = 1$. When $[1, 2, \ldots, n]$ is permuted to a new order $\sigma = [\sigma_1, \sigma_2, \ldots, \sigma_n]$, the initial 1 either stays put or goes to some other slot. In fact, we can break the set $\mathfrak{S}_n$ or permutations into n disjoint sets

$$\mathfrak{S}_{n,i} = \{\sigma \in \mathfrak{S}_n \mid \sigma_i = 1\}$$

This decomposition in turn allows us to decompose the sum over $\mathfrak{S}_n$ in Definition 15.4

$$\begin{aligned} \det(\mathbf{M}) &= \sum_{j=1}^{n} \sum_{\sigma \in \mathfrak{S}_{n,j}} \varepsilon(\sigma)\, m_{1\sigma_1} m_{2\sigma_2} \cdots m_{n\sigma_n} \\ &= \sum_{\sigma \in \mathfrak{S}_{n,1}} \varepsilon(\sigma)\, m_{11} m_{2\sigma_2} \cdots m_{n\sigma_n} \\ &\quad + \sum_{\sigma \in \mathfrak{S}_{n,2}} \varepsilon(\sigma)\, m_{1\sigma_1} m_{21} \cdots m_{n\sigma_n} \\ &\quad + \cdots \\ &\quad + \sum_{\sigma \in \mathfrak{S}_{n,i}} \varepsilon(\sigma)\, m_{1\sigma_1} m_{2\sigma_2} \cdots m_{n1} \end{aligned}$$

$$\begin{aligned}
&= m_{11}\sum_{\sigma\in\mathfrak{S}_{n,1}}\varepsilon(\sigma)\,m_{2\sigma_2}\cdots m_{n\sigma_n}\\
&\quad +m_{21}\sum_{\sigma\in\mathfrak{S}_{n,2}}\varepsilon(\sigma)\,m_{1\sigma_1}\cdots\widehat{m_{21}}\cdots m_{n\sigma_n}\\
&\quad +\cdots\\
&\quad +m_{n1}\sum_{\sigma\in\mathfrak{S}_{n,i}}\varepsilon(\sigma)\,m_{1\sigma_1}m_{2\sigma_2}\cdots m_{n-1,\sigma_{n-1}}
\end{aligned}$$

(Here and below a $\widehat{\ }$ over a symbol indicates that the term below the hat **does not appear** in the pattern indicated by the ellipses $\cdots$.) What I'll show below is that

$$(**)\qquad \sum_{\sigma\in\mathfrak{S}_{n,i}}\varepsilon(\sigma)\,m_{1\sigma_1}\cdots\widehat{m_{i\sigma_i}}\cdots m_{n\sigma_n}=(-1)^{1+i}\det(\mathbf{M}_{i1})$$

Indeed, the left hand side already looks a lot like our formula for a determinant; however, since $\mathbf{M}_{i1}$ is an $(n-1)\times(n-1)$ matrix, we need to relate the left hand side to a sum over the permutations of $[1,\ldots,n-1]$ and verify that it actually coincides with

$$(***)\qquad (-1)^{1+i}\sum_{\sigma'\in\mathfrak{S}_{n-1}}\varepsilon(\sigma')\,(\mathbf{M}_{i1})_{1\sigma'_1}(\mathbf{M}_{i1})_{2\sigma'_2}\cdots(\mathbf{M}_{i1})_{n-1,\sigma'_{n-1}}$$

(which is the right hand side of (***) written out explicitly via our definition of $\det(\mathbf{M}_{i1})$).

So let me first establish an explicit connection between the permutations in $\mathfrak{S}_{n-1}$ and those in $\mathfrak{S}_{n,i}\subset\mathfrak{S}_n$. For $\sigma'=\left[\sigma'_1,\ldots,\sigma'_{n-1}\right]\in\mathfrak{S}_{n-1}$, define $\eta_i(\sigma')\in\mathfrak{S}_{n,i}$ by

$$(\eta_i(\sigma'))_j=\begin{cases}\sigma'_j+1 & \text{if } \sigma'_j<i\\ 1 & \text{if } \sigma'_j=i\\ \sigma'_j+1 & \text{if } \sigma'_j>i\end{cases}$$

By construction, $\eta_i(\sigma')$ is always an arrangement of the numbers $1,\ldots,n$ with a 1 in the i^{th} slot, thus. $\eta_i(\sigma')$ is an element of $\mathfrak{S}_{n,i}$. In fact, η_i provides a bijection between $\mathfrak{S}_{n-1}$ and $\mathfrak{S}_{n,i}$ and

$$\begin{aligned}
\sum_{\sigma\in\mathfrak{S}_{n,i}}\varepsilon(\sigma)\,m_{1\sigma_2}\cdots\widehat{m_{i1}}\cdots m_{n\sigma_n} &= \sum_{\sigma'\in\mathfrak{S}_{n-1}}\varepsilon(\eta_i(\sigma'))\,m_{1\eta(\sigma')_1}\cdots\widehat{m_{i1}}\cdots m_{n\eta(\sigma')_n}\\
&= \sum_{\sigma'\in\mathfrak{S}_{n-1}}\varepsilon(\eta_i(\sigma'))\,m_{1,\sigma'_1+1}\cdots\widehat{m_{i1}}\cdots m_{n,\sigma'_n+1}\\
&= \sum_{\sigma'\in\mathfrak{S}_{n-1}}\varepsilon(\eta_i(\sigma'))\,(\mathbf{M}_{i1})_{1,\sigma'_1}\cdots(\mathbf{M}_{i1})_{n-1,\sigma'_{n-1}}
\end{aligned}$$

(I'm just using the fact that every element of $\mathfrak{S}_{n,i}$ is the image by η_i of an element $\sigma'\in\mathfrak{S}_{n-1}$, and then relating the factors $m_{1\eta(\sigma')_1}\cdots\widehat{m_{i1}}\cdots m_{n\eta(\sigma')}$ to products of the entries of the minor matrix $\mathbf{M}_{i1}$). If I can show that

$$(****)\qquad \varepsilon(\eta_i(\sigma'))=(-1)^{1+i}\varepsilon(\sigma')$$

then the identity that we need, (**), will follow.

OK, so let's fix i and an particular $\sigma'\in\mathfrak{S}_{n-1}$ and set $\sigma=\eta_i(\sigma')$. Then the way we have things set up

$$\sigma=\left[\sigma'_1+1,\sigma'_2+1,\ldots,\sigma_{i-1}+1,1,\sigma_i+1,\ldots,\sigma'_{n-1}+1\right]$$

We have

$$\begin{aligned}
\varepsilon(\sigma) &= (-1)^{\#\text{ inversions in the list }\sigma}\\
\varepsilon(\sigma') &= (-1)^{\#\text{ inversions in the list }\sigma'}
\end{aligned}$$

Now notice that to every inversion in $\left[\sigma'_1,\sigma'_2,\ldots,\sigma_{i-1},\sigma'_i,\ldots,\sigma'_{n-1}\right]$ there will be a corresponding inversion of $\left[\sigma'_1+1,\sigma'_2+1,\ldots,\sigma_{i-1}+1,\widehat{1},\sigma_i+1,\ldots,\sigma'_{n-1}+1\right]$ and hence an inversion of $\left[\sigma'_1+1,\sigma'_2+1,\ldots,\sigma_{i-1}+1,1,\sigma_i+1,\ldots,\right.$

In fact, we'll have a 1-1 correspondence between the inversions in σ' and the inversions of σ that don't involve that fixed 1 entry in the i^{th} place. On the other hand, there will be exactly $i-1$ inversions in $\left[\sigma_1'+1,\sigma_2'+1,\ldots,\sigma_{i-1}+1,1,\sigma_i+1,\ldots,\sigma_{n-1}'+1\right]$ that **do** involve that fixed 1 entry; because each since $\sigma_j'+1$ is always greater than 1, each of the $i-1$ entries of σ that occur the i^{th} entry 1 will lead to an inversion. Hence,

$$\#\text{ inversions in the list }\sigma=\#\text{ inversions in the list }\sigma'+(i-1)$$

Thus,

$$\varepsilon(\sigma)=(-1)^{\#\text{ inversions in the list }\sigma}=(-1)^{\#\text{ inversions in the list }\sigma'+(i-1)}=(-1)^{i+1}\varepsilon(\sigma')$$

We have thus, demonstrated (****), hence (**) and the proposition follow. □

□

Proposition 16.10. *Let $\mathbf{M}^t$ be the transpose of an $n\times n$ matrix $\mathbf{M}$. Then*

$$\det\left(\mathbf{M}^t\right)=\det(\mathbf{M})$$

Proof. This formula could probably be proved by staring hard enough at the formula for $\det(\mathbf{M})$, however, I'm just as likely to go cross-eyed trying. So rather than prove it using only the tools at hand, let me try to indicate the outlines of a simple proof using representation theory (which I do not expect to be all that understandable, but it will serve as a nice advertisement for my own field). The function $\sigma:\mathfrak{S}_n\to\pm1$ is an example of what's known as a *group character*. More generally, a group character is a function $\chi:G\to\mathbb{C}$ on a group G such that

$$\chi(gg')=\chi(g)\,\chi(g')$$

(equivalently, χ is a group homomorphism from G to $\mathbb{C}$).

Consider a typical term in our definition of the determinant

$$\varepsilon(\sigma)\,M_{1\sigma_1}M_{2\sigma_2}\cdots M_{n\sigma_n}$$

Note that each integer between 1 and n occurs as one of the column indices on the right, so in principle we could order the factors by their column indices rather than their row indices. But then row indices would appear in some other order τ.

$$M_{1\sigma_1}M_{2\sigma_2}\cdots M_{n\sigma_n}=M_{\tau_11}M_{\tau_22}\cdots M_{\tau_nn}$$

It turns out that $\tau=\sigma^{-1}$ as elements of the permutation group. In other words, if we write

$$e=[1,2,\ldots,n]$$

for the identity permutation, then τ and σ will satisfy $\sigma\tau=e$. But then since $\varepsilon(e)=\varepsilon([1,\ldots,n])=1$, we'll have

$$1=\varepsilon(e)=\varepsilon(\sigma\tau)=\varepsilon(\sigma)\,\varepsilon(\tau)$$

which will imply $\varepsilon(\sigma)=\varepsilon(\tau)$ since ε only has the values ±1. Thus,

$$\varepsilon(\sigma)\,M_{1\sigma_1}M_{2\sigma_2}\cdots M_{n\sigma_n}=\varepsilon(\sigma)\,M_{\tau_11}\cdots M_{\tau_nn}=\varepsilon(\tau)\,M_{\tau_11}\cdots M_{\tau_nn}$$

And so now we have a fomula

$$\det(\mathbf{M})=\sum_{\tau\in\mathfrak{S}_n}\varepsilon(\tau)\,M_{\tau_11}\cdot M_{\tau_nn}\tag{(***)}$$

But then

$$\det\left(\mathbf{M}^t\right)=\sum_{\tau\in\mathfrak{S}_n}\varepsilon(\tau)\,M_{1,\tau_1}\cdot M_{n,\tau_n}=\det(\mathbf{M})$$

□

Corollary 16.11. *Let $\mathbf{M}$ be an $n\times n$ matrix and let j be any column index.*

$$\det(\mathbf{M})=\sum_{i=1}^{n}(-1)^{i+j}\det(\mathbf{M}_{ij})$$

Proof. This follows readily from the forrmula (***) by essentially the same arguments we used to prove Proposition 15.7. □

1. Computing Determinants via Row Reduction

THEOREM 16.12. *Suppose* $\mathbf{A}$ *is an* $n \times n$ *matrix in row echelon form. Then*

$$\det(\mathbf{A}) = \prod_{i=1}^{n} A_{ii}$$

that is to say, the determinant of $\mathbf{A}$ *will be the product of its entries along its main diagonal.*

Proof. We will proceed by induction on n. The statement is certainly true for a 1×1 matrix.

Now suppose it is true for $(n-1) \times (n-1)$ matrices. If we do a cofactor expansion of $\det(\mathbf{A})$ along the last row, we'll have

$$\det(\mathbf{A}) = \sum_{j=1}^{n} (-1)^{n+j} A_{nj} \det(\mathbf{A}_{nj})$$

Now because $\mathbf{A}$ is assumed to be in row echelon form only last term will be non-zero (as we'll have $A_{nj} = 0$ when $j < n$). Thus,

$$\det(\mathbf{A}) = (-1)^{n+n} A_{nn} \det(\mathbf{A}_{nn}) = A_{nn} \det(\mathbf{A}_{nn}) \quad .$$

But then if $\mathbf{A}$ is in row echelon form, so will be its n, n minor $\mathbf{A}_{nn}$. Since this minor is an $(n-1) \times (n-1)$ matrix we can apply the inductive hypothesis

$$\det(\mathbf{A}_{nn}) = \text{product of its diagonal entries } = \prod_{i=1}^{n-1} A_{ii}$$

Thus,

$$\det(\mathbf{A}) = A_{nn} \left(\prod_{i=1}^{n-1} A_{ii} \right) = \prod_{i=1}^{n} A_{ii} \quad .$$

□

We can now combine the preceding theorem with Proposition 15.5 and Corollary 15.6 to yield the following algorithm for computing the determinant of an $n \times n$ matrix.

THEOREM 16.13. *Suppose* $\mathbf{A}$ *is an* $n \times n$ *matrix and* $\mathbf{A}'$ *is a matrix in row echelon form obtained from* $\mathbf{A}$ *by a sequence* $\mathcal{R}_1, \mathcal{R}_2, \ldots, \mathcal{R}_k$ *of elementary row operations. Then*

$$\det(\mathbf{A}) = (-1)^s \left(\prod_{i=1}^{t} \lambda_i \right) \left(\prod_{j=1}^{n} A'_{jj} \right)$$

where

(i) *s is the number of row interchanges amongst the operations* $\mathcal{R}_1, \ldots, \mathcal{R}_k$*;*
(ii) *there is a factor λ_i for each elementary row operation that is scalar multiplication of a row by* λ_i*;*
(iii) *the A'_{jj} are the diagonal entries of the row echelon form* $\mathbf{A}'$*.*

This theorem in turn gives us three more important properties of the determinant.

THEOREM 16.14. *Suppose* $\mathbf{A}$ *is an* $n \times n$ *matrix and rank* $(\mathbf{A}) < n$*, then* $\det(\mathbf{A}) = 0$.

Proof. If $\mathbf{A}$ has rank $< n$, then any row echelon form $\mathbf{A}'$ of $\mathbf{A}$ will have all 0's in the last row. But then $A'_{nn} = 0$. Since this A'_{nn} is one of the factors of in the formula of the preceding theorem, we can conclude that $\det(\mathbf{A}) = 0$. □

Finally we come to an extremely useful formula for the determinant of a product of two $n \times n$ matrices.

THEOREM 16.15. *Suppose* $\mathbf{A}$ *and* $\mathbf{B}$ *are invertible* $n \times n$ *matrices. Then*

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}) \quad .$$

Proof. Suppose first that $\mathbf{A}$ has rank n. Then $\mathbf{A}$ is reducible to the $n \times n$ identity matrix by a sequence of elementary row operations

$$\mathbf{A} \xrightarrow{\mathcal{R}_1} \mathbf{A}_1 \xrightarrow{\mathcal{R}_2} \mathbf{A}_2 \xrightarrow{\mathcal{R}_3} \cdots \xrightarrow{\mathcal{R}_k} \mathbf{I}_n$$

These same elementary row operations could be implemented via multiplication by the corresponding elementary matrices:

$$\mathbf{A} \to \mathcal{E}_{\mathcal{R}_1}\mathbf{A} \to \mathcal{E}_{\mathcal{R}_2}\mathcal{E}_{\mathcal{R}_1}\mathbf{A} \to \cdots \to \mathcal{E}_{\mathcal{R}_k}\cdots\mathcal{E}_{\mathcal{R}_2}\mathcal{E}_{\mathcal{R}_1}\mathbf{A} = \mathbf{I}_n \quad .$$

On the other hand, this same sequence or elementary row operation could be applied to the product $\mathbf{AB}$:

$$\mathbf{AB} \xrightarrow{\mathcal{R}_1} (\mathbf{AB})_1 \xrightarrow{\mathcal{R}_2} (\mathbf{AB})_2 \xrightarrow{\mathcal{R}_3} \cdots \xrightarrow{\mathcal{R}_k} (\mathbf{AB})_k$$

or via multiplication via the same sequence of elementatry matrices

$$\mathbf{AB} \to \mathcal{E}_{\mathcal{R}_1}\mathbf{AB} \to \mathcal{E}_{\mathcal{R}_2}\mathcal{E}_{\mathcal{R}_1}\mathbf{AB} \to \cdots \to \mathcal{E}_{\mathcal{R}_k}\cdots\mathcal{E}_{\mathcal{R}_2}\mathcal{E}_{\mathcal{R}_1}\mathbf{AB} = \mathbf{I}_n\mathbf{B} = \mathbf{B} \quad .$$

In other words, the sequence of elementary row operations that converts $\mathbf{A}$ to $\mathbf{I}_n$, will convert the product $\mathbf{AB}$ to $\mathbf{B}$.

Now suppose this sequence elementary row operations $\mathcal{R}_1, \ldots, \mathcal{R}_k$ has s row interchanges and has t row rescalings by factors $\lambda_1, \lambda_2, \ldots, \lambda_t$. Then, on the one hand,

$$\det(\mathbf{AB}) = (-1)^s \left(\prod_{i=1}^{t} \lambda_i\right) \det(\mathbf{B})$$

while, on the other hand, we'll also have

$$\det(\mathbf{A}) = (-1)^s \left(\prod_{i=1}^{t} \lambda_i\right) \det(\mathbf{I}_n) = (-1)^s \left(\prod_{i=1}^{t} \lambda_i\right) \quad .$$

Thus,

$$\det(\mathbf{AB}) = \det(\mathbf{A})\det(\mathbf{B}) \quad .$$

Finally, in the situation where $rank(\mathbf{A}) < n$, it easy to see that $rank(\mathbf{AB}) < n$ (because the columns of the product $\mathbf{AB}$ will be linear combinations of the columns of $\mathbf{A}$ and so if the rows of $\mathbf{A}$ are not linearly independent, the row of $\mathbf{AB}$ can't be linearly indepdent either)

$$0 = \det(\mathbf{AB}) = 0 \cdot \det(\mathbf{B}) = \det(\mathbf{A})\det(\mathbf{B})$$

and thus the stated formula holds in this situation as well. □

COROLLARY 16.16. *Suppose* $\mathbf{A}$ *is an invertible* $n \times n$ *matrix. Then*

$$\det\left(\mathbf{A}^{-1}\right) = \frac{1}{\det(\mathbf{A})} \quad .$$

Proof. We have

$$1 = \det(\mathbf{I}_n) = \det\left(\mathbf{A}^{-1}\mathbf{A}\right) = \det\left(\mathbf{A}^{-1}\right)\det(\mathbf{A}) \quad \Rightarrow \quad \det\left(\mathbf{A}^{-1}\right) = \frac{1}{\det(\mathbf{A})} \quad .$$

□

2. Cramer's Rule

Determinants provide a way of solving certain $n \times n$ linear systems.

THEOREM 16.17 (Cramer's Rule). *Suppose the coefficient matrix* $\mathbf{A}$ *of an* $n \times n$ *linear system* $\mathbf{Ax} = \mathbf{b}$ *has a non-zero determinant. Then the solution of this system is given by*

$$x_i = \frac{\det(\mathbf{A}_i)}{\det(\mathbf{A})} \qquad , \qquad i = 1, 2, \ldots, n$$

where $\mathbf{A}_i$ *is the matrix obtained from* $\mathbf{A}$ *by replacing its* i^{th} *column with the column vector* $\mathbf{b}$.

Proof. Suppose $\mathbf{x} = [x_1, \ldots, x_n]$ is a solution.

$$\mathbf{b} = \mathbf{Ax} = x_1\mathbf{c}_1 + x_2\mathbf{c}_2 + \cdots + x_n\mathbf{c}_n$$

where $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are the columns of $\mathbf{A}$. By assumption, $\det(\mathbf{A}) \neq 0$, and so the column vectors $\mathbf{c}_1, \ldots, \mathbf{c}_n$ are linearly independent. On the other hand,

$$\begin{aligned}
\det(\mathbf{A}_i) &= \det([\mathbf{c}_1, \ldots, \mathbf{c}_{i-1}, \mathbf{b}, \mathbf{c}_{i+1}, \ldots, \mathbf{c}_n]) \\
&= \det(([\mathbf{c}_1, \ldots, \mathbf{c}_{i-1}, \mathbf{x}_1\mathbf{c}_1 + \cdots + x_n\mathbf{c}_n, \mathbf{c}_{i+1}, \ldots, \mathbf{c}_n])) \\
&= \sum_{j=1}^{n} x_j \det([\mathbf{c}_1, \ldots, \mathbf{c}_{i-1}, \mathbf{c}_j, \mathbf{c}_{i+1}, \ldots, \mathbf{c}_n])
\end{aligned}$$

Now whenever $j \neq i$ in the sum we'll be taking the determinant of a matrix with two identical columns - such a determinant will vanish identically since the columns won't be linearly independent. Thus, only the term where $j = i$ will contribut to the sum:

$$\begin{aligned}
\det(\mathbf{A}_i) &= 0 + 0 + \cdots + 0 + x_i \det(\mathbf{c}_1, \ldots, \mathbf{c}_{i-1}, \mathbf{c}_i, \mathbf{c}_{i+1}, \ldots, \mathbf{c}_n) + 0 + \cdots + 0 \\
&= x_i \det(\mathbf{A})
\end{aligned}$$

Since $\det(\mathbf{A}) \neq 0$ we can solve divide both sides of this last equation by $\det(\mathbf{A})$ to get

$$x_i = \frac{\det(\mathbf{A}_i)}{\det(\mathbf{A})} \quad .$$

□

This result also gives a new way of computing the inverse of a matrix. For if $\mathbf{A}$ is invertible, we must also have

$$\mathbf{x} = \mathbf{A}^{-1}\mathbf{b}$$

as the unique solution to $\mathbf{Ax} = \mathbf{b}$. Thus

(*)
$$\mathbf{A}^{-1}\mathbf{b} = \begin{pmatrix} \frac{\det(\mathbf{A}_1)}{\det(\mathbf{A})} \\ \vdots \\ \frac{\det(\mathbf{A}_n)}{\det(\mathbf{A})} \end{pmatrix} = \frac{1}{\det(\mathbf{A})} \begin{pmatrix} \det(\mathbf{A}_1) \\ \vdots \\ \det(\mathbf{A}_n) \end{pmatrix}$$

Now

(**)
$$\det \mathbf{A}_i = \det \begin{pmatrix} a_{11} & b_1 & a_{1n} \\ & b_i & \\ a_{n1} & b_n & a_{nn} \end{pmatrix} = \sum_{j=1}^{n} (-1)^{i+j} b_j \det(\mathbf{A}_{ij})$$

If we define the *adjoint matrix* $\mathbf{C}$ of $\mathbf{A}$ by

(***)
$$(\mathbf{C})_{ij} = (-1)^{i+j} (\det(\mathbf{A}_{ji}))$$

then (**) can be rewritten as

$$\det(\mathbf{A}_i) = \sum_{j=1}^{n} (-1)^{i+j} b_j \det(\mathbf{A}_{ij}) = \sum_{j=1}^{n} (\mathbf{C}_{ij}) b_j = (\mathbf{Cb})_i$$

Inserting these expressions for $\det(\mathbf{A}_i)$ inot the right hand side of (*), then yields

$$\mathbf{A}^{-1}\mathbf{b} = \frac{1}{\det(\mathbf{A})}\mathbf{Cb}$$

We conclude

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\mathbf{C}$$

where $\mathbf{C}$ is the adjoint matrix defined by (***).

Example 16.18. Compute the inverse of a general (invertible) 2×2 matrix.

- Let

$$\mathbf{A} = \begin{pmatrix} a & b \\ c & d \end{pmatrix}$$

We have

$$\begin{array}{rcl} \mathbf{A}_{11} & = & [d] \quad \Rightarrow \quad \mathbf{C}_{11} = (-1)^{1+1} \det([d]) = d \\ \mathbf{A}_{12} & = & [c] \quad \Rightarrow \quad \mathbf{C}_{21} = (-1)^{1+2} \det([c]) = -c \\ \mathbf{A}_{21} & = & [b] \quad \Rightarrow \quad \mathbf{C}_{12} = (-1)^{2+1} \det([b]) = -b \\ \mathbf{A}_{22} & = & [a] \quad \Rightarrow \quad \mathbf{C}_{22} = (-1)^{2+2} \det[a] = a \end{array}$$

So

$$\mathbf{C} = \begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

and thus

$$\mathbf{A}^{-1} = \frac{1}{\det(\mathbf{A})}\mathbf{C} = \frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}$$

Indeed

$$\begin{aligned} \begin{pmatrix} a & b \\ c & d \end{pmatrix}\left(\frac{1}{ad-bc}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix}\right) &= \frac{1}{ad-bc}\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} d & -b \\ -c & a \end{pmatrix} \\ &= \frac{1}{ad-bc}\begin{pmatrix} ad-bc & -ab+ab \\ cd-dc & -cb+da \end{pmatrix} \\ &= \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \end{aligned}$$

3. Key Facts About Determinants

We've proved a number of things about determinants in this lecture. I'd thought it wise to collect here the most salient points.

Fact 16.19. *There are a number of ways to compute determinants:*

- *Via explicit formulas when $n \leq 3$*
- *Using the recursive definition.*
- *Using the sum over permutations of column indices (useful theoretically, rarely practical though)*
- *Using cofactor expansions (useful when a particular row or column has lots of 0's)*
- *Via row reduction to an upper triangular matrix*

Theorem 16.20. *Suppose $\mathbf{A}$ is an $n \times n$ matrix. Then the following statements are equivalent.*

- $\det(\mathbf{A}) \neq 0$
- *The columns of $\mathbf{A}$ are linearly independent.*
- *The rows of $\mathbf{A}$ are linearly independent.*
- *$\mathbf{A}$ has rank n.*
- *For each vector $\mathbf{b} \in \mathbb{F}^n$, there is a unique solution ot $\mathbf{Ax} = \mathbf{b}$.*
- *There are no non-trivial solutions of $\mathbf{Ax} = \mathbf{0}$.*
- *$\mathbf{A}$ is an invertible matrix.*
- *Multiplication by $\mathbf{A}$ is an automorphism of $\mathbb{F}^n$.*

Determinants provide a nice way of circumventing **ugly** row reduction computations

Theorem 16.21. *Solution of $\mathbf{Ax} = \mathbf{b}$ is given by Crammers rule*

$$x_i = \frac{\det(\mathbf{A}_i)}{\det(\mathbf{A})} \quad .$$

Theorem 16.22. *The inverse of an invertible matrix $\mathbf{A}$ is given by*

$$\left(\mathbf{A}^{-1}\right)_{ij} = \frac{(-1)^{i+j}}{\det(\mathbf{A})} \det(\mathbf{A}_{ji})$$

Finally, we have the following very useful (theoretically) facts.

Theorem 16.23. *If $\mathbf{A}$ and $\mathbf{B}$ are $n \times n$ matrices, then*

$$\det(\mathbf{AB}) = \det(\mathbf{A}) \det(\mathbf{B}) \det(\mathbf{B})$$

Corollary 16.24. *If $\mathbf{A}$ is an invertible $n \times n$ matrix, then*

$$det\left(\mathbf{A}^{-1}\right) = \frac{1}{\det(\mathbf{A})}$$