TLDR

Quick summary of today's notes. Lecture starts on next page.

- Let n be any positive integer. The *determinant* is a function $\det : \{n \times n \text{ matrices}\} \to \mathbb{R}$.
- The value of the function det at the $n \times n$ identity matrix is one, so $\det I_n = 1$.
- If A is an $n \times n$ matrix and B is formed by swapping two columns of A, then $\det A = -\det B$:

$$\det \begin{bmatrix} 1 & 2 & 3 \\ 0 & 4 & 5 \\ 0 & 0 & 6 \end{bmatrix} = 24 \qquad \text{and} \qquad \det \begin{bmatrix} 1 & 3 & 2 \\ 0 & 5 & 4 \\ 0 & 6 & 0 \end{bmatrix} = -24.$$

- Choose any vectors $a_2, a_3, \dots, a_n \in \mathbb{R}^n$.

  Then the function $f : \mathbb{R}^n \to \mathbb{R}$ with the formula $f(v) = \det \begin{bmatrix} v & a_2 & a_3 & \dots & a_n \end{bmatrix}$ is linear.

  For example, it turns out that

$$\det \begin{bmatrix} v_1 & 2 & 3 \\ v_2 & 4 & 5 \\ v_3 & 0 & 6 \end{bmatrix} = 24v_1 - 2(6v_2 - 5v_3) + 3(-4v_3) = \underbrace{\begin{bmatrix} 24 & -12 & -2 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ v_3 \end{bmatrix}}_{\text{formula for a linear function}}.$$

- There is only one function $\{n \times n \text{ matrices}\} \to \mathbb{R}$ that has the preceding properties, and we define det to be this function. Although it's not obvious, these properties lead to a formula for det.
- For 2×2 matrices, $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc$.
- For 3×3 matrices, $\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a(ei - fh) - b(di - fg) + c(dh - eg)$.
- Permutation matrices are square matrices formed by rearranging the columns in I_n. For example:

$$\underset{\substack{\det=1 \\ \mathsf{inv}=0}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \quad \underset{\substack{\det=-1 \\ \mathsf{inv}=1}}{\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}} \quad \underset{\substack{\det=-1 \\ \mathsf{inv}=1}}{\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}} \quad \underset{\substack{\det=1 \\ \mathsf{inv}=2}}{\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}} \quad \underset{\substack{\det=1 \\ \mathsf{inv}=2}}{\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}} \quad \underset{\substack{\det=-1 \\ \mathsf{inv}=3}}{\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}}.$$

  Determinants of $n \times n$ permutation matrices are always ± 1.

  Suppose A is a permutation matrix.

  If $\mathsf{inv}(A)$ is the number of 2×2 submatrices of A equal to $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, then $\det A = (-1)^{\mathsf{inv}(A)}$.
- There is a general formula for $\det A$ given by a sum over all $n \times n$ permutation matrices.

  The $n = 2$ and $n = 3$ formulas for $\det A$ are special cases of this formula.

  The general formula is sometimes useful but is probably not worth memorizing.

  We will describe a more efficient way to compute $\det A$ next time.
- If A is not invertible then $\det A = 0$.

  Next time, we will show that if A is invertible then $\det A \neq 0$.

1 Last time: theorems about bases and rank

A *subspace* of $\mathbb{R}^n$ is a nonempty subset H with the property that $u + v \in H$ and $cv \in H$ whenever $u, v \in H$ and $c \in \mathbb{R}$. (Requiring H that be nonempty is equivalent to requiring that $0 \in H$.)

A *basis* of a subspace is a linearly independent set of vectors whose span is the whole subspace.

The plural of "basis" is "bases." Two crucial facts about bases:

- Every subspace has a basis.
- Every basis of a given subspace has the same number of elements.

The *dimension* of a subspace is the common size of all of its bases.

If H is a subspace with $\dim H = p$ then any set of p vectors in H that are linearly independent, or that span H, form a basis for H. The dimension of $\mathbb{R}^n$ is n, while the dimension of $\{0\}$ is 0.

Every subspace H of $\mathbb{R}^n$ that is not $\{0\}$ or $\mathbb{R}^n$ has $0 < \dim H < n$.

Be sure to know how to (1) construct a basis of $\operatorname{Nul} A$ and (2) construct a basis of $\operatorname{Col} A$.

Theorem (Rank theorem)**.** Let A be an $m \times n$ matrix.

1. The dimension of $\operatorname{Nul} A = \{v \in \mathbb{R}^n : Av = 0\}$ is the number of free variables in $Ax = 0$.
2. The dimension of $\operatorname{Col} A =$ (the span of the columns of A) is the number of pivot columns in A.
3. It holds that $\operatorname{rank} A + \dim \operatorname{Nul} A = n$, where we define $\operatorname{rank} A = \dim \operatorname{Col} A$.

(Exercise: why does the third statement follow from the first two?)

Corollary. For an $n \times n$ matrix A, the following are equivalent:

1. A is invertible.
2. $\operatorname{rank} A = n$.
3. $\dim \operatorname{Nul} A = 0$.

If U and V are two sets then $U \subset V$ means that every element of U is also an element of V.

Therefore, the only way that we can have both $U \subset V$ and $V \subset U$ is if $U = V$.

Proposition. Suppose $U \subset V$ are subspaces of $\mathbb{R}^n$. Then $\dim U \leq \dim V$.

Moreover, if $U \neq V$ then $\dim U < \dim V$. Equivalently, if $\dim U = \dim V$, then $U = V$.

Another way of defining a basis of a subspace H of $\mathbb{R}^n$ is as a set of vectors $b_1, b_2, \ldots, b_k$ with the property that if m is any positive integer and $v_1, v_2, \ldots, v_k$ are any vectors in $\mathbb{R}^m$, there there is a unique linear transformation $T : H \to \mathbb{R}^m$ with $T(b_i) = v_i$ for $i = 1, 2, \ldots, k$.

For our applications, it is not essential to know how to prove this. But if you wanted to try: first show that the existence of such a linear transformation T follows from the linear independence of $b_1, b_2, \ldots, b_k$. Then check that T is unique exactly when $b_1, b_2, \ldots, b_k$ span H.

If $\mathcal{B} = (b_1, b_2, \ldots, b_k)$ is an ordered basis of a k-dimensional subspace H, then we define

$$[\cdot]_{\mathcal{B}} : H \to \mathbb{R}^k$$

as the unique linear function with $[b_i]_{\mathcal{B}} = e_i \in \mathbb{R}^k$ for $i = 1, 2, \ldots, k$.

Recall that $e_i \in \mathbb{R}^k$ is the vector with 1 in row i and 0 in all other rows.

We call $[v]_{\mathcal{B}}$ the *coordinate vector* of $v \in H$ in the basis $\mathcal{B}$.

2 Determinants

The subject of the next few lectures is the *determinant* of a square matrix. The determinant is important in various parts of math and physics, for example, in computing integrals in multivariable calculus.

The following theorem says that a set of three special properties uniquely identifies the determinant among all functions on $n \times n$ matrices. This result tells us some important facts about the determinant, but it's not clear at first how we are supposed to compute this function.

Theorem. Let n be any positive integer. There exists a unique function

$$\det : \{n \times n \text{ matrices}\} \to \mathbb{R},$$

called the *determinant*, with the following properties:

(1) $\det I_n = 1$. In words: the determinant of the identity matrix is 1.

(2) If $a_1, a_2, \dots, a_n \in \mathbb{R}^n$ and $1 \leq i < j \leq n$ then

$$\det \begin{bmatrix} a_1 & \cdots & a_i & \cdots & a_j & \cdots & a_n \end{bmatrix} = -\det \begin{bmatrix} a_1 & \cdots & a_j & \cdots & a_i & \cdots & a_n \end{bmatrix}$$

In words: interchanging two columns in an $n \times n$ matrix reverses the sign of the determinant.

(3) Choose any vectors $a_2, a_3, \dots, a_n \in \mathbb{R}^n$. If $u, v \in \mathbb{R}^n$ and $c \in \mathbb{R}$ then

$$\det \begin{bmatrix} u+v & a_2 & a_3 & \cdots & a_n \end{bmatrix} = \det \begin{bmatrix} u & a_2 & a_3 & \cdots & a_n \end{bmatrix} + \det \begin{bmatrix} v & a_2 & a_3 & \cdots & a_n \end{bmatrix},$$
$$\det \begin{bmatrix} cv & a_2 & a_3 & \cdots & a_n \end{bmatrix} = c \cdot \det \begin{bmatrix} v & a_2 & a_3 & \cdots & a_n \end{bmatrix}.$$

In words: if all but the first column of an $n \times n$ matrix are fixed, and the determinant is viewed as a function of the first column, then we get a linear function $\mathbb{R}^n \to \mathbb{R}$.

This is a very abstract way of defining a function. However, the upshot is that later this description will make it easy to derive several useful, but not obviously equivalent, formulas for the determinant.

We spend the rest of this lecture proving the theorem. To do this, we start by *assuming there exists a function* det *with the given properties*. We will use these properties to narrow the possibilities for det down to one function given by a certain formula, and then check that this formula does satisfy (1)-(3).

Let $A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$ be an $n \times n$ matrix with columns $a_1, a_2, \dots, a_n \in \mathbb{R}^n$.

Lemma. If A has two equal columns then $\det A = 0$.

Proof. Suppose $a_i = a_j$ for $i < j$.

Then $\det A = -\det \begin{bmatrix} a_1 & \cdots & a_j & \cdots & a_i & \cdots & a_n \end{bmatrix} = -\det A$ so $2(\det A) = 0$ and $\det A = 0$. □

Corollary. If the columns of A are not linearly independent, then $\det A = 0$.

Proof. Assume the columns of A are not linearly independent.

This means that one column a_i of A is a linear combination of the others.

If $a_1 = c_2a_2 + c_3a_3 + \cdots + c_na_n$ for some $c_2, c_3, \dots, c_n \in \mathbb{R}$, then property (3) implies that

$$\det A = c_2 \det \begin{bmatrix} a_2 & a_2 & \dots & a_n \end{bmatrix} + c_3 \det \begin{bmatrix} a_3 & a_2 & a_3 & \dots & a_n \end{bmatrix} + \cdots + c_n \det \begin{bmatrix} a_n & a_2 & \dots & a_n \end{bmatrix}.$$

Each determinant in the sum is zero by the previous lemma so $\det A = 0$.

If a different column of A is a linear combination of the others, then define B by swapping that column and the first column of A. Then $\det A = -\det B$ and the argument just given shows that $\det B = 0$. □

Corollary. If A is not invertible then $\det A = 0$.

Proof. If A is not invertible then its columns are not linearly independent. $\square$

Recall that $A = \begin{bmatrix} a_1 & a_2 & \dots & a_n \end{bmatrix}$ where each $a_i \in \mathbb{R}^n$.

Lemma. Suppose $1 \le i \le n$ and $a_i = pu + qv$ for some $p, q \in \mathbb{R}$ and $u, v \in \mathbb{R}^n$. Define

$$B = \begin{bmatrix} a_1 & \dots & a_{i-1} & u & a_{i+1} & \dots & a_n \end{bmatrix} \quad \text{and} \quad C = \begin{bmatrix} a_1 & \dots & a_{i-1} & v & a_{i+1} & \dots & a_n \end{bmatrix}.$$

Then $\det A = p \cdot \det B + q \cdot \det C$.

In other words, det is linear as a function of any single column in a matrix, not just the first.

Proof. If $i = 1$ then this follows by property (3) in the theorem defining det.

If $i > 1$ then form A', B', and C' by swapping columns 1 and i in A, B, and C.

Then $\det A' = p \cdot \det B' + q \cdot \det C'$ by property (3).

Now substitute $\det A' = -\det A$ and $\det B' = -\det B$ and $\det C' = -\det C$ and cancel signs. $\square$

Example. For 1×1 matrices we have $\det \begin{bmatrix} a \end{bmatrix} = a \det \begin{bmatrix} 1 \end{bmatrix} = a$.

Example. For 2×2 matrices we have

$$\begin{aligned}
\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} &= \det \begin{bmatrix} \begin{bmatrix} a \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ c \end{bmatrix} & \begin{bmatrix} b \\ 0 \end{bmatrix} + \begin{bmatrix} 0 \\ d \end{bmatrix} \end{bmatrix} \\
&= \det \begin{bmatrix} a & b \\ 0 & 0 \end{bmatrix} + \det \begin{bmatrix} a & 0 \\ 0 & d \end{bmatrix} + \det \begin{bmatrix} 0 & b \\ c & 0 \end{bmatrix} + \det \begin{bmatrix} 0 & 0 \\ c & d \end{bmatrix} \\
&= ab \underbrace{\det \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}}_{=0} + ad \underbrace{\det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}}_{=\det I_2 = 1} + bc \underbrace{\det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}}_{=-\det I_2 = -1} + cd \underbrace{\det \begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}}_{=0} = ad - bc.
\end{aligned}$$

The first equality just rewrites the two columns of our first matrix as sums of simpler vectors.

The second and third equalities follow by extensive use of the previous lemma.

A formula to remember: $\boxed{\det \begin{bmatrix} a & b \\ c & d \end{bmatrix} = ad - bc}$

3 Permutation matrices

To continue, we need to discuss a family of square matrices whose determinants are easy to compute.

A *permutation matrix* is an $n \times n$ matrix whose entries are all 0 or 1, and which has exactly one nonzero entry in each row and in each column. Let S_n be the set of $n \times n$ permutation matrices.

Example. The elements of S_2 are $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ and $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

Example. The elements of S_3 are

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \quad \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix} \quad \begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}.$$

Let R_n be the set of $n \times n$ matrices whose entries are all 0 or 1, and which have exactly one nonzero entry in each column (but possibly multiple nonzero entries in a given row).

Example. The elements of R_2 are $\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$, $\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix}$, $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$, and $\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix}$.

Note that $S_n \subset R_n$. The size of S_n is $n!$ while the size of R_n is n^n.

Lemma. If $X \in R_n$ but $X \notin S_n$ then $\det X = 0$.

Proof. In this case X must have two equal columns. □

Given $X \in S_n$, define $\mathsf{inv}(X)$ to be the number of 2×2 submatrices of X equal to $\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$.

To form a 2×2 submatrix of X, choose any two rows and any two columns, not necessarily adjacent, and then take the 4 entries in those rows and columns.

Equivalently, $\mathsf{inv}(X)$ is the number of pairs of 1s in X with one 1 below and to the left of the other:

$$\mathsf{inv}\left(\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}\right) = 2, \qquad \mathsf{inv}\left(\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}\right) = 0, \qquad \mathsf{inv}\left(\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}\right) = 3.$$

Lemma. If $X \in S_n$ then $\det X = (-1)^{\mathsf{inv}(X)}$.

Proof. If $X \in S_n$ and $\mathsf{inv}(X) > 0$, then X must have two adjacent columns where the 1 in the left column is below the 1 in the right column. Form Y by interchanging these two columns.

Drawing a picture of X and Y shows that $\mathsf{inv}(Y) = \mathsf{inv}(X) - 1$. We know that $\det Y = -\det X$.

If $\mathsf{inv}(Y) > 0$, then construct a permutation matrix Z from Y in the same way. Continuing this process will eventually give a permutation matrix $A \in S_n$ with $\det X = (-1)^{\mathsf{inv}(X)} \det A$ and $\mathsf{inv}(A) = 0$.

The only permutation matrix $A \in S_n$ with $\mathsf{inv}(A) = 0$ is $A = I_n$, so $\det(X) = (-1)^{\mathsf{inv}(X)}$. □

4 A formula for $\det A$

Given a matrix $X \in R_n$ and an arbitrary $n \times n$ matrix A, define

$$\mathsf{prod}(X, A) = \text{ the product of the entries of } A \text{ in the nonzero positions of } X.$$

For example, $\mathsf{prod}\left(\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}, \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}\right) = cdh.$

We can now give the first *concrete* (though still complicated) description of the determinant.

Theorem. Suppose A is an $n \times n$ matrix. Then $\boxed{\det A = \sum_{X \in S_n} \mathsf{prod}(X, A)(-1)^{\mathsf{inv}(X)}}$

Here the notation $\sum_{X \in S_n}$ means "compute $\mathsf{prod}(X, A)(-1)^{\mathsf{inv}(X)}$ for each $n \times n$ permutation matrix X and then take the sum of all of the resulting numbers."

The function given by this formula has the defining properties of the determinant. This confirms our first theorem: the only function with the properties we originally ascribed to the determinant is this formula.

Example. We can use the general formula for $\det A$ to compute the determinant of a 3×3 matrix.

Suppose $A = \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix}$. Then our formula becomes

$$\begin{aligned}\det A = \ & \mathsf{prod}\left(\begin{bmatrix} 1 & & \\ & 1 & \\ & & 1 \end{bmatrix}, A\right)(-1)^0 + \mathsf{prod}\left(\begin{bmatrix} 1 & & \\ & & 1 \\ & 1 & \end{bmatrix}, A\right)(-1)^1 + \\ & \mathsf{prod}\left(\begin{bmatrix} & 1 & \\ 1 & & \\ & & 1 \end{bmatrix}, A\right)(-1)^1 + \mathsf{prod}\left(\begin{bmatrix} & 1 & \\ & & 1 \\ 1 & & \end{bmatrix}, A\right)(-1)^2 + \\ & \mathsf{prod}\left(\begin{bmatrix} & & 1 \\ 1 & & \\ & 1 & \end{bmatrix}, A\right)(-1)^2 + \mathsf{prod}\left(\begin{bmatrix} & & 1 \\ & 1 & \\ 1 & & \end{bmatrix}, A\right)(-1)^3 = aei - afh - bdi + bfg + cdh - ceg.\end{aligned}$$

The 0s are omitted in the permutation matrices to improve readability. We can rewrite this as

$$\boxed{\det \begin{bmatrix} a & b & c \\ d & e & f \\ g & h & i \end{bmatrix} = a(ei - fh) - b(di - fg) + c(dh - eg)}$$

Note that each term in parentheses is the determinant of a 2×2 submatrix of A.

Proof of theorem. The most difficult part of the proof is our notation, which gets fairly complicated.

Suppose $A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \vdots & & & \vdots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}$. Then $A = \left[\ \sum_{i=1}^n a_{i1}e_i \quad \sum_{j=1}^n a_{j2}e_j \quad \dots \quad \sum_{k=1}^n a_{kn}e_k\ \right]$.

In words: express each column of A as a linear combination of the basis vectors $e_1, e_2, \dots, e_n$ of $\mathbb{R}^n$.

Using the fact that the determinant is linear as a function of each column of A, it follows that

$$\begin{aligned}\det A &= \det \left[\ \sum_{i=1}^n a_{i1}e_i \quad \sum_{j=1}^n a_{j2}e_j \quad \dots \quad \sum_{k=1}^n a_{kn}e_k\ \right] \\ &= \sum_{i=1}^n a_{i1}\cdot \det \left[\ e_i \quad \sum_{j=1}^n a_{j2}e_j \quad \dots \quad \sum_{k=1}^n a_{kn}e_k\ \right] \\ &= \sum_{i=1}^n \sum_{j=1}^n a_{i1}a_{j2}\cdot \det \left[\ e_i \quad e_j \quad \dots \quad \sum_{k=1}^n a_{kn}e_k\ \right] \\ &\ \vdots \\ &= \underbrace{\sum_{i=1}^n \sum_{j=1}^n \cdots \sum_{k=1}^n}_{n \text{ summations}} \underbrace{a_{i1}a_{j2}\cdots a_{kn}}_{=\mathsf{prod}(X,A)} \det \underbrace{\left[\ e_i \quad e_j \quad \dots \quad e_k\ \right]}_{\text{this is a matrix } X \in R_n}.\end{aligned}$$

If this sequence of equalities is confusing, try to see if the corresponding step in our calculation of $\det \begin{bmatrix} a & b \\ c & d \end{bmatrix}$ makes more sense. We are really just generalising that calculation from 2 to n dimensions.

Key observation: the matrix $\left[\ e_i \quad e_j \quad \dots \quad e_k\ \right]$ varies over all elements of R_n as the indices $i, j, \dots, k$ vary in the summations $\sum_{i=1}^n \sum_{j=1}^n \cdots \sum_{k=1}^n$.

This means we can rewrite the last formula as $\det A = \sum_{X\in R_n} \mathsf{prod}(X, A) \det X$.

Let $X \in R_n$. Then $\det X = (-1)^{\mathsf{inv}(X)}$ if $X \in S_n$ and otherwise $\det X = 0$. Therefore, we actually have

$$\det A = \sum_{X \in S_n} \mathsf{prod}(X, A)(-1)^{\mathsf{inv}(X)}. \qquad (*)$$

This formula was computed *under the assumption that a function* det *exists with the properties in our first theorem*. This means that if our first theorem is true, then the determinant must be given by the formula (*). To finish, we just need to check *that the function (*) actually has properties (1)-(3).*

This is not too hard, and mostly involves some exercises in algebra manipulating the expression (*):

(1) We have $\det I_n = \sum_{X \in S_n} \mathsf{prod}(X, I_n)(-1)^{\mathsf{inv}(X)} = 1$.

Proof. This holds since $\mathsf{prod}(X, I_n) = 0$ unless $X = I_n$ if $X \in S_n$. □

(2) If we interchange two columns in A then $\det A$ changes by a factor of -1.

Proof. Let $\tilde{X}$ be the matrix given by interchanging columns i and j in X.

If $X \in S_n$ then $\tilde{X} \in S_n$ and $\mathsf{inv}(\tilde{X}) - \mathsf{inv}(X)$ is an odd number.

This is not obvious but can be shown by drawing a picture of X compared to $\tilde{X}$.

Hence $(-1)^{\mathsf{inv}(X)} = -(-1)^{\mathsf{inv}(\tilde{X})}$.

If $X \in S_n$ then $\mathsf{prod}(X, A) = \mathsf{prod}(\tilde{X}, \tilde{A})$ for all matrices A. (Why?)

Thus $\det A = \sum_{X \in S_n} \mathsf{prod}(X, A)(-1)^{\mathsf{inv}(X)} = -\sum_{X \in S_n} \mathsf{prod}(\tilde{X}, \tilde{A})(-1)^{\mathsf{inv}(\tilde{X})} = -\det \tilde{A}$. □

(3) The formula (*) is linear as a function $\mathbb{R}^n \to \mathbb{R}$ of the first column of A.

Proof. Suppose the first column of A is the vector $v = \begin{bmatrix} v_1 \\ v_2 \\ \vdots \\ v_n \end{bmatrix}$ where each v_i is a variable.

Assume all the other columns of A are fixed vectors.

Let $X \in S_n$. Suppose the one in the first column of X is in row i.

Form $Y \in S_{n-1}$ from X by deleting the first column and ith row.

Form B from A by deleting the first column and ith row.

Then $\mathsf{prod}(X, A)(-1)^{\mathsf{inv}(X)} = cv_i$ where $c = (-1)^{\mathsf{inv}(X)}\mathsf{prod}(Y, B) \in \mathbb{R}$.

The formula $v \mapsto cv_i$ is a linear function $\mathbb{R}^n \to \mathbb{R}$ whenever $c \in \mathbb{R}$.

Sums of linear functions $\mathbb{R}^n \to \mathbb{R}$ are linear.

Hence the formula (*) is linear as function of the first column of A. □

This confirms that (*) does have the properties we stated in our first theorem. □

The formula $\det A = \sum_{X \in S_n} \mathsf{prod}(X, A)(-1)^{\mathsf{inv}(A)}$ is not an efficient way of computing the determinant of most matrices since the sum involves a huge number of terms if n is large.

There are 2 terms for $n = 2$, 6 for $n = 3$, 24 terms for $n = 4$, and 120 terms for $n = 5$.

Next time: more properties of determinants and how to compute them efficiently.

5 Vocabulary

Keywords from today's lecture:

1. **Permutation matrix**.

   A square matrix P whose entries are each 0 or 1, that has exactly one nonzero entry equal to 1 in each row and each column.

   If P is an $n \times n$ permutation matrix and A is a matrix with n rows then PA is a matrix formed by rearranging ("permuting") the rows of A. If A is a matrix with n columns then AP is a matrix formed by rearranging the columns of A.

   Example: $\left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{array}\right], \left[\begin{array}{ccc} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{array}\right], \left[\begin{array}{ccc} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{array}\right], \left[\begin{array}{ccc} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{array}\right], \left[\begin{array}{ccc} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{array}\right]$, or $\left[\begin{array}{ccc} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{array}\right]$.

2. **Determinant**.

   The unique function
   $$\det : \{n \times n \text{ matrices}\} \to \mathbb{R}$$
   with $\det I_n = 1$, with the property that interchanging two columns in an $n \times n$ matrix A reverses the sign of $\det A$, and with the property that if all but the first column in an $n \times n$ matrix A are fixed, then $\det A$ is a linear function of the first column.

   Example: $\det \left[\begin{array}{cc} a & b \\ c & d \end{array}\right] = ad - bc$.