Lecture 20

Convention:

For any integer n, $\mathbb{R}^n$ will denote the set $M_{n\times 1}(\mathbb{R})$.

We will also call this the set of column vectors with n entries.

(Elements of $M_{1\times n}(\mathbb{R})$ are called row vectors with n entries).

$\mathbb{R}^n$ is our prototypical example of something called a vector space.

Recall that the set $\mathbb{R}^n$ has a binary operation called addition under which it is an abelian group.

$\mathbb{R}^n$ also has an operation called <u>scaler multiplication</u> which is a function

$$\mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$$

$$\left(\alpha, \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}\right) \longmapsto \begin{bmatrix} \alpha x_1 \\ \alpha x_2 \\ \vdots \\ \alpha x_n \end{bmatrix}$$

(This is <u>not</u> a binary operation on $\mathbb{R}^n$.)

Vector space

A (real) vector space, or a vector space over $\mathbb{R}$ is a triple $(V, +, \cdot)$ where

- V is a set
- $+$ is a binary operation on V
- $\cdot$ is a function $\mathbb{R} \times V \to V$, such that

(a) V is an abelian group under $+$.

(b) $1 \cdot v = v \quad \forall \quad v \in V$.

(c) $(\alpha+\beta)\cdot v = \alpha v + \beta v$
$\forall \ \alpha, \beta \in \mathbb{R}$ and $\forall v \in V$

(d) $\alpha \cdot (v_1 + v_2) = \alpha \cdot v_1 + \alpha \cdot v_2$
$\forall \ \alpha \in \mathbb{R}$ and $\forall \ v_1, v_2 \in V$.

Convention

If the operations $+$ and $\cdot$ are known or implicitly understood, we may refer to "the vector space V" instead of writing the triple $(V, +, \cdot)$.

Remark: We may replace $\mathbb{R}$ by $\mathbb{C}$ or $\mathbb{Q}$ to define vector spaces over $\mathbb{C}$ or $\mathbb{Q}$ respectively.

More generally, we can define vector spaces over any "field". (Fields are sets with two binary operations $+$ and $\times$ satisfying certain properties.)

When we study groups, we study functions $G_1 \longrightarrow G_2$ between groups which "respect the binary operation."

When we study vector spaces, we study functions $V_1 \longrightarrow V_2$ which "respect" addition and scaler multiplication.

Linear transformation

Let V_1 and V_2 be vector spaces. A linear map or a linear transformation is a function $\varphi : V_1 \to V_2$ such that the following two conditions hold:

(1) φ is a group homomorphism with respect to addition.

(2) $\varphi(\alpha \cdot v) = \alpha \cdot \varphi(v)$
for all $\alpha \in \mathbb{R}$ and $v \in V$.

Example

Let A be an $m \times n$ matrix.
Then, the function $\varphi : \mathbb{R}^n \longrightarrow \mathbb{R}^m$
defined by $\varphi(v) = A \cdot v$
is a linear transformation.

We will come back to linear transformations later...

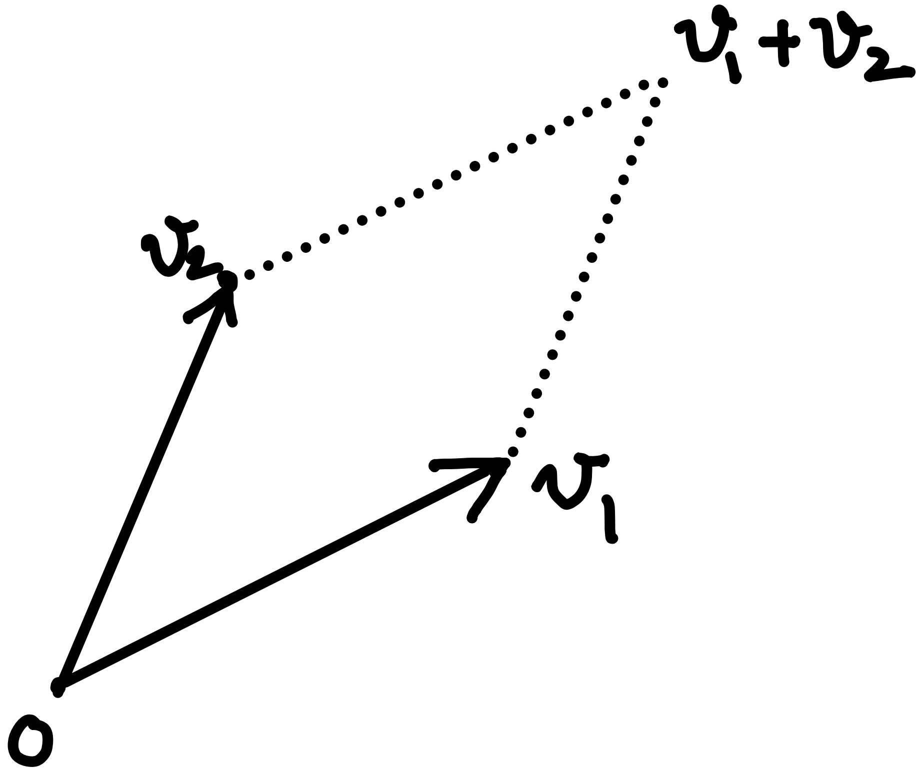
Today we will use the concept of "linearity" of a function as we study <u>determinants</u>.

An $n \times n$ matrix A can be viewed as an <u>ordered</u> collection of n column vectors (an "n-tuple" of column vectors).

$$\begin{bmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{n1} & a_{n2} & & a_{nn} \end{bmatrix}$$

For example, when $n=2$, a 2×2 matrix gives us a pair of vectors in $\mathbb{R}^2$, which we visualize as a plane. A vector in $\mathbb{R}^2$ is just a point in $\mathbb{R}^2$, but we represent it as an arrow from 0 to the point.

$$A = \begin{bmatrix} 2 & 1 \\ 1 & 2 \end{bmatrix} \qquad v_1 = \begin{bmatrix} 2 \\ 1 \end{bmatrix} \qquad v_2 = \begin{bmatrix} 1 \\ 2 \end{bmatrix}$$



Consider the parallelogram with vertices 0, v_1, v_2 and $v_1 + v_2$.

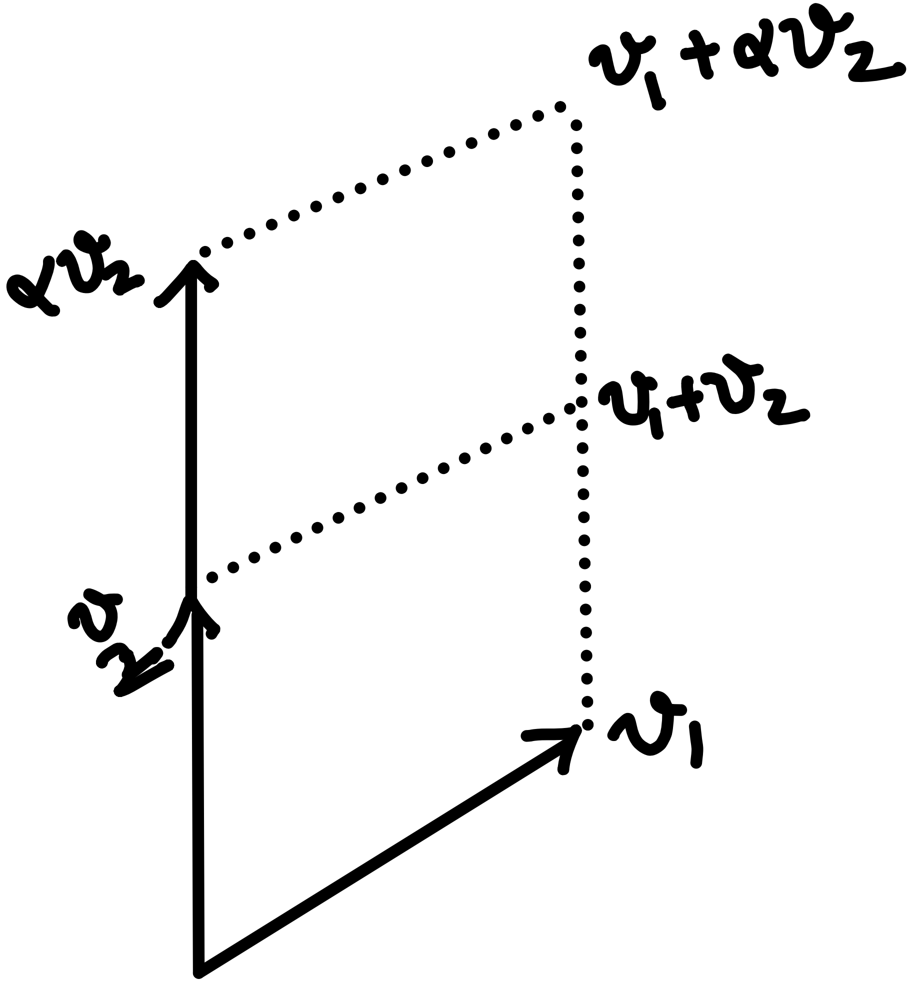
Area of parallelogram
$=$ (length of a side) $\times$
(distance from that side to the opposite side).

Let us define a function

$$\text{Area} : M_{2\times 2}(\mathbb{R}) \to \mathbb{R}$$

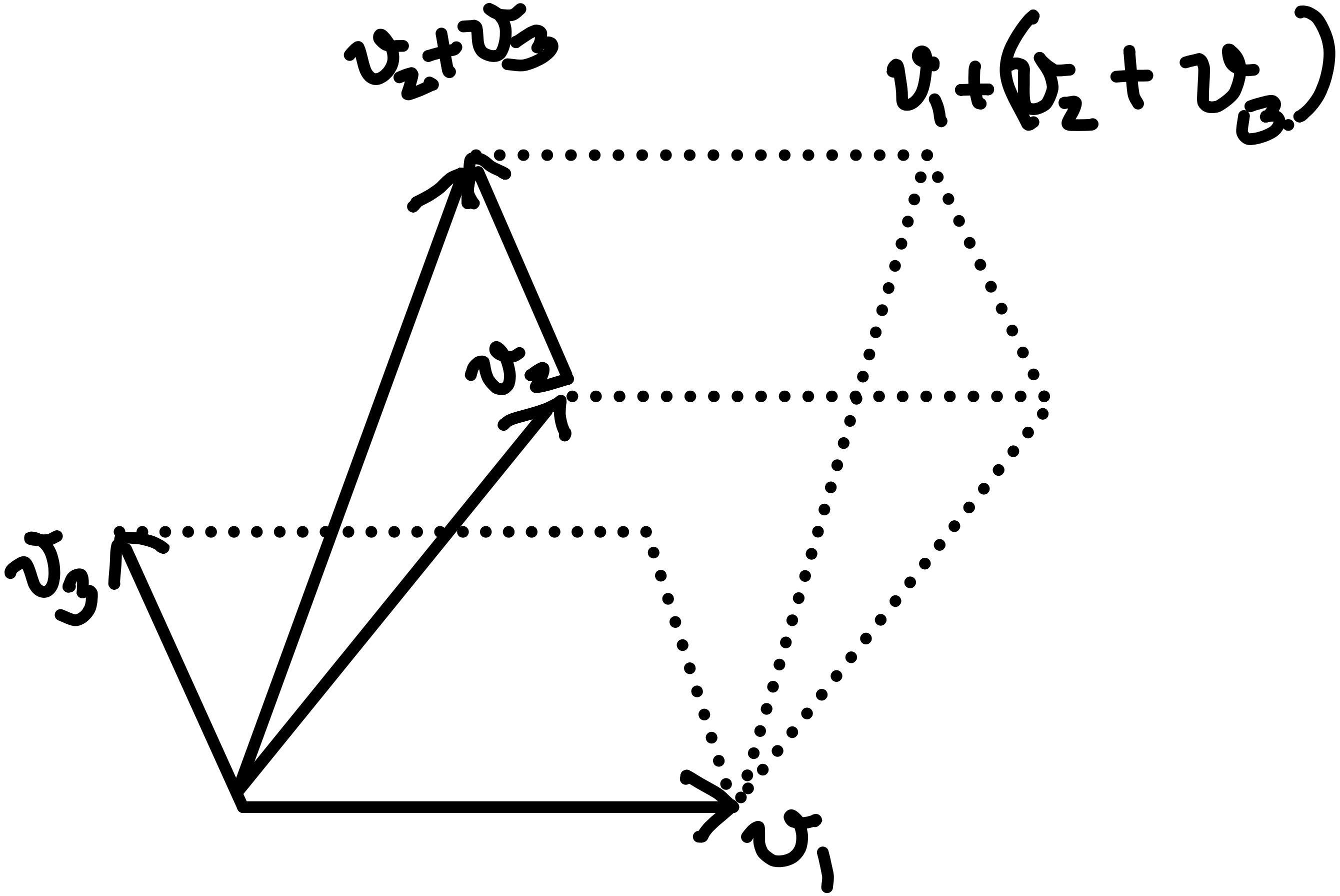
$A = [v_1, v_2] \longmapsto$ area of parallelogram with vertices 0, v_1, v_2 and $v_1 + v_2$.

This function seems to be "linear" in the variables v_1 and v_2.



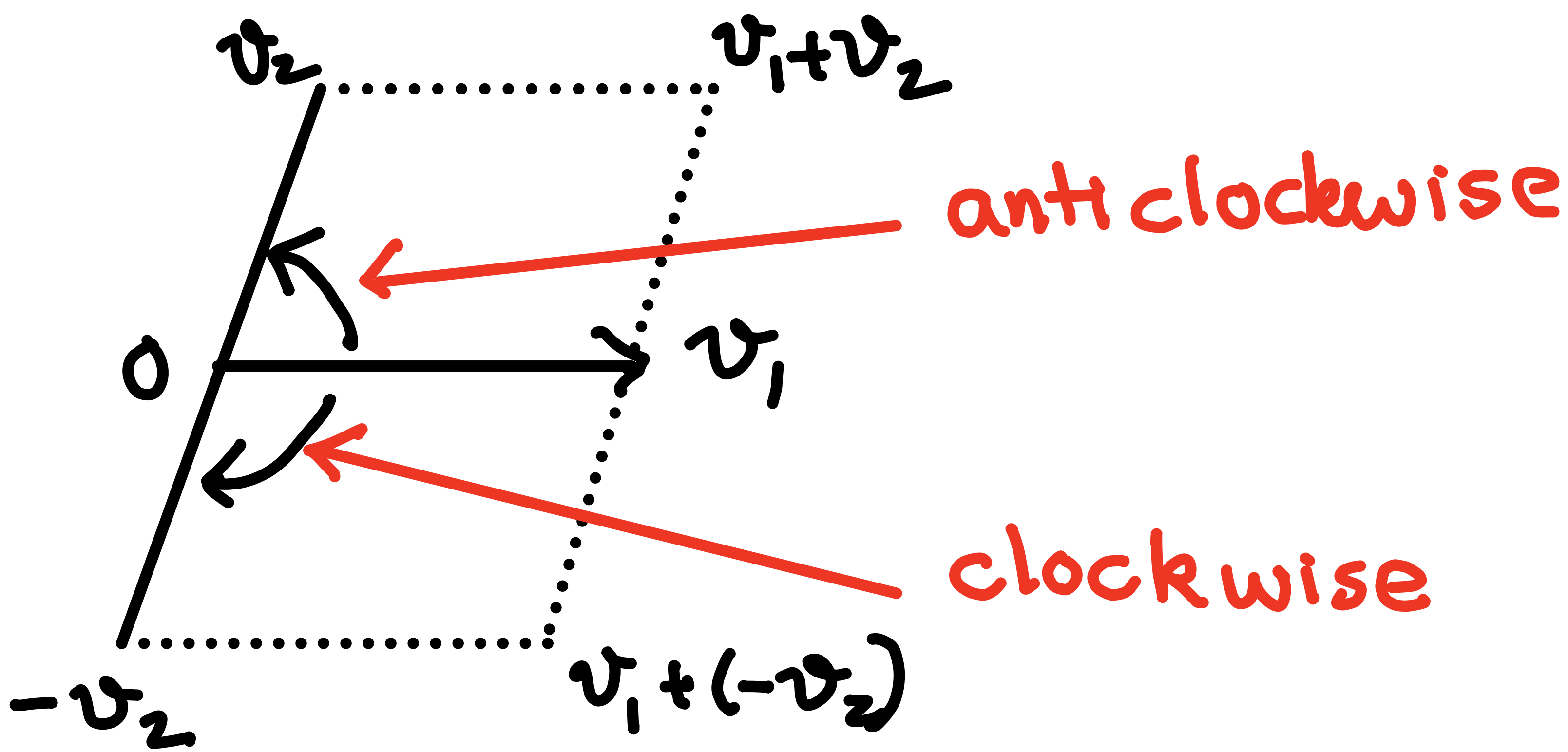
$$\text{Area}(v_1, \alpha v_2) = \alpha \cdot \text{Area}(v_1, v_2)$$

$$\mathrm{Area}(v_1, v_2 + v_3) = \mathrm{Area}(v_1, v_2) + \mathrm{Area}(v_1, v_3).$$



However, for area to be genuinely linear in each variable, it cannot always be a positive number. Area will have a "sign". In other words, some pairs of vectors v_1, v_2 will enclose parallelograms with negative area.

In fact, for linearity, we need $\text{Area}(v_1, -v_2) = -\text{Area}(v_1, v_2)$



Also, for any v we should have $\text{Area}(v, v) = 0$

(The "parallelogram" with vertices 0, v, v and $v+v$ is "flat".)

So for any vectors v and w, we have

$$\text{Area}(v+w, v+w) = 0.$$

But

$$\begin{aligned}
&\text{Area}(v+w,\ v+w) \\
&= \text{Area}(v,\ v+w) + \text{Area}(w,\ v+w) \\
&= \text{Area}(v,v) + \text{Area}(v,w) \\
&\quad + \text{Area}(w,v) + \text{Area}(w,w).
\end{aligned}$$

$\text{Area}(v,v) = 0$ and $\text{Area}(w,w) = 0$

So $\text{Area}(v,w) = -\text{Area}(w,v)$.

So, a "good" area function should satisfy:

(1) $\text{Area}(v_1 + v_2, w) = \text{Area}(v_1, w) + \text{Area}(v_2, w)$

(2) $\text{Area}(v, w_1 + w_2) = \text{Area}(v, w_1) + \text{Area}(v, w_2)$

(3) $\text{Area}(\alpha v, w) = \alpha \, \text{Area}(v, w)$

$\text{Area}(v, \alpha w) = \alpha \, \text{Area}(v, w)$

(4) $\text{Area}(v, w) = - \text{Area}(w, v)$.

Similar arguments hold in higher dimensions.

Given an $n \times n$ matrix, we view it as an n-tuple of column vectors $v_1, v_2, \cdots v_n$. These vectors enclose an "n-dimensional box" with vertices $a_1 v_1 + a_2 v_2 + \cdots + a_n v_n$ where $a_i = 0$ or 1.

We want a "volume" function

$$\mathrm{Vol}_n : M_{n\times n}(\mathbb{R}) \longrightarrow \mathbb{R}$$

$$[v_1, v_2 \cdots v_n] \longmapsto \mathrm{Vol}_n(v_1, \cdots v_n).$$

This function should be <u>linear in each variable.</u>

If two vectors v_i, v_j interchange positions, the volume should change sign.

Unit of measurement

An n-dimensional cube with each side equal to 1 has volume 1.

So if $e_i = \begin{bmatrix} 0 \\ \vdots \\ 1 \\ \vdots \\ 0 \end{bmatrix}$ (1 in row i and 0 in othe rows)

then $Vol_n(e_1, e_2 \cdots e_n) = 1$.

So, does such a function exist?

Yes.

We will assume that it exists and derive a formula. (We should then check if the formula has the required properties. — omitted for this course.)

This volume function is called the determinant.

For any n, we have the function $\det : M_{n \times n}(\mathbb{R}) \longrightarrow \mathbb{R}$.

Let us assume that this is a function with all the properties we listed and try to get a formula.

$\underline{n=2}$

$$A = \begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix}$$

$$\det(A) = \det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & 0 \end{bmatrix} + \det\begin{bmatrix} a_{11} & 0 \\ a_{21} & a_{22} \end{bmatrix}$$

$$= \det\begin{bmatrix} a_{11} & a_{12} \\ 0 & 0 \end{bmatrix} + \det\begin{bmatrix} 0 & a_{12} \\ a_{21} & 0 \end{bmatrix}$$

$$+ \det\begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} + \det\begin{bmatrix} 0 & 0 \\ a_{21} & a_{22} \end{bmatrix}$$

$$\det\begin{bmatrix} a_{11} & a_{12} \\ 0 & 0 \end{bmatrix} = a_{11}a_{12} \det\begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} = 0$$

$$\det\begin{bmatrix} 0 & a_{12} \\ a_{21} & 0 \end{bmatrix} = a_{12}a_{21} \det\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$$

$$\det\begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix} = a_{11}a_{22} \det\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

$$\det\begin{bmatrix} 0 & 0 \\ a_{21} & a_{22} \end{bmatrix} = a_{21}a_{22} \det\begin{bmatrix} 0 & 0 \\ 1 & 1 \end{bmatrix} = 0$$

So $\det\begin{bmatrix} a_{11} & a_{12} \\ a_{21} & a_{22} \end{bmatrix} = \det\begin{bmatrix} a_{11} & 0 \\ 0 & a_{22} \end{bmatrix}$

$+\det\begin{bmatrix} 0 & a_{12} \\ a_{21} & 0 \end{bmatrix}$

(only matrices with one entry in each row and column survive)

$= a_{11}a_{22}\det\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + a_{12}a_{21}\det\begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix}$

But we want $\det\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = 1$

So $\det \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} = -\det \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} = -1$

So $\det A = a_{11} a_{22} - a_{12} a_{21}$.

(You can easily check this has all required properties.)

Similarly

$$\det\begin{bmatrix} a_{11} & a_{12} & a_{13} \\ a_{21} & a_{22} & a_{23} \\ a_{31} & a_{32} & a_{33} \end{bmatrix} = a_{11}a_{22}a_{33}\det\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$+ a_{11}a_{23}a_{32}\det\begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix}$$

$$+ a_{12}a_{23}a_{31}\det\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

$$+ a_{12}a_{21}a_{33}\det\begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

$$+ a_{13}a_{21}a_{32} \cdot \det\begin{bmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{bmatrix}$$

$$+ a_{13}\, a_{22}\, a_{31} \cdot \det\begin{bmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$$

Let us understand each term.

Pick a permutation σ of $\{1,2,3\}$.

For example: $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$

Take terms of the form $a_{i\sigma(i)}$

So: a_{12}, a_{23} and a_{31}.

Use σ to permute colums of I_3: $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

Term in det A : $a_{12}a_{23}a_{31} \det \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$

A matrix obtained by permuting the columns of I_n is called a permutation matrix.

How do we calculate the determinant of a permutation matrix?

Easy observation

Any permutation can be written as a product of 2-cycles (also called transposition).

To see this, write any permutation as a product of cycles. Given a cycle $(a_1, a_2, \cdots a_r)$ we see that

$$(a_1, a_2 \cdots a_r) = (a_1, a_r) \cdots (a_1, a_3)(a_1, a_2)$$

So, we start with I_n and switch columns two at a time. Each such switch contributes a minus sign.

So, if total transpositions used is k, the determinant is $(-1)^k$.

We say that $(-1)^k$ is the sign of σ (denoted $\mathrm{sgn}(\sigma)$)

For example. $\sigma = \begin{pmatrix} 1 & 2 & 3 \\ 2 & 3 & 1 \end{pmatrix}$

$\sigma = (1,3)(1,2)$

$$\begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \xrightarrow[(1,2)]{} \begin{bmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

det = 1   det = −1

$$\xrightarrow[(1,3)]{} \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

det = 1

So $\text{sgn}(\sigma) = 1$.

So, we see that given an $n \times n$ matrix $A = (a_{ij})_{i,j}$

$$\det(A) = \sum_{\sigma \in S_n} \left(\operatorname{sgn}(\sigma)\, a_{1\sigma(1)} a_{2\sigma(2)} \cdots a_{n\sigma(n)} \right)$$

(We have skipped a lot of proofs! But... that's okay.)