Lecture 2

Identity: A is a 3×3 matrix. Then

$$\varepsilon_{mnp}(\det A) = \varepsilon_{ijk} A_{im} A_{jn} A_{kp} \quad \text{— (1)}$$

$$= \varepsilon_{ijk} A_{mi} A_{nj} A_{pk} \quad \text{— (2)}$$

(1) is a column expansion of the determinant, while (2) is a row " " " ".

For example, taking $m=1$, $n=2$, $p=3$

$$\varepsilon_{ijk} A_{i1} A_{j2} A_{k3} = \varepsilon_{123} A_{11} A_{22} A_{33} + \varepsilon_{132} A_{11} A_{23} A_{32}$$
$$+ \varepsilon_{213} A_{21} A_{12} A_{33} + \varepsilon_{231} A_{21} A_{32} A_{13}$$
$$+ \varepsilon_{312} A_{31} A_{12} A_{23} + \varepsilon_{321} A_{31} A_{22} A_{13}$$
$$= A_{11}(A_{22}A_{33} - A_{23}A_{32}) - A_{21}(A_{12}A_{33} - A_{32}A_{13})$$
$$+ A_{31}(A_{12}A_{23} - A_{22}A_{13})$$
$$= \det A.$$

Derivative of the determinant with respect to matrix entries

We will now prove $\dfrac{\partial}{\partial A}(\det A) = A^{-T}\det A$, if $\det A \neq 0$.

Before we prove the result, we collect the following facts about determinant and inverse of a square matrix: We assume all matrices are of size 3x3.

- $\det(AB) = \det A \times \det B$ — 3

- A matrix, denoted as A^{-1} is called a inverse of A if $AA^{-1} = A^{-1}A = I$. — 4

Ex: Show that if A^{-1} exists, then it is unique. In other words, the phrase "a inverse of A" can be replaced by "the inverse of A".

Ex: Show that A^{-1} exists $\iff$ (if and only if) $\det A \neq 0$

Solution: ($\Rightarrow$) If A^{-1} exists, then from 3 and 4, $(\det A)(\det A^{-1}) = 1 \Rightarrow \det A \neq 0$.

($\Leftarrow$) Recall (1.7) $\varepsilon_{mnp}(\det A) = \varepsilon_{ijk} A_{mi} A_{nj} A_{pk}$.

Multiplying by ε_{mnq}, and using (1.4) we have

$$2\delta_{pq} = \frac{1}{(\det A)} \varepsilon_{ijk} \varepsilon_{mnq} A_{mi} A_{nj} A_{pk} \quad (\because \det A \neq 0)$$

$$\Rightarrow \delta_{pq} = A_{pk} \left(\frac{1}{2 \det A} \varepsilon_{ijk} \varepsilon_{mnq} A_{mi} A_{nj} \right)$$

$$=: A_{pk} A^{-1}_{kq}$$

$$\Rightarrow \boxed{A^{-1}_{kq} = \frac{1}{2 \det A} \varepsilon_{ijk} \varepsilon_{mnq} A_{mi} A_{nj}} \quad \text{—} (5)$$

Note: Matrix multiplication is not commutative $AB \neq BA$.

Exercise: Is it possible that $AB = I$, but $BA \neq I$. ?

Solution: No. Assume $AB = I$. Then

$B(AB) = B$. Matrix multiplication is associative.

$$\Rightarrow (BA)B - B = 0$$

$$\Rightarrow (BA - I)B = 0 \quad \text{—} (*)$$

From 3, we know $\det B \cdot \det A = 1 \Rightarrow \det B \neq 0$

From previous exercise, we know that since $\det B \neq 0$, B^{-1} exists. Multiplying $(*)$ with B^{-1} from the right,

$$BA - I = 0 \Rightarrow BA = I.$$

We now get back to deriving the derivative of det.

$$\frac{\partial}{\partial A_{rs}}(\det A) = \frac{1}{6}\varepsilon_{mnp}\varepsilon_{ijk}\frac{\partial(A_{im}A_{jn}A_{kp})}{\partial A_{rs}}$$

$$= \frac{1}{6}\varepsilon_{mnp}\varepsilon_{ijk}\Big[\delta_{ir}\delta_{ms}A_{jn}A_{kp} + \delta_{jr}\delta_{ns}A_{im}A_{kp} + \delta_{kr}\delta_{ps}A_{im}A_{jn}\Big]$$

$$= \frac{1}{2}\varepsilon_{snp}\varepsilon_{rjk}A_{jn}A_{kp}, \quad \text{— ⑥}$$

where we have used the fact that the three terms in the second equality are the same by appropriately renaming the dummy indices.

To get to the eventual answer, we note that

$$I = AA^{-1},$$

$$\delta_{rt} = A_{ri}A^{-1}_{it}.$$

Multiplying ⑥ with the identity,

$$\delta_{rt}\frac{\partial}{\partial A_{rs}}(\det A) = \frac{1}{2}\varepsilon_{snp}\varepsilon_{rjk}A_{jn}A_{kp}A_{ri}A^{-1}_{it}$$

$$\frac{\partial}{\partial A_{ts}}(\det A) = \frac{1}{2}\varepsilon_{snp}\varepsilon_{inp}(\det A)A^{-1}_{it},$$

where we used 1 to arrive at the previous equality. Further, using (1·4) we obtain

$$\frac{\partial}{\partial A_{ts}}(\det A) = \frac{1}{2}(2\delta_{is})(\det A)A^{-1}_{it}$$

$$= (\det A)\, A^{-1}_{st}$$

$$\Rightarrow \frac{\partial(\det A)}{\partial A} = (\det A)\, A^{-T}. \qquad \square$$

2.1 Vectors

A common definition of vector one often hears is that "anything that has a magnitude and direction" $\longrightarrow$ This is not a complete definition.

Definition: A real vector space is a set that satisfies

1) vector addition: For any $\underset{\sim}{a}, \underset{\sim}{b} \in V$, $\exists$ (there exists) a vector $\underset{\sim}{a} + \underset{\sim}{b} \in V$

2) scalar multiplication: For every $\lambda \in \mathbb{R}$ (real), and $\underset{\sim}{v} \in V$, $\exists$ $\alpha \underset{\sim}{v} \in V$.

"in"

with the following properties $\forall\ \underset{\sim}{a}, \underset{\sim}{b}, \underset{\sim}{c} \in V$ and $\alpha, \mu \in \mathbb{R}$
"for all".

1. $\underset{\sim}{a} + \underset{\sim}{b} = \underset{\sim}{b} + \underset{\sim}{a}$ — Addition is commutative
2. $\underset{\sim}{a} + (\underset{\sim}{b} + \underset{\sim}{c}) = (\underset{\sim}{a} + \underset{\sim}{b}) + \underset{\sim}{c}$ — " " associative
3. $\exists\ \underset{\sim}{0} \in V$ such that $\underset{\sim}{a} + \underset{\sim}{0} = \underset{\sim}{a}$
4. $\exists\ -\underset{\sim}{a} \in V$ such that $\underset{\sim}{a} + (-\underset{\sim}{a}) = \underset{\sim}{0}$

} Existence of zero and inverse.

5. $(\alpha\mu)\underset{\sim}{a} = \alpha(\mu\underset{\sim}{a})$
6. $(\alpha + \mu)\underset{\sim}{a} = \alpha\underset{\sim}{a} + \mu\underset{\sim}{a}$
7. $\alpha(\underset{\sim}{a} + \underset{\sim}{b}) = \alpha\underset{\sim}{a} + \alpha\underset{\sim}{b}$
8. $1\underset{\sim}{a} = \underset{\sim}{a}$.

By "vector space" we will always refer to "real vector space". Of course, we have complex vector space as well, where $\alpha, \mu \in \mathbb{C}$. We never use vector spaces derived from other "fields" besides reals.

The above abstract definition is quite general. For example, the set of all continuously differentiable function in the interval $[0,1]$ such that $f(0) = f(1) = 0$ is actually a vector space!

— convince yourself!.

A familiar example: $\mathbb{R}^3$ — three dimensional real space.

Finite dimensional spaces and basis vectors

V - real vector space.

We say vectors $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_n \in V$ are <u>linearly dependent</u> if $\exists\ \alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ <u>not all</u> equal to zero such that

$$\alpha_1 \underset{\sim}{v}_1 + \cdots + \overbrace{\alpha_n \underset{\sim}{v}_n}^{\text{no sum}} = \underset{\sim}{0}.$$

Otherwise, the vectors are said to be <u>linearly independent</u>.

<u>Dimension of V</u>: The <u>largest</u> possible number of linearly independent vectors of V.

<u>Basis of V</u>: If V is a n-dimensional vector space, then <u>any</u> set of n linearly independent vectors is a basis of V.

<u>Exercise</u>: V - n dimensional vector space. Let $\{\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, v_n\}$, where $\underset{\sim}{v}_i$ $(i=1,\ldots,n) \in V$ describe a basis of V. Then show that for every $\underset{\sim}{v} \in V$, $\exists!$ (there exists unique) $\alpha_1, \alpha_2, \ldots, \alpha_n \in \mathbb{R}$ such that

$$\alpha_1 \underset{\sim}{v}_1 + \alpha_2 \underset{\sim}{v}_2 + \cdots + \alpha_n \underset{\sim}{v}_n = \underset{\sim}{v}.$$
(no sum)

Solution: Since $\underset{\sim}{v}_i$'s form a basis, and by the definition of basis, $\{\underset{\sim}{v}, \underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_n\}$ cannot be linearly independent.

$\Rightarrow \exists\ \lambda_0, \lambda_1, \ldots, \lambda_n \in \mathbb{R}$, not all zero, such that

$$\lambda_0 \underset{\sim}{v} + \lambda_1 \underset{\sim}{v}_1 + \cdots + \lambda_n \underset{\sim}{v}_n = \underset{\sim}{0} \quad (7)$$
(no sum)

Moreover $\lambda_0 \neq 0$, since $\underset{\sim}{v}_1, \underset{\sim}{v}_2, \ldots, \underset{\sim}{v}_n$ are linearly independent. Therefore, we can divide (7) by λ_0 to obtain

$$\underset{\sim}{v} = -\frac{\lambda_1}{\lambda_0}\underset{\sim}{v}_1 - \frac{\lambda_2}{\lambda_0}\underset{\sim}{v}_2 - \cdots - \frac{\lambda_n}{\lambda_0}\underset{\sim}{v}_n.$$

We have shown the existence of $\alpha_1 = -\frac{\lambda_1}{\lambda_0}, \ldots, \alpha_n = -\frac{\lambda_n}{\lambda_0}$.

How about uniqueness? — DO IT YOURSELF.

Inner product and norm

Recall that $\mathbb{R}^n$, the n-dimensional real coordinate space is an example of a vector space.

An arbitrary element of $\mathbb{R}^n$ looks like $(\alpha_1, \cdots, \alpha_n)$, where $\alpha_i \in \mathbb{R}$. A basis of $\mathbb{R}^n$:

$$\underset{\sim}{e_1} = (1, 0, \cdots, 0), \ \underset{\sim}{e_2} = (0, 1, 0, \cdots, 0), \ \cdots\cdots, \underset{\sim}{e_n} = (0, 0, \cdots, 0, 1).$$

<u>Addition</u>: If $\underset{\sim}{v} = (\alpha_1, \cdots, \alpha_n)$ and $\underset{\sim}{u} = (\beta_1, \cdots, \beta_n)$, then $\underset{\sim}{v} + \underset{\sim}{u} = (\alpha_1 + \beta_1, \alpha_2 + \beta_2, \cdots, \alpha_n + \beta_n)$.

<u>Multiplication</u>: If $\alpha \in \mathbb{R}$, $\underset{\sim}{v} = (\alpha_1, \cdots, \alpha_n)$, then $\alpha \underset{\sim}{v} := (\alpha\alpha_1, \alpha\alpha_2, \cdots, \alpha\alpha_n)$.

With the above definitions for addition and multiplication operations, convince yourself that $\mathbb{R}^n$ is indeed a real vector space, i.e it satisfies all the necessary axioms.

There is no mention of "angles" between vectors or "size" of a vector so far. It turns out that these notions require some more structure to be defined. Inner product generalizes the notion of angles.

<u>Definition</u>: V - vector space. An inner product on V is a <u>mapping</u> from $V \times V$ to $\mathbb{R}$, denoted by $\langle \underset{\sim}{u}, \underset{\sim}{v} \rangle$, where $\underset{\sim}{u} \in V$, $\underset{\sim}{v} \in V$, such that

a) $\langle \alpha \underset{\sim}{u} + \beta \underset{\sim}{v}, \underset{\sim}{w} \rangle = \alpha \langle \underset{\sim}{u}, \underset{\sim}{w} \rangle + \beta \langle \underset{\sim}{v}, \underset{\sim}{w} \rangle \quad \forall\, \alpha, \beta \in \mathbb{R},$
$\underset{\sim}{u}, \underset{\sim}{v}, \underset{\sim}{w} \in V.$

b) $\langle \underset{\sim}{u}, \underset{\sim}{v} \rangle = \langle \underset{\sim}{v}, \underset{\sim}{u} \rangle \quad \forall\, \underset{\sim}{u}, \underset{\sim}{v} \in V$

c) $\langle \underset{\sim}{u}, \underset{\sim}{u} \rangle > 0 \quad \forall\, \underset{\sim}{u} \neq \underset{\sim}{0}$

The quantity $\|\underset{\sim}{u}\| := \sqrt{\langle \underset{\sim}{u}, \underset{\sim}{u} \rangle}$ is referred to as the norm of $\underset{\sim}{u}$. Norm generalizes the notion of size of a vector.

Enough of abstraction! Let's now see a concrete example of an <u>inner-product space</u>, i.e a vector space equipped with an inner product.

<u>Euclidean vector space</u>: A finite dimensional real vector space with an inner product. An example is a real coordinate space with the inner product

$$\langle \underset{\sim}{u}, \underset{\sim}{v} \rangle = u_1 v_1 + u_2 v_2 + \cdots + u_n v_n ,$$

where $\underset{\sim}{u} = (u_1, \cdots, u_n) \in \mathbb{R}^n$ and $\underset{\sim}{v} = (v_1, \cdots, v_n) \in \mathbb{R}^n$.