Chapter 0 Preliminaries

Introduction

- *Vector space*
- *Inner product on vector space*
- *Linear transformation*
- *Lines, planes, and spheres*
- *Einstein Convention*
- *Vector Calculus*

0.1 Vector Spaces

Definition 0.1. Vector Space

A vector space V is a nonempty set with two binary operations "+" and scalar multiplication "·" satisfying the following eight axioms: let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $r, s \in \mathbb{R}$, we have

- $\mathbf{u} + \mathbf{v} = \mathbf{v} + \mathbf{u}$;
- $\mathbf{u} + (\mathbf{v} + \mathbf{w}) = (\mathbf{u} + \mathbf{v}) + \mathbf{w}$;
- $\mathbf{0} + \mathbf{u} = \mathbf{u} + \mathbf{0} = \mathbf{u}$ *for a vector called* $\mathbf{0}$;
- $(rs) \cdot \mathbf{u} = r \cdot (s \cdot \mathbf{u})$;
- $(r + s) \cdot \mathbf{u} = r \cdot \mathbf{u} + s \cdot \mathbf{u}$;
- $r \cdot (\mathbf{u} + \mathbf{v}) = r \cdot \mathbf{u} + r \cdot \mathbf{v}$;
- $0 \cdot \mathbf{u} = \mathbf{0}$;
- $1 \cdot \mathbf{u} = \mathbf{u}$.

As is well-known, the set V satisfying the first three axioms form an *Abel semi-group*.
The existence of the inverse of a vector $\mathbf{u}$ can be verified by using the scalar multiplication.
Let $\mathbf{u}$ be a vector, we claim that $(-1) \cdot \mathbf{u}$ is the inverse of $\mathbf{u}$ because

$$\mathbf{u} + (-1) \cdot \mathbf{u} = (1 + (-1)) \cdot \mathbf{u} = \mathbf{0}.$$

Similarly, we can define the subtraction by

$$\mathbf{u} - \mathbf{v} = \mathbf{u} + (-1) \cdot \mathbf{v}.$$

Note *We sometimes omit the $\cdot$ and write, for example, $r\mathbf{u}$ for $r \cdot \mathbf{u}$.*

External Link. *Here is the video explanation of vector space (linear indepedence).* *Useful!*

External Link. *Here is the Math 162A pre-requisite videos.*

In linear algebra, we restrict ourselves to finite dimensional vector space. But, a lot of results in finite dimensional case can be extended to infinite dimensional case as well as *abstract* vector space cases.

In the following, we give some examples of vector spaces.

Example 0.1 $\mathbb{R}^n$, the n-dimensional *Euclidean space*, is the set of all n-vectors

$$\mathbb{R}^n = \left\{ \quad \mathbf{x} = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} \mid x_i \in \mathbb{R} \right\}.$$

Example 0.2 The set of all $m \times n$ matrices for an (mn)-dimensional vector space.

Example 0.3 The space of polynomials of degree no more than n, where n is a nonnegative integer, is a vector space.

A single-variable polynomial of degree no more than n can be expressed as

$$p(t) = a_0 + a_1 t + \cdots + a_n t^n.$$

Let

$$q(t) = b_0 + b_1 t + \cdots + b_n t^n$$

be another polynomial. Define the addition to be

$$(p+q)(t) = (a_0 + b_0) + (a_1 + b_1)t + \cdots + (a_n + b_n)t^n,$$

and the scalar multipliction to be: let $\lambda \in \mathbb{R}$

$$(\lambda p)(t) = (\lambda a_0) + (\lambda a_1)t + \cdots + (\lambda a_n)t^n.$$

With respect to the addition and scalar multiplication, the space is a vector space.

In the above two examples, the dimensions of the vector spaces are *finite*. Let's show some examples of infinite dimensional vector spaces.

Note *The set of all polynomials of degree equal to n is not a vector space.*

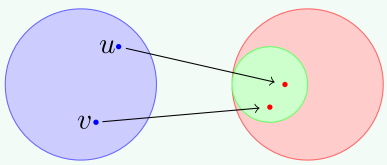
Example 0.4 Moreover, the space of all real-valued functions on a set is a (infinite dimensional) vector space.

Let's recall the definition of a function.

Definition 0.2

A function $f : X \to Y$ is a triple (f, X, Y), where X, Y are sets, and f is an assignment, or a rule, that for any element x in X, there is a unique $y = f(x)$ in Y attached to it. X is called the domain of f, and Y is called the codomain of f. The range is the subset of the codomain Y consists of all $f(x)$ when X is running through X. The assignment sometimes is written as $x \mapsto f(x)$. Thus a complete description of a function can be given as

$$f : X \to Y, \quad x \mapsto f(x).$$



As above, we can define the addition and scalar multiplication as

$$(f+g)(x) = f(x) + g(x), \qquad (\lambda f)(x) = \lambda f(x).$$

Remark The vector space we study in differential geometry are the "abstract" vector space, which is on the contrary to the vector spaces we studied in Math 3A.

The following concepts are defined in abstract vector spaces similar to those in $\mathbb{R}^n$.

1. linear combination, span;
2. linear dependence and independence;
3. basis and dimension.

Remark In infinite dimensional space, a basis is defined by a set of vectors

$$\mathfrak{I} = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n, \cdots\}$$

satisfying the following

1. any finite subset of $\mathfrak{I}$ is linearly independent;
2. any element can be expressed as a (finite) linear combination of element in $\mathfrak{I}$.

Let $(V, +, \cdot)$ be a vector space. We can endow geometric structure onto it by defining the concept of inner product.

Remark Let

$$\mathfrak{S} = \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}$$

be a finite set. Let

$$V = \operatorname{Span} \mathfrak{S} = \operatorname{Span} \{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_n\}.$$

Then we say $\mathfrak{S}$ spans V, and $\mathfrak{S}$ is a *spanning* set of V.

0.2 Inner Product

The addition and scalar multiplication define the *algebraic structure* of a vector space. In order to introduce geometry to linear algebra, we can endow geometric structure onto it by defining the concept called *inner product*.

Definition 0.3

An inner product on a vector space V is a function $\langle\ ,\ \rangle$

$$V \times V \to \mathbb{R}$$

satisfying the following properties: let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in V$ and $r, s \in \mathbb{R}$, we have

$(a).\ \langle \mathbf{u}, \mathbf{v} \rangle = \langle \mathbf{v}, \mathbf{u} \rangle$ *Symmetry*

$(b).\ \langle \mathbf{u}, r\mathbf{v} + s\mathbf{w} \rangle = r\langle \mathbf{u}, \mathbf{v} \rangle + s\langle \mathbf{u}, \mathbf{w} \rangle$ *Linearity*

$(c).\langle \mathbf{u}, \mathbf{u} \rangle \geq 0$ *and equality is true if and only if* $\mathbf{u} = \mathbf{0}$ *Positivity*

Once we introduce the inner product, we introduce geometry into vector space. For example

Definition 0.4

(Length of a vector) We can define the length, or the norm, of a vector to be

$$\|\mathbf{u}\| = \sqrt{\langle \mathbf{u}, \mathbf{u} \rangle}.$$

(Distance) Let $\mathbf{u}, \mathbf{v}$ be two vectors. Then their distance is defined to be

$$\operatorname{dist}(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\| = \sqrt{\langle \mathbf{u} - \mathbf{v}, \mathbf{u} - \mathbf{v} \rangle}.$$

Note in the definition of distance, we used both the geometric structure (inner product) and algebraic structrue (subtraction is defined by $\mathbf{u} - \mathbf{v} = \mathbf{u} + (-1) \cdot \mathbf{v}$).

Example 0.5 In $\mathbb{R}^3$ (and in $\mathbb{R}^n$), the ordinary *dot product*

$$\langle (a^1, a^2, a^3), (b^1, b^2, b^3) \rangle = a^1 b^1 + a^2 b^2 + a^3 b^3$$

is an inner product.

However, there are many other ways one can define inner product on $\mathbb{R}^3$. For example, we can define

$$\langle (a^1, a^2, a^3), (b^1, b^2, b^3) \rangle = 2a^1 b^1 + 3a^2 b^2 + 4a^3 b^3.$$

It is an inner product. In general, let

$$R = \begin{bmatrix} r_{11} & r_{12} & r_{13} \\ r_{21} & r_{22} & r_{23} \\ r_{31} & r_{32} & r_{33} \end{bmatrix}$$

be a positive definite matrix. Then

$$\langle (a^1, a^2, a^3), (b^1, b^2, b^3) \rangle = \sum_{i,j=1}^{3} r_{ij} a^i b^j$$

is an inner product.

Example 0.6 We let $R[t]$ be the vector space of all polynomials. Define the inner product

$$\langle p, q \rangle = \int_{-1}^{1} p(x) q(x) \, dx$$

More general, if $\rho(x) > 0$ be a positive continuous function, then

$$\langle p, q \rangle = \int_{-1}^{1} p(x) q(x) \rho(x) \, dx$$

is an inner product.

One of the most important inequality in mathematics is called the *Cauchy-Schwarz Inequality*.

Theorem 0.1

Let $\mathbf{u}, \mathbf{v}$ be two vectors. Then we have

$$|\langle \mathbf{u}, \mathbf{v} \rangle| \leq \|\mathbf{u}\| \cdot \|\mathbf{v}\|,$$

and the equality holds if and only if $\mathbf{u}, \mathbf{v}$ are linearly dependent.

♡

Proof. This is the standard proof. Let t be a real number. Then by the positivity of the inner product, we have

$$\langle \mathbf{u} + t\mathbf{v}, \mathbf{u} + t\mathbf{v} \rangle \geq 0.$$

Using the linearity, we get

$$t^2 \langle \mathbf{v}, \mathbf{v} \rangle + 2t \langle \mathbf{u}, \mathbf{v} \rangle + \langle \mathbf{u}, \mathbf{u} \rangle \geq 0.$$

Since the above inequality is true for any real number t, the discriminant

$$\Delta = 4|\langle \mathbf{u}, \mathbf{v} \rangle|^2 - 4\langle \mathbf{u}, \mathbf{u} \rangle \cdot \langle \mathbf{v}, \mathbf{v} \rangle \leq 0,$$

which proves the inequality. ■

Proof. [Second Proof] Assume that $\mathbf{v} \neq \mathbf{0}$. Then we have

$$\langle \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v}, \mathbf{u} - \frac{\langle \mathbf{u}, \mathbf{v} \rangle}{\langle \mathbf{v}, \mathbf{v} \rangle} \mathbf{v} \rangle \geq 0.$$

Expanding the above inequality, we obtain the Cauchy-Schwarz Inequality. ■

Remark The Cauchy-Schwarz inequality is also called the Cauchy-Bunyakovsky-Schwarz inequality.

The inequality for sums was published by Augustin-Louis Cauchy (1821), while the corresponding inequality for integrals was first proved by Viktor Bunyakovsky (1859). The modern proof of the integral version was given by Hermann Schwarz (1888).

Definition 0.5

Two vectors $\mathbf{u}, \mathbf{v}$ *are called* *orthogonal*, *if* $\langle \mathbf{u}, \mathbf{v} \rangle = 0$.

Example 0.7 In $\mathbb{R}^3$, the vector $(1, 2, 3)$ is orthogonal to $(4, -5, 2)$ with respect to the dot product, because

$$(1, 2, 3) \cdot (4, -5, 2) = 1 \cdot 4 + 2 \cdot (-5) + 3 \cdot 2 = 0.$$

Example 0.8 Under the inner product

$$\langle p, q \rangle = \int_{-1}^{1} p(x) q(x) \, dx.$$

The function x and $x^2 + 1$ are orthogonal because $p(x)$ is an odd function.

An *orthonormal basis* of an n-dimensional vector space is a set $\{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$ such that

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \begin{cases} 1, & \text{if} \quad i = j \\ 0, & \text{if} \quad i \neq j \end{cases}.$$

We shall introduce the *Kronecker symbol* δ_{ij} as follows

$$\delta_{ij} = \begin{cases} 1, & \text{if} \quad i = j \\ 0, & \text{if} \quad i \neq j \end{cases}.$$

Under this notation, we have

$$\langle \mathbf{v}_i, \mathbf{v}_j \rangle = \delta_{ij}.$$

0.3 Linear Transformation

Definition 0.6

Given two vector spaces, V and W, a linear transformation T from V to W is a mapping

$$T : V \to W$$

such that

$$T(\alpha \mathbf{u} + \beta \mathbf{v}) = \alpha T(\mathbf{u}) + \beta T(\mathbf{v})$$

for all $\mathbf{u}, \mathbf{v} \in V$ and $\alpha, \beta \in \mathbb{R}$.

A linear transformation of a vector space to itself is called an *endomorphism*. One of the most important concept of an endomorphism is its eigenvalue and eigenvector.

Definition 0.7

Let $T : V \to V$ be an endomorphism. Assume that there is a $\mathbf{v} \in V$ and $\mathbf{v} \neq 0$ such that

$$T(\mathbf{v}) = \lambda \mathbf{v}$$

for some complex number λ. Then λ is called an eigenvalue of T and v is an eigenvector of λ.

Note *In the following, we need to prove that the definition of eigenvalue is equivalent to the definition of an eigenvalue of a matrix.*

Let V be a finite dimensional space and let

$$\mathfrak{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}$$

be a basis of V. We use the notation $[\mathbf{x}]_{\mathfrak{B}}$ to represent the coordinates of $\mathbf{x} \in V$, that is, when we write x in terms of the linear combination of the basis,

$$\mathbf{x} = c_1 \mathbf{v}_1 + \cdots + c_n \mathbf{v}_n,$$

we have

$$[\mathbf{x}]_{\mathfrak{B}} = \begin{bmatrix} c_1 \\ \vdots \\ c_n \end{bmatrix}.$$

Now let $T : V \to W$ be a linear transformation. As above, we assume that $\mathfrak{B}$ is the basis of V, and let $\mathfrak{C} = \{\mathbf{w}_1, \cdots, \mathbf{w}_m\}$ be a basis of W. Then the matrix representation M of T is

$$M = [[T(\mathbf{v}_1)]_{\mathfrak{C}}, \cdots, [T(\mathbf{v}_n)]_{\mathfrak{C}}]$$

in the sense that

$$[T(\mathbf{x})]_{\mathfrak{C}} = M \cdot [\mathbf{x}]_{\mathfrak{B}}.$$

Now we specialize the above result to the following case. Let

$$T : V \to V$$

be a linear transformation from V to itself. Let

$$\mathfrak{B} = \{\mathbf{v}_1, \cdots, \mathbf{v}_n\}, \qquad \mathfrak{C} = \{\mathbf{w}_1, \cdots, \mathbf{w}_n\}$$

be two bases. Let M_1, M_2 be the matrix representatives with respect to the two bases respectively. We thus have

$$[T(\mathbf{x})]_{\mathfrak{B}} = M_1 \cdot [\mathbf{x}]_{\mathfrak{B}};$$
$$[T(\mathbf{x})]_{\mathfrak{C}} = M_2 \cdot [\mathbf{x}]_{\mathfrak{C}}.$$

Let A be the invertible matrix such that

$$[\mathbf{x}]_{\mathfrak{B}} = A \cdot [\mathbf{x}]_{\mathfrak{C}}.$$

Such a matrix is called a *transition matrix*. Then

$$[T(\mathbf{x})]_{\mathfrak{B}} = A \cdot [T(\mathbf{x})]_{\mathfrak{C}}.$$

As a result,

$$A \cdot M_2 \cdot [\mathbf{x}]_{\mathfrak{C}} = A \cdot [T(\mathbf{x})]_{\mathfrak{C}} = [T(\mathbf{x})]_{\mathfrak{B}} = M_1 \cdot A \cdot [\mathbf{x}]_{\mathfrak{C}}.$$

Thus we have

$$AM_2 = M_1 A,$$

or

$$M_2 = A^{-1} M_1 A.$$

Thus M_1, M_2 are similar, having the same eigenvalue set.

Now we talk about *orientation and cross product*.

Let $\mathfrak{B}$ and $\mathfrak{C}$ be two bases. We say that $\mathfrak{B}$ and $\mathfrak{C}$ are having the same orientation, if when we write

$$\mathbf{v}_i = \sum_{j=1}^{n} a_{ij} \mathbf{w}_j$$

for $i = 1, \cdots, n$, then we have $\det(A) = \det(a_{ij}) > 0$. They give the opposite orientation if $\det(a_{ij}) < 0$.

Example 0.9 The left hand and the right hand define two opposite orientations of $\mathbb{R}^3$.

Example 0.10 Let

$$\mathfrak{B} = \{(1, 0, 0), (0, 1, 0), (0, 0, 1)\}$$

and

$$\mathfrak{C} = \{(1, 1, 0), (1, 0, -1), (2, 1, 3)\}$$

are of the opposite orientation.

Definition 0.8

Let $\{\mathbf{e}_1, \mathbf{e}_2, \mathbf{e}_3\}$ *be the standard basis of* $\mathbb{R}^3$*. If*

$$\mathbf{u} = \sum_{i=1}^{3} a_i \mathbf{e}_i, \qquad \mathbf{v} = \sum_{j=1}^{3} b_j \mathbf{e}_j$$

are two vectors in $\mathbb{R}^3$*, the cross product of* $\mathbf{u}, \mathbf{v}$ *is given by*

$$\mathbf{u} \times \mathbf{v} = (a_2 b_3 - a_3 b_2)\mathbf{e}_1 + (a_3 b_1 - a_1 b_3)\mathbf{e}_2 + (a_1 b_2 - a_2 b_1)\mathbf{e}_3,$$

or we can write

$$\mathbf{u} \times \mathbf{v} = \det \begin{bmatrix} \mathbf{e}_1 & \mathbf{e}_2 & \mathbf{e}_3 \\ a_1 & a_2 & a_3 \\ b_1 & b_2 & b_3 \end{bmatrix}.$$

The cross product satisfies the following properties:

Lemma 0.1

Let $\mathbf{u}, \mathbf{v}, \mathbf{w} \in \mathbb{R}^3$ *and* $r \in \mathbb{R}$*. Then*

- $\mathbf{u} \times \mathbf{v} = -\mathbf{v} \times \mathbf{u}$
- $(r\mathbf{u}) \times \mathbf{v} = r(\mathbf{u} \times \mathbf{v})$
- $\mathbf{u} \times \mathbf{v} = 0$ *if and only if* $\mathbf{u}$ *and* $\mathbf{v}$ *are linearly dependent*
- $(\mathbf{u} + \mathbf{v}) \times \mathbf{w} = \mathbf{u} \times \mathbf{w} + \mathbf{v} \times \mathbf{w}$
- $\mathbf{u} \times \mathbf{v}$ *is perpendicular to both* $\mathbf{u}$ *and* $\mathbf{v}$ *under the usual dot product*
- $\|\mathbf{u} \times \mathbf{v}\| = \|\mathbf{u}\| \cdot \|\mathbf{v}\| \sin\theta$*, where* θ *is the angle between* $\mathbf{u}$ *and* $\mathbf{v}$*.*
- $\{\mathbf{u}, \mathbf{v}, \mathbf{u} \times \mathbf{v}\}$ *gives a right hand orientation to* $\mathbb{R}^3$ *if* $\{\mathbf{u}, \mathbf{v}\}$ *is linearly independent*

♡

As a result, we have the relationship between the inner product and outer product (cross product)

$$|\langle \mathbf{u}, \mathbf{v} \rangle|^2 + \|\mathbf{u} \times \mathbf{v}\|^2 = \|\mathbf{u}\|^2 \cdot \|\mathbf{v}\|^2,$$

which implies the Cauchy-Schwarz Inequality in the three dimensional space.

Definition 0.9

The mixed (or triple) product of $\mathbf{u}, \mathbf{v}, \mathbf{w}$ *is*

$$[\mathbf{u}, \mathbf{v}, \mathbf{w}] = \langle \mathbf{u} \times \mathbf{v}, \mathbf{w} \rangle.$$

A geometric interpretation of the norm of the cross product is that it is the area of the parallelogram spanned by $\mathbf{u}, \mathbf{v}$. A geometric interpretation of the mixed scalar product is that it is the volume of the parallelepiped spanned by $\mathbf{u}, \mathbf{v}, \mathbf{w}$.

0.4 Lines, planes, and spheres

In this section, we use vector notations to express some basic objects in analytic geometry.

Definition 0.10

The line through $\mathbf{x}_0 \in \mathbb{R}^3$ *and parallel to a vector* $\mathbf{v} \neq 0$ *has the equation*

$$\alpha(t) = \mathbf{x}_0 + t\mathbf{v}.$$

Remark This is a vector notation of parametrization of a line.

Example 0.11 Let $\mathbf{x}_1, \mathbf{x}_2 \in \mathbb{R}^3$ be two points. Then the line through $\mathbf{x}_1$ and $\mathbf{x}_2$ in $\mathbb{R}^3$ has the equation.

$$\alpha(t) = \mathbf{x}_1 + t(\mathbf{x}_2 - \mathbf{x}_1).$$

Definition 0.11

The plane through $\mathbf{x}_0$ *perpendicular to* $\mathbf{n} \neq 0$ *has the equation*

$$\langle \mathbf{x} - \mathbf{x}_0, \mathbf{n} \rangle = 0.$$

Lemma 0.2

Let $\{\mathbf{u}, \mathbf{v}\}$ *be two linearly independent vectors. Then the place through* $\mathbf{x}_0$ *and parallel to the subspace spanned by* $\{\mathbf{u}, \mathbf{v}\}$ *has the equation*

$$[\mathbf{u}, \mathbf{v}, \mathbf{x} - \mathbf{x}_0] = \langle \mathbf{x} - \mathbf{x}_0, \mathbf{u} \times \mathbf{v} \rangle = 0.$$

♡

Definition 0.12

The sphere in $\mathbb{R}^3$ *with center* $\mathbf{m}$ *and radius* $r > 0$ *has equation*

$$\langle \mathbf{x} - \mathbf{m}, \mathbf{x} - \mathbf{m} \rangle = \|\mathbf{x} - \mathbf{m}\|^2 = r^2. \tag{1}$$

Remark Let

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix}, \qquad \mathbf{m} = \begin{bmatrix} m_1 \\ m_2 \\ m_3 \end{bmatrix}.$$

Then we get the usual equation of a sphere

$$(x_1 - m_1)^2 + (x_2 - m_2)^2 + (x_3 - m_3)^2 = r^2.$$

Example 0.12 (Kelvin Transformation)

We consider the equation of a sphere (1). Let $\mathbf{x}_0$ be a point on the sphere, that is, we have

$$\langle \mathbf{x}_0 - \mathbf{m}, \mathbf{x}_0 - \mathbf{m} \rangle = r^2.$$

The Kelvin Transformation is a map

$$K : \mathbb{R}^3 \to \mathbb{R}^3, \qquad \mathbf{x} \mapsto \mathbf{x}_0 + \frac{\mathbf{x} - \mathbf{x}_0}{\|\mathbf{x} - \mathbf{x}_o\|^2}$$

By a straightforward computation, we have $K^2 = id$. Assume

$$\langle K(x) - \mathbf{m}, K(x) - \mathbf{m} \rangle = r^2.$$

We get

$$1 + 2\langle \mathbf{x} - \mathbf{x}_0, \mathbf{x}_0 - \mathbf{m} \rangle = 0.$$

So the Kelvin transformation maps a sphere to a plane.

External Link. *The detailed computation can be found here.*

Example 0.13 (Ptolemy Inequality) Let $\mathbf{u}, \mathbf{v}, \mathbf{w}, \mathbf{x}$ be four vectors in the Euclidean plane. Then we have

$$\|\mathbf{u} - \mathbf{w}\| \cdot \|\mathbf{v} - \mathbf{x}\| \leq \|\mathbf{u} - \mathbf{v}\| \cdot \|\mathbf{x} - \mathbf{w}\| + \|\mathbf{u} - \mathbf{x}\| \cdot \|\mathbf{v} - \mathbf{w}\|.$$

The equality is valid if and only if these four vectors are concyclic.

External Link. *The Ptolemy Inequality is closely related to the Ptolemy Theorem. For details of the Ptolemy and his theorem, see Wikepedia of Ptolemy Theorem*

0.5 Vector Calculus

In differential geometry, in addition to study functions of several variables. We also need to study vector-valued functions.

We can define derivatives, indefinite integral and definite integrable in similar ways to those of multi-variable functions.

Let V, W be finite dimensional vector spaces. Let

$$F : V \to W$$

be a *differentiable* function, with definition as follows.

Definition 0.13

We fix a basis of V and using that basis, we identify V to $\mathbb{R}^n$. Similarly, and we fix a basis of W and identify it to $\mathbb{R}^m$. Then we can identify $F : V \to W$ by $F : \mathbb{R}^n \to \mathbb{R}^m$. So F is differentiable if and only if F is a differentiable as a mapping of $\mathbb{R}^n \to \mathbb{R}^m$.

External Link. *Here is a video clip for the details of the above definition.*

Let $f : \mathbb{R} \to \mathbb{R}^n$ be a single variable vector-valued function. We can define

$$\frac{df}{dt} = \begin{bmatrix} \frac{df_1}{dt} \\ \vdots \\ \frac{df_n}{dt} \end{bmatrix}$$

if

$$f = \begin{bmatrix} f_1 \\ \vdots \\ f_n \end{bmatrix}$$

Likewise, we can define

$$\int f(t)\,dt, \qquad \int_a^b f(t)\,dt$$

in similar ways.

If $f : \mathbb{R} \to V$ be an abstract vector-valued function, we can identity V with $\mathbb{R}^n$ under a fixed basis, and define the derivative, integral, etc, similarly.

Lemma 0.3

Let $f : \mathbb{R} \to V, g : \mathbb{R} \to V$ and let $\langle\ ,\ \rangle$ be an inner product on V. Then if f and g are differentiable, so is $\langle f, g\rangle$, which is a function of one variable. Moreover, we have

$$\frac{d}{dt}\langle f, g\rangle = \langle \frac{df}{dt}, g\rangle + \langle f, \frac{dg}{dt}\rangle.$$

♡

Similarly, we have

Lemma 0.4

Using the notations as in the above lemma, we have

$$\frac{d}{dt}(f \times g) = \frac{df}{dt} \times g + f \times \frac{dg}{dt}.$$

♡

Both of the above two lemmas can be proved directly. Moreover, we can generalize the above results into the following.

Let V, W, S be vector spaces (probably of infinite dimensional) and let

$$K : V \times W \to S$$

be a map. We say K is *bilinear*, if K is linear with respect to each component.

Both the inner product and cross product are bilinear mappings.

Let $f : \mathbb{R} \to V$, $g : \mathbb{R} \to W$ be differentiable functions. Then the function

$$h(t) = K(f(t), g(t))$$

is differentiable, and

$$\frac{dh}{dt} = K(\frac{df}{dt}, g(t)) + K(f(t), \frac{dg}{dt}).$$

Definition 0.14

Let $f : \mathbb{R} \to \mathbb{R}$ be a continuous function. We say f is of class C^k, if all derivatives up through order k exist and are continuous.

$f : \mathbb{R}^n \to \mathbb{R}$ is of class C^k if all its (mixed) partial derivatives of up through order k exist

and are continuous. A vector-valued function is of class $\mathcal{C}^k$ if all of its components with respect to a given basis are of class $\mathcal{C}^k$.

If f is of class $\mathcal{C}^k$ for any k, we say f is of $\mathcal{C}^\infty$, or we say f is smooth.

We assume most of the functions we shall study in this course are smooth, or at least of $\mathcal{C}^3$.

Finally, we review the chain rule. Let x be a function of $(u_1, \cdots, u_n)$, and if each u_i are functions of $(v_1, \cdots, v_m)$, say,

$$u_i = u_i(v_1, \cdots, v_m).$$

for $i = 1, \cdots, n$. Then we have the chain rule

$$\frac{\partial x}{\partial v_\alpha} = \sum_{i=1}^{n} \frac{\partial x}{\partial u_i} \cdot \frac{\partial u_i}{\partial v_\alpha} \tag{2}$$

for $\alpha = 1, \cdots, m$.

0.6 Einstein Convention

Definition 0.15. Einstein Convention

When an index variable appears twice in a single term and is not otherwise defined, it implies summation of that term over all the values of the index. When an index variable appears only once, it implies that the equation is valid for every value of such an index.

For example, Equation (2) can be written as

$$\frac{\partial x}{\partial v_\alpha} = \frac{\partial x}{\partial u_i} \cdot \frac{\partial u_i}{\partial v_\alpha}.$$

In the above equation, the index i in the right appears twice, so we assume the expression is summing over all possible i. On the other hand, the index α appears only once, so we assume the equation is valid for all range of α.

Example 0.14 Let $A = (a_{ij})$, $B = (b_{ij})$, and $C = (c_{ij})$ be matrices. Then the matrix multiplication,

$$C = AB,$$

can be written using the Einstein Convention as

$$c_{ij} = a_{ik} b_{kj}.$$

The Einstein Convention gives another way to express and generalize linear algebra.

Note *Let's discuss the representation of a matrix. In linear algebra, there are three ways to represent a matrix*

$$A = [\mathbf{a}_1, \cdots, \mathbf{a}_n] = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & & \vdots \\ a_{m1} & \cdots & a_{mn} \end{bmatrix},$$

where $\mathbf{a}_1, \cdots, \mathbf{a}_n$ are column vectors. The easiest way to represent a matrix is to a capital letter, say A. But this would contain the least amount of information about the matrix. On the other extreme, if we represent a matrix by providing all the details, it would be too clumsy to write down.

Here we give the fourth method of representing a matrix, by writing it as (a_{ij}), which takes care of both simplicity and information.

Example 0.15 Prove the associativity of matrix multiplication.

Proof. Let $A = (a_{ij})$, $B = (b_{ij})$ be matrices. Let $D = (d_{ij})$ be the matrix $D = AB$. In terms of the Einstein Convention, $D = AB$ is equivalent to

$$d_{ij} = a_{ik}b_{kj}. \tag{3}$$

Here the index k is called a *dummy* index in the sense that we can replace it with other indices without changing the equations:

$$d_{ij} = a_{ik}b_{kj} = a_{it}b_{tj} = a_{i\alpha}b_{\alpha j}. \tag{4}$$

Now let $C = (c_{ij})$, $E = BC = (e_{ij})$, $F = (AB)C = (f_{ij})$ and $G = A(BC) = (g_{ij})$. Then the entries for $(AB)C = DC$ would be

$$f_{ij} = d_{ik}c_{kj} = d_{it}c_{tj}.$$

From (4), we know that $d_{it} = a_{ik}b_{kt}$. Thus

$$f_{ij} = a_{ik}b_{kt}c_{tj}.$$

The reason we use t as the dummy index in (4) is because k has been used in (3) so we need to use a different one, keeping indices repeated at most twice.

Similarly, we have

$$g_{ij} = a_{it}b_{tk}c_{kj}.$$

Thus $f_{ij} = g_{ij}$ and hence

$$(AB)C = A(BC),$$

proving the associativity. ∎

Example 0.16 Using the Einstein Convention to prove the following version of the Cauchy inequality. Let $\mathbf{x} = (x_1, \cdots, x_n)$ and $\mathbf{y} = (y_1, \cdots, y_n)$. Then

$$|\mathbf{x} \cdot \mathbf{y}|^2 \le \|\mathbf{x}\|^2 \cdot \|\mathbf{y}\|^2.$$

Proof. Using the Einstein Convention, we can write

$$\mathbf{x} \cdot \mathbf{y} = \sum_{i=1}^{n} x_i y_i = x_i y_i.$$

Thus

$$|\mathbf{x} \cdot \mathbf{y}|^2 = \left(\sum_{i=1}^{n} x_i y_i\right)^2 = \left(\sum_{i=1}^{n} x_i y_i\right) \cdot \left(\sum_{j=1}^{n} x_j y_j\right) = x_i y_i x_j y_j.$$

On the other hands, we can write

$$\|\mathbf{x}\|^2 \cdot \|\mathbf{y}\|^2 = \left(\sum_{i=1}^{n} x_i^2\right) \cdot \left(\sum_{j=1}^{n} y_j^2\right) = x_i^2 y_j^2.$$

Thus the Cauchy inequality, written under the Einstein Convention, is

$$x_i^2 y_j^2 - x_i y_i x_j y_j = \frac{1}{2}(x_i^2 y_j^2 + x_j^2 y_i^2) - x_i y_i x_j y_j = \frac{1}{2}(x_i y_j - x_j y_i)^2 \geq 0.$$

This completes the proof. ■

Note *If $n = 3$, then we can define*

$$(\mathbf{x}, \mathbf{y}) \mapsto (x_2 y_3 - x_2 y_2, x_3 y_1 - x_1 y_3, x_1 y_2 - x_2 y_1) = \mathbf{x} \times \mathbf{y}$$

which is the cross product. If $n \neq 3$, then the vector $(x_i y_j - x_j y_i)$ for $i < j$ is of dimension $n(n-1)/2 \neq 3$. This explains why we can only define the cross product in 3 dimensional vector space. A more general algebraic product, called wedge product, will be used in any dimensional vector spaces to catch in the excess of the Cauchy inequality.

External Link. *As fun reading, you can find the Shoelace Formula in the Wikipedia, which is related to both the cross product and wedge product.*