A JOURNEY

IN

PURE MATHEMATICS

A JOURNEY

IN

PURE MATHEMATICS

MAT3006 & 3040 & 4002 Notebook

Prof. Daniel Wong

The Chinese University of Hongkong, Shenzhen

香港中文大學(深圳)
The Chinese University of Hong Kong, Shenzhen

Contents

Acknowledgments

This book is from the MAT3006,MAT3040,MAT4002 in spring semester, 2018-2019.

CUHK(SZ)

Notations and Conventions

$\mathbb{R}^n$	n-dimensional real space
$\mathbb{C}^n$	n-dimensional complex space
$\mathbb{R}^{m\times n}$	set of all $m \times n$ real-valued matrices
$\mathbb{C}^{m\times n}$	set of all $m \times n$ complex-valued matrices
x_i	ith entry of column vector $\boldsymbol{x}$
a_{ij}	(i,j)th entry of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i$	ith column of matrix $\boldsymbol{A}$
$\boldsymbol{a}_i^{\mathrm{T}}$	ith row of matrix $\boldsymbol{A}$
$\mathbb{S}^n$	set of all $n \times n$ real symmetric matrices, i.e., $\boldsymbol{A} \in \mathbb{R}^{n\times n}$ and $a_{ij} = a_{ji}$ for all i, j
$\mathbb{H}^n$	set of all $n \times n$ complex Hermitian matrices, i.e., $\boldsymbol{A} \in \mathbb{C}^{n\times n}$ and $\bar{a}_{ij} = a_{ji}$ for all i, j
$\boldsymbol{A}^{\mathrm{T}}$	transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{T}}$ means $b_{ji} = a_{ij}$ for all i, j
$\boldsymbol{A}^{\mathrm{H}}$	Hermitian transpose of $\boldsymbol{A}$, i.e, $\boldsymbol{B} = \boldsymbol{A}^{\mathrm{H}}$ means $b_{ji} = \bar{a}_{ij}$ for all i, j
$\mathrm{trace}(\boldsymbol{A})$	sum of diagonal entries of square matrix $\boldsymbol{A}$
$\mathbf{1}$	A vector with all 1 entries
$\mathbf{0}$	either a vector of all zeros, or a matrix of all zeros
$\boldsymbol{e}_i$	a unit vector with the nonzero element at the ith entry
$\mathcal{C}(\boldsymbol{A})$	the column space of $\boldsymbol{A}$
$\mathcal{R}(\boldsymbol{A})$	the row space of $\boldsymbol{A}$
$\mathcal{N}(\boldsymbol{A})$	the null space of $\boldsymbol{A}$
$\mathrm{Proj}_{\mathcal{M}}(\boldsymbol{A})$	the projection of $\boldsymbol{A}$ onto the set $\mathcal{M}$

Chapter 1

Week1

1.1. Monday for MAT3040

1.1.1. Introduction to Advanced Linear Algebra

Advanced Linear Algebra is one of the most important course in MATH major, with pre-request MAT2040. This course will offer the really linear algebra knowledge.

What the content will be covered?.

- In MAT2040 we have studied the space $\mathbb{R}^n$; while in MAT3040 we will study the general vector space V.
- In MAT2040 we have studied the *linear transformation* between Euclidean spaces, i.e., $T:\mathbb{R}^n \to \mathbb{R}^m$; while in MAT3040 we will study the linear transformation from vector spaces to vector spaces: $T: V \to W$
- In MAT2040 we have studied the eigenvalues of $n\times n$ matrix $\boldsymbol{A}$; while in MAT3040 we will study the eigenvalues of a **linear operator** $T: V \to V$.
- In MAT2040 we have studied the dot product $\boldsymbol{x}\cdot\boldsymbol{y} = \sum_{i=1}^n x_i y_i$; while in MAT3040 we will study the **inner product** $\langle \boldsymbol{v}_1, \boldsymbol{v}_2\rangle$.

Why do we do the generalization?. We are studying many other spaces, e.g., $\mathcal{C}(\mathbb{R})$ is called the space of all functions on $\mathbb{R}$, $\mathcal{C}^\infty(\mathbb{R})$ is called the space of all infinitely differentiable functions on $\mathbb{R}$, $\mathbb{R}[x]$ is the space of polynomials of one-variable.

■ **Example 1.1**

1. Consider the Laplace equation $\Delta f = 0$ with linear operator Δ:

$$\Delta : \mathcal{C}^\infty(\mathbb{R}^3) \to \mathcal{C}^\infty(\mathbb{R}^3) \quad f \mapsto (\frac{\partial^2}{\partial x^2} + \frac{\partial^2}{\partial y^2} + \frac{\partial^2}{\partial z^2})f$$

   The solution to the PDE $\Delta f = 0$ is the 0-eigenspace of Δ.

2. Consider the Schrödinger equation $\hat{H}f = Ef$ with the linear operator

$$\hat{H} : \mathcal{C}^\infty(\mathbb{R}^3) \to \mathbb{R}^3, \quad f \to \left[\frac{-\hbar^2}{2\mu}\nabla^2 + V(x,y,z)\right]f$$

   Solving the equation $\hat{H}f = Ef$ is equivalent to finding the eigenvectors of $\hat{H}$. In fact, the eigenvalues of $\hat{H}$ are **discrete**.

■

1.1.2. Vector Spaces

Definition 1.1 [Vector Space] A **vector space** over a field $\mathbb{F}$ (in particular, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$) is a set of objects V equipped with vector addiction and scalar multiplication such that

1. the vector addiction $+$ is closed with the rules:
   (a) **Commutativity**: $\forall \boldsymbol{v}_1, \boldsymbol{v}_2 \in V$, $\boldsymbol{v}_1 + \boldsymbol{v}_2 = \boldsymbol{v}_2 + \boldsymbol{v}_1$.
   (b) **Associativity**: $\boldsymbol{v}_1 + (\boldsymbol{v}_2 + \boldsymbol{v}_3) = (\boldsymbol{v}_1 + \boldsymbol{v}_2) + \boldsymbol{v}_3$.
   (c) **Addictive Identity**: $\exists \mathbf{0} \in V$ such that $\mathbf{0} + \boldsymbol{v} = \boldsymbol{v}$, $\forall \boldsymbol{v} \in V$.
2. the **scalar multiplication** is closed with the rules:
   (a) **Distributive**: $\alpha(\boldsymbol{v}_1 + \boldsymbol{v}_2) = \alpha\boldsymbol{v}_1 + \alpha\boldsymbol{v}_2, \forall \alpha \in \mathbb{F}$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$
   (b) **Distributive**: $(\alpha_1 + \alpha_2)\boldsymbol{v} = \alpha_1\boldsymbol{v} + \alpha_2\boldsymbol{v}$
   (c) **Compatibility**: $a(b\boldsymbol{v}) = (ab)\boldsymbol{v}$ for $\forall a, b \in \mathbb{F}$ and $\boldsymbol{b} \in V$.
   (d) $0\boldsymbol{v} = \mathbf{0}$, $1\boldsymbol{v} = \boldsymbol{v}$.

■

Here we study several examples of vector spaces:

■ **Example 1.2** For $V = \mathbb{F}^n$, we can define

1. Addictive Identity:
$$\mathbf{0} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix}$$
2. Scalar Multiplication:
$$\alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}$$
3. Vector Addiction:
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y_n \end{pmatrix} = \begin{pmatrix} x_1 + y_1 \\ \vdots \\ x_n + y_n \end{pmatrix}$$

■

■ **Example 1.3**

1. It is clear that the set $V = M_{n\times n}(\mathbb{F})$ (the set of all $m \times n$ matrices) is a vector space as well.
2. The set $V = \mathcal{C}(\mathbb{R})$ is a vector space:
   (a) Vector Addiction:
   $$(f+g)(x) = f(x) + g(x), \forall f, g \in V$$
   (b) Scalar Multiplication:
   $$(\alpha f)(x) = \alpha f(x), \forall \alpha \in \mathbb{R}, f \in V$$
   (c) Addictive Identity is a zero function, i.e., $\mathbf{0}(x) = 0$ for all $x \in \mathbb{R}$.

■

Definition 1.2 A sub-collection $W \subseteq V$ of a vector space V is called a **vector subspace of** V if W itself forms a vector space, denoted by $W \le V$.

Example 1.4

1. For $V = \mathbb{R}^3$, we claim that $W = \{(x,y,0) \mid x,y \in \mathbb{R}\} \le V$
2. $W = \{(x,y,1) \mid x,y \in \mathbb{R}\}$ is not the vector subspace of V.

Proposition 1.1 $W \subseteq V$ is a **vector subspace** of V iff for $\forall \boldsymbol{w}_1, \boldsymbol{w}_2 \in W$, we have $\alpha \boldsymbol{w}_1 + \beta \boldsymbol{w}_2 \in W$, for $\forall \alpha, \beta \in \mathbb{F}$.

Example 1.5

1. For $V = M_{n\times n}(\mathbb{F})$, the subspace $W = \{A \in V \mid \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}\} \le V$
2. For $V = \mathcal{C}^\infty(\mathbb{R})$, define $W = \{f \in V \mid \frac{\mathrm{d}^2}{\mathrm{d}x^2} f + f = 0\} \le V$. For $f, g \in W$, we have

$$(\alpha f + \beta g)'' = \alpha f'' + \beta g'' = \alpha(-f) + \beta(-g) = -(\alpha f + \beta g),$$

   which implies $(\alpha f + \beta g)'' + (\alpha f + \beta g) = 0$.

1.2. Monday for MAT3006

1.2.1. Overview on uniform convergence

Definition 1.3 [Convergence] Let $f_n(x)$ be a sequence of functions on an interval $I = [a,b]$. Then $f_n(x)$ converges **pointwise** to $f(x)$ (i.e., $f_n(x_0) \to f(x_0)$) for $\forall x_0 \in I$, if

$$\forall \varepsilon > 0, \exists N_{x_0,\varepsilon} \text{ such that } |f_n(x_0) - f(x_0)| < \varepsilon, \forall n \ge N_{x_0,\varepsilon}$$

We say $f_n(x)$ converges uniformly to $f(x)$, (i.e., $f_n(x) \rightrightarrows f(x)$) for $\forall x_0 \in I$, if

$$\forall \varepsilon > 0, \exists N_{\varepsilon} \text{ such that } |f_n(x_0) - f(x_0)| < \varepsilon, \forall n \ge N_{\varepsilon}$$

■ **Example 1.6** It is clear that the function $f_n(x) = \frac{n}{1+nx}$ converges pointwise into $f(x) = \frac{1}{x}$ on $[0,\infty)$, and it is uniformly convergent on $[1,\infty)$.

Proposition 1.2 If $\{f_n\}$ is a sequence of continuous functions on I, and $f_n(x) \rightrightarrows f(x)$, then the following results hold:

1. $f(x)$ is continuous on I.
2. f is (Riemann) integrable with $\int_a^b f_n(x)\,\mathrm{d}x \to \int_a^b f(x)\,\mathrm{d}x$.
3. Suppose furthermore that $f_n(x)$ is **continuously differentiable**, and $f_n'(x) \rightrightarrows g(x)$, then $f(x)$ is differentiable, with $f_n'(x) \to f'(x)$.

We can put the discussions above into the content of series, i.e., $f_n(x) = \sum_{k=1}^n S_k(x)$.

Proposition 1.3 If $S_k(x)$ is continuous for $\forall k$, and $\sum_{k=1}^n S_k \rightrightarrows \sum_{k=1}^\infty S_k$, then

1. $\sum_{k=1}^\infty S_k(x)$ is continuous,
2. The series $\sum_{k=1}^\infty S_k$ is (Riemann) integrable, with $\sum_{k=1}^\infty \int_a^b S_k(x)\,\mathrm{d}x = \int_a^b \sum_{k=1}^\infty S_k(x)\,\mathrm{d}x$
3. If $\sum_{k=1}^n S_k$ is continuously differentiable, and the derivative of which is uniform

convergent, then the series $\sum_{k=1}^{\infty} S_k$ is differentiable, with

$$\left(\sum_{k=1}^{\infty} S_k(x)\right)' = \sum_{k=1}^{\infty} S'_k(x)$$

Then we can discuss the properties for a special kind of series, say power series.

Proposition 1.4 Suppose the power series $f(x) = \sum_{k=1}^{\infty} a_k x^k$ has radius of convergence R, then

1. $\sum_{k=1}^{n} a_k x^k \rightrightarrows f(x)$ for any $[-L, L]$ with $L < R$.
2. The function $f(x)$ is continuous on $(-R, R)$, and moreover, is differentiable and (Riemann) integrable on $[-L, L]$ with $L < R$:

$$\int_0^x f(t)\,\mathrm{d}t = \sum_{k=1}^{\infty} \frac{a_k}{k+1} x^{k+1}$$

$$f'(x) = \sum_{k=1}^{\infty} k a_k x^{k-1}$$

1.2.2. Introduction to MAT3006

What are we going to do.

1. (a) Generalize our study of (sequence, series, functions) on $\mathbb{R}^n$ into a metric space.

   (b) We will study spaces outside $\mathbb{R}^n$.

   Remark:

   - For (a), different metric may yield different kind of convergence of sequences. For (b), one important example we will study is $X = \mathcal{C}[a, b]$ (all continuous functions defined on $[a, b]$.) We will generalize X into $\mathcal{C}_b(E)$, which means the set of bounded continuous functions defined on $E \subseteq \mathbb{R}^n$.
   - The insights of analysis is to find a **unified** theory to study sequences/series on a metric space X, e.g., $X = \mathbb{R}^n, \mathcal{C}[a, b]$. In particular, for $\mathcal{C}[a, b]$, we will see that
     - most functions in $\mathcal{C}[a, b]$ are nowhere differentiable. (repeat part of

content in MAT2006)

  - We will prove the existence and uniqueness of ODEs.
  - the set poly$[a,b]$ (the set of polynomials on $[a,b]$) is dense in $\mathcal{C}[a,b]$. (analogy: $\mathbb{Q} \subseteq \mathbb{R}$ is dense)

2. Introduction to the Lebesgue Integration.

   For convergence of integration $\int_a^b f_n(x)\,dx \to \int_a^b f(x)$, we need the pre-conditions (a) $f_n(x)$ is continuous, and (b) $f_n(x) \rightrightarrows f(x)$. The natural question is that can we relax these conditions to

   (a) $f_n(x)$ is integrable?

   (b) $f_n(x) \to f(x)$ pointwisely?

   The answer is yes, by using the tool of Lebesgue integration. If $f_n(x) \to f(x)$ and $f_n(x)$ is Lebesgue integrable, then $\int_a^b f_n(x)\,dx \to \int_a^b f(x)\,dx$, which is so called the **dominated convergence**.

1.2.3. Metric Spaces

We will study the **length** of an element, or the **distance** between two elements in an arbitrary set X. First let's discuss the length defined on a well-structured set, say vector space.

Definition 1.4 [Normed Space] Let X be a vector space. A **norm** on X is a function $\|\cdot\| : X \to \mathbb{R}$ such that

1. $\|\boldsymbol{x}\| \geq 0$ for $\forall \boldsymbol{x} \in X$, with equality iff $\boldsymbol{x} = \boldsymbol{0}$
2. $\|\alpha\boldsymbol{x}\| = |\alpha|\|\boldsymbol{x}\|$, for $\forall \alpha \in \mathbb{R}$ and $\boldsymbol{x} \in X$.
3. $\|\boldsymbol{x} + \boldsymbol{y}\| \leq \|\boldsymbol{x}\| + \|\boldsymbol{y}\|$ (triangular inequality)

Any vector space equipped with $\|\cdot\|$ is called a **normed space**.

Example 1.7 1. For $X = \mathbb{R}^n$, define

$$\|\boldsymbol{x}\|_2 = \left(\sum_{i=1}^n x_i^2\right)^{1/2} \quad \text{(Euclidean Norm)}$$

$$\|\boldsymbol{x}\|_p = \left(\sum_{i=1}^n |x_i|^p\right)^{1/p} \quad (p\text{-norm})$$

2. For $X = \mathcal{C}[a,b]$, define

$$\|f\|_\infty = \max_{x\in[a,b]} |f(x)|$$

$$\|f\|_p = \left(\int_a^b |f(x)|^p \,\mathrm{d}x\right)^{1/p}$$

Exercise: check the norm defined above are well-defined.

Here we can define the distance in an arbitrary set:

Definition 1.5 A set X is a **metric space** with metric (X,d) if there exists a (distance) function $d : X \times X \to \mathbb{R}$ such that

1. $d(\boldsymbol{x},\boldsymbol{y}) \geq 0$ for $\forall \boldsymbol{x},\boldsymbol{y} \in X$, with equality iff $\boldsymbol{x} = \boldsymbol{y}$.
2. $d(\boldsymbol{x},\boldsymbol{y}) = d(\boldsymbol{y},\boldsymbol{x})$.
3. $d(\boldsymbol{x},\boldsymbol{z}) \leq d(\boldsymbol{x},\boldsymbol{y}) + d(\boldsymbol{y},\boldsymbol{z})$.

Example 1.8 1. If X is a normed space, then define $d(\boldsymbol{x},\boldsymbol{y}) = \|\boldsymbol{x}-\boldsymbol{y}\|$, which is so called the metric induced from the norm $\|\cdot\|$.

2. Let X be any (non-empty) set with $\boldsymbol{x},\boldsymbol{y} \in X$, the discrete metric is given by:

$$d(\boldsymbol{x},\boldsymbol{y}) = \begin{cases} 0, & \text{if } x = y \\ 1, & \text{if } x \neq y \end{cases}$$

Exercise: check the metric space defined above are well-defined.

R Adopting the infinite norm discussed in Example (1.7), we can define a metric on $\mathcal{C}[a,b]$ by

$$d_\infty(f,g) = \|f-g\|_\infty := \max_{x\in[a,b]} |f(x)-g(x)|$$

which is the correct metric to study the uniform convergence for $\{f_n\} \subseteq C[a,b]$.

Definition 1.6 Let (X,d) be a metric space. An **open ball** centered at $\boldsymbol{x} \in X$ of radius r is the set

$$B_r(\boldsymbol{x}) = \{\boldsymbol{y} \in X \mid d(\boldsymbol{x},\boldsymbol{y}) < r\}.$$

■ **Example 1.9** 1. For $X = \mathbb{R}^2$, we can draw the $B_1(\boldsymbol{0})$ with respect to the metrics d_1, d_2:

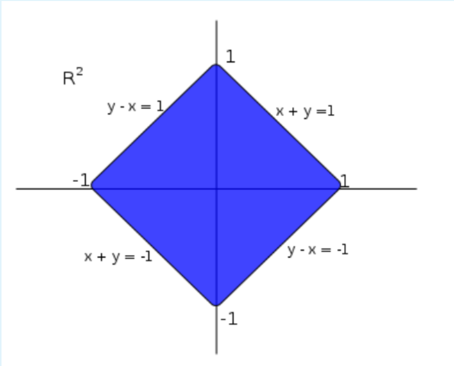


Figure 1.1: $B_1(\boldsymbol{0})$ w.r.t. the metric d_1

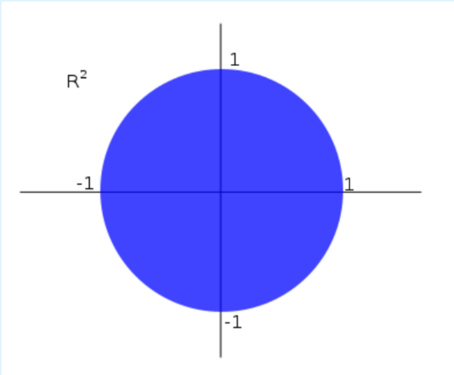
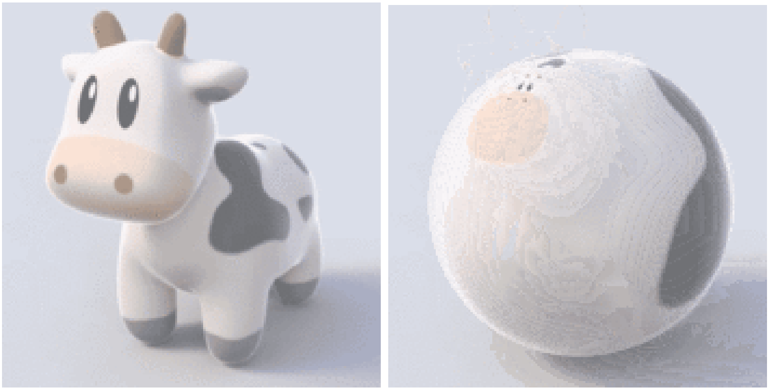


Figure 1.2: $B_1(\boldsymbol{0})$ w.r.t. the metric d_2

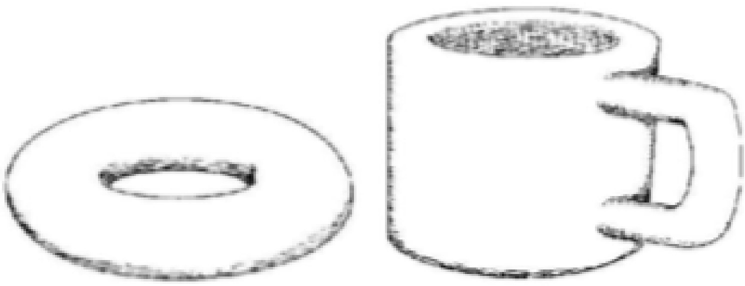
1.3. Monday for MAT4002

1.3.1. Introduction to Topology

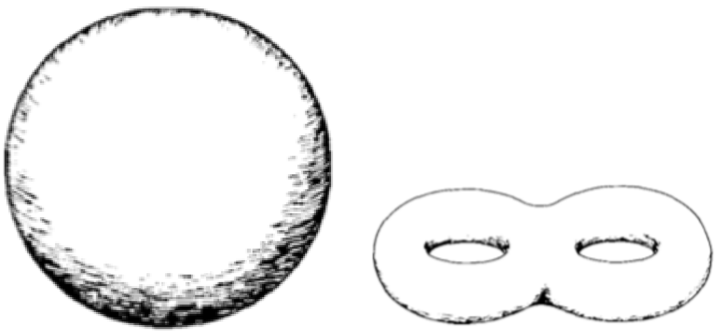
We will study global properties of a geometric object, i.e., *the distrance between 2 points in an object is totally ignored*. For example, the objects shown below are essentially invariant under a certain kind of transformation:



Another example is that the coffee cup and the donut have the same topology:



However, the two objects below have the intrinsically different topologies:



In this course, we will study the phenomenon described above mathematically.

1.3.2. Metric Spaces

In order to ingnore about the distances, we need to learn about distances first.

Definition 1.7 [Metric Space] Metric space is a set X where one can measure distance between any two objects in X.

Specifically speaking, a metric space X is a non-empty set endowed with a function (distance function) $d : X \times X \to \mathbb{R}$ such that

1. $d(\boldsymbol{x},\boldsymbol{y}) \geq 0$ for $\forall \boldsymbol{x},\boldsymbol{y} \in X$ with equality iff $\boldsymbol{x} = \boldsymbol{y}$
2. $d(\boldsymbol{x},\boldsymbol{y}) = d(\boldsymbol{y},\boldsymbol{x})$
3. $d(\boldsymbol{x},\boldsymbol{z}) \leq d(\boldsymbol{x},\boldsymbol{y}) + d(\boldsymbol{y},\boldsymbol{z})$ (triangular inequality)

■

■ **Example 1.10** 1. Let $X = \mathbb{R}^n$, with

$$d_2(\boldsymbol{x},\boldsymbol{y}) = \sqrt{\sum_{i=1}^{n}(x_i - y_i)^2}$$

$$d_\infty(\boldsymbol{x},\boldsymbol{y}) = \max_{i=1,\dots,n} |x_i - y_i|$$

2. Let X be any set, and define the discrete metric

$$d(\boldsymbol{x},\boldsymbol{y}) = \begin{cases} 0, & \text{if } x = y \\ 1, & \text{if } x \neq y \end{cases}$$

Homework: Show that (1) and (2) defines a metric. ■

Definition 1.8 [Open Ball] An **open ball** of radius r centered at $\boldsymbol{x} \in X$ is the set

$$B_r(\boldsymbol{x}) = \{\boldsymbol{y} \in X \mid d(\boldsymbol{x},\boldsymbol{y}) < r\}$$

■

■ **Example 1.11** 1. The set $B_1(0,0)$ defines an open ball under the metric $(X = \mathbb{R}^2, d_2)$, or the metric $(X = \mathbb{R}^2, d_\infty)$. The corresponding diagram is shown below:

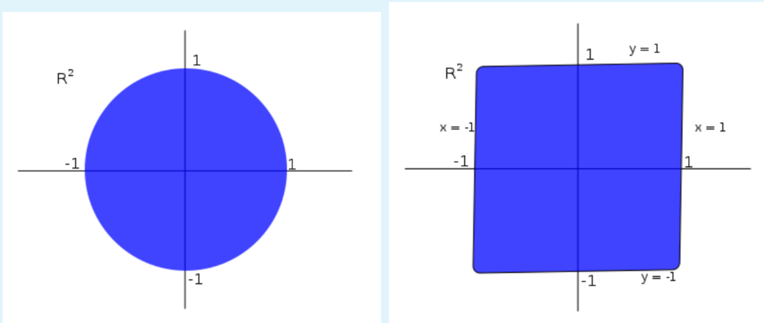


Figure 1.3: Left: under the metric $(X = \mathbb{R}^2, d_2)$; Right: under the metric $(X = \mathbb{R}^2, d_\infty)$

2. Under the metric $(X = \mathbb{R}^2, \text{discrete metric})$, the set $B_1(0,0)$ is one single point, also defines an open ball.

■

Definition 1.9 [Open Set] Let X be a metric space, $U \subseteq X$ is an open set in X if $\forall u \in U$, there exists $\epsilon_u > 0$ such that $B_{\epsilon_u}(u) \subseteq U$. ■

Definition 1.10 The **topology** induced from (X, d) is the collection of all open sets in (X, d), denoted as the symbol $\mathcal{T}$. ■

Proposition 1.5 All open balls $B_r(\boldsymbol{x})$ are open in (X, d).

Proof. Consider the example $X = \mathbb{R}$ with metric d_2. Therefore $B_r(x) = (x - r, x + r)$. Take $\boldsymbol{y} \in B_r(\boldsymbol{x})$ such that $d(\boldsymbol{x}, \boldsymbol{y}) = q < r$ and consider $B_{(r-q)/2}(\boldsymbol{y})$: for all $z \in B_{(r-q)/2}(\boldsymbol{y})$, we have

$$d(\boldsymbol{x}, \boldsymbol{z}) \leq d(\boldsymbol{x}, \boldsymbol{y}) + d(\boldsymbol{y}, \boldsymbol{z}) < q + \frac{r-q}{2} < r,$$

which implies $\boldsymbol{z} \in B_r(x)$. ■

Proposition 1.6 Let $(X, \boldsymbol{d})$ be a metric space, and $\mathcal{T}$ is the topology induced from (X, d), then

1. let the set $\{G_\alpha \mid \alpha \in \mathcal{A}\}$ be a collection of (uncountable) open sets, i.e., $G_\alpha \in \mathcal{T}$,

then $\bigcup_{\alpha \in \mathcal{A}} G_\alpha \in \mathcal{T}$.

2. let $G_1, \ldots, G_n \in \mathcal{T}$, then $\bigcap_{i=1}^n G_i \in \mathcal{T}$. The finite intersection of open sets is open.

Proof. 1. Take $x \in \bigcup_{\alpha \in \mathcal{A}} G_\alpha$, then $x \in G_\beta$ for some $\beta \in \mathcal{A}$. Since G_β is open, there exists $\epsilon_x > 0$ s.t.

$$B_{\epsilon_x}(x) \subseteq G_\beta \subseteq \bigcup_{\alpha \in \mathcal{A}} G_\alpha$$

2. Take $x \in \bigcap_{i=1}^n G_i$, i.e., $x \in G_i$ for $i = 1, \ldots, n$, i.e., there exists $\epsilon_i > 0$ such that $B_{\epsilon_i}(x) \subseteq G_i$ for $i = 1, \ldots, n$. Take $\epsilon = \min\{\epsilon_1, \ldots, \epsilon_n\}$, which implies

$$B_\epsilon(x) \subseteq B_{\epsilon_i}(x) \subseteq G_i, \forall i$$

which implies $B_\epsilon(x) \subseteq \bigcap_{i=1}^n G_i$

■

Exercise.

1. let $\mathcal{T}_2, \mathcal{T}_\infty$ be topologies induced from the metrices d_2, d_∞ in $\mathbb{R}^2$. Show that $J_2 = J_\infty$, i.e., every open set in $(\mathbb{R}^2, d_2)$ is open in $(\mathbb{R}^2, d_\infty)$, and every open set in $(\mathbb{R}^2, d_\infty)$ is open in $(\mathbb{R}_2, d_2)$.
2. Let $\mathcal{T}$ be the topology induced from the discrete metric (X, d_{discrete}). What is $\mathcal{T}$?

1.4. Wednesday for MAT3040

1.4.1. Review

1. Vector Space: e.g., $\mathbb{R}, M_{n\times n}(\mathbb{R}), \mathcal{C}(\mathbb{R}^n), \mathbb{R}[x]$.
2. Vector Subspace: $W \le V$, e.g.,
   (a) $V = \mathbb{R}^2$, the set $W := \mathbb{R}^2_+$ is not a vector subspace since W is not closed under scalar multiplication;
   (b) the set $W = \mathbb{R}^2_+ \bigcup \mathbb{R}^2_-$ is not a vector subspace since it is not closed under addition.
   (c) For $V = \mathbb{M}_{3\times 3}(\mathbb{R})$, the set of invertible 3×3 matrices is not a vector subspace, since we cannot define zero vector inside.
   (d) Exercise: How about the set of all singular matrices? Answer: it is not a vector subspace since the vector addition does not necessarily hold.

1.4.2. Spanning Set

Definition 1.11 [Span] Let V be a vector space over $\mathbb{F}$:

1. A linear combination of a subset S in V is of the form

$$\sum_{i=1}^{n} \alpha_i \boldsymbol{s}_i, \quad \alpha_i \in \mathbb{F}, \boldsymbol{s}_i \in S$$

   Note that the summation should be finite.
2. The **span** of a subset $S \subseteq V$ is

$$\operatorname{span}(S) = \left\{ \sum_{i=1}^{n} \alpha_i \boldsymbol{s}_i \,\middle|\, \alpha_i \in \mathbb{F}, \boldsymbol{s}_i \in S \right\}$$

3. S is a spanning set of V, or say S spans V, if

$$\operatorname{span}(S) = V.$$

■ **Example 1.12** For $V = \mathbb{R}[x]$, define the set

$$S = \{1, x^2, x^4, \dots, x^6\},$$

then $2 + x^4 + \pi x^{106} \in \operatorname{span}(S)$, while the series $1 + x^2 + x^4 + \cdots \notin \operatorname{span}(S)$.

It is clear that $\operatorname{span}(S) \neq V$, but S is the spanning set of $W = \{p \in V \mid p(x) = p(-x)\}$.

■ **Example 1.13** For $V = M_{3\times 3}(\mathbb{R})$, let $W_1 = \{\boldsymbol{A} \in V \mid \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}\}$ and $W_2 = \{\boldsymbol{B} \in V \mid \boldsymbol{B}^{\mathrm{T}} = -\boldsymbol{B}\}$ (the set of skew-symmetric matrices) be two vector subspaces. Define the set

$$\boldsymbol{S} := W_1 \bigcup W_2$$

Exercise: $\boldsymbol{S}$ spans V.

Proposition 1.7 Let S be a subset in a vector space V.

1. $S \subseteq \operatorname{span}(S)$
2. $\operatorname{span}(S) = \operatorname{span}(\operatorname{span}(S))$
3. If $\boldsymbol{w} \in \operatorname{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\} \setminus \operatorname{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$, then

$$\boldsymbol{v}_1 \in \operatorname{span}\{\boldsymbol{w}, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\} \setminus \operatorname{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$$

Proof. 1. For each $\boldsymbol{s} \in S$, we have

$$\boldsymbol{s} = 1 \cdot \boldsymbol{s} \in \operatorname{span}(S)$$

2. From (1), it's clear that $\operatorname{span}(S) \subseteq \operatorname{span}(\operatorname{span}(S))$, and therefore suffices to show $\operatorname{span}(\operatorname{span}(S)) \subseteq \operatorname{span}(S)$:

Pick $\boldsymbol{v} = \sum_{i=1}^{n} \alpha_i \boldsymbol{v}_i \in \operatorname{span}(\operatorname{span}(S))$, where $\boldsymbol{v}_i \in \operatorname{span}(S)$. Rewrite

$$\boldsymbol{v}_i = \sum_{j=1}^{n_i} \beta_{ij} \boldsymbol{s}_j, \quad \boldsymbol{s}_j \in S,$$

which implies

$$\begin{aligned}\boldsymbol{v} &= \sum_{i=1}^{n} \alpha_i \sum_{j=1}^{n_i} \beta_{ij} \boldsymbol{s}_j \\ &= \sum_{i=1}^{n} \sum_{j=1}^{n_i} (\alpha_i \beta_{ij}) \boldsymbol{s}_j,\end{aligned}$$

i.e., $\boldsymbol{v}$ is the finite combination of elements in S, whcih implies $\boldsymbol{v} \in \operatorname{span}(S)$.

3. By hypothesis, $\boldsymbol{w} = \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_n \boldsymbol{v}_n$ with $\alpha_1 \neq 0$, which implies

$$\boldsymbol{v}_1 = -\frac{\alpha_2}{\alpha_1} \boldsymbol{v}_2 + \cdots + \left(-\frac{1}{\alpha_1} \boldsymbol{w} \right)$$

which implies $\boldsymbol{v}_1 \in \operatorname{span}\{\boldsymbol{w}, \boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$. It suffices to show $\boldsymbol{v}_1 \notin \operatorname{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$. Suppose on the contrary that $\boldsymbol{v}_1 \in \operatorname{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$. It's clear that $\operatorname{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\} = \operatorname{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_n\}$. (left as exercise). Therefore,

$$\emptyset = \operatorname{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\} \setminus \operatorname{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_n\},$$

which is a contradiction.

■

1.4.3. Linear Independence and Basis

Definition 1.12 [Linear Independence] Let S be a (not necessarily finite) subset of V. Then S is **linearly independent** (l.i.) on V if for any finite subset $\{\boldsymbol{s}_1, \dots, \boldsymbol{s}_k\}$ in S,

$$\sum_{i=1}^{k} \alpha_i \boldsymbol{s}_i = 0 \Longleftrightarrow \alpha_i = 0, \forall i$$

■ **Example 1.14** For $V = \mathcal{C}(\mathbb{R})$,

1. let $S_1 = \{\sin x, \cos x\}$, which is l.i., since

$$\alpha \sin x + \beta \cos x = \mathbf{0}(\text{means zero function})$$

Taking $x = 0$ both sides leads to $\beta = 0$; taking $x = \frac{\pi}{2}$ both sides leads to $\alpha = 0$.

2. let $S_2 = \{\sin^2 x, \cos^2 x, 1\}$, which is linearly dependent, since

$$1 \cdot \sin^2 x + 1 \cdot \cos^2 x + (-1) \cdot 1 = 0, \forall x$$

3. Exercise: For $V = \mathbb{R}[x]$, let $S = \{1, x, x^2, x^3, \dots, \}$, which is l.i.:

Pick $x^{k_1}, \dots, x^{k_n} \in S$ with $k_1 < \cdots < k_n$. Consider that the euqation

$$\alpha_1 x^{k_1} + \cdots + \alpha_n x^{k_n} = \mathbf{0}$$

holds for all x, and try to solve for $\alpha_1, \dots, \alpha_n$ (one way is differentation.)

Definition 1.13 [Basis] A subset S is a **basis** of V if

(a) S spans V;

(b) S is l.i.

■ **Example 1.15**

1. For $V = \mathbb{R}^n$, $S = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}$ is a basis of V
2. For $V = \mathbb{R}[x]$, $S = \{1, x, x^2, \dots\}$ is a basis of V
3. For $V = M_{2\times 2}(\mathbb{R})$,

$$S = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix} \right\}$$

is a basis of V

(R) Note that there can be many basis for a vector space V.

Proposition 1.8 Let $V = \text{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$, then there exists a subset of $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$, which is a basis of V.

Proof. If $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$ is l.i., the proof is complete.

Suppose not, then $\alpha_1\boldsymbol{v}_1 + \cdots + \alpha_m\boldsymbol{v}_m = \boldsymbol{0}$ has a non-trivial solution. w.l.o.g., $\alpha_1 \neq 0$, which implies

$$\boldsymbol{v}_1 = -\frac{\alpha_2}{\alpha_1}\boldsymbol{v}_2 + \cdots + \left(\frac{\alpha_m}{\alpha_1}\right)\boldsymbol{v}_m \implies \boldsymbol{v}_1 \in \text{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_m\}$$

By the proof in (c), Proposition (1.7),

$$\text{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\} = \text{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_m\},$$

which implies $V = \text{span}\{\boldsymbol{v}_2, \dots, \boldsymbol{v}_m\}$.

Continuse this argument finitely many times to guarantee that $\{\boldsymbol{v}_i, \boldsymbol{v}_{i+1}, \dots, \boldsymbol{v}_m\}$ is l.i., and spans V. The proof is complete. ∎

Corollary 1.1 If $V = \text{span}\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$ (i.e., V is finitely generated), then V has a basis. (The same holds for non-finitely generated V).

Proposition 1.9 If $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ is a basis of V, then every $\boldsymbol{v} \in V$ can be expressed uniquely as

$$\boldsymbol{v} = \alpha_1\boldsymbol{v}_1 + \cdots + \alpha_n\boldsymbol{v}_n$$

Proof. Since $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ spans V, so $\boldsymbol{v} \in V$ can be written as

$$\boldsymbol{v} = \alpha_1\boldsymbol{v}_1 + \cdots + \alpha_n\boldsymbol{v}_n \tag{1.1}$$

Suppose further that

$$\boldsymbol{v} = \beta_1 \boldsymbol{v}_1 + \cdots + \beta_n \boldsymbol{v}_n, \tag{1.2}$$

it suffices to show that $\alpha_i = \beta_i$ for $\forall i$:

Subtracting (1.1) into (1.2) leads to

$$(\alpha_1 - \beta_1)\boldsymbol{v}_1 + \cdots + (\alpha_n - \beta_n)\boldsymbol{v}_n = 0.$$

By the hypothesis of linear independence, we have $\alpha_i - \beta_i = 0$ for $\forall i$, i.e., $\alpha_i = \beta_i$. ■

1.5. Wednesday for MAT3006

Reviewing.

- Normed Space: a norm on a vector space
- Metric Space
- Open Ball

1.5.1. Convergence of Sequences

Since $\mathbb{R}^n$ and $\mathcal{C}[a,b]$ are both metric spaces, we can study the convergence in $\mathbb{R}^n$ and the functions defined on $[a,b]$ at the same time.

Definition 1.14 [Convergence] Let (X,d) be a metric space. A sequence $\{x_n\}$ in X is **convergent** to x if $\forall \varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$d(x_n, x) < \varepsilon, \forall n \geq N.$$

We can denote the convergence by

$$x_n \to x, \quad \text{or} \quad \lim_{n\to\infty} x_n = x, \quad \text{or} \quad \lim_{n\to\infty} d(x_n, x) = 0$$

Proposition 1.10 If the limit of $\{x_n\}$ exists, then it is unique.

(R) Note that the proposition above does not necessarily hold for topology spaces.

Proof. Suppose $x_n \to x$ and $x_n \to y$, which implies

$$0 \leq d(x,y) \leq d(x,x_n) + d(x_n,y), \forall n$$

Taking the limit $n \to \infty$ both sides, we imply $d(x,y) = 0$, i.e., $x = y$. ∎

■ **Example 1.16** 1. Consider the metric space $(\mathbb{R}^k, d_\infty)$ and study the convergence

$$\lim_{n\to\infty} \boldsymbol{x}_n = \boldsymbol{x} \iff \lim_{n\to\infty}\left(\max_{i=1\dots,k}|x_{n_i} - x_i|\right) = 0$$

$$\iff \lim_{n\to\infty}|x_{n_i} - x_i| = 0, \forall i = 1,\dots,k$$

$$\iff \lim_{n\to\infty} x_{n_i} = x_i,$$

i.e., the convergence defined in d_∞ is the same as the convergence defined in d_2.

2. Consider the convergence in the metric space $(\mathcal{C}[a,b], d_\infty)$:

$$\lim_{n\to\infty} f_n = f \iff \lim_{n\to\infty}\left(\max_{[a,b]}|f_n(x) - f(x)|\right) = 0$$

$$\iff \forall \varepsilon > 0, \forall x \in [a,b], \exists N_\varepsilon \text{ such that } |f_n(x) - f(x)| < \varepsilon, \forall n \geq N_\varepsilon$$

which is equivalent to the uniform convergence of functions, i.e., the convergence defined in d_2.

■

Definition 1.15 [Equivalent metrics] Let d and ρ be metrics on X.

1. We say ρ is **stronger** than d (or d is **weaker** than ρ) if

$$\exists K > 0 \text{ such that } d(x,y) \leq K\rho(x,y), \forall x,y \in X$$

2. The metrics d and ρ are equivalent if there exists $K_1, K_2 > 0$ such that

$$d(x,y) \leq K_1\rho(x,y) \leq K_2 d(x,y)$$

■

(R) The strongerness of ρ than d is depicted in the graph below:

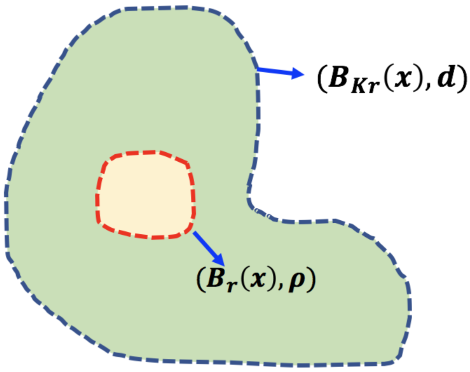


Figure 1.4: The open ball $(B_r(x), \rho)$ is contained by the open ball $(B_{Kr}(x), d)$

For each $x \in X$, consider the open ball $(B_r(x), \rho)$ and the open ball $(B_{Kr}(x), d)$:

$$B_r(x) = \{y \mid \rho(x,y) < r\}, \quad B_{Kr}(x) = \{z \mid d(x,z) < Kr\}.$$

For $y \in (B_r(x), \rho)$, we have $d(x,y) < K\rho(x,y) < Kr$, which implies $y \in (B_{Kr}(x), d)$, i.e, $(B_r(x), \rho) \subseteq (B_{Kr}(x), d)$ for any $x \in X$ and $r > 0$.

■ **Example 1.17** 1. d_1, d_2, d_∞ in $\mathbb{R}^n$ are equivalent

$$d_1(\boldsymbol{x}, \boldsymbol{y}) \le d_\infty(\boldsymbol{x}, \boldsymbol{y}) \le n d_1(\boldsymbol{x}, \boldsymbol{y})$$
$$d_2(\boldsymbol{x}, \boldsymbol{y}) \le d_\infty(\boldsymbol{x}, \boldsymbol{y}) \le \sqrt{n} d_2(\boldsymbol{x}, \boldsymbol{y})$$

We use two relation depiected in the figure below to explain these two inequalities:

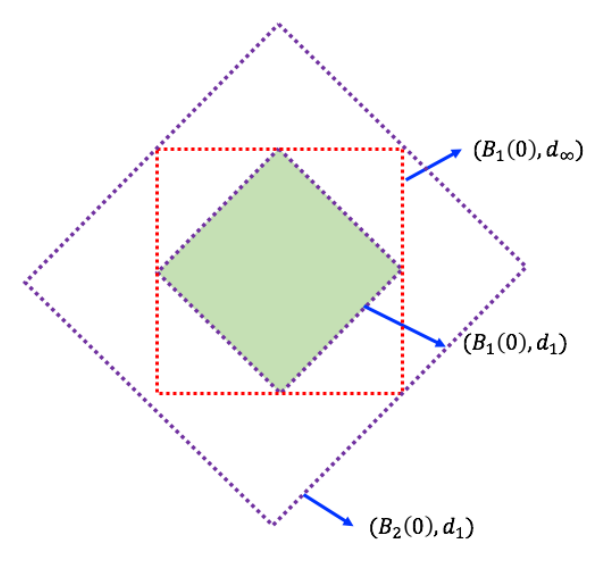


Figure 1.5: The diagram for the relation $(B_1(x), d_1) \subseteq (B_\infty(x), d_\infty) \subseteq (B_2(x), d_1)$ on $\mathbb{R}^2$

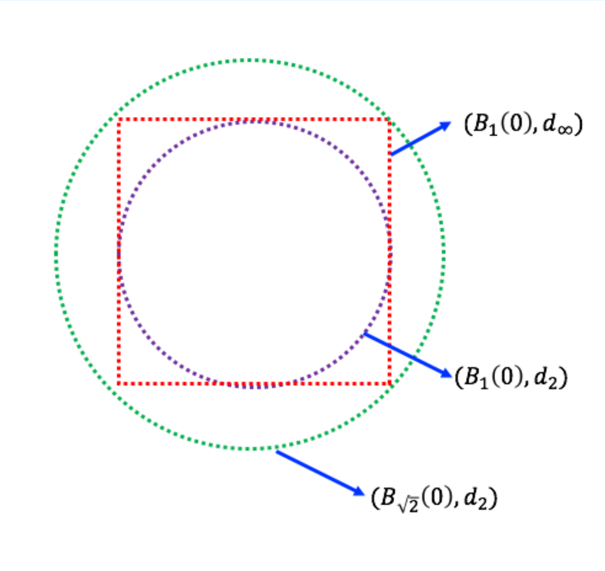


Figure 1.6: The diagram for the relation $(B_1(x), d_2) \subseteq (B_\infty(x), d_\infty) \subseteq (B_{\sqrt{2}}(x), d_2)$ on $\mathbb{R}^2$

It's easy to conclude the simple generalization for example (1.16):

Proposition 1.11 If d and ρ are equivalent, then

$$\lim_{n\to\infty} d(x_n, x) = 0 \Longleftrightarrow \lim_{n\to\infty} \rho(x_n, x) = 0$$

Note that this does not necessarily hold for topology spaces.

2. Consider d_1, d_∞ in $\mathcal{C}[a,b]$:

$$d_1(f,g) := \int_a^b |f-g|\,\mathrm{d}x \le \int_a^b \sup_{[a,b]} |f-g|\,\mathrm{d}x = (b-a)d_\infty(f,g),$$

i.e., d_∞ is stronger than d_1. Question: Are they equivalent? **No**.

Justification. Consider $f_n(x) = n^2x^n(1-x)$. Check that

$$\lim_{n\to\infty} d_1(f_n(x), 1) = 0, \quad \text{but } d_\infty(f_n(x), 1) \to \infty$$

The peak of f_n may go to infinite, while the integration converges to zero. Therefore d_1 and d_∞ have different limits. We will discuss this topic at Lebsegue integration again. ■

1.5.2. Continuity

Definition 1.16 [Continuity] Let $f : (X,d) \to (Y,d)$ be a function and $x_0 \in X$. Then f is continuous at x_0 if $\forall \varepsilon > 0$, there exists $\delta > 0$ such that

$$d(x,x_0) < \delta \implies \rho(f(x), f(x_0)) < \varepsilon$$

The function f is continuous in X if f is continuous for all $x_0 \in X$.

Proposition 1.12 The function f is continuous at x if and only if for all $\{x_n\} \to x$ under d, $f(x_n) \to f(x)$ under ρ.

Proof. Necessity: Given $\varepsilon > 0$, by continuity,

$$d(x,x') < \delta \implies \rho(f(x'), f(x)) < \varepsilon. \tag{1.3}$$

Consider the sequence $\{x_n\} \to x$, then there exists N such that $d(x_n, x) < \delta$ for $\forall n \ge N$. By applying (1.3), $\rho(f(x_n), f(x)) < \varepsilon$ for $\forall n \ge N$, i.e., $f(x_n) \to f(x)$.

Sufficiency: Assume that f is not continuous at x, then there exists ε_0 such that for $\delta_n = \frac{1}{n}$, there exists x_n such that

$$d(x_n, x) < \delta_n, \text{ but } \rho(f(x_n), f(x)) > \varepsilon_0.$$

Then $\{x_n\} \to x$ by our construction, while $\{f(x_n)\}$ does not converge to $f(x)$, which is a contradiction. ■

Corollary 1.2 If the function $f : (X, d) \to (Y, \rho)$ is continuous at x, the function $g : (Y, \rho) \to (Z, m)$ is continuous at $f(x)$, then $g \circ f : (X, d) \to (Z, m)$ is continuous at x.

Proof. Note that

$$\{x_n\} \to x \xRightarrow{(a)} \{f(x_n)\} \to f(x) \xRightarrow{(b)} \{g(f(x_n))\} \to g(f(x)) \xRightarrow{(c)} g \circ f \text{ is continuous at } x.$$

where $(a), (b), (c)$ are all by proposition (1.12). ■

1.5.3. Open and Closed Sets

We have open/closed intervals in $\mathbb{R}$, and they are important in some theorems (e.g, continuous functions bring closed intervals to closed intervals).

Definition 1.17 [Open] Let (X, d) be a metric space. A set $U \subseteq X$ is open if for each $x \in U$, there exists $\rho_x > 0$ such that $B_{\rho_x}(x) \subseteq U$. The empty set $\emptyset$ is defined to be open.

■

■ **Example 1.18** Let $(\mathbb{R}, d_2 \text{ or } d_\infty)$ be a metric space. The set $U = (a, b)$ is open. ■

Proposition 1.13
1. Let (X, d) be a metric space. Then all open balls $B_r(x)$ are open
2. All open sets in X can be written as a union of open balls.

Proof. 1. Let $y \in B_r(x)$, i.e., $d(x, y) := q < r$. Consider the open ball $B_{(r-q)/2}(y)$. It suffices to show $B_{(r-q)/2}(y) \subseteq B_r(x)$. For any $z \in B_{(r-q)/2}(y)$,

$$d(x,z) \leq d(x,y) + d(y,z) < q + \frac{r-q}{2} = \frac{r+q}{2} < r.$$

The proof is complete.

2. Let $U \subseteq X$ be open, i.e., for $\forall x \in U$, there exists $\varepsilon_x > 0$ such that $B_{\varepsilon_x}(x) \subseteq U$. Therefore

$$\{x\} \subseteq B_{\varepsilon_x}(x) \subseteq U, \forall x \in U$$

which implies

$$U = \bigcup_{x \in U} \{x\} \subseteq \bigcup_{x \in U} B_{\varepsilon_x}(x) \subseteq U,$$

i.e., $U = \bigcup_{x \in U} B_{\varepsilon_x}(x)$.

■

1.6. Wednesday for MAT4002

Reviewing.

- Metric Space (X, d)
- Open balls and open sets (note that the emoty set $\emptyset$ is open)
- Define the collection of open sets in X, say $\mathcal{T}$ is the topology.

Exercise.

1. Show that the $\mathcal{T}_2$ under $(X = \mathbb{R}^2, d_2)$ and $\mathcal{T}_\infty$ under $(X = \mathbb{R}^2, d_\infty)$ are the same.

   Ideas. Follow the procedure below:

   An open ball in d_2-metric is open in d_∞;

   Any open set in d_2-metric is open in d_∞;

   Switch d_2 and d_∞. ■

2. Describe the topology $\mathcal{T}_{\text{discrete}}$ under the metric space $(X = \mathbb{R}^2, d_{\text{discrete}})$.

   Outlines. Note that $\{x\} = B_{1/2}(x)$ is an open set.

   For any subset $W \subseteq \mathbb{R}^2$, $W = \bigcup_{w \in W}\{w\}$ is open.

   Therefore $\mathcal{T}_{\text{discrete}}$ is all subsets of $\mathbb{R}^2$. ■

1.6.1. Forget about metric

Next, we will try to define closedness, compactness, etc., without using the tool of metric:

Definition 1.18 [closed] A subset $V \subseteq X$ is **closed** if $X \setminus V$ is **open**.

Example 1.19 Under the metric space $(\mathbb{R}, d_1)$,

$$\mathbb{R} \setminus [b, a] = (a, \infty) \bigcup (-\infty, b) \text{ is open} \implies [b, a] \text{ is closed}$$

Proposition 1.14 Let X be a metric space.

1. $\emptyset, X$ is closed in X
2. If F_α is closed in X, so is $\bigcap_{\alpha \in A} F_\alpha$.
3. If $F_1, \dots, F_k$ is closed, so is $\bigcup_{i=1}^k F_i$.

Proof.

1. Note that X is open in X, which implies $\emptyset = X \setminus X$ is closed in X; Similarly, $\emptyset$ is open in X, which implies $X = X \setminus \emptyset$ is closed in X;
2. The set F_α is closed implies there exists open $U_\alpha \subseteq X$ such that $F_\alpha = X \setminus U_\alpha$. By De Morgan's Law,
$$\bigcap_{\alpha \in A} F_\alpha = \bigcap_{\alpha \in A} (X \setminus U_\alpha) = X \setminus (\bigcup_{\alpha \in A} U_\alpha).$$
By part (a) in proposition (1.6), the set $\bigcup_{\alpha \in A} U_\alpha$ is openm which implies $\bigcap_{\alpha \in A} F_\alpha$ is closed.
3. The result follows from part (b) in proposition (1.6) by taking complements.

■

We illustrate examples where open set is used to define convergence and continuity.

1. Convergence of sequences:

   Definition 1.19 [Convergence] Let (X, d) be a metric space, then $\{x_n\} \to x$ means
   $$\forall \varepsilon > 0, \exists N \text{ such that } d(x_n, x) < \varepsilon, \forall n \geq N.$$

   We will study the convergence by using open sets instead of metric.

   Proposition 1.15 Let X be a metric space, then $\{x_n\} \to x$ if and only if for $\forall$ open set $U \ni x$, there exists N such that $x_n \in U$ for $\forall n \geq N$.

   Proof. *Necessity*: Since $U \ni x$ is open, there exists $\varepsilon > 0$ such that $B_\varepsilon(x) \subseteq U$. Since $\{x_n\} \to x$, there exists N such that $d(x_n, x) < \varepsilon$, i.e., $x_n \in B_\varepsilon(x) \subseteq U$ for $\forall n \geq N$.

Sufficiency: Let $\varepsilon > 0$ be given. Take the open set $U = B_\varepsilon(x) \ni x$, then there exists N such that $x_n \in U = B_\varepsilon(x)$ for $\forall n \geq N$, i.e., $d(x_n, x) < \varepsilon$, $\forall n \geq N$.

■

2. Continuity:

Definition 1.20 [Continuity] Let (X, d) and (Y, ρ) be given metric spaces. Then $f : X \to Y$ is continuous at $x_0 \in X$ if

$$\forall \varepsilon > 0, \exists \delta > 0 \text{ such that } d(x, x_0) < \delta \implies \rho(f(x), f(x_0)) < \varepsilon.$$

The function f is continuous on X if f is continous for all $x_0 \in X$.

We can get rid of metrics to study continuity:

Proposition 1.16 (a) The function f is continuous at x if and only if for all open $U \ni f(x)$, there exists $\delta > 0$ such that the set $B(x, \delta) \subseteq f^{-1}(U)$.

(b) The function f is continuous on X if and only if $f^{-1}(U)$ is open in X for each open set $U \subseteq Y$.

During the proof we will apply a small lemma:

Proposition 1.17 f is continuous at x if and only if for all $\{x_n\} \to x$, we have $\{f(x_n)\} \to f(x)$.

Proof. (a) *Necessity*:

Due to the openness of $U \ni f(x)$, there exists a ball $B(f(x), \varepsilon) \subseteq U$.

Due to the continuity of f at x, there exists $\delta > 0$ such that $d(x, x') < \delta$ implies $d(f(x), f(x')) < \varepsilon$, which implies

$$f(B(x, \delta)) \subseteq B(f(x), \varepsilon) \subseteq U,$$

which implies $B(x, \delta) \subseteq f^{-1}(U)$.

Sufficiency:

Let $\{x_n\} \to x$. It suffices to show $\{f(x_n)\} \to f(x)$. For each open $U \ni f(x)$, by hypothesis, there exists $\delta > 0$ such that $B_\delta(x) \subseteq f^{-1}(U)$.

Since $\{x_n\} \to x$, there exists N such that

$$x_n \in B_\delta(x) \subseteq f^{-1}(U), \forall n \geq N \implies f(x_n) \in U, \forall n \geq N$$

Let $\varepsilon > 0$ be given, and then construct the $U = B_\varepsilon(f(x))$. The argument above shows that $f(x_n) \in B_\varepsilon(f(x))$ for $\forall n \geq N$, which implies $\rho(f(x_n), f(x)) < \varepsilon$, i.e., $\{f(x_n)\} \to f(x)$.

(b) For the forward direction, it suffices to show that each point x of $f^{-1}(U)$ is an interior point of $f^{-1}(U)$, which is shown by part (a); the converse follows trivially by applying (a).

■

R As illustracted above, convergence, continuity, (and compactness) can be defined by using open sets $\mathcal{T}$ only.

1.6.2. Topological Spaces

Definition 1.21 A **topological space** $(X, \mathcal{T})$ consists of a (non-empty) set X, and a family of subsets of X ("open sets" $\mathcal{T}$) such that

1. $\emptyset, X \in \mathcal{T}$
2. $U, V \in \mathcal{T}$ implies $U \bigcap V \in \mathcal{T}$
3. If $U_\alpha \in \mathcal{T}$ for all $\alpha \in \mathcal{A}$, then $\bigcup_{\alpha \in \mathcal{A}} U_\alpha \in \mathcal{T}$.

The elements in $\mathcal{T}$ are called **open subsets** of X. The $\mathcal{T}$ is called a **topology** on X. ■

■ **Example 1.20** 1. Let (X, d) be any metric space, and

$$\mathcal{T} = \{\text{all open subsets of } X\}$$

It's clear that $\mathcal{T}$ is a topology on X.

2. Define the discrete topology

$$\mathcal{T}_{\text{dis}} = \{\text{all subsets of } X\}$$

It's clear that $\mathcal{T}_{\text{dis}}$ is a topology on X, (which also comes from the discrete metric (X, d_{discrete})).

R We say $(X, \mathcal{T})$ is induced from a metric (X,d) (or it is **metrizable**) if $\mathcal{T}$ is the faimly of open subsets in (X,d).

3. Consider the indiscrete topology $(X, \mathcal{T}_{\text{indis}})$, where X contains more than one element:

$$\mathcal{T}_{\text{indis}} = \{\emptyset, X\}.$$

Question: is $(X, \mathcal{T}_{\text{indis}})$ metrizable? No. For any metric d defined on X, let x,y be distinct points in X, and then $\varepsilon := d(x,y) > 0$, hence $B_{\frac{1}{2}\varepsilon}(x)$ is a open set belonging to the corresponding induced topology. Since $x \in B_{\frac{1}{2}\varepsilon}(x)$ and $y \notin B_{\frac{1}{2}\varepsilon}(y)$, we conclude that $B_{\frac{1}{2}\varepsilon}(x)$ is neither $\emptyset$ nor X, i.e., the topology induced by any metric d is not the indiscrete topology.

4. Consider the cofinite topology $(X, \mathcal{T}_{\text{cofin}})$:

$$\mathcal{T}_{\text{cofin}} = \{U \mid X \setminus U \text{ is a finite set}\} \bigcup \{\emptyset\}$$

Question: is $(X, \mathcal{T}_{\text{cofin}})$ metrizable?

Definition 1.22 [Equivalence] Two metric spaces are **topologically equivalent** if they give rise to the same topology.

Example 1.21 Metrics d_1, d_2, d_∞ in $\mathbb{R}^n$ are topologically equivalent.

1.6.3. Closed Subsets

Definition 1.23 [Closed] Let $(X, \mathcal{T})$ be a topology space. Then $V \subseteq X$ is **closed** if $X \setminus V \in J$ ■

■ **Example 1.22** Under the topology space $(\mathbb{R}, \mathcal{T}_{\text{usual}})$, $(b, \infty) \bigcup (-\infty, a) \in \mathcal{T}$. Therefore,

$$[a,b] = \mathbb{R} \setminus \Big((b, \infty) \bigcup (-\infty, a) \Big)$$

is closed in $\mathbb{R}$ under usual topology. ■

R It is important to say that V is **closed in** X**.** You need to specify the underlying the space X.

Chapter 2

Week2

2.1. Monday for MAT3040

Reviewing.

1. Linear Combination and Span
2. Linear Independence
3. Basis: a set of vectors $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is called a **basis** for V if $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$ is linearly independent, and $V = \text{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$.

   Lemma: Given $V = \text{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_k\}$, we can find a basis for this set. Here V is said to be **finitely generated**.
4. Lemma: The vector $\boldsymbol{w} \in \text{span}\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\} \setminus \text{span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ implies that

$$\boldsymbol{v}_1 \in \text{span}\{\boldsymbol{w}, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\} \setminus \text{span}\{\boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$$

2.1.1. Basis and Dimension

Theorem 2.1 Let V be a finitely generated vector space. Suppose $\{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_m\}$ and $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ are two basis of V. Then $m = n$. (where m is called the **dimension**)

Proof. Suppose on the contrary that $m \neq n$. Without loss of generality (w.l.o.g.), assume that $m < n$. Let $\boldsymbol{v}_1 = \alpha_1 \boldsymbol{w}_1 + \cdots + \alpha_n \boldsymbol{w}_n$, with some $\alpha_i \neq 0$. w.l.o.g., assume $\alpha_1 \neq 0$. Therefore,

$$\boldsymbol{v}_1 \in \text{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\} \setminus \text{span}\{\boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\} \tag{2.1}$$

which implies that $\boldsymbol{w}_1 \in \text{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\} \setminus \text{span}\{\boldsymbol{w}_2, \ldots, \boldsymbol{w}_n\}$.

Then we claim that $\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$ is a basis of V:

1. Note that $\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$ is a spannning set:

$$\begin{aligned}\boldsymbol{w}_1 \in \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} &\implies \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} \subseteq \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} \\ &\implies \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} \subseteq \operatorname{span}\{\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}\} \subseteq \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}\end{aligned}$$

   Since $V = \operatorname{span}\{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$, we have $\operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} = V$.

2. Then we show the linear independence of $\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$. Consider the equation

$$\beta_1 \boldsymbol{v}_1 + \beta_2 \boldsymbol{v}_2 + \cdots + \beta_n \boldsymbol{w}_n = \mathbf{0}$$

   (a) When $\beta_1 \neq 0$, we imply

$$\boldsymbol{v}_1 = \left(-\frac{\beta_2}{\beta_1}\right)\boldsymbol{w}_2 + \cdots + \left(-\frac{\beta_n}{\beta_1}\right)\boldsymbol{w}_n \in \operatorname{span}\{\boldsymbol{w}_2, \dots, \boldsymbol{w}_n\},$$

   which contradicts (2.1).

   (b) When $\beta_1 = 0$, then $\beta_2 \boldsymbol{w}_2 + \cdots + \beta_n \boldsymbol{w}_n = \mathbf{0}$, which implies $\beta_2 = \cdots = \beta_n = 0$, due to the independence of $\{\boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$.

Therefore, $\boldsymbol{v}_2 \in \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$, i.e.,

$$\boldsymbol{v}_2 = \gamma_1 \boldsymbol{v}_1 + \cdots + \gamma_n \boldsymbol{v}_n,$$

where $\gamma_2, \dots, \gamma_n$ cannot be all zeros, since otherwise $\{\boldsymbol{v}_1, \boldsymbol{v}_2\}$ are linearly dependent, i.e., $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}$ cannot form a basis. w.l.o.g., assume $\gamma_2 \neq 0$, which implies

$$\boldsymbol{w}_2 \in \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{w}_3, \dots, \boldsymbol{w}_n\} \setminus \operatorname{span}\{\boldsymbol{v}_1, \boldsymbol{w}_3, \dots, \boldsymbol{w}_n\}.$$

Following the simlar argument above, $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \boldsymbol{w}_3, \dots, \boldsymbol{w}_n\}$ forms a basis of V.

Continuing the argument above, we imply $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m, \boldsymbol{w}_{m+1}, \dots, \boldsymbol{w}_n\}$ is a basis of V.

Since $\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_m\}$ is a basis as well, we imply

$$\boldsymbol{w}_{m+1} = \delta_1\boldsymbol{v}_1 + \cdots + \delta_m\boldsymbol{v}_m$$

for some $\delta_i \in \mathbb{F}$, i.e., $\{\boldsymbol{v}_1,\dots,\boldsymbol{v}_m,\boldsymbol{w}_{m+1}\}$ is linearly dependent, which is a contradction.

■

■ **Example 2.1** A vector space may have more than one basis.

Suppose $V = \mathbb{F}^n$, it is clear that $\dim(V) = n$, and

$$\{\boldsymbol{e}_1,\dots,\boldsymbol{e}_n\} \text{ is a basis of } V\text{, where } \boldsymbol{e}_i \text{ denotes a unit vector.}$$

There could be other basis of V, such as

$$\left\{\begin{pmatrix}1\\0\\\vdots\\0\end{pmatrix},\begin{pmatrix}1\\1\\\vdots\\0\end{pmatrix},\dots,\begin{pmatrix}1\\1\\\vdots\\1\end{pmatrix},\right\}$$

Actually, the columns of any invertible $n \times n$ matrix forms a basis of V. ■

■ **Example 2.2** Suppose $V = M_{m\times n}(\mathbb{R})$, we claim that $\dim(V) = mn$:

$$\left\{E_{ij}\;\middle|\;\begin{matrix}1\le i\le m\\1\le j\le n\end{matrix}\right\} \text{ is a basis of } V,$$

where E_{ij} is $m \times n$ matrix with 1 at (i,j)-th entry, and 0s at the remaining entries. ■

■ **Example 2.3** Suppose $V = \{\text{all polynomials of degree} \le \text{n}\}$, then $\dim(V) = n+1$. ■

■ **Example 2.4** Supppose $V = \{\boldsymbol{A} \in M_{n\times n}(\mathbb{R}) \mid \boldsymbol{A}^{\mathrm{T}} = \boldsymbol{A}\}$, then $\dim(V) = \frac{n(n+1)}{2}$. ■

■ **Example 2.5** Let $W = \{\boldsymbol{B} \in M_{n\times n}(\mathbb{R}) \mid \boldsymbol{B}^{\mathrm{T}} = -\boldsymbol{B}\}$, then $\dim(V) = \frac{n(n-1)}{2}$. ■

(R) Sometimes it should be classified the field $\mathbb{F}$ for the scalar multiplication to define a vector space. Conside the example below:

1. Let $V = \mathbb{C}$, then $\dim(\mathbb{C}) = 1$ for the scalar multiplication defined under the field $\mathbb{C}$.
2. Let $V = \operatorname{span}\{1, i\} = \mathbb{C}$, then $\dim(\mathbb{C}) = 2$ for the scalar multiplication defined under the field $\mathbb{R}$, since all $z \in V$ can be written as $z = a + bi$, $\forall a, b \in \mathbb{R}$.
3. Therefore, to aviod confusion, it is safe to write

$$\dim_{\mathbb{C}}(\mathbb{C}) = 1, \quad \dim_{\mathbb{R}}(\mathbb{C}) = 2.$$

2.1.2. Operations on a vector space

Note that the basis for a vector space is characterized as the **maximal linearly independent set**.

Theorem 2.2 — **Basis Extension.** Let V be a finite dimensional vector space, and $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ be a linearly independent set on V, Then we can extend it to the basis $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_n\}$ of V.

Proof.

- Suppose $\dim(V) = n > k$, and $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ is a basis of V. Consider the set $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\} \bigcup \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$, which is linearly dependent, i.e.,

$$\alpha_1 \boldsymbol{w}_1 + \cdots + \alpha_n \boldsymbol{w}_n + \beta_1 \boldsymbol{v}_1 + \cdots + \beta_k \boldsymbol{v}_k = \mathbf{0},$$

with some $\alpha_i \neq 0$, since otherwise this equation will only have trivial solution. w.l.o.g., assume $\alpha_1 \neq 0$.

- Therefore, consider the set $\{\boldsymbol{w}_2, \dots, \boldsymbol{w}_n\} \bigcup \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$. We keep removing elements from $\{\boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$ until we first get the set

$$S \bigcup \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\},$$

with $S \subseteq \{\boldsymbol{w}_1, \boldsymbol{w}_2, \dots, \boldsymbol{w}_n\}$ and $S \bigcup \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ is linearly independent, i.e., S is a maximal subset of $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ such that $S \bigcup \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ is linearly independent.

- Rewrite $S = \{\boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_m\}$ and therefore $S' = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_m\}$ are linearly independent. It suffices to show S' spans V.
  - Indeed, for all $\boldsymbol{w}_i \in \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$, $\boldsymbol{w}_i \in \operatorname{span}(S')$, since otherwise the equation

    $$\alpha \boldsymbol{w}_i + \beta_1 \boldsymbol{v}_1 + \cdots + \beta_m \boldsymbol{v}_m = \mathbf{0} \implies \alpha = 0,$$

    which implies that $\beta_1 \boldsymbol{v}_1 + \cdots + \beta_m \boldsymbol{v}_m = \mathbf{0}$ admits only trivial solution, i.e.,

    $$\{\boldsymbol{w}_i\} \bigcup S' = \{\boldsymbol{w}_i\} \bigcup S \bigcup \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\} \text{ is linearly independent,}$$

    which violetes the maximality of S.

  Therefore, all $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\} \subseteq \operatorname{span}(S')$, which implies $\operatorname{span}(S') = V$.

  Therefore, S' is a basis of V.

■

R Start with a spanning set, we keep removing something to form a basis; start with independent set, we keep adding something to form a basis.

In other words, the basis is both the minimal spanning set, and the maximal linearly independent set.

Definition 2.1 [Direct Sum] Let W_1, W_2 be two vector subspaces of V, then

1. $W_1 \bigcap W_2 := \{\boldsymbol{w} \in V \mid \boldsymbol{w} \in W_1, \text{ and } \boldsymbol{w} \in W_2\}$
2. $W_1 + W_2 := \{\boldsymbol{w}_1 + \boldsymbol{w}_2 \mid \boldsymbol{w}_i \in W_i\}$

3. If furthermore that $W_1 \bigcap W_2 = \{\mathbf{0}\}$, then $W_1 + W_2$ is denoted as $W_1 \oplus W_2$, which is called **direct sum**.

Proposition 2.1 $W_1 \bigcap W_2$ and $W_1 + W_2$ are vector subspaces of V.

2.2. Monday for MAT3006

Reviewing.

1. Equivalent Metric:
$$d_1(\boldsymbol{x},\boldsymbol{y}) \le Kd_2(\boldsymbol{x},\boldsymbol{y}) \le K'd_1(\boldsymbol{x},\boldsymbol{y})$$
In $\mathcal{C}[0,1]$, the metric d_1 and d_∞ are not equivalent:
For $f_n(x) = x^n n^2(1-x)$, $d_1(f_n,0) \to 1$ and $d_\infty(f_n,0) \to \infty$. Suppose on contrary that
$$d_1(f_n,0) \le Kd_\infty(\boldsymbol{x},\boldsymbol{y}) \le K'd_1(\boldsymbol{x},\boldsymbol{y}).$$
Taking limit both sides, we imply the immediate term goes to infinite, which is a contradiction.
2. Continuous functions: the function f is continuous is equivalent to say for $\forall x_n \to x$, we have $f(x_n) \to f(x)$.
3. Open sets: Let (X,d) be a metric space. A set $U \subseteq X$ is open if for each $x \in U$, there exists $\rho_x > 0$ such that $B_{\rho_x}(x) \subseteq U$.

R Unless stated otherwise, we assume that
$$\mathcal{C}[a,b] \longleftrightarrow (\mathcal{C}[a,b],d_\infty)$$
$$\mathbb{R}^n \longleftrightarrow (\mathbb{R}^n,d_2)$$

2.2.1. Remark on Open and Closed Set

■ **Example 2.6** Let $X = \mathcal{C}[a,b]$, show that the set
$$U := \{f \in X \mid f(x) > 0, \forall x \in [a,b]\} \quad \text{is open.}$$

Take a point $f \in U$, then

$$\inf_{[a,b]} f(x) = m > 0.$$

Consider the ball $B_{m/2}(f)$, and for $\forall g \in B_{m/2}(f)$,

$$\begin{aligned} |g(x)| &\geq |f(x)| - |f(x) - g(x)| \\ &\geq \inf_{[a,b]} |f(x)| - \sup_{[a,b]} |f(x) - g(x)| \\ &\geq m - \frac{m}{2} \\ &= \frac{m}{2} > 0, \ \forall x \in [a,b] \end{aligned}$$

Therefore, we imply $g \in U$, i.e., $B_{m/2}(f) \subseteq U$, i.e., U is open in X. ■

Proposition 2.2 Let (X,d) be a metric space. Then

1. $\emptyset, X$ are open in X
2. If $\{U_\alpha \mid \alpha \in \mathcal{A}\}$ are open in X, then $\bigcup_{\alpha \in \mathcal{A}}$ is also open in X
3. If $U_1, \dots, U_n$ are open in X, then $\bigcap_{i=1}^{n} U_i$ are open in X

R Note that $\bigcap_{i=1}^{\infty} U_i$ is not necessarily open if all U_i's are all open:

$$\bigcap_{i=1}^{\infty} \left(-\frac{1}{i}, 1 + \frac{1}{i}\right) = [0,1]$$

Definition 2.2 [Closed] The closed set in metric space (X,d) are the complement of open sets in X, i.e., any closed set in X is of the form $V = X \setminus U$, where U is oepn. ■

For example, in $\mathbb{R}$,

$$[a,b] = \mathbb{R} \setminus \{(-\infty, a) \bigcup (b, \infty)\}$$

Proposition 2.3

1. $\emptyset, X$ are closed in X
2. If $\{V_\alpha \mid \alpha \in \mathcal{A}\}$ are closed subsets in X, then $\bigcap_{\alpha \in \mathcal{A}} V_\alpha$ is also closed in X
3. If $V_1, \dots, V_n$ are closed in X, then $\bigcup_{i=1}^{n} V_i$ is also closed in X.

R Whenever you say U is open or V is closed, you need to specify the underlying space, e.g.,

$$\textbf{Wrong}: U \text{ is open}$$

$$\textbf{Right}: U \text{ is open in } X$$

Proposition 2.4 The following two statements are equivalent:

1. The set V is closed in metric space (X,d).
2. If the sequence $\{v_n\}$ in V converges to x, then $x \in V$

Proof. Necessity.

Suppose on the contrary that $\{v_n\} \to x \notin V$. Since $X \setminus V \ni x$ is open, there exists an open ball $B_\varepsilon(x) \subseteq X \setminus V$.

Due to the convergence of sequence, there exists N such that $d(v_n, x) < \varepsilon$ for $\forall n \geq N$, i.e., $v_n \in B_\varepsilon(x)$, i.e., $v_n \notin V$, which contradicts to $\{v_n\} \subseteq V$.

Sufficiency.

Suppose on the contrary that V is not closed in X, i.e., $X \setminus V$ is not open, i.e., there exists $x \notin V$ such that for all open $U \ni x$, $U \bigcap V \neq \emptyset$. In particular, take

$$U_n = B_{1/n}(x), \implies \exists v_n \in B_{1/n}(x) \bigcap V,$$

i.e., $\{v_n\} \to x$ but $x \notin V$, which is a contradiction. ∎

Proposition 2.5 Given two metric space (X,d) and (Y,ρ), the following statements are equivalent:

1. A function $f : (X,d) \to (Y,\rho)$ is continuous on X
2. For $\forall U \subseteq Y$ open in Y, $f^{-1}(U)$ is open in X.
3. For $\forall V \subseteq Y$ closed in Y, $f^{-1}(V)$ is closed in X.

■ **Example 2.7** The mapping $\Psi : \mathcal{C}[a,b] \to \mathbb{R}$ is defined as:

$$f \mapsto f(c)$$

where Ψ is called a **functional**.

Show that Ψ is continuous by using d_∞ metric on $\mathcal{C}[a,b]$:

1. Any open set in $\mathbb{R}$ can be written as countably union of open disjoint intervals, and therefore suffices to consider the pre-image $\Psi^{-1}(a,b) = \{f \mid f(c) \in (a,b)\}$. Following the similar idea in Example (2.6), it is clear that $\Psi^{-1}(a,b)$ is open in $(\mathcal{C}[a,b],d_\infty)$. Therefore, Ψ is continuous.
2. Another way is to apply definition.

We now study open sets in a subspace $(Y,d_Y) \subseteq (X,d_X)$, i.e.,

$$d_Y(y_1,y_2) := d_X(y_1,y_2).$$

Therefore, the open ball is defined as

$$\begin{aligned}
B_\varepsilon^Y(y) &= \{y' \in Y \mid d_Y(y,y') < \varepsilon\} \\
&= \{y' \in Y \mid d_X(y,y') < \varepsilon\} \\
&= \{y' \in X \mid d_X(y,y') < \varepsilon, y' \in Y\} \\
&= B_\varepsilon^X(y) \bigcap Y
\end{aligned}$$

Proposition 2.6 All open sets in the subspace $(Y,d_Y) \subseteq (X,d_X)$ are of the form $U \bigcap Y$, where U is open in X.

Corollary 2.1 For the subspace $(Y,d_Y) \subseteq (X,d_X)$, the mapping $i : (Y,d_Y) \to (X,d_X)$ with $i(y) = y, \forall y \in Y$ is continuous.

Proof. $i^{-1}(U) = U \bigcap Y$ for any subset $U \subseteq X$. The results follows from proposition (2.5). ∎

R It's important to specify the underlying space to describe an open set.

For example, the interval $[0,\frac{1}{2})$ is not open in $\mathbb{R}$, while $[0,\frac{1}{2})$ is open in $[0,1]$, since

$$[0, \frac{1}{2}) = (-\frac{1}{2}, \frac{1}{2}) \bigcap [0, 1].$$

2.2.2. Boundary, Closure, and Interior

Definition 2.3 Let (X, d) be a metric space, then

1. A point x is a **boundary point** of $S \subseteq X$ (denoted as $x \in \partial S$) if for any open $U \ni x$, then both $U \bigcap S, U \setminus S$ are non-empty.
   (one can replace U by $B_{1/n}(x)$, with $n = 1, 2, \dots$)
2. The **closure** of S is defined as $\overline{S} = S \bigcup \partial S$.
3. A point x is an **interior point** of S (denoted as $x \in S^\circ$) if there $\exists U \ni x$ open such that $U \subseteq S$. We use S° to denote the set of interior points.

Proposition 2.7

1. The closure of S can be equivalently defined as

$$\overline{S} = \bigcap \{C \in X \mid C \text{ is closed and } C \supseteq S\}$$

   Therefore, $\overline{S}$ is the smallest closed set containing S.
2. The interior set of S can be equivalently defined as

$$S^\circ = \bigcup \{U \subseteq X \mid U \text{ is open and } U \subseteq S\}$$

   Therefore, S° is the largest open set contained in S.

Example 2.8 For $S = [0, \frac{1}{2}] \subseteq X$, we have

1. $\partial S = \{0, \frac{1}{2}\}$
2. $\overline{S} = [0, \frac{1}{2}]$
3. $S^\circ = (0, \frac{1}{2})$

Proof. 1. (a) Firstly, we show that $\overline{S}$ is closed, i.e., $X \setminus \overline{S}$ is open.

- Take $x \notin \overline{S}$. Since $x \notin \partial S$, there $\exists B_r(x) \ni x$ such that

$$B_r(x) \bigcap S, \quad \text{or} \quad B_r(x) \setminus S \text{ is } \emptyset.$$

- Since $x \notin S$, the set $B_r(x) \setminus S$ is not empty. Therefore, $B_r(x) \bigcap S = \emptyset$.
- It's clear that $B_{r/2}(x) \bigcap S = \emptyset$. We claim that $B_{r/2}(x) \bigcap \overline{S}$ is empty. Suppose on the contrary that

$$y \in B_{r/2}(x) \bigcap \partial S,$$

which implies that $B_{r/2}(y) \bigcap S \neq \emptyset$. Therefore,

$$B_{r/2}(y) \subseteq B_r(x) \implies B_r(x) \bigcap S \supseteq B_{r/2}(y) \bigcap S \neq \emptyset,$$

which is a contradiction.

Therefore, $x \in X \setminus \overline{S}$ implies $B_{r/2}(x) \bigcap \overline{S} = \emptyset$, i.e., $X \setminus \overline{S}$ is open, i.e., $\overline{S}$ is closed.

(b) Secondly, we show that $\overline{S} \subseteq C$, for any closed $C \supseteq S$, i.e., suffices to show $\partial S \subseteq C$.

Take $x \in \partial S$, and construct a sequence

$$x_n \in B_{1/n}(x) \bigcap S.$$

Here $\{x_n\}$ is a sequence in $S \subseteq C$ converging to x, which implies $x \in C$, due to the closeness of C in X.

Combining (a) and (b), the result follows naturally. (Question: do we need to show the well-defineness?)

2. Exercise. Show that

$$S^\circ = S \setminus \partial S = X \setminus (\overline{X \setminus S}).$$

Then it's clear that S° is open, and contained in S.

$\blacksquare$

The next lecture we will talk about compactness and sequential compactness.

2.3. Monday for MAT4002

Reviewing.

1. Topological Space $(X, \mathcal{J})$: a special class of topological space is that induced from metric space (X, d):

$$(X, \mathcal{T}), \quad \text{with } \mathcal{T} = \{\text{all open sets in } (X, d)\}$$

2. Closed Sets $(X \setminus U)$ with U open.

Proposition 2.8 Let $(X, \mathcal{T})$ be a topological space,

1. $\emptyset, X$ are closed in X
2. V_1, V_2 closed in X implies that $V_1 \bigcup V_2$ closed in X
3. $\{V_\alpha \mid \alpha \in \mathcal{A}\}$ closed in X implies that $\bigcap_{\alpha \in \mathcal{A}} V_\alpha$ closed in X

Proof. Applying the De Morgan's Law

$$(X \setminus \bigcup_{i \in I} U_i) = \bigcap_{i \in I} (X \setminus U_i)$$

■

2.3.1. Convergence in topological space

Definition 2.4 [Convergence] A sequence $\{x_n\}$ of a topological space $(X, \mathcal{T})$ converges to $x \in X$ if $\forall U \ni x$ is open, there $\exists N$ such that $x_n \in U, \forall n \geq N$.

Example 2.9 1. The topology for the space $(X = \mathbb{R}^n, d_2) \to (X, \mathcal{T})$ (i.e., a topological space induced from meric space $(X = \mathbb{R}^n, d_2)$) is called a **usual topology** on $\mathbb{R}^n$.

When I say $\mathbb{R}^n$ (or subset of $\mathbb{R}^n$) is a topological space, it is equipeed with usual topology.

Convergence of sequence in $(\mathbb{R}^n, \mathcal{T})$ is the usual convergence in analysis.

For $\mathbb{R}^n$ or metric space, the limit of sequence (if exists) is unique.

2. Consider the topological space $(X, \mathcal{T}_{\text{indiscrete}})$. Take any sequence $\{x_n\}$ in X, it is convergent to any $x \in X$. Indeed, for $\forall U \ni x$ open, $U = X$. Therefore,

$$x_n \in U(= X), \forall n \geq 1.$$

3. Consider the topological space $(X, \mathcal{T}_{\text{cofinite}})$, where X is infinite. Consider $\{x_n\}$ is a sequence satisfying $m \neq n$ implies $x_m \neq x_n$. Then $\{x_n\}$ is convergent to any $x \in X$. (Question: how to define openness for $\mathcal{T}_{\text{cofinite}}$ and $\mathcal{T}_{\text{indiscrete}}$)?
4. Consider the topological space $(X, \mathcal{T}_{\text{discrete}})$, the sequence $\{x_n\} \to x$ is equivalent to say $x_n = x$ for all sufficiently large n.

■

R The limit of sequences may not be unique. The reason is that "$\mathcal{T}$ is not big enough". We will give a criterion to make sure the limit is unique in the future. (Hausdorff)

Proposition 2.9 If $F \subseteq (X, \mathcal{T})$ is closed, then for any convergent sequence $\{x_n\}$ in F, the limit(s) are also in F.

Proof. Let $\{x_n\}$ be a sequence in F with limit $x \in X$. Suppose on the contrary that $x \notin F$ (i.e., $x \in X \setminus F$ that is open). There exists N such that

$$x_n \in X \setminus F, \forall n \geq N,$$

i.e., $x_n \notin F$, which is a contradiction. ■

R The converse may not be true. If the $(X, \mathcal{T})$ is metrizable, the converse holds. Counter-example: Consider the co-countable topological space $(X, \mathcal{T}_{\text{co-co}})$, where

$$\mathcal{T}_{\text{co-co}} = \{U \mid X \setminus U \text{ is a countable set}\} \bigcup \{\varnothing\},$$

and X is uncontable. Let $F \subsetneq$ be an un-countable set such that is closed under limits, e.g., $[0,1]$. It's clear that $X \setminus F \notin \mathcal{T}_{\text{co-co}}$, i.e., F is not closed.

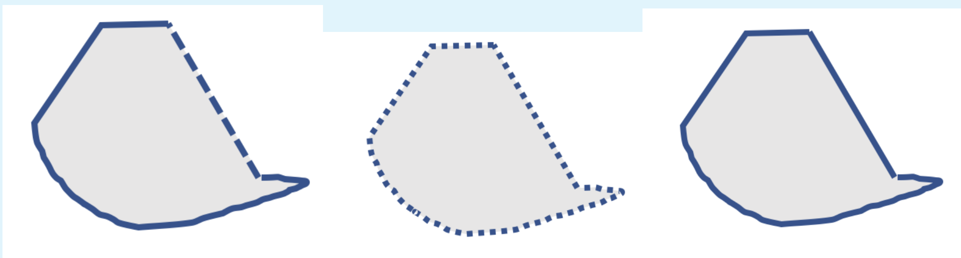
2.3.2. Interior, Closure, Boundary

Definition 2.5 Let $(X,\mathcal{T})$ be a topological space, and $A \subseteq X$ a subset.

1. The **interior** of A is
$$A^\circ = \bigcup_{U \subseteq A, U \text{ is open}} U$$
2. The **closure** of A is
$$\overline{A} = \bigcap_{A \subseteq V, V \text{ is closed}} V$$

If $\overline{A} = X$, we say that A is dense in X.

The graph illustration of the definition above is as follows:



(a) Illustration of A (b) Illustration of A° (c) Illustration of $\overline{A}$

Figure 2.1: Graph Illustrations

■ **Example 2.10**

1. For $[a,b) \subseteq \mathbb{R}$, we have:
$$[a,b)^\circ = (a,b), \quad \overline{[a,b)} = [a,b]$$
2. For $X = \mathbb{R}$, $\mathbb{Q}^\circ = \emptyset$ and $\overline{\mathbb{Q}} = \mathbb{R}$.
3. Consider the discrete topology $(X, \mathcal{T}_{\text{discrete}})$, we have
$$S^\circ = S, \quad \overline{S} = S$$

The insights behind the definition (2.5) is as follows

Proposition 2.10 1. A° is the largest open subset of X contained in A;

$\overline{A}$ is the smallest closed subset of X containing A.

2. If $A \subseteq B$, then $A^\circ \subseteq B$ and $\overline{A} \subseteq \overline{B}$
3. A is open in X is equivalent to say $A^\circ = A$; A is closed in X is equivalent to say $\overline{A} = A$.

■ **Example 2.11** Let (X,d) be a metric space. What's the closure of an open ball $B_r(x)$?

The direct intuition is to define the closed ball

$$\bar{B}_r(x) = \{y \in X \mid d(x,y) \le r\}.$$

Question: is $\bar{B}_r(x) = \overline{B_r(x)}$?

1. Since $\bar{B}_r(x)$ is a closed subset of X, and $B_r(x) \subseteq \bar{B}_r(x)$, we imply that

$$\overline{B_r(x)} \subseteq \bar{B}_r(x)$$

2. However, we may find an example such that $\overline{B_r(x)}$ is a proper subset of $\bar{B}_r(x)$: Consider the discrete metric space (X, d_{discrete}) and for $\forall x \in X$,

$$B_1(x) = \{x\} \implies \overline{B_1(x)} = \{x\}, \quad \bar{B}_1(x) = X$$

The equality $\bar{B}_r(x) = \overline{B_r(x)}$ holds when (X,d) is a normed space.

Here is another characterization of $\overline{A}$:

Proposition 2.11

$$\overline{A} = \{x \in X \mid \forall \text{open } U \ni x, U \bigcap A \ne \emptyset\}$$

Proof. Define

$$S = \{x \in X \mid \forall \text{open } U \ni x, U \bigcap A \ne \emptyset\}$$

It suffices to show that $\overline{A} = S$.

1. First show that S is closed:

$$X \setminus S = \{x \in X \mid \exists U_x \ni x \text{ open s.t. } U_x \bigcap A = \emptyset\}$$

Take $x \in X \setminus S$, we imply there exists open $U_x \ni x$ such that $U_x \bigcap A = \emptyset$. We claim $U_x \subseteq X \setminus S$:

- For $\forall y \in U_x$, note that $U_x \ni y$ that is open, such that $U_x \bigcap A = \emptyset$. Therefore, $y \in X \setminus S$.

Therefore, we have $x \in U_x \subseteq X \setminus S$ for any $\forall x \in X \setminus S$.

Note that

$$X \setminus S = \bigcup_{x \in X \setminus S} \{x\} \subseteq \bigcup_{x \in X \setminus S} U_x \subseteq X \setminus S,$$

which implies $X \setminus S = \bigcup_{x \in X \setminus S} U_x$ is open, i.e., S is closed in X.

2. By definition, it is clear that $A \subseteq S$:

$$\forall a \in A, \forall \text{open } U \ni a, U \bigcap A \supseteq \{a\} \neq \emptyset \implies a \in S.$$

Therefore, $\overline{A} \subseteq \overline{S} = S$.

3. Suppose on the contrary that there exists $y \in S \setminus \overline{A}$.

Since $y \notin \overline{A}$, by definition, there exists $F \supseteq A$ closed such that $y \notin F$.

Therefore, $y \in X \setminus F$ that is open, and

$$(X \setminus F) \bigcap A \subseteq (X \setminus A) \bigcap A = \emptyset \implies y \notin S,$$

which is a contradiction. Therefore, $S = \overline{A}$.

■

Definition 2.6 [accumulation point] Let $A \subseteq X$ be a subset in a topological space. We call $x \in X$ are an **accumulation point (limit point)** of A if

$$\forall U \subseteq X \text{ open s.t. } U \ni x, (U \setminus \{x\}) \bigcap A \neq \emptyset.$$

The set of accumulation points of A is denoted as A' ■

Proposition 2.12 $\overline{A} = A \bigcup A'$.

2.4. Wednesday for MAT3040

Reviewing.

- Basis, Dimension
- Basis Extension
- $W_1 \bigcap W_2 = \emptyset$ implies $W_1 \oplus W_2 = W_1 + W_2$ (Direct Sum).

2.4.1. Remark on Direct Sum

Proposition 2.13 The set $W_1 + W_2 = W_1 \oplus W_2$ iff any $\boldsymbol{w} \in W_1 + W_2$ can be uniquely expressed as

$$\boldsymbol{w} = \boldsymbol{w}_1 + \boldsymbol{w}_2,$$

where $\boldsymbol{w}_i \in W_i$ for $i = 1,2$.

(R) We can also define addiction among finite set of vector spaces $\{W_1, \dots, W_k\}$.

If $\boldsymbol{w}_1 + \cdots + \boldsymbol{w}_k = \mathbf{0}$ implies $\boldsymbol{w}_i = 0, \forall i$, then we can write $W_1 + \cdots + W_k$ as

$$W_1 \oplus \cdots \oplus W_k$$

Proposition 2.14 — **Complementation.** Let $W \le V$ be a vector subspace of a fintie dimension vector space V. Then there exists $W' \le V$ such that

$$W \oplus W' = V.$$

Proof. It's clear that $\dim(W) := k \le n := \dim(V)$. Suppose $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ is a basis of W.

By the basis extension proposition, we can extend it into $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k, \boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_n\}$, which is a basis of V.

Therefore, we take $W' = \operatorname{span}\{\boldsymbol{v}_{k+1}, \dots, \boldsymbol{v}_n\}$, which follows that

1. $W + W' = V$: $\forall \boldsymbol{v} \in V$ has the form

$$\boldsymbol{v} = (\alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_k \boldsymbol{v}_k) + (\alpha_{k+1} \boldsymbol{v}_{k+1} + \cdots + \alpha_n \boldsymbol{v}_n),$$

   where $\alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_k \boldsymbol{v}_k \in W$ and $\alpha_{k+1} \boldsymbol{v}_{k+1} + \cdots + \alpha_n \boldsymbol{v}_n \in W'$.

2. $W \bigcap W' = \{\mathbf{0}\}$: Suppose $\boldsymbol{v} \in W \bigcap W'$, i.e.,

$$\begin{aligned} \boldsymbol{v} &= (\beta_1 \boldsymbol{v}_1 + \cdots + \beta_k \boldsymbol{v}_k) + (0\boldsymbol{v}_{k+1} + \cdots + 0\boldsymbol{v}_n) \in W \\ &= (0\boldsymbol{v}_1 + \cdots + 0\boldsymbol{v}_k) + (\beta_{k+1} \boldsymbol{v}_{k+1} + \cdots + \beta_n \boldsymbol{v}_n) \in W'. \end{aligned}$$

   By the uniqueness of coordinates, we imply $\beta_1 = \cdots = \beta_n = 0$, i.e., $\boldsymbol{v} = \mathbf{0}$.

Therefore, we conclude that $W \oplus W' = V$. ■

2.4.2. Linear Transformation

Definition 2.7 [Linear Transformation] Let V, W be vector spaces. Then $T : V \to W$ is a **linear transformation** if

$$T(\alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2) = \alpha T(\boldsymbol{v}_1) + \beta T(\boldsymbol{v}_2),$$

for $\forall \alpha, \beta \in \mathbb{F}$ and $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$.

■ **Example 2.12**

1. The transformation $T : \mathbb{R}^n \to \mathbb{R}^m$ defined as $\boldsymbol{x} \mapsto \boldsymbol{A}\boldsymbol{x}$ (where $\boldsymbol{A} \in \mathbb{R}^{m \times n}$) is a linear transformation.
2. The transformation $T : \mathbb{R}[x] \to \mathbb{R}[x]$ defined as

$$p(x) \mapsto T(p(x)) = p'(x), \quad p(x) \mapsto T(p(x)) = \int_0^x p(t) \, \mathrm{d}t$$

   is a linear transformation

3. The transformation $T : M_{n\times n}(\mathbb{R}) \to \mathbb{R}$ defined as

$$\boldsymbol{A} \mapsto \text{trace}(\boldsymbol{A}) := \sum_{i=1}^{n} a_{ii}$$

is a linear transformation.

However, the transformation

$$\boldsymbol{A} \mapsto \det(\boldsymbol{A})$$

is not a linear transformation.

Definition 2.8 [Kernel/Image] Let $T : V \to W$ be a linear transfomation.

1. The **kernel** of T is

$$\ker(T) = T^{-1}(\mathbf{0}) = \{\boldsymbol{v} \in V \mid T(\boldsymbol{v}) = \mathbf{0}\}$$

2. The **image** (or range) of T is

$$\text{Im}(T) = T(\boldsymbol{v}) = \{T(\boldsymbol{v}) \in W \mid \boldsymbol{v} \in V\}$$

Example 2.13 1. Let $T : \mathbb{R}^n \to \mathbb{R}^n$ be a linear transformation with $T(\boldsymbol{x}) = \boldsymbol{A}\boldsymbol{x}$, then

$$\ker(T) = \{\boldsymbol{x} \in \mathbb{R}^n \mid \boldsymbol{A}\boldsymbol{x} = \mathbf{0}\} = \text{Null}(\boldsymbol{A}) \qquad \text{Null Space}$$

and

$$\text{Im}(T) = \{\boldsymbol{A}\boldsymbol{x} \mid \boldsymbol{x} \in \mathbb{R}^n\} = \text{Col}(\boldsymbol{A}) = \text{span}\{\text{columns of } \boldsymbol{A}\} \qquad \text{Column Space}$$

2. For $T(p(x)) = p'(x)$, $\ker(T) = \{\text{constant polynomials}\}$ and $\text{Im}(T) = \mathbb{R}[x]$.

Proposition 2.15 The kernel or image for a linear transformation $T : V \to W$ also forms a vector subspace:

$$\ker(T) \leq V, \quad \mathrm{Im}(T) \leq W$$

Proof. For $\boldsymbol{v}_1, \boldsymbol{v}_2 \in \ker(T)$, we imply

$$T(\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2) = \mathbf{0},$$

which implies $\alpha\boldsymbol{v}_1 + \beta\boldsymbol{v}_2 \in \ker(T)$.

The remaining proof follows similarly. ■

Definition 2.9 [Rank/Nullity] Let V, W be finite dimensional vector spaces and $T : V \to W$ a linear transformation. Then we define

$$\mathrm{rank}(T) = \dim(\mathrm{im}(T))$$

$$\mathrm{nullity}(T) = \dim(\ker(T))$$

R Let

$$\mathrm{Hom}_{\mathbb{F}}(V, W) = \{\text{all linear transformations } T : V \to W\},$$

and we can define the addiction and scalar multiplication to make it a vector space:

1. For $T, S \in \mathrm{Hom}_{\mathbb{F}}(V, W)$, define

$$(T + S)(\boldsymbol{v}) = T(\boldsymbol{v}) + S(\boldsymbol{v}),$$

which implies $T + S \in \mathrm{Hom}_{\mathbb{F}}(V, W)$.

2. Also, define

$$(\gamma T)(\boldsymbol{v}) = \gamma T(\boldsymbol{v}), \qquad \text{for } \forall \gamma \in \mathbb{F},$$

which implies $\gamma T \in \mathrm{Hom}_{\mathbb{F}}(V, W)$.

In particular, if $V = \mathbb{R}^n, W = \mathbb{R}^m$, then

$$\mathrm{Hom}_{\mathbb{F}}(V,W) = M_{m\times n}(\mathbb{R}).$$

Proposition 2.16 If $\dim(V) = n, \dim(W) = m$, then $\dim(\mathrm{Hom}_{\mathbb{F}}(V,W)) = mn$.

Proposition 2.17 There are anternative characterizations for the injectivity and surjectivity of lienar transformation T:

1. The linear transformation T is injective if and only if

$$\ker(T) = 0, \Longleftrightarrow \mathrm{nullity}(T) = 0.$$

2. The linear transformation T is surjective if and only if

$$\mathrm{im}(T) = W, \Longleftrightarrow \mathrm{rank}(T) = \dim(W).$$

3. If T is bijective, then T^{-1} is a linear transformation.

Proof. 1. (a) For the forward direction of (1),

$$\boldsymbol{x} \in \ker(T) \implies T(\boldsymbol{x}) = 0 = T(\boldsymbol{0}) \implies \boldsymbol{x} = \boldsymbol{0}$$

(b) For the reverse direction of (1),

$$T(\boldsymbol{x}) = T(\boldsymbol{y}) \implies T(\boldsymbol{x}-\boldsymbol{y}) = \boldsymbol{0} \implies \boldsymbol{x}-\boldsymbol{y} \in \ker(T) = \boldsymbol{0} \implies \boldsymbol{x} = \boldsymbol{y}$$

2. The proof follows similar idea in (1).
3. Let $T^{-1} : W \to V$. For all $\boldsymbol{w}_1, \boldsymbol{w}_2 \in W$, there exists $\boldsymbol{v}_1, \boldsymbol{v}_2 \in V$ such that $T(\boldsymbol{v}_i) = \boldsymbol{w}_i$, i.e., $T^{-1}(\boldsymbol{w}_i) = \boldsymbol{v}_i$ $i = 1,2$.

Consider the mapping

$$\begin{aligned}T(\alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2) &= \alpha T(\boldsymbol{v}_1) + \beta T(\boldsymbol{v}_2) \\ &= \alpha \boldsymbol{w}_1 + \beta \boldsymbol{w}_2,\end{aligned}$$

which implies $\alpha \boldsymbol{v}_1 + \beta \boldsymbol{v}_2 = T^{-1}(\alpha \boldsymbol{w}_1 + \beta \boldsymbol{w}_2)$, i.e.,

$$\alpha T^{-1}(\boldsymbol{w}_1) + \beta T^{-1}(\boldsymbol{w}_2) = T^{-1}(\alpha \boldsymbol{w}_1 + \beta \boldsymbol{w}_2).$$

■

Definition 2.10 [isomorphism]

We say the vector subspaces V and W are isomorphic if there exists a bijective linear transfomation $T : V \to W$. ($V \cong W$)

This mapping T is called an **isomorphism** from V to W.

R If $\dim(V) = \dim(W) = n < \infty$, then $V \cong W$:

Take $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}, \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ as basis of V and W, respectively. Then one can construct $T : V \to W$ satisfying $T(\boldsymbol{v}_i) = \boldsymbol{w}_i$ for $\forall i$ as follows:

$$T(\alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_n \boldsymbol{v}_n) = \alpha_n \boldsymbol{w}_1 + \cdots + \alpha_n \boldsymbol{w}_n \ \forall \alpha_i \in \mathbb{F}$$

It's clear that our constructed T is a linear transformation.

R $V \cong W$ doesn't imply any linear transformations $T : V \to W$ is an isomorphism. e.g., $T(\boldsymbol{v}) = \boldsymbol{0}$ is not an isomorphic if $W \neq \{\boldsymbol{0}\}$.

Theorem 2.3 — Rank-Nullity Theorem. Let $T : V \to W$ be a linear transformation with $\dim(V) < \infty$. Then

$$\operatorname{rank}(T) + \operatorname{nullity}(T) = \dim(V).$$

Proof. Since $\ker(T) \le V$, by proposition (2.14), there exists $V_1 \le V$ such that

$$V = \ker(T) \oplus V_1.$$

1. Consider the transformation $T|_{V_1}: V_1 \to T(V_1)$, which is an isomorphism, since:
   - Surjectivity is immediate
   - For $\boldsymbol{v} \in \ker(T|_{V_1})$,
     $$T(\boldsymbol{v}) = \mathbf{0} \implies \boldsymbol{v} \in \ker(T),$$
     which implies $\boldsymbol{v} = \mathbf{0}$ since $\boldsymbol{v} \in \ker(T) \cap V_1 = 0$, i.e., the injectivity follows.

   Therefore, $\dim(V_1) = \dim(T(V_1))$.
2. Secondly, given an isomorphism T from X to Y with $\dim(X) < \infty$, then $\dim(X) = \dim(T(X))$. The reason follows from assignment 1 questions (8-9):
   $$\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\} \text{ is a basis of } X \implies \{T(\boldsymbol{v}_1), \dots, T(\boldsymbol{v}_k)\} \text{ is a basis of } Y$$
3. Note that $T(V_1) = T(V) = \operatorname{im}(T)$, since:
   - for $\forall \boldsymbol{v} \in V$, $\boldsymbol{v} = \boldsymbol{v}_k + \boldsymbol{v}_1$, where $\boldsymbol{v}_k \in \ker(T), \boldsymbol{v}_1 \in V_1$, which implies
     $$T(\boldsymbol{v}) = T(\boldsymbol{v}_k) + T(\boldsymbol{v}_1) = \mathbf{0} + T(\boldsymbol{v}_1),$$
     i.e., $T(V) \subseteq T(V_1) \subseteq T(V)$, i.e., $T(V) = T(V_1)$.
4. By the proof of complementation,
   $$\begin{aligned}\dim(V) &= \dim(\ker(T)) + \dim(V_1) \\ &= \operatorname{nullity}(T) + \dim(T(V_1)) \\ &= \operatorname{nullity}(T) + \dim(T(V)) \\ &= \operatorname{nullity}(T) + \dim(\operatorname{im}(T)) \\ &= \operatorname{nullity}(T) + \operatorname{rank}(T).\end{aligned}$$

2.5. Wednesday for MAT3006

2.5.1. Compactness

This lecture will talk about the generalization of closeness and boundedness property in $\mathbb{R}^n$. First let's review some simple definitions:

Definition 2.11 [Compact] Let (X,d) be a metric space, and $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ a collection of open sets.

1. $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ is called an **open cover** of $E\subseteq X$ if $E\subseteq\cup_{\alpha\in\mathcal{A}}U_\alpha$
2. A **finite subcover** of $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ is a finite sub-collection $\{U_{\alpha_1},\dots,U_{\alpha_n}\}\subseteq\{U_\alpha\}$ covering E.
3. The set $E\subseteq X$ is **compact** if every open cover of E has a finite subcover.

A well-known result is talked in MAT2006:

Theorem 2.4 — Heine-Borel Theorem. The set $E\subseteq\mathbb{R}^n$ is **compact** if and only if E is closed and bounded.

However, there's a notion of sequentially compact, and we haven't identify its gap and relation with compactness.

Definition 2.12 [Sequentially Compact] Let (X,d) be a metric space. Then $E\subseteq X$ is **sequentially compact** if every sequence in E has a convergent subsequence with limit in E.

A well-known result is talked in MAT2006:

Theorem 2.5 — Bolzano-Weierstrass Theorem. The set $E\subseteq\mathbb{R}^n$ is closed and bounded if and only if E is sequentially compact.

Actually, the definitions of comapctness and the sequential compactness are equivalent under a metric space.

Theorem 2.6 Let (X,d) be a metric space, then $E \subseteq X$ is compact if and only if E is sequentially compact.

Proof. *Necessity*

Suppose $\{x_n\}$ is a sequence in E, it suffices to show it has a convergent subsequence. Consider the tail of $\{x_n\}$, say

$$F_n = \{x_k \mid k \geq n\} \implies F_1 \supseteq F_2 \supseteq \cdots.$$

- Note that $\cap_{i=1}^{\infty} F_i \neq \emptyset$. Assume not, then we imply $\cup_{i=1}^{\infty}(E \setminus F_i) = E$, i.e., $\{E \setminus F_i\}_{i=1}^{\infty}$ a open cover of E. By the compactness of E, we imply there exists a finite subcover of E:

$$E = \bigcup_{j=1}^{r}(E \setminus F_{i_j}) \implies \bigcap_{j=1}^{r} F_{i_j} = \emptyset \implies F_{i_j} = \emptyset, \forall j$$

  which is a contradiction, and there must exist an element $x \in \cap_{n=1}^{\infty} F_i$.

- For any $n \geq 1$, the open ball $B_{1/n}(x)$ must intersect with the n-th tail of the sequence $\{x_n\}$:

$$B_{1/n}(x) \cap \{x_k \mid k \geq n\} \neq \emptyset$$

  Pick the r-th intersection, say x_{n_r}, which implies that the subsequence $x_{n_r} \to x$ as $r \to \infty$. The proof for necessity is complete.

Sufficiency

Firstly, let's assume the claim below hold (which will be shown later):

Proposition 2.18 If $E \subseteq X$ is sequentially compact, then for any$\varepsilon > 0$, there exists finitely many open balls, say $\{B_\varepsilon(x_1), \ldots, B_\varepsilon(x_n)\}$, covering E.

Suppose on the contrary that there exists an open cover $\{U_\alpha\}_{\alpha \in \mathcal{A}}$ of E, that has no finite subcover.

- By proposition (2.18), for $n \geq 1$, there are finitely many balls of radius $1/n$ covering E. Due to our assumption, there exists a open ball $B_{1/n}(y_n)$ such that $B_{1/n}(y) \cap E$ cannot be covered by finitely many members in $\{U_\alpha\}_{\alpha \in \mathcal{A}}$.

- Pick $x_n \in B_{1/n}(y_n)$ to form a sequence. Due to the sequential compactness of E, there exists a subsequence $\{x_{n_j}\} \to x$ for some $x \in E$.

- Since $\{U_\alpha\}_{\alpha\in\mathcal{A}}$ covers E, there exists a U_β containing x. Since U_β is open and the radius of $B_{1/n_j}(y_{n_j})$ tends to 0, we imply that, for sufficiently large n_j, the set $B_{1/n_j}(y_{n_j}) \cap E$ is contained in U_β.

  In oteher words, U_β forms a **single** subcover of $B_{1/n}(y) \cap E$, which contradicts to our choice of $B_{1/n_j}(y_{n_j}) \cap E$. The proof for sufficiency is complete.

■

Proof for proposition (2.18). Pick $B_\varepsilon(x_1)$ for some $x_1 \in E$. Suppose $E \setminus B_\varepsilon(x_1) \neq \varnothing$. We can find $x_2 \notin B_\varepsilon(x_1)$ such that $d(x_2, x_1) \geq \varepsilon$.

Suppose $E \setminus (B_\varepsilon(x_1) \bigcup B_\varepsilon(x_2))$ is non-empty, then we can find $x_3 \notin B_\varepsilon(x_1) \bigcup B_\varepsilon(x_2)$ so that $d(x_j, x_3) \geq \varepsilon$, $j = 1,2$.

Keeping this procedure, we obtain a sequence $\{x_n\}$ in E such that

$$E \setminus \bigcup_{j=1}^{n} B_\varepsilon(x_j) \neq \varnothing, \qquad \text{and} \qquad d(x_j, x_n) \geq \varepsilon, j = 1,2,\dots,n-1.$$

By the sequential compactness of E, there exists $\{x_{n_j}\}$ and $x \in E$ so that $x_{n_j} \to x$ as $j \to \infty$. But then $d(x_{n_j}, x_{n_k}) < d(x_{n_j}, x) + d(x_{n_k}, x) \to 0$, which contradicts that $d(x_j, x_n) \geq \varepsilon$ for $\forall j < n$.

Therefore, one must have $E \setminus \bigcup_{j=1}^{N} B_\varepsilon(x_j) = \varnothing$ for some finite N.

The proof is complete. ■

R

1. Given the condition metric space,

$$\text{Sequential Compactness} \iff \text{Compactness}$$

2. Given the condition metric space, we will show that

$$\text{Compactness} \implies \text{Closed and Bounded}$$

However, the converse may not necessarily hold. Given the condition the metric space is $\mathbb{R}^n$, then

$$\text{Compactness} \iff \text{Closed and Bounded}$$

Proposition 2.19 Let (X,d) be a metric space. Then $E \subseteq X$ is compact implies that E is closed and bounded.

Proof.
1. Let $\{x_n\}$ be a convergent sequence in E. By sequential compactness, $\{x_{n_j}\} \to x$ for some $x \in E$. By the uniqueness of limits, under metric space, $\{x_n\} \to x$ for $x \in E$. The closeness is shown
2. Take $x \in E$ and consider the open cover $\bigcup_{n=1}^{\infty} B_n(x)$ of E. By compactness,

$$E \subseteq \bigcup_{i=1}^{k} B_{n_i}(x) = B_{n_k}(x),$$

which implies that for any $y,z \in E$,

$$d(y,z) \leq d(y,x) + d(x,z) \leq n_k + n_k = 2n_k.$$

The boundness is shown.

$\blacksquare$

Here we raise several examples to show that the coverse does not necessarily hold under a metric space.

Example 2.14 Given the metric space $\mathcal{C}[0,1]$ and a set $E = \{f \in \mathcal{C}[0,1] \mid 0 \leq f(x) \leq 1\}$. Notice that E is closed and bounded:

- $E = \cap_{x\in[0,1]} \Psi_x^{-1}([0,1])$, where $\Psi_x(f) = f(x)$, which implies that E is closed.
- Note that $E \subseteq B_2(\mathbf{0}) = \{f \mid |f| < 2\}$, i.e., E is bounded.

However, E may not be compact. Consider a sequence $\{f_n\}$ with

$$f_n(x) = \begin{cases} nx, & 0 \le x \le \frac{1}{n} \\ 1, & \frac{1}{n} \le x \le 1 \end{cases}$$

Suppose on the contrary that E is sequentially compact, therefore there exists a subsequence $\{f_{n_k}\} \to f$ under d_∞ metric, which implies, $\{f_{n_k}\}$ uniformly converges to f. By the definition of $f_n(x)$, we imply

$$f(x) = \begin{cases} 0, & x = 0 \\ 1, & x \in (0,1] \end{cases}$$

However, since d_∞ indicates uniform convergence, the limit for $\{f_{n_k}\}$, say f, must be continuous, which is a contradiction.

■

Theorem 2.7 Let the set E be compact in (X,d) and the function $f : (X,d) \to (Y,\rho)$ is continuous. Then $f(E)$ is compact in Y.

Note that the technique to show compactness by using the sequential compactness is very useful. However, this technique only applies to the metric space, but fail in general topological spaces.

Proof. Let $\{y_n\} = \{f(x_n)\}$ be any sequence in $f(E)$.

- By the compactness of X, $\{x_n\}$ has a convergent subsequence $\{x_{n_r}\} \to x$ as $r \to \infty$.
- Therefore, $\{y_{n_r}\} := \{f(x_{n_r})\} \to f(x)$ by the continuity of f.
- Therefore, $f(E)$ is sequentially compact, i.e., compact.

■

R The Theorem (2.7) is a generalization of the statement that *a continuous function on* $\mathbb{R}^n$ *admits its minimum and maximum*. Note that such an extreme value property no longer holds for arbitrary closed, bounded sets in a general metric space, but it continues to hold when the sets are strengthened to compact ones.

Another characterization of compactness in $\mathcal{C}[a,b]$ is shown in the Ascoli-Arzela Theorem (see Theorem (14.1) in MAT2006 Notebook).

2.5.2. Completeness

Definition 2.13 [Complete] Let (X,d) be metric space.

1. A sequence $\{x_n\}$ in (X,d) is a **Cauchy sequence** if for every $\varepsilon > 0$, there exists some N such that $d(x_n, x_m) < \varepsilon$ for all $n, m \geq N$.
2. A subset $E \subseteq X$ is said to be **complete** if every Cauchy sequence in E is convergent.

■ **Example 2.15** The set $X = \mathcal{C}[a,b]$ is complete:

- Suppose $\{f_n\}$ is Cauchy in $\mathcal{C}[a,b]$, i.e., $\{f_n(x)\}$ is Cauchy in $\mathbb{R}$ for $\forall x \in [a,b]$.
- By the compactness of $\mathbb{R}$, the sequence $f_n(x) \to f(x)$ for some $f(x) \in \mathbb{R}, \forall x \in [a,b]$. It suffices to show $f_n \to f$ uniformly:
  - For fixed $\varepsilon > 0$, there exists $N > 0$ such that

$$d_\infty(f_n, f_{n+k}) < \frac{\varepsilon}{2}, \qquad \forall n \geq N, k \in \mathbb{N}$$

which implies that for $\forall x \in [a,b]$, $\forall n \geq N, k \in \mathbb{N}$,

$$|f_n(x) - f_{n+k}(x)| < \frac{\varepsilon}{2} \implies \lim_{k \to \infty} |f_n(x) - f_{n+k}(x)| \leq \frac{\varepsilon}{2}$$

Therefore, we imply

$$|f_n(x) - f(x)| = \lim_{k\to\infty} |f_n(x) - f_{n+k}(x)| \le \frac{\varepsilon}{2} < \varepsilon, \qquad \forall n \ge N, x \in [a,b]$$

The proof is complete.

∎

2.6. Wednesday for MAT4002

Reviewing.

1. Interior, Closure:

$$\overline{A} = \{x \mid \forall U \ni x \text{ open}, U \bigcap A \neq \emptyset\}$$

2. Accumulation points

2.6.1. Remark on Closure

Definition 2.14 [Sequential Closure] Let A_S be the set of limits of any convergent sequence in A, then A_S is called the **sequential closure** of A.

Definition 2.15 [Accumulation/Cluster Points] The set of accumulation (limit) points is defined as

$$A' = \{x \mid \forall U \ni x \text{ open }, (U \setminus \{x\}) \bigcap A \neq \emptyset\}$$

R

1. (a) There exists some point in A but not in A':

$$A = \{1, 2, 3, \dots, n, \dots\}$$

Then any point in A is not in A'

(b) There also exists some point in A' but not in A:

$$A = \{\frac{1}{n} \mid n \geq 1\}$$

Then the point 0 is in A' but not in A.

2. The closure $\overline{A} = A \bigcup A'$.
3. The size of the sequencial closure A_S is between A and $\overline{A}$, i.e., $A \subseteq A_S \subseteq \overline{A}$:

It's clear that $A \subseteq A_S$, since the sequence $\{a_n := a\}$ is convergent to a for $\forall a \in A$.

For all $a \in A_S$, we have $\{a_n\} \to a$. Then for any open $U \ni a$, there exists N such that $\{a_N, a_{N+1}, \dots\} \subseteq U \bigcap A \neq \emptyset$. Therefore, $a \in \overline{A}$, i.e., $A_S \subseteq \overline{A}$.

Question: Is $A_S = \overline{A}$?

Proposition 2.20 Let (X, d) be a metric space, then $A_S = \bar{A}$.

Proof. Let $a \in \overline{A}$, then there exists $a_n \in B_{1/n}(a) \bigcap A$, which implies $\{a_n\} \to a$, i.e., $a \in A_S$. ■

R If $(X, \mathcal{T})$ is metrizable, then $A_S = \overline{A}$. The same goes for first countable topological spaces. However, A_S is a proper subset of A in general:

Let $A \subseteq X$ be the set of continuous functions, where $X = \mathbb{R}^{\mathbb{R}}$ denotes the set of all real-valued functions on $\mathbb{R}$, with the topology of pointwise convergence.

Then $A_S = B_1$, the set of all functions of first Baire-Category on $\mathbb{R}$; and $[A_S]_S = B_2$, the set of all functions of second Baire-Category on $\mathbb{R}$. Since $B_1 \neq B_2$, we have $[A_S]_S = A_S$. Note that $\overline{\overline{A}} = \overline{A}$. We conclude that A_S cannot equal to $\overline{A}$, since the sequential closure operator cannot be idemotenet.

Definition 2.16 [Boundary] The **boundary** of $\boldsymbol{A}$ is defined as

$$\partial \boldsymbol{A} = \overline{A} \setminus A^\circ$$

Proposition 2.21 Let $(X, \mathcal{T})$ be a topological space with $A, B \subseteq X$.

$$\overline{X \setminus A} = X \setminus A^\circ, \quad (X \setminus B)^\circ = X \setminus \overline{B} \quad \partial A = \overline{A} \bigcap (\overline{X \setminus A})$$

Proof.

$$X \setminus A^\circ = X \setminus \left(\bigcup_{U \text{ is open}, U \subseteq A} U \right) \tag{2.2a}$$

$$= \bigcap_{U \text{ is open}, U \subseteq A} (X \setminus U) \tag{2.2b}$$

$$= \bigcap_{V \text{ is closed}, F \supseteq X \setminus A} F \tag{2.2c}$$

$$= \overline{X \setminus A} \tag{2.2d}$$

Denoting $X \setminus A$ by B, we obtain:

$$(X \setminus B)^\circ = A^\circ \tag{2.3a}$$

$$= X \setminus (X \setminus A^\circ) \tag{2.3b}$$

$$= X \setminus \overline{X \setminus A} \tag{2.3c}$$

$$= X \setminus \overline{B} \tag{2.3d}$$

By definition of ∂A,

$$\partial A = \overline{A} \setminus A^\circ \tag{2.4a}$$

$$= \overline{A} \bigcap (X \setminus A^\circ) \tag{2.4b}$$

$$= \overline{A} \bigcap (\overline{X \setminus A}) \tag{2.4c}$$

■

2.6.2. Functions on Topological Space

Definition 2.17 [Continuous] Let $f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y)$ be a map. Then the function f is continuous, if

$$U \in \mathcal{T}_Y \implies f^{-1}(U) \in \mathcal{T}_X$$

■ **Example 2.16**

1. The identity map $\mathrm{id} : (X, \mathcal{T}) \to (X, \mathcal{T})$ defined as $x \mapsto x$ is continuous
2. The identity map $\mathrm{id} : (X, \mathcal{T}_{\text{discrete}}) \to (X, \mathcal{T}_{\text{indiscrete}})$ defined as $x \mapsto x$ is continuous. Since $\mathrm{id}^{-1}(\emptyset) = \emptyset$ and $\mathrm{id}^{-1}(X) = X$
3. The identity map $\mathrm{id} : (X, \mathcal{T}_{\text{indiscrete}}) \to (X, \mathcal{T}_{\text{discrete}})$ defined as $x \mapsto x$ is not continuous.

Proposition 2.22 If $f : X \to Y$, and $g : Y \to Z$ be continuous, then $g \circ f$ is continuous

Proof. For given $U \in \mathcal{T}_Z$, we imply

$$g^{-1}(U) \in \mathcal{T}_Y \implies f^{-1}(g^{-1}(U)) \in \mathcal{T}_X,$$

i.e., $(g \circ f)^{-1}(U) \in \mathcal{T}_X$ ■

Proposition 2.23 Suppose $f : X \to Y$ is continuous between two topological spaces. Then $\{x_n\} \to X$ implies $\{f(x_n)\} \to f(x)$.

Proof. Take open $U \ni f(x)$, which implies $f^{-1}(U) \ni x$. Since $f^{-1}(U)$ is open, we imply there exists N such that

$$\{x_n \mid n \geq N\} \subseteq f^{-1}(U),$$

i.e., $\{f(x_n) \mid n \geq N\} \subseteq U$ ■

We use the notion of Homeomorphism to describe the equivalence between two topological spaces.

Definition 2.18 [Homeomorphism] A **homeomorphism** between spaces topological spaces $(X, \mathcal{T}_X)$ and $(Y, \mathcal{T}_Y)$ is a bijection

$$f : (X, \mathcal{T}_X) \to (Y, \mathcal{T}_Y),$$

such that both f and f^{-1} are continuous.

2.6.3. Subspace Topology

Definition 2.19 Let $A \subseteq X$ be a non-empty set. The **subspace topology** of A is defined as:

1. $\mathcal{T}_A := \{U \bigcap A \mid U \in \mathcal{T}_A\}$
2. The **coarsest** topology on A such that the **inclusion map**

$$i : (A, \mathcal{T}_A) \to (X, \mathcal{T}_X), \quad i(x) = x$$

   is continuous.

   (We say the topology $\mathcal{T}_1$ is **coarser** than $\mathcal{T}_2$, or $\mathcal{T}_2$ is **finer** than $\mathcal{T}_1$, if $\mathcal{T}_1 \subseteq \mathcal{T}_2$ e.g., $\mathcal{T}_{\text{discrete}}$ is the finest topology, and $\mathcal{T}_{\text{indiscrete}}$ is coarsest topology.)

3. The (**unique**) topology such that for any $(Y, \mathcal{T}_Y)$,

$$f : (Y, \mathcal{T}_Y) \to (A, \mathcal{T}_A)$$

   is continuous iff $i \circ f : (Y, \mathcal{T}_Y) \to (X, \mathcal{T}_X)$ (where i is the inclusion map) is continuous.

Proposition 2.24 The definition (1) and (2) in (2.19) are equivalent.

Outline. The proof is by applying

$$i^{-1}(S) = S \bigcap A, \quad \forall S$$

∎

Example 2.17 Let all English and numerical letters be subset of $\mathbb{R}^2$:

P,6

The homeomorphism can be constrcuted between these two English letters.

Proposition 2.25 The definition (2) and (3) in (2.19) are equivalent.

Proof. Necessity.

- For $\forall U \in \mathcal{T}_X$, consider that

$$(i \circ f)^{-1}(U) = f^{-1}(i^{-1}(U)) = f^{-1}(U \bigcap A)$$

  since $U \bigcap A \in \mathcal{T}_A$ and f is continuous, we imply $(i \circ f)^{-1}(U) \in \mathcal{T}_Y$

- For $\forall U' \in \mathcal{T}_A$, we have $U' = U \bigcap A$ for some $U \in \mathcal{T}_X$. Therefore,

$$f^{-1}(U') = f^{-1}(U \bigcap A) = f^{-1}(i^{-1}(U)) = (i \circ f)^{-1}(U) \in \mathcal{T}_Y.$$

The sufficiency is left as exercise. ■

Proposition 2.26 1. The definition (1) in (2.19) does define a topology of A

2. Closed sets of A under subspace topology are of the form $V \bigcap A$, where V is closed in X

Proposition 2.27 Suppose $(A, \mathcal{T}_A) \subseteq (X, \mathcal{T}_X)$ is a subspace topology, and $B \subseteq A \subseteq X$. Then

1. $\bar{B}^A = \bar{B}^X \bigcap A$.
2. $B^{\circ A} \supseteq B^{\circ X}$

Proof. By proposition (2.26), $\bar{B}^X \bigcap A$ is closed in A, and $\bar{B}^X \bigcap A \supset B$, which implies

$$\bar{B}^A \subseteq \bar{B}^X \bigcap A$$

Note that $\bar{B}^A \supset B$ is closed in A, which implies $\bar{B}^A = V \bigcap A \subseteq V$, where V is closed in X. Therefore,

$$\bar{B}^X \subseteq V \implies \bar{B}^X \bigcap A \subseteq V \bigcap A = \bar{B}^A$$

Therefore, $\bar{B}^A = \bar{B}^X \subseteq V$

■

Can we have $B^{\circ X} = B^{\circ A}$?

2.6.4. Basis (Base) of a topology

Roughly speaking, a basis of a topology is a family of "generators" of the topology.

Definition 2.20 Let $(X, \mathcal{T})$ be a topological space. A family of subsets $\mathcal{B}$ in X is a **basis** for $\mathcal{T}$ if

1. $\mathcal{B} \subseteq \mathcal{T}$, i.e., everything in $\mathcal{B}$ is open
2. Every $U \in \mathcal{T}$ can be written as union of elements in $\mathcal{B}$.

■ **Example 2.18** 1. $\mathcal{B} = \mathcal{T}$ is a basis.

2. For $X = \mathbb{R}^n$,

$$\mathcal{B} = \{B_r(\boldsymbol{x}) \mid \boldsymbol{x} \in \mathbb{Q}^n, r \in \mathbb{Q} \bigcap (0, \infty)\}$$

Exercise: every $(a, b) = \bigcup_{i \in I} (p_i, q_i)$ for $p_i, q_i \in \mathbb{Q}$.

Therefore, $\mathcal{B}$ is countable.

Proposition 2.28 If $(X, \mathcal{T})$ has a countable basis, e.g., $\mathbb{R}^n$, then $(X, \mathcal{T})$ has a second-countable space.

Chapter 3

Week3

3.1. Monday for MAT3040

Reviewing.

1. Complementation. Suppose $\dim(V) = n < \infty$, then $W \leq V$ implies that there exists W' such that
$$W \oplus W' = V.$$
2. Given the linear transformation $T : V \to W$, define the set $\ker(T)$ and $\mathrm{Im}(T)$.
3. Isomorphism of vector spaces: $T : V \cong W$
4. Rank-Nullity Theorem

3.1.1. Remarks on Isomorphism

Proposition 3.1 If $T : V \to W$ is an isomorphism, then

1. the set $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_k\}$ is linearly independent in V if and only if $\{T\boldsymbol{v}_1, \dots, T\boldsymbol{v}_k\}$ is linearly independent.
2. The same goes if we replace the linearly independence by spans.
3. If $\dim(V) = n$, then $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ forms a basis of V if and only if $\{T\boldsymbol{v}_1, \dots, T\boldsymbol{v}_n\}$ forms a basis of W. In particular, $\dim(V) = \dim(W)$.
4. Two vector spaces with finite dimensions are isomorphic if and only if they have the same dimension:

Proof. It suffices to show the reverse direction. Let $\{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $\{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$ be two basis of V, W, respectively. Define the linear transformation $T : V \to W$ by

$$T(a_1\boldsymbol{v}_1 + \cdots + a_n\boldsymbol{v}_n) = a_1\boldsymbol{w}_1 + \cdots + a_n\boldsymbol{w}_n$$

Then T is surjective since $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ spans W; T is injective since $\{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ is linearly independent. ■

3.1.2. Change of Basis and Matrix Representation

Definition 3.1 [Coordinate Vector] Let V be a finite dimensional vector space and $B = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ an **ordered** basis of V. Any vector $\boldsymbol{v} \in V$ can be uniquely written as

$$\boldsymbol{v} = \alpha_1\boldsymbol{v}_1 + \cdots + \alpha_n\boldsymbol{v}_n,$$

Therefore we define the map $[\cdot]_{\mathcal{B}} : V \to \mathbb{F}^n$, which maps any vector in $\boldsymbol{v}$ into its **coordinate vector**:

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

R Note that $\{\boldsymbol{v}_1, \boldsymbol{v}_2, \ldots, \boldsymbol{v}_n\}$ and $\{\boldsymbol{v}_2, \boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ are distinct ordered basis.

■ **Example 3.1** Given $V = M_{2\times 2}(\mathbb{F})$ and the ordered basis

$$\mathcal{B} = \left\{ \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \right\}$$

Any matrix has the coordinate vector w.r.t. $\mathcal{B}$, i.e.,

$$\left[\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}\right]_{\mathcal{B}} = \begin{pmatrix} 1 \\ 4 \\ 2 \\ 3 \end{pmatrix}$$

However, if given another ordered basis

$$\mathcal{B}_1 = \left\{ \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 1 & 0 \end{pmatrix}, \begin{pmatrix} 0 & 0 \\ 0 & 1 \end{pmatrix}, \right\},$$

the matrix may have the different coordinate vector w.r.t. $\mathcal{B}_1$:

$$\left[\begin{pmatrix} 1 & 4 \\ 2 & 3 \end{pmatrix}\right]_{\mathcal{B}_1} = \begin{pmatrix} 4 \\ 1 \\ 2 \\ 3 \end{pmatrix}$$

Theorem 3.1 The mapping $[\cdot]_{\mathcal{B}} : V \to \mathbb{F}^n$ is an isomorphism.

Proof. 1. First show the operator $[\cdot]_{\mathcal{B}}$ is well-defined, i.e., the same input gives the same output. Suppose that

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} \quad [\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} \alpha_1' \\ \vdots \\ \alpha_n' \end{pmatrix},$$

then we imply

$$\begin{aligned} \boldsymbol{v} &= \alpha_1 \boldsymbol{v}_1 + \cdots + \alpha_n \boldsymbol{v}_n \\ &= \alpha_1' \boldsymbol{v}_1 + \cdots + \alpha_n' \boldsymbol{v}_n. \end{aligned}$$

By the uniqueness of coordinates, we imply $\alpha_i = \alpha_i'$ for $i = 1, \dots, n$.

2. It's clear that the operator $[\cdot]_{\mathcal{B}}$ is a linear transformation, i.e.,

$$[p\boldsymbol{v} + q\boldsymbol{w}]_{\mathcal{B}} = p[\boldsymbol{v}]_{\mathcal{B}} + q[\boldsymbol{w}]_{\mathcal{B}} \quad \forall p, q \in \mathbb{F}$$

3. The operator $[\cdot]_{\mathcal{B}}$ is surjective:

$$[\boldsymbol{v}]_{\mathcal{B}} = \begin{pmatrix} 0 \\ \vdots \\ 0 \end{pmatrix} \implies \boldsymbol{v} = 0\boldsymbol{v}_1 + \cdots + 0\boldsymbol{v}_n = \mathbf{0}.$$

4. The injective is clear, i.e., $[\boldsymbol{v}]_{\mathcal{B}} = [\boldsymbol{w}]_{\mathcal{B}}$ implies $\boldsymbol{v} = \boldsymbol{w}$.

Therefore, $[\cdot]_{\mathcal{B}}$ is an isomorphism. ■

We can use the Theorem (3.1) to simplify computations in vector spaces:

▪ **Example 3.2** Given a vector sapce $V = P_3[x]$ and its basis $B = \{1, x, x^2, x^3\}$.

To check if the set $\{1 + x^2, 3 - x^3, x - x^3\}$ is linearly independent, by part (1) in Proposition (3.1) and Theorem (3.1), it suffices to check whether the corresponding coordinate vectors

$$\left\{ \begin{pmatrix} 1 \\ 0 \\ 1 \\ 0 \end{pmatrix}, \begin{pmatrix} 3 \\ 0 \\ 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 0 \\ 1 \\ 0 \\ -1 \end{pmatrix} \right\}$$

is linearly independent, i.e., do Gaussian Elimination and check the number of pivots. ▪

Here gives rise to the question: if $\mathcal{B}_1, \mathcal{B}_2$ form two basis of V, then how are $[\boldsymbol{v}]_{\mathcal{B}_1}, [\boldsymbol{v}]_{\mathcal{B}_2}$ related to each other?

Here we consider an easy example first:

■ **Example 3.3** Consider $V = \mathbb{R}^n$ and its basis $\mathcal{B}_1 = \{\boldsymbol{e}_1, \ldots, \boldsymbol{e}_n\}$. For any $\boldsymbol{v} \in V$,

$$\boldsymbol{v} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = \alpha_n \boldsymbol{e}_1 + \cdots + \alpha_n \boldsymbol{e}_n \implies [\boldsymbol{v}]_{\mathcal{B}_1} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix}$$

Also, we can construct a different basis of V:

$$\mathcal{B}_2 = \left\{ \begin{pmatrix} 1 \\ 0 \\ \vdots \\ 0 \end{pmatrix}, \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 0 \end{pmatrix}, \ldots, \begin{pmatrix} 1 \\ 1 \\ \vdots \\ 1 \end{pmatrix} \right\},$$

which gives a different coordinate vector of $\boldsymbol{v}$:

$$[\boldsymbol{v}]_{\mathcal{B}_2} = \begin{pmatrix} \alpha_1 - \alpha_2 \\ \alpha_2 - \alpha_3 \\ \vdots \\ \alpha_{n-1} - \alpha_n \\ \alpha_n \end{pmatrix}$$

■

Proposition 3.2 — Change of Basis. Let $\mathcal{A} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ and $\mathcal{A}' = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$ be two ordered basis of a vector space V. Define the **change of basis** matrix from $\mathcal{A}$ to $\mathcal{A}'$, say $\mathcal{C}_{\mathcal{A}',\mathcal{A}} := [\alpha_{ij}]$, where

$$\boldsymbol{v}_j = \sum_{i=1}^{m} \alpha_{ij} \boldsymbol{w}_i$$

Then for any vector $\boldsymbol{v} \in V$, the *change of basis amounts to left-multiplying the change of basis matrix*:

$$\mathcal{C}_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}]_A = [\boldsymbol{v}]_{A'} \tag{3.1}$$

Define matrix $\mathcal{C}_{\mathcal{A},\mathcal{A}'} := [\beta_{ij}]$, where

$$\boldsymbol{w}_j = \sum_{i=1}^{n} \beta_{ij}\boldsymbol{v}_i$$

Then we imply that

$$(\mathcal{C}_{\mathcal{A},\mathcal{A}'})^{-1} = \mathcal{C}_{\mathcal{A}',\mathcal{A}}$$

Proof. 1. First show (3.1) holds for $\boldsymbol{v} = \boldsymbol{v}_j$, $j = 1, \dots, n$:

$$\text{LHS of (3.1)} = [\alpha_{ij}]\boldsymbol{e}_j = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

$$\text{RHS of (3.1)} = [\boldsymbol{v}_j]_{\mathcal{A}'} = \left[\sum_{i=1}^{n} \alpha_i \boldsymbol{w}_i \right]_{\mathcal{A}'} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

Therefore,

$$\mathcal{C}_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}_j]_{\mathcal{A}} = [\boldsymbol{v}_j]_{\mathcal{A}'}, \quad \forall j = 1, \dots, n. \tag{3.2}$$

2. Then for any $\boldsymbol{v} \in V$, we imply $\boldsymbol{v} = r_1\boldsymbol{v}_1 + \cdots + r_n\boldsymbol{v}_n$, which implies that

$$\begin{aligned}
\mathcal{C}_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}]_{\mathcal{A}} &= \mathcal{C}_{\mathcal{A}',\mathcal{A}}[r_1\boldsymbol{v}_1 + \cdots + r_n\boldsymbol{v}_n]_{\mathcal{A}} && (3.3a) \\
&= \mathcal{C}_{\mathcal{A}',\mathcal{A}}\left(r_1[\boldsymbol{v}_1]_{\mathcal{A}} + \cdots + r_n[\boldsymbol{v}_n]_{\mathcal{A}}\right) && (3.3b) \\
&= \sum_{j=1}^{n} r_j \mathcal{C}_{\mathcal{A}',\mathcal{A}}[\boldsymbol{v}_j]_{\mathcal{A}} && (3.3c) \\
&= \sum_{j=1}^{n} r_j [\boldsymbol{v}_j]_{\mathcal{A}'} && (3.3d) \\
&= \left[\sum_{j=1}^{n} r_j \boldsymbol{v}_j \right]_{\mathcal{A}'} && (3.3e) \\
&= [\boldsymbol{v}]_{\mathcal{A}'} && (3.3f)
\end{aligned}$$

where (3.3a) and (3.3e) is by applying the lineaity of $[\cdot]_{\mathcal{A}}$ and $[\cdot]_{\mathcal{A}'}$; (3.3d) is by applying the result (3.12). Therefore (3.1) is shown for $\forall \boldsymbol{v} \in V$.

3. Now we show that $(\mathcal{C}_{\mathcal{A}\mathcal{A}'}\mathcal{C}_{\mathcal{A}'\mathcal{A}}) = \boldsymbol{I}_n$. Note that

$$\begin{aligned}\boldsymbol{v}_j &= \sum_{i=1}^{n} \alpha_{ij}\boldsymbol{w}_i \\ &= \sum_{i=1}^{n} \alpha_{ij} \sum_{k=1}^{n} \beta_{ki}\boldsymbol{v}_k \\ &= \sum_{k=1}^{n} \left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) \boldsymbol{v}_i\end{aligned}$$

By the uniqueness of coordinates, we imply

$$\left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) = \delta_{jk} := \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

By the matrix multiplication, the (k, j)-th entry for $\mathcal{C}_{\mathcal{A}\mathcal{A}'}\mathcal{C}_{\mathcal{A}'\mathcal{A}}$ is

$$[\mathcal{C}_{\mathcal{A}\mathcal{A}'}\mathcal{C}_{\mathcal{A}'\mathcal{A}}]_{kj} = \left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) = \delta_{jk} \implies (\mathcal{C}_{\mathcal{A}\mathcal{A}'}\mathcal{C}_{\mathcal{A}'\mathcal{A}}) = \boldsymbol{I}_n$$

Noew, suppose

$$\begin{aligned}\boldsymbol{v}_j &= \sum_{i=1}^{n} \alpha_{ij}\boldsymbol{w}_i \\ &= \sum_{i=1}^{n} \alpha_{ij} \sum_{k=1}^{n} \beta_{ki}\boldsymbol{v}_k \\ &= \sum_{k=1}^{n} \left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) \boldsymbol{v}_i\end{aligned}$$

By the uniqueness of coordinates, we imply

$$\left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) = \begin{cases} 1, & j = k \\ 0, & j \neq k \end{cases}$$

where

$$\left(\sum_{i=1}^{n} \beta_{ki}\alpha_{ij} \right) = (\mathcal{C}_{AA'}\mathcal{C}_{A'A}).$$

Therefore, $(\mathcal{C}_{AA'}\mathcal{C}_{A'A}) = \boldsymbol{I}_n$. ■

■ **Example 3.4** Back to Example (3.3), write $\mathcal{B}_1, \mathcal{B}_2$ as

$$\mathcal{B}_1 = \{\boldsymbol{e}_1, \dots, \boldsymbol{e}_n\}, \quad \mathcal{B}_2 = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_n\}$$

and therefore $\boldsymbol{w}_i = \boldsymbol{e}_1 + \dots + \boldsymbol{e}_i$. The change of basis matrix is given by

$$\mathcal{C}_{\mathcal{B}_1,\mathcal{B}_2} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix}$$

which implies that for $\boldsymbol{v}$ in the example,

$$\mathcal{C}_{\mathcal{B}_1,\mathcal{B}_2}[\boldsymbol{v}]_{\mathcal{B}_2} = \begin{pmatrix} 1 & 1 & \cdots & 1 \\ 0 & 1 & \cdots & 1 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{pmatrix} \begin{pmatrix} \alpha_1 - \alpha_2 \\ \vdots \\ \alpha_{n-1} - \alpha_n \\ \alpha_n \end{pmatrix} = \begin{pmatrix} \alpha_1 \\ \vdots \\ \alpha_n \end{pmatrix} = [\boldsymbol{v}]_{\mathcal{B}_1}$$

■

Definition 3.2 Let $T: V \to W$ be a linear transformation, and

$$\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_m\}, \quad \mathcal{B} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$$

be basis of V and W, respectively. The **matrix representation** of T with respect to (w.r.t.) $\mathcal{A}$ and $\mathcal{B}$ is defined as $(T)_{\mathcal{B}\mathcal{A}} \in M_{m\times m}(\mathbb{F})$, where

$$T(\boldsymbol{v}_j) = \sum_{i=1}^{m} \alpha_{ij}\boldsymbol{w}_j$$

■

3.2. Monday for MAT3006

Reviewing.

1. Compactness/Sequential Compactness:
   - Equivalence for metric space
   - Stronger than closed and bounded
2. Completeness:
   - The metric space (E,d) is complete if every Cauchy sequence on E is convergent.
   - $\mathbb{P}[a,b] \subseteq \mathcal{C}[a,b]$ is not complete:

$$f_N(x) = \sum_{n=0}^{N} (-1)^n \frac{x^{2n}}{(2n)!} \to \cos x,$$

     while $\cos x \notin \mathcal{P}[a,b]$.

3.2.1. Remarks on Completeness

Proposition 3.3 Let (X,d) be a metric space.

1. If X is complete and $E \subseteq X$ is closed, then E is complete.
2. If $E \subseteq X$ is complete, then E is closed in X.
3. If $E \subseteq X$ is compact, then E is complete.

Proof.

1. Every Cauchy sequence $\{e_n\}$ in $E \subseteq X$ is also a Cauchy sequence in E. Therefore we imply $\{e_n\} \to x \in X$, due to the completeness of X. Due to the closedness of E, the limit $x \in E$, i.e., E is complete.
2. Consider any convergent sequence $\{e_n\}$ in E, with some limit $x \in X$. We imply $\{e_n\}$ is Cauchy and thus $\{e_n\} \to e \in E$, due to the completeness of E. By the uniqueness of limits, we must have $x = z \in E$, i.e., E is closed.
3. Consider a Cauchy sequence $\{e_n\}$ in E. There exists a subsequence $\{e_{n_j}\} \to e \in E$, due to the sequential compactness of E.

It follows that for large n and j,

$$d(e_n,e) \overset{\text{(a)}}{\leq} d(e_n,e_{n_j}) + d(e_{n_j},e) \overset{\text{(b)}}{<} \varepsilon$$

where (a) is due to triangle inequality and (b) is due to the Cauchy property of $\{e_n\}$ and the convergence of $\{e_{n_j}\}$.

Therefore, we imply $\{e_n\} \to e \in E$, i.e., E is complete.

■

R Given any metric space that may not be necessarily complete, we can make the union of it with another space to make it complete, e.g., just like the completion from $\mathbb{Q}$ to $\mathbb{R}$.

3.2.2. Contraction Mapping Theorem

The motivation of the contraction mapping theorem comes from solving an equation $f(x)$. More precisely, such a problem can be turned into a problem for fixed points, i.e., it suffices to find the fixed points for $g(x)$, with $g(x) = f(x) + x$.

Definition 3.3 Let (X,d) be a metric space. A map $T:(X,d) \to (X,d)$ is a **contraction** if there exists a constant $\tau \in (0,1)$ such that

$$d(T(x),T(y)) < \tau \cdot d(x,y), \quad \forall x,y \in X$$

A point x is called a fixed point of T if $T(x) = x$.

R All contractions are continuous: Given any convergence sequence $\{x_n\} \to x$, for $\varepsilon > 0$, take N such that $d(x_n,x) < \frac{\varepsilon}{\tau}$ for $n \geq N$. It suffices to show the convergence of $\{T(x_n)\}$:

$$d(T(x_n),T(x)) < \tau \cdot T(x_n,x) < \tau \cdot \frac{\varepsilon}{\tau} = \varepsilon.$$

Therefore, the contraction is Lipschitz continuous with Lipschitz constant τ.

Theorem 3.2 — **Contraction Mapping Theorem / Banach Fixed Point Theorem.** Every contraction T in a **complete** metric space X has a unique fixed point.

■ **Example 3.5**

1. The mapping $f(x) = x + 1$ is not a contraction in $X = \mathbb{R}$, and it has no fixed point.
2. Consider an in-complete space $X = (0, 1)$ and a contraction $f(x) = \frac{x+1}{2}$. It doesn't admit a fixed point on X as well.

■

Proof. Pick any $x_0 \in X$, and define a sequence recursively by setting $x_{n+1} = T(x_n)$ for $n \geq 0$.

1. Firstly show that the sequence $\{x_n\}$ is Cauchy.

   We can upper bound the term $d(T^n(x_0), T^{n-1}(x_0))$:

$$d(T^n(x_0), T^{n-1}(x_0)) \leq \tau d(T^{n-1}(x_0), T^{n-2}(x_0)) \leq \cdots \leq \tau^{n-1} d(T(x_0), x_0) \tag{3.4}$$

   Therefore for any $n \geq m$, where m is going to be specified later,

$$\begin{aligned}
d(x_n, x_m) &= d(T^n(x_0), T^m(x_0)) && \text{(3.5a)}\\
&\leq \tau d(T^{n-1}(x_0), T^{m-1}(x_0)) \leq \cdots \leq \tau^m d(T^{n-m}(x_0), x_0) && \text{(3.5b)}\\
&\leq \tau^m \sum_{j=1}^{n-m} \tau^{n-m-j} d(T(x_0), x_0) && \text{(3.5c)}\\
&< \frac{\tau^m}{1-\tau} d(T(x_0), x_0) && \text{(3.5d)}\\
&\leq \varepsilon && \text{(3.5e)}
\end{aligned}$$

   where (3.5b) is by repeatedly applying contraction property of d; (3.5c) is by applying the triangle inequality and (3.4); (3.5e) is by choosing sufficiently large m such that $\frac{\tau^m}{1-\tau} d(T(x_0), x_0) < \varepsilon$.

   Therefore, $\{x_n\}$ is Cauchy. By the completeness of X, we imply $\{x_n\} \to x \in X$.

2. Therefore, we imply

$$x = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} T(x_n) = T(\lim_{n\to\infty} x_n) = T(x),$$

i.e., x is a fixed point of T.

Now we show the uniqueness of the fixed point. Suppose that there is another fixed point $y \in X$, then

$$d(x,y) = d(T(x),T(y)) < \tau \cdot d(x,y) \implies d(x,y) < \tau d(x,y), \quad \tau \in (0,1),$$

and we conclude that $d(x,y) = 0$, i.e., $x = y$.

■

■ **Example 3.6** [Convergence of Newton's Method] The Newton's method aims to find the root of $f(x)$ by applying the iteration

$$x_{n+1} = x_n - \frac{f(x_n)}{f'(x)}$$

Suppose r is a root for f, the pre-assumption for the convergence of Newton's method is:

1. $f'(r) \neq 0$
2. $f \in \mathcal{C}^2$ on some neighborhood of r

Proof. 1. We first show that there exists $[r-\varepsilon, r+\varepsilon]$ such that the mapping

$$T : \mathcal{C}[r-\varepsilon, r+\varepsilon] \to \mathbb{R}, \quad f(x) \mapsto x - \frac{f(x)}{f'(x)}$$

satisfies $|T'(x)| < 1$ for $\forall x \in [r-\varepsilon, r+\varepsilon]$:

Note that $T'(x) = \frac{f(x)}{[f'(x)]^2} f''(x)$, and we define $h(x) = |T'(x)|$.

It's clear that $h(r) = 0$ and $h(x)$ is continuous, which implies

$$r \in h^{-1}((-1,1)) \implies B_\rho(r) \subseteq h^{-1}((-1,1)) \text{ for some } \rho > 0.$$

Or equivalently, $h((r-\rho,r+\rho)) \subseteq (-1,1)$. Take $\varepsilon = \frac{\rho}{2}$, and the result is obvious.

2. Therefore, any $x,y \in [r-\varepsilon, r+\varepsilon]$,

$$\begin{aligned} d(T(x),T(y)) &:= |T(x)-T(y)| && (3.6a)\\ &= |T'(\xi)||x-y| && (3.6b)\\ &\leq \max_{\xi\in[r-\varepsilon,r+\varepsilon]} |T'(\xi)||x-y| && (3.6c)\\ &< m\cdot|x-y| && (3.6d) \end{aligned}$$

where (3.6b) is by applying MVT, and ξ is some point in $[r-\varepsilon, r+\varepsilon]$; we assume that $\max_{\xi\in[r-\varepsilon,r+\varepsilon]} |T'(\xi)| < m$ for some $m<1$ in (3.6d).

Therefore, $T \in \mathcal{C}[r-\varepsilon, r+\varepsilon]$ is a contraction. By applying the contraction mapping theorem, there exists a unique fixed point near $[r-\varepsilon, r+\varepsilon]$:

$$x - \frac{f(x)}{f'(x)} = x \implies \frac{f(x)}{f'(x)} = 0 \implies f(x) = 0,$$

i.e., we obtain a root $x=r$. ■

Summary: if we use Newton's method on any point between $[r-\varepsilon, r+\varepsilon]$ where $f(r)=0$ and ε is sufficiently small, then we will eventually get close to r. ▪

3.2.3. Picard Lindelof Theorem

We will use Banach fixed point theorem to show the existence and uniqueness of the solution of ODE

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = f(x,y(x)) \\ y(x_0) = y_0 \end{cases} \qquad \textbf{Initial Value Problem, IVP} \tag{3.7}$$

■ **Example 3.7** Consider the IVP

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = x^2 y^{1/5} \\ y(x_0) = c > 0 \end{cases} \implies y = \left(\frac{4x^3}{15} + c^{4/5}\right)^{5/4}$$

which can be solved by the separation of variables:

$$c > 0 \implies y = \left(\frac{4x^3}{15} + c^{4/5}\right)^{5/4}.$$

However, when $c = 0$, the ODE does not have a unique solution. One can verify that y_1, y_2 given below are both solutions of this ODE:

$$y_1 = (\frac{4x^3}{15})^{5/4}, \qquad y_2 = 0$$

This example shows that even when f is very nice, the IVP may not have unique solution. The Picard-Lindelof theorem will give a clean condition on f ensuring the unique solvability of the IVP (3.7). ■

3.3. Monday for MAT4002

3.3.1. Remarks on Basis and Homeomorphism

Reviewing.

1. $A \subseteq A_S \subseteq \overline{A}$, where A_S is sequential closure and $\overline{A}$ denotes closure.
2. Subspace topology.
3. Homeomorphism. Consider the mapping $f : X \to Y$ with the topogical space X, Y shown below, with the standard topology, the question is whether f is continuous?

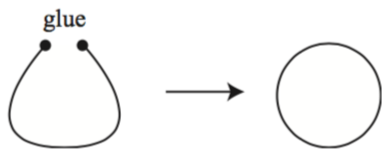


Figure 3.1: Diagram for mapping f

The answer is no, since the left in (3.1) can be isomorphically mapped into $(0,1)$; the right can be isomorphically mapped into $[0,1]$, and the mapping $(0,1) \to [0,1]$ cannot be isomorphism:

Proof. Assume otherwise the mapping $g : (0,1) \to [0,1]$ is isomorphism, and therefore $f^{-1}(U)$ is open for any open set U in the space $[0,1]$.
Construct $U = (1-\delta, 1]$ for $\delta \leq 1$, and therfore $f^{-1}((1-\delta,1])$ is open, and therfore for the point $x = f^{-1}(1)$, there exists $\varepsilon > 0$ such that

$$B_\varepsilon(x) \subseteq f^{-1}((1-\delta,1]) \implies [x-\varepsilon, x) \subseteq f^{-1}((1-\delta,1)), \text{ and } (x, x+\varepsilon] \subseteq f^{-1}((1-\delta,1)).$$

which implies that there exists a, b such that $[x-\varepsilon, x) = f^{-1}((a,1))$ and $(x, x+\varepsilon] = f^{-1}((b,1))$, i.e., $f^{-1}((a,b) \cap (b,1))$ admits into two values in $[x-\varepsilon, x)$ and $(x, x+\varepsilon]$, which is a contradiction. ■

4. Basis of a topology $\mathcal{B} \subseteq (X, \mathcal{T})$ is a collection of open sets in the space such that the whole space can be recovered, or equivalently

(a) $\mathcal{B} \subseteq \mathcal{T}$

(b) Every set in $\mathcal{T}$ can expressed as a union of sets in $\mathcal{B}$

Example: Let $\mathbb{R}^n$ be equipped with usual topology, then

$$B = \{B_q(x) \mid x \in \mathbb{Q}^n, q \in \mathbb{Q}^+\} \text{ is a basis of } \mathbb{R}^n.$$

It suffices to show $U \subseteq \mathbb{R}^n$ can be written as

$$U = U_{x \in \mathbb{Q}} B_{q_x}(x)$$

Proposition 3.4 Let X, Y be topological spaces, and $\mathcal{B}$ a basis for topology on Y. Then

$$f : X \to Y \text{ is continuous} \iff f^{-1}(B) \text{ is open in X, } \forall B \in \mathcal{B}$$

Therefore checking $f^{-1}(U)$ is open for all $U \in \mathcal{T}_Y$ suffices to checking $f^{-1}(N)$ is open for all $B \in \mathcal{B}$.

Proof. The forward direction follows from the fact $B \subseteq \mathcal{T}_Y$.

To show the reverse direction, let $U \in \mathcal{T}_Y$, then $U = \bigcup_{i \in I} B_i$, where $B_i \in \mathcal{B}$, which implies

$$f^{-1}(U) = f^{-1}\left(\bigcup_{i \in I} B_i\right) = \bigcup_{i \in I} f^{-1}(B_i)$$

which is open in X by our hypothesis. ■

Corollary 3.1 Let $f : X \to Y$ be a bijection. Suppose there is a basis $\mathcal{B}_X$ of $\mathcal{T}_X$ such that $\{f(B) \mid B \in \mathcal{B}_X\}$ forms a basis of $\mathcal{T}_Y$. Then $X \cong Y$.

Proof. Suppose $W \in \mathcal{T}_Y$, then by our hypothesis,

$$W = \bigcup_{i \in I} f(B_i),\ B_i \in \mathcal{B}_X \implies f^{-1}(W) = \bigcup_{i \in I} B_i \in \mathcal{T}_X,$$

which implies f is continuous.

Suppose $U \in \mathcal{T}_X$, then

$$U = \bigcup_{i \in I} B_i \implies f(U) = \bigcup_{i \in I} f(B_i) \in \mathcal{T}_Y \implies [f^{-1}]^{-1}(U) \in \mathcal{T}_Y,$$

i.e., f is continuous. ■

Question: *how to recognise whether a family of subsets is a basis for some given topology?*

Proposition 3.5 Let X be a set, $\mathcal{B}$ is a collection of subsets satisfying

1. X is a union of sets in $\mathcal{B}$, i.e., every $x \in X$ lies in some $B_x \in \mathcal{B}$
2. The intersection $B_1 \cap B_2$ for $\forall B_1, B_2 \in \mathcal{B}$ is a union of sets in $\mathcal{B}$, i.e., for each $B_1, B_2 \in \mathcal{B}$, and $x \in B_1 \cap B_2$, then there exists $B_3 \in \mathcal{B}$ such that $x \in B_3 \subseteq B_1 \cap B_2$.

Then the collection of subsets $\mathcal{T}_{\mathcal{B}}$, formed by taking any union of sets in $\mathcal{B}$, is a topology, and $\mathcal{B}$ is a basis for $\mathcal{T}_{\mathcal{B}}$.

Proof. 1. $\emptyset \in \mathcal{T}_{\mathcal{B}}$ (taking nothing from $\mathcal{B}$); for $x \in X, B_x \in \mathcal{B}$, by hypothesis (1),

$$X = \bigcup_{x \in X} B_x \in \mathcal{T}_{\mathcal{B}}$$

2. Suppose $T_1, T_2 \in \mathcal{T}_{\mathcal{B}}$. Let $x \in T_1 \cap T_2$, where T_i is a union of subsets in $\mathcal{B}$. Therefore,

$$\begin{cases} x \in B_1 \subseteq T_1, & B_1 \in \mathcal{B} \\ x \in B_2 \subseteq T_2, & B_2 \in \mathcal{B} \end{cases}$$

which implies $x \in B_1 \cap B_2$, i.e., $x \in B_x \subseteq B_1 \cap B_2$ for some $B_x \in \mathcal{B}$. Therefore,

$$\bigcup_{x \in B_1 \cap B_2} \{x\} \subseteq \bigcup_{x \in B_1 \cap B_2} B_x \subseteq B_1 \cap B_2,$$

i.e., $B_1 \cap B_2 = \bigcup_{x \in B_1 \cap B_2} B_x$, i.e., $B_1 \cap B_2 \in \mathcal{T}_{\mathcal{B}}$.

3. The property that $\mathcal{T}_{\mathcal{B}}$ is closed under union operations can be checked directly.

The proof is complete. ■

3.3.2. Product Space

Now we discuss how to construct new topological spaces out of given ones is by taking Cartesian products:

Definition 3.4 Let $(X, \mathcal{T}_X), (Y, \mathcal{T}_Y)$ be topological spaces. Consider the family of subsets in $X \times Y$:

$$\mathcal{B}_{X\times Y} = \{U \times V \mid U \in \mathcal{T}_X, V \in \mathcal{T}_y\}$$

This $\mathcal{B}_{X\times Y}$ forms a basis of a topology on $X \times Y$. The induced topology from $\mathcal{B}_{X\times Y}$ is called **product topology**.

For example, for $X = \mathbb{R}, Y = \mathbb{R}$, the elements in $\mathcal{B}_{X\times Y}$ are rectangles.

Proof for well-definedness in definition (3.4). We apply proposition (3.5) to check whether $B_{X\times Y}$ forms a basis:

1. For any $(x,y) \in X \times Y$, we imply $x \in X, y \in Y$. Note that $X \in \mathcal{T}_X, Y \in \mathcal{T}_Y$, we imply $(x,y) \in X \times Y \in \mathcal{B}_{X\times Y}$.
2. Suppose $U_1 \times V_1, U_2 \times V_2 \in \mathcal{B}_{X\times Y}$, then

$$(U_1 \times V_1) \cap (U_2 \times V_2) = (U_1 \cap U_2) \times (V_1 \cap V_2),$$

   where $U_1 \cap U_2 \in \mathcal{T}_X, V_1 \cap V_2 \in \mathcal{T}_Y$. Therefore, $(U_1 \times V_1) \cap (U_2 \times V_2) \in \mathcal{B}_{X\times Y}$.

∎

R However, the product topology may not necessarily become the largest topology in the space $X \times Y$. Consider $X = \mathbb{R}, Y = \mathbb{R}$, the open set in the space $X \times Y$ may not necessarily be rectangles. However, all elements in $\mathcal{B}_{X\times Y}$ are rectangles.

Example 3.8 The space $\mathbb{R} \times \mathbb{R}$ is isomorphic to $\mathbb{R}^2$, where the product topology is defined on $\mathbb{R} \times \mathbb{R}$ and the standard topology is defined on $\mathbb{R}^2$:

Construct the function $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2$ with $(a,b) \to (a,b)$.

Obviously, $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}^2$ is a bijection.

Take the basis of the topology on $\mathbb{R}$ as open intervals,

$$B_X = \{(a,b) \mid a < b \text{ in } \mathbb{R}\}$$

Therefore, one can verify that the set $\mathcal{B} := \{(a,b) \times (c,d) \mid a < b, c < d\}$ forms a basis for the product topology, and

$$\{f(B) \mid B \in \mathcal{B}\} = \{(a,b) \times (c,d) \mid a < b, c < d\}$$

forms a basis of the usual topology in $\mathbb{R}^2$.

By Corollary (3.1), we imply $\mathbb{R} \times \mathbb{R} \cong \mathbb{R}^2$. ■

We also raise an example on the homeomorphism related to product spaces:

■ **Example 3.9** Let $S^1 = \{(\cos x, \sin x \mid x \in [0, 2\pi])\}$ be a unit circle on $\mathbb{R}^2$.

Consider $f : S^1 \times (0, \infty) \to \mathbb{R}^2 \setminus \{\mathbf{0}\}$ defined as

$$f(\cos x, \sin x, r) \mapsto (r \cos x, r \sin x)$$

It's clear that f is a bijection, and f is continuous. Moreover, the inverse $g := f^{-1}$ is defined as

$$g(a,b) = (\frac{a}{\sqrt{a^2+b^2}}, \frac{b}{\sqrt{a^2+b^2}}, \sqrt{a^2+b^2})$$

which is continuous as well. Therefore, the $f : \mathcal{S}^1 \times (0, \infty) \to \mathbb{R}^2 \setminus \{\mathbf{0}\}$ is a homeomorphism.

■

3.4. Wednesday for MAT3040

3.4.1. Remarks for the Change of Basis

Reviewing.

- $[\cdot]_{\mathcal{A}} : V \to \mathbb{F}^n$ denotes coordinate vector mapping
- Change of Basis matrix: $\mathcal{C}_{\mathcal{A}',\mathcal{A}}$
- $T : V \to W$, $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$ and $\boldsymbol{B} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$.

  $\mathrm{Hom}_{\mathbb{F}}(V,W) \to M_{m\times n}(\mathbb{F})$

■ **Example 3.10** Let $V = \mathbb{P}_3[x]$ and $\mathcal{A} = \{1, x, x^2, x^3\}$.

Let $T : V \to V$ defined as $p(x) \mapsto p'(x)$:

$$\begin{cases} T(1) = 0\cdot 1 + 0\cdot x + 0\cdot x^2 + 0\cdot x^3 \\ T(x) = 1\cdot 1 + 0\cdot x + 0\cdot x^2 + 0\cdot x^3 \\ T(x^2) = 0\cdot 1 + 2\cdot x + 0\cdot x^2 + 0\cdot x^3 \\ T(x^3) = 0\cdot 1 + 0\cdot x + 3\cdot x^2 + 0\cdot x^3 \end{cases}$$

We can define the change of basis matrix for a linear transformation T as well, w.r.t. $\mathcal{A}$ and $\mathcal{A}$:

$$\mathcal{C}_{\mathcal{A},\mathcal{A}} = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix}$$

Also, we can define a different basis $\mathcal{A}' = \{x^3, x^2, x, 1\}$ for the output space for T, say $T : V_{\mathcal{A}} \to V_{\mathcal{A}'}$:

$$(T)_{\mathcal{A},\mathcal{A}'} = \begin{pmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 2 & 0 \\ 0 & 1 & 0 & 0 \end{pmatrix}$$

Our observation is that the corresponding coordinate vectors before and after linear transformation admits a matrix multiplication:

$$(2x^2 + 4x^3) \xrightarrow{T} ((4x + 12x^2))$$

$$(2x^2 + 4x^3)_{\mathcal{A}} = \begin{pmatrix} 0 \\ 0 \\ 2 \\ 4 \end{pmatrix} \qquad (4x + 12x^2)_{\mathcal{A}} = \begin{pmatrix} 0 \\ 4 \\ 12 \\ 0 \end{pmatrix}$$

$$\begin{pmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 2 & 0 \\ 0 & 0 & 0 & 3 \\ 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} 0 \\ 0 \\ 2 \\ 4 \end{pmatrix} = \begin{pmatrix} 0 \\ 4 \\ 12 \\ 0 \end{pmatrix}$$

$$\mathcal{C}_{\mathcal{A}\mathcal{A}} \cdot (2x^2 + 4x^3)_{\mathcal{A}} = (4x + 12x^2)_{\mathcal{A}}$$

■

Theorem 3.3 — Matrix Representation. Let $T : V \to W$ be a linear transformation of finite dimensional vector sapces. Let $\mathcal{A}, \mathcal{B}$ the ordered basis of V, W, respectively. Then the following diagram holds:

$$\begin{tikzcd} V \arrow[r, "T"] \arrow[d, "{[\cdot]_{\mathcal{A}}}"'] & W \arrow[d, "{[\cdot]_{\mathcal{B}}}"] \\ \mathbb{F}^n \arrow[r, "\text{Left-multiplying}", "T_{\mathcal{A}\mathcal{B}}"'] & \mathbb{F}^m \end{tikzcd}$$

Figure 3.2: Diagram for the matrix reprentation, where $n := \dim(V)$ and $m := \dim(W)$

namely, for any $\boldsymbol{v} \in V$,

$$(T)_{\mathcal{B},\mathcal{A}}(\boldsymbol{v})_{\mathcal{A}} = (T\boldsymbol{v})_{\mathcal{B}}$$

Therefore, we can compute $T\boldsymbol{v}$ by matrix multiplication.

R Linear transformation corresponds to coordinate matrix multiplication.

Proof. Suppose $\mathcal{A} = \{\boldsymbol{v}_1, \ldots, \boldsymbol{v}_n\}$ and $\mathcal{B} = \{\boldsymbol{w}_1, \ldots, \boldsymbol{w}_n\}$. The proof of this theorem follows the same procedure of that in Theorem (3.1)

1. We show this result for $\boldsymbol{v} = \boldsymbol{v}_j$ first:

$$\text{LHS} = [\alpha_{ij}]\boldsymbol{e}_j = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

$$\text{RHS} = (T\boldsymbol{v}_j)_{\mathcal{B}} = \left(\sum_{i=1}^{m} \alpha_{ij}\boldsymbol{w}_i\right)_{\mathcal{B}} = \begin{pmatrix} \alpha_{1j} \\ \vdots \\ \alpha_{nj} \end{pmatrix}$$

2. Then we show the theorem holds for any $\boldsymbol{v} := \sum_{j=1}^{n} r_j \boldsymbol{v}_j$ in V:

$$(T)_{\mathcal{B}\mathcal{A}}(\boldsymbol{v})_{\mathcal{A}} = (T)_{\mathcal{B}\mathcal{A}}\left(\sum_{j=1}^{n} r_j\boldsymbol{v}_j\right)_{\mathcal{A}} \tag{3.8a}$$

$$= (T)_{\mathcal{B}\mathcal{A}}\left(\sum_{j=1}^{n} r_j(\boldsymbol{v}_j)_{\mathcal{A}}\right) \tag{3.8b}$$

$$= \sum_{j=1}^{n} r_j (T)_{\mathcal{B}\mathcal{A}}(\boldsymbol{v}_j)_{\mathcal{A}} \tag{3.8c}$$

$$= \sum_{j=1}^{n} r_j (T\boldsymbol{v}_j)_{\mathcal{B}} \tag{3.8d}$$

$$= \left(\sum_{j=1}^{n} r_j (T\boldsymbol{v}_j)\right)_{\mathcal{B}} \tag{3.8e}$$

$$= \left[T(\sum_{j=1}^{n} r_j\boldsymbol{v}_j)\right]_{\mathcal{B}} \tag{3.8f}$$

$$= (T\boldsymbol{v})_{\mathcal{B}} \tag{3.8g}$$

The justification for (3.8a) is similar to that shown in Theorem (3.1). The proof is complete.

∎

R Consider a special case for Theorem (3.3), i.e., $T = \mathrm{id}$ and $\mathcal{A}, \mathcal{A}'$ are two ordered basis for the input and output space, respectively. Then the result in Theorem (3.3) implies

$$\mathcal{C}_{\mathcal{A}',\mathcal{A}}(\boldsymbol{v})_{\mathcal{A}} = (\boldsymbol{v})_{\mathcal{A}'}$$

i.e., the matrix representation theorem (3.3) is a general case for the change of basis theorem (3.1)

Proposition 3.6 — Functionality. Suppose V, W, U are finite dimensional vector spaces, and let $\mathcal{A}, \mathcal{B}, \mathcal{C}$ be the ordered basis for V, W, U, respectively. Suppose that

$$T : V \to W, \quad S : W \to U$$

are given two linear transformations, then

$$(S \circ T)_{\mathcal{C},\mathcal{A}} = (S)_{\mathcal{C},\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}$$

Composition of linear transformation corresponds to the multiplication of change of basis matrices.

Proof. Suppose the ordered basis $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}$, $\mathcal{B} = \{\boldsymbol{w}_1, \dots, \boldsymbol{w}_m\}$, $\mathcal{C} = \{\boldsymbol{u}_1, \dots, \boldsymbol{u}_p\}$. By defintion of change of basis matrices,

$$T(\boldsymbol{v}_j) = \sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \boldsymbol{w}_i$$

$$S(\boldsymbol{w}_i) = \sum_k (S_{\mathcal{C},\mathcal{B}})_{ki} \boldsymbol{u}_k$$

We start from the j-th column of $(S \circ T)_{\mathcal{C},\mathcal{A}}$ for $j = 1, \dots, n$, namely

$$\begin{aligned}
(S \circ T)_{\mathcal{C},\mathcal{A}}(\boldsymbol{v}_j)_{\mathcal{A}} &= (S \circ T(\boldsymbol{v}_j))_{\mathcal{C}} && \text{(3.9a)}\\
&= \left[S \circ \left(\sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \boldsymbol{w}_i \right) \right]_{\mathcal{C}} && \text{(3.9b)}\\
&= \sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \, (S(\boldsymbol{w}_i))_{\mathcal{C}} && \text{(3.9c)}\\
&= \sum_i (T_{\mathcal{B},\mathcal{A}})_{ij} \left(\sum_k (S_{\mathcal{C},\mathcal{B}})_{ki} \boldsymbol{u}_k \right)_{\mathcal{C}} && \text{(3.9d)}\\
&= \sum_k \sum_i (S_{\mathcal{C},\mathcal{B}})_{ki} (T_{\mathcal{B},\mathcal{A}})_{ij} (\boldsymbol{u}_k)_{\mathcal{C}} && \text{(3.9e)}\\
&= \sum_k (S_{\mathcal{C},\mathcal{B}} T_{\mathcal{B},\mathcal{A}})_{kj} (\boldsymbol{u}_k)_{\mathcal{C}} && \text{(3.9f)}\\
&= \sum_k (S_{\mathcal{C},\mathcal{B}} T_{\mathcal{B},\mathcal{A}})_{kj} \boldsymbol{e}_k && \text{(3.9g)}\\
&= j\text{-th column of } [S_{\mathcal{C}\mathcal{B}} T_{\mathcal{B},\mathcal{A}}] && \text{(3.9h)}
\end{aligned}$$

where (3.9a) is by the result in theorem (3.3); (3.9b) and (3.9d) follows from definitions of $T(\boldsymbol{v}_j)$ and $S(\boldsymbol{w}_i)$; (3.9c) and (3.9e) follows from the linearity of $\mathcal{C}$; (3.9f) follows from the matrix multiplication definition; (3.9g) is because $(\boldsymbol{u}_k)_{\mathcal{C}} = \boldsymbol{e}_k$.

Therefore, $(S \circ T)_{\mathcal{C}\mathcal{A}}$ and $(S_{\mathcal{C},\mathcal{B}})(T_{\mathcal{B},\mathcal{A}})$ share the same j-th column, and thus equal to each other. ■

Corollary 3.2 Suppose that S and T are two identity mappings $V \to V$, and consider $(S)_{\mathcal{A}'\mathcal{A}}$ and $(T)_{\mathcal{A},\mathcal{A}'}$ in proposition (3.6), then

$$(S \circ T)_{\mathcal{A}',\mathcal{A}'} = (S)_{\mathcal{A}'\mathcal{A}} (T)_{\mathcal{A},\mathcal{A}'}$$

Therefore,

$$\text{Identity matrix} = \mathcal{C}_{\mathcal{A}',\mathcal{A}} \mathcal{C}_{\mathcal{A},\mathcal{A}'}$$

Proposition 3.7 Let $T : V \to W$ with $\dim(V) = n, \dim(W) = m$, and let

- $\mathcal{A}, \mathcal{A}'$ be ordered basis of V
- $\mathcal{B}, \mathcal{B}'$ be ordered basis of W

then the change of basis matrices admit the relation

$$(T)_{\mathcal{B}',\mathcal{A}'} = \mathcal{C}_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}\mathcal{C}_{\mathcal{A}\mathcal{A}'} \tag{3.10}$$

Here note that $(T)_{\mathcal{B}',\mathcal{A}'}, (T)_{\mathcal{B},\mathcal{A}} \in \mathbb{F}^{m\times n}$; $\mathcal{C}_{\mathcal{B}',\mathcal{B}} \in \mathbb{F}^{m\times m}$; and $\mathcal{C}_{\mathcal{A}\mathcal{A}'} \in \mathbb{F}^{n\times n}$.

Proof. Let $\mathcal{A} = \{\boldsymbol{v}_1, \dots, \boldsymbol{v}_n\}, \mathcal{A}' = \{\boldsymbol{v}_1', \dots, \boldsymbol{v}_n'\}$. Consider simplifying the j-th column for the LHS and RHS of (3.10) and showing they are equal:

$$\begin{aligned}
\text{LHS} &= (T)_{\mathcal{B}',\mathcal{A}'}\boldsymbol{e}_j \\
&= (T)_{\mathcal{B}',\mathcal{A}'}(\boldsymbol{v}_j')_{\mathcal{A}'} \\
&= (T\boldsymbol{v}_j')_{\mathcal{B}'}
\end{aligned}$$

$$\begin{aligned}
\text{RHS} &= \mathcal{C}_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}\mathcal{C}_{\mathcal{A}\mathcal{A}'}\boldsymbol{e}_j \\
&= \mathcal{C}_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}\mathcal{C}_{\mathcal{A}\mathcal{A}'}(\boldsymbol{v}_j')_{\mathcal{A}'} \\
&= \mathcal{C}_{\mathcal{B}',\mathcal{B}}(T)_{\mathcal{B},\mathcal{A}}(\boldsymbol{v}_j')_{\mathcal{A}} \\
&= \mathcal{C}_{\mathcal{B}',\mathcal{B}}(T\boldsymbol{v}_j')_{\mathcal{B}} \\
&= (T\boldsymbol{v}_j')_{\mathcal{B}'}
\end{aligned}$$

■

R Let $T : V \to V$ be a linear operator with $\mathcal{A}, \mathcal{A}'$ being two ordered basisof V, then

$$(T)_{\mathcal{A}'\mathcal{A}'} = \mathcal{C}_{\mathcal{A}',\mathcal{A}}(T)_{\mathcal{A}\mathcal{A}}\mathcal{C}_{\mathcal{A},\mathcal{A}'} = (\mathcal{C}_{\mathcal{A},\mathcal{A}'})^{-1}(T)_{\mathcal{A}\mathcal{A}}\mathcal{C}_{\mathcal{A},\mathcal{A}'}$$

Therefore, the change of basis matrices $(T)_{\mathcal{A}'\mathcal{A}'}$ and $(T)_{\mathcal{A}\mathcal{A}}$ are similar to each other, which means they share the same eigenvalues, determinant, trace.

Therefore, two similar matrices cooresponds to same linear transformation using different basis.

3.5. Wednesday for MAT3006

3.5.1. Remarks on Contraction

Reviewing.

- Suppose $E \subseteq X$ with X being complete, then E is closed in X iff E is complete
- Suppose $E \subseteq X$, then E is closed in X if E is complete.
- Contraction Mapping Theorem
- Classification for the Convergence of Newton's method: the Newton's method aims to find the fixed point of T.

  $$T : \mathbb{R} \to \mathbb{R}, \quad T(x) = x - \tfrac{f(x)}{f''(x)}$$

  In the last lecture we claim that there exists $[r-\varepsilon, r+\varepsilon]$ such that $\sup_{[r-\varepsilon,r+\varepsilon]} |T'(x)| < 1$.

  Note that we doesn't make our statement rigorous enough. we need to furthermore show that $T(X) \subseteq X$:

  - $T : [r-\varepsilon, r+\varepsilon] \to [r-\varepsilon, r+\varepsilon]$, since

    $$|T(x) - r| = |T(x) - T(r)| = |T'(s)||x-r| \le \sup_{[r-\varepsilon,r+\varepsilon]} |T'(s)||x-r| < |x-r|$$

    Therefore, if $x \in [r-\varepsilon, r+\varepsilon]$, then $T(x) \in [r-\varepsilon, r+\varepsilon]$.
  - T is a contraction:

    $$|T(x) - T(y)| < \tau \cdot |x-y|$$

  Therefore, applying contraction mapping theorem gives the desired result.

3.5.2. Picard-Lindelof Theorem

Consider solving the the initival value problem given below

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = f(x,y) \\ y(\alpha) = \beta \end{cases} \implies y(x) = \beta + \int_\alpha^x f(t, y(t))\,\mathrm{d}t \tag{3.11}$$

Definition 3.5 Let $R = [\alpha - a, \alpha + a] \times [\beta - b, \beta + b]$. Then the function $f(x,y)$ satisfies the **Lipschitz condition** on R if there exists $L > 0$ such that

$$|f(x,y_1) - f(x,y_2)| < L \cdot |y_1 - y_2|, \quad \forall (x, y_i) \in R \tag{3.12}$$

The smallest number $L^* = \inf\{L \mid \text{The relation (3.12) holds for } L\}$ is called the **Lipschitz constant** for f.

Example 3.11 A $\mathcal{C}^1$-function f(x,y) in a rectangle automatically satisfies the Lipschitz condition:

$$f(x,y_1) - f(x,y_2) \overset{\text{Appling MVT}}{=} \frac{\partial f}{\partial y}(x, \tilde{y})(y_1 - y_2)$$

Since $\frac{\partial f}{\partial y}$ is continuous on R and thus bounded, we imply

$$|f(x,y_1) - f(x,y_2)| < L \cdot |y_1 - y_2|, \quad \forall (x, y_i) \in R$$

where

$$L = \max\left\{ \text{abs}\left(\frac{\partial f}{\partial y}\right) \middle| (x,y) \in R \right\}$$

Theorem 3.4 — Picard-Lindelof Theorem (existence part). Suppose $f \in \mathcal{C}(R)$ be such that f satisfies the Lipschitz condition, then there exists $a'' \in (0, a]$ such that (**??**) is solvable with $y(x) \in \mathcal{C}([\alpha - a'', \alpha + a''])$.

Proof. Consider the complete metric space

$$X = \{y(x) \in \mathcal{C}([\alpha - a, \alpha + a]) \mid \beta - b \leq y(x) \leq \beta + b\},$$

with the mapping $T : X \to X$ defined as

$$(Ty)(x) = \beta + \int_{\alpha}^{x} f(t, y(t))\,\mathrm{d}t$$

It suffices to show that T is a contraction, but here we need to estrict a a smaller number as follows:

1. Well-definedness of T: Take $M := \sup\{f(x,y) \mid (x,y) \in R\}$ and construct $a' = \min\{b/M, a\}$. Consider the complete matric space

   $$X = \{y(x) \in \mathcal{C}([\alpha - a', \alpha + a']) \mid \beta - b \leq y(x) \leq \beta + b\}$$

   which implies that

   $$|(Ty)(x) - \beta| \leq \left| \int_{\alpha}^{x} f(t, y(t))\,\mathrm{d}t \right| \leq M|x - \alpha| \leq Ma' \leq b,$$

   i.e., $T(X) \subseteq X$, and therefore $T : X \to X$ is well-defined.
2. Contraction for T: Construct $a'' \in \min\{a', \frac{1}{2L^*}\}$, where L^* is the Lipschitz constant for f. and consider the complete metric space

   $$X = \{y(x) \in \mathcal{C}([\alpha - a'', \alpha + a'']) \mid \beta - b \leq y(x) \leq \beta + b\}$$

   Therefore for $\forall x \in [\alpha - a'', \alpha + a'']$ and the mapping $T : X \to X$,

   $$\begin{aligned} |[T(y_1) - T(y_2)](x)| &\leq \left| \int_{\alpha}^{x} [f(t, y_2(t)) - f(t, y_1(t))]\,\mathrm{d}t \right| \\ &\leq \int_{\alpha}^{x} |f(t, y_2) - f(t, y_1)|\,\mathrm{d}t \leq \int_{\alpha}^{x} L^*|y_2(t) - y_1(t)|\,\mathrm{d}t \\ &\leq L^*|x - \alpha| \sup |y_2(t) - y_1(t)| \leq L^* a'' d_\infty(y_2, y_1) \leq \frac{1}{2} d_\infty(y_2, y_1) \end{aligned}$$

Therefore, we imply $d_\infty(Ty_2,Ty_1) \leq \frac{1}{2}d_\infty(y_2,y_1)$, i.e., T is a contraction.

Applying contraction mapping theorem, there exists $y(x) \in X$ such that $Ty = y$, i.e.,

$$y = \beta + \int_\alpha^x f(t,y(t))\,\mathrm{d}t$$

Thus y is a solution for the IVP (3.11). ■

R On $[\alpha - a'', \alpha + a'']$, we can solve the IVP (3.11) by recursively applying T:

$$\begin{aligned} y_0(x) &= \beta, \qquad \forall x \in [\alpha - a'', \alpha + a''] \\ y_1 &= T(y_0) = \beta + \int_\alpha^x f(t,\beta)\,\mathrm{d}t \\ y_2 &= T(y_1) \\ &\cdots\cdots\cdots\cdots\cdots\cdots\cdots \end{aligned}$$

By studying (3.11) on different rectangles, we are able to show the uniqueness of our solution:

Proposition 3.8 Suppose f satisfies the Lipschitz conditon, and y_1, y_2 are two solutions of (3.11), where y_1 is defined on $x \in I_1$, and y_2 is defined on $x \in I_2$. Suppose $I_1 \cap I_2 \neq \emptyset$ and y_1, y_2 share the same initial value condition $y(\alpha) = \beta$. Then $y_1(x) = y_2(x)$ on $I_1 \cap I_2$.

Proof. Suppose $I_1 \cap I_2 = [p,q]$ and let $z := \sup\{x \mid y_1 \equiv y_2 \text{ on } [\alpha, x]\}$. It suffices to show $z = q$. Now suppose on the contrary that $z < q$, and consider the subtraction $|y_1 - y_2|$:

$$y_i = \beta + \int_\alpha^x f(t,y_i)\,\mathrm{d}t \implies |y_1 - y_2| = \left| \int_z^x f(t,y_1) - f(t,y_2)\,\mathrm{d}t \right|.$$

Construct an interval $I^* = [z - \frac{1}{2L^*}, z + \frac{1}{2L^*}] \cap [p,q]$, and let $x_m = \arg\max_{x \in I^*} |y_1(x) -$

$y_2(x)|$, which implies for $\forall x \in I^*$,

$$\begin{aligned}
|y_1(x) - y_2(x)| &= \left|\int_z^x f(t,y_1) - f(t,y_2)\,\mathrm{d}t\right| \\
&\leq \int_z^x |f(t,y_1(t)) - f(t,y_2(t))|\,\mathrm{d}t \\
&\leq L^* \int_z^x |y_1(x) - y_2(x)|\,\mathrm{d}t \\
&\leq L^* |x-z||y_1(x_m) - y_2(x_m)| \\
&\leq \frac{1}{2}|y_1(x_m) - y_2(x_m)|.
\end{aligned}$$

Taking $x = x_m$, we imply $y_1 \equiv y_2$ for $\forall x \in I^*$, which contradicts the maximality of z. ∎

Combining Theorem (3.4) and proposition (3.8), we imply the existence of a unique "maximal" solution for the IVP (3.11), i.e., the unique solution is defined on a maximal interval.

Corollary 3.3 Let $U \subseteq \mathbb{R}^2$ be an open set such that $f(x,y)$ satisfies the Lipschitz condition for any $[a,b] \times [c,d] \subseteq U$, then there exists x_m and x_M in $\overline{\mathbb{R}}$ such that

- The IVP (3.11) admits a solution $y(x)$ for $x \in (x_m, x_M)$, and if y^* is another solution of (3.11) on some interval $I \subseteq (x_m, x_M)$, then $y \equiv y^*$ on I.
- Therefore $y(x)$ is maximally defined; and $y(x)$ is unique.

■ **Example 3.12** Consider the IVP

$$\begin{cases} \dfrac{\mathrm{d}y}{\mathrm{d}x} = x^2 y^{1/5} \\ y(0) = C \end{cases} \implies \frac{\partial f}{\partial y} = \frac{x^2}{5y^{4/5}}.$$

- Taking $U = \mathbb{R} \times (0,\infty)$ implies $y = \left(\frac{4x^3}{15} + c^{4/5}\right)^{5/4}$, defined on $(\sqrt[3]{-15/4c^{4/5}}, \infty)$.
- When $c = 0$, $f(x,y)$ does not satisfy the Lipschitz condition. The uniqueness of solution does not hold.

3.6. Wednesday for MAT4002

3.6.1. Remarks on product space

Reviewing.

- Product Topology: For topological space $(X, \mathcal{T}_X)$ and $(Y, \mathcal{Y})$, define the basis

$$\mathcal{B}_{X\times Y} = \{U \times V \mid U \in \mathcal{T}_X, V \in \mathcal{T}_Y\}$$

  and the family of union of subsets in $\mathcal{B}_{X\times Y}$ forms a product topology.

Proposition 3.9 a ring torus is homeomorphic to the Cartesian product of two circles, say $S^1 \times S^1 \cong T$.

Proof. Define a mapping $f : [0, 2\pi] \times [0, 2\pi] \to T$ as

$$f(\theta, \phi) = \Big((R + r\cos\theta)\cos\phi, \quad (R + r\cos\theta)\sin\phi, \quad r\sin\theta\Big)$$

Define $i : T \to \mathbb{R}^3$, we imply

$$i \circ f : [0, 2\pi] \times [0, 2\pi] \to \mathbb{R}^3 \text{ is continuous}$$

Therefore we imply $f : [0, 2\pi] \times [0, 2\pi] \to T$ is continuous. Together with the condition that

$$\begin{cases} f(0, y) = f(2\pi, y) \\ f(x, 0) = f(x, 2\pi) \end{cases}$$

we imply the function $f : S^1 \to S^1 \to T$ is continuous. We can also show it is bijective. We can also show f^{-1} is continuous. ■

Proposition 3.10 1. Let $X \times Y$ be endowed with product topology. The projection mappings defined as

$$p_X : X \times Y \to X, \text{ with } p_X(x,y) = x$$
$$p_Y : X \times Y \to Y, \text{ with } p_Y(x,y) = y$$

are continuous.

2. (an equivalent definition for product topology) The product topology is the **coarest topology** on $X \times Y$ such that p_X and p_Y are both continuous.
3. (an equivalent definition for product topology) Let Z be a topological space, then the product topology is the unique topology that the red and the blue line in the diagram commutes:

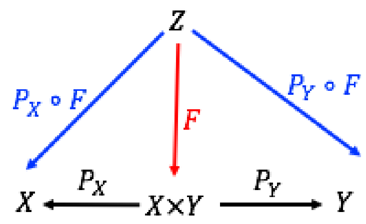


Figure 3.3: Diagram summarizing the statement (*)

namely,

the mapping $F : Z \to X \times Y$ is continuous iff both $P_X \circ F : Z \to X$ and $P_Y \circ F : Z \to Y$ are continuous. (*)

Proof. 1. For any open U, we imply $p_X^{-1}(U) = U \times Y \in \mathcal{B}_{X\times Y} \subseteq \mathcal{T}_{X\times Y}$, i.e., $p_X^{-1}(U)$ is open. The same goes for p_Y.

2. It suffices to show any topology $\mathcal{T}$ that meets the condition in (2) must contain $\mathcal{T}_{\text{product}}$. We imply that for $\forall U \in \mathcal{T}_X, V \in \mathcal{T}_Y$,

$$\begin{cases} p_X^{-1}(U) = U \times X \in \mathcal{T} \\ p_Y^{-1}(V) = X \times V \in \mathcal{T} \end{cases} \implies (U \times Y) \cap (X \times V) = (U \cap X) \times (Y \cap V) = U \times V \in \mathcal{T},$$

which implies $\mathcal{B}_{X\times Y} \subseteq \mathcal{T}$. Since $\mathcal{T}$ is closed for union operation on subsets, we imply $\mathcal{T}_{\text{product topology}} \subseteq \mathcal{T}$.

3. (a) Firstly show that $\mathcal{T}_{\text{product}}$ satisfies (*).

   - For the forward direction, by (1) we imply both $p_X \circ F$ and $p_Y \circ F$ are continuous, since the composition of continuous functions are continuous as well.
   - For the reverse direction, for $\forall U \times \mathcal{T}_X, V \in \mathcal{T}_Y$,

   $$F^{-1}(U \times V) = (p_X \circ F)^{-1}(X) \cap (p_Y \circ F)^{-1}(Y),$$

   which is open due to the continuity of $p_X \circ F$ and $p_Y \circ F$.

   (b) Then we show the uniqueness of $\mathcal{T}_{\text{product}}$. Let $\mathcal{T}$ be another topology $X \times Y$ satisfying (*).

   - Take $Z = (X \times Y, \mathcal{T})$, and consider the identity mapping $F = \text{id} : Z \to Z$, which is continuous. Therefore $p_X \circ \text{id}$ and $p_Y \circ \text{id}$ are continuous, i.e., p_X and p_Y are continuous. By (2) we imply $\mathcal{T}_{\text{product}} \subseteq \mathcal{T}$.
   - Take $Z = (X \times Y, \mathcal{T}_{\text{product}})$, and consider the identity mapping $F = \text{id} : Z \to Z$. Note that $p_X \circ F = p_X$ and $p_Y \circ F = p_Y$, which is continuous by (1). Therefore, the identity mapping $F : (X \times Y, \mathcal{T}_{\text{product}}) \to (X \times Y, \mathcal{T})$ is continuous, which implies

   $$U = \text{id}^{-1}(U) \subseteq \mathcal{T}_{\text{product}} \text{ for } \forall U \in \mathcal{T},$$

   i.e., $\mathcal{T} \subseteq \mathcal{T}_{\text{product}}$.

   The proof is complete.

■

Definition 3.6 [Disjoint Union] Let $X \times Y$ be two topological spaces, then the **disjoint union** of X and Y is

$$X \sqcup Y := (X \times \{0\}) \cup (Y \times \{1\})$$

R

1. We define that U is open in $X \sqcup Y$ if

    (a) $U \cap (X \times \{0\})$ is open in $X \times \{0\}$; and

    (b) $U \cap (Y \times \{1\})$ is open in $Y \times \{1\}$.

    We also need to show the weill-definedness for this definition.

2. S is open in $X \perp Y$ iff S can be expressed as

$$S = (U \times \{0\}) \cup (V \times \{1\})$$

where $U \subseteq X$ is open and $V \subseteq Y$ is open.

3.6.2. Properties of Topological Spaces

3.6.2.1. Hausdorff Property

Definition 3.7 [First Separation Axiom] A topological space X satisfies the **first separation axiom** if for any two distinct points $x \neq y \in X$, there exists open $U \ni x$ but not including y.

Proposition 3.11 A topological space X has first separation property if and only if for $\forall x \in X$, $\{x\}$ is closed in X.

Proof. Sufficiency. Suppose that $x \neq y$, then construct $U := X \setminus \{y\}$, which is a open set that contains x but not includes y.

Necessity. Take any $x \in X$, then for $\forall y \neq x$, there exists $y \in U_y$ that is open and $x \notin U_y$. Thus

$$\{y\} \subseteq U_y \subseteq X \setminus \{x\}$$

which implies

$$\bigcup_{y \in X \setminus \{x\}} \{y\} \subseteq \bigcup_{y \in X \setminus \{x\}} U_y \subseteq X \setminus \{x\},$$

i.e., $X \setminus \{x\} = \bigcup_{y \in X \setminus \{x\}} U_y$ is open in X, i.e., $\{x\}$ is closed in X. ■

Definition 3.8 [Second separation Axiom] A topological space satisfies the **second separation axiom** (or X is Hausdorff) if for all $x \neq y$ in X, there exists open sets U, V such that

$$x \in U, \quad y \in V, \quad U \cap V = \emptyset$$

■ **Example 3.13** All metrizable topological spaces are Hausdorff.

Suppose $d(x,y) = r > 0$, then take $B_{r/2}(x)$ and $B_{r/2}(y)$

■ **Example 3.14** Note that a topological space that is **first separable** may not necessarily be **second separable**:

Consider $\mathcal{T}_{\text{co-finite}}$, then X is first separable but not Hausdorff:

Suppose on the contrary that for given $x \neq y$, there exists open sets U, V such that $x \in U, y \in V$, and

$$U \cap V = \emptyset \implies X = X \setminus (U \cap V) = (X \setminus U) \cup (X \setminus V),$$

implying that the union of two finite sets equals X, which is infinite, which is a contradiction.