A FIRST COURSE

IN

ANALYSIS

A FIRST COURSE

IN

ANALYSIS

MAT2006 Notebook

Lecturer

Prof. Weiming Ni

The Chinese University of Hongkong, Shenzhen

Tex Written By

Mr. Jie Wang

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 taken notes from the MAT2006 in fall semester, 2018. These lecture notes were taken and compiled in LaTeX by Jie Wang, an undergraduate student in Fall 2018. Prof. Weiming Ni has not edited this document. Students taking this course may use the notes as part of their reading and reference materials. This version of the lecture notes were revised and extended for many times, but may still contain many mistakes and typos, including English grammatical and spelling errors, in the notes. It would be greatly appreciated if those students, who will use the notes as their reading or reference material, tell any mistakes and typos to Jie Wang for improving this notebook.

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 8

Week8

8.1. Friday

This lecture will discuss the multi-variable calculus.

8.1.1. Introduction to metric space

The multi-variable calculus aims to study the function $f:\mathbb{R}^m \mapsto \mathbb{R}^n$:

$$f(\underbrace{x_1,x_2,\dots,x_m}_{\boldsymbol{x}}) = (f_1(\boldsymbol{x}),f_2(\boldsymbol{x}),\dots,f_n(\boldsymbol{x}))$$

To begin with, let's assume $n=1$, i.e., we study only one component in RHS first. The preliminaries for this concept are limit points or something else. Let's define them in high dimension first.

Generalization from $\mathbb{R}$ to $\mathbb{R}^2$. The distance between two points in $\mathbb{R}^2$ is usually defined as follows:

$$\begin{aligned}\boldsymbol{x}&=(x_1,x_2)\\ \boldsymbol{y}&=(y_1,y_2)\end{aligned} \implies d(\boldsymbol{x},\boldsymbol{y}) = \sqrt{(x_1-y_1)^2+(x_2-y_2)^2} \quad (L_2 \text{ norm})$$

- Sometimes another distance measrue is $d_1(\boldsymbol{x},\boldsymbol{y}) = |x_1-y_1|+|x_2-y_2|$, which is called L_1 norm
- Or more generally, $d_\infty(\boldsymbol{x},\boldsymbol{y}) = \max_{1\le i\le n}|x_i-y_i|$, which is called L_∞ norm.

Those distance measures are essentially the same in order, i.e.,

$$\frac{1}{\sqrt{2}}d(\boldsymbol{x},\boldsymbol{y}) \le d_1(\boldsymbol{x},\boldsymbol{y}) \le \sqrt{2}d(\boldsymbol{x},\boldsymbol{y}), \qquad \forall x,y \in \mathbb{R}^2$$
$$d(\boldsymbol{x},\boldsymbol{y}) \le d_\infty(\boldsymbol{x},\boldsymbol{y}) \le \sqrt{2}d(\boldsymbol{x},\boldsymbol{y}), \qquad \forall x,y \in \mathbb{R}^2$$

R

- For those distance measures with the same order, the corresponding properties defined with those measures are also nearly the same. We always define the L_2 norm as our distance measure by default.
- However, one distance measure is different in order from those above:

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

Definition 8.1 [Matric Space] The binary operation $d : \mathcal{H} \times \mathcal{H} \mapsto \mathbb{R}$ is called a matric if the following are satisfied:

1. $d(\boldsymbol{x},\boldsymbol{y}) \ge 0, \forall x,y \in \mathcal{H}$;
2. $d(\boldsymbol{x},\boldsymbol{y}) = d(\boldsymbol{y},\boldsymbol{x}), \forall x,y \in \mathcal{H}$;
3. $d(\boldsymbol{x},\boldsymbol{z}) \le d(\boldsymbol{x},\boldsymbol{y}) + d(\boldsymbol{y},\boldsymbol{z})$, $\forall x,y,z \in \mathcal{H}$,

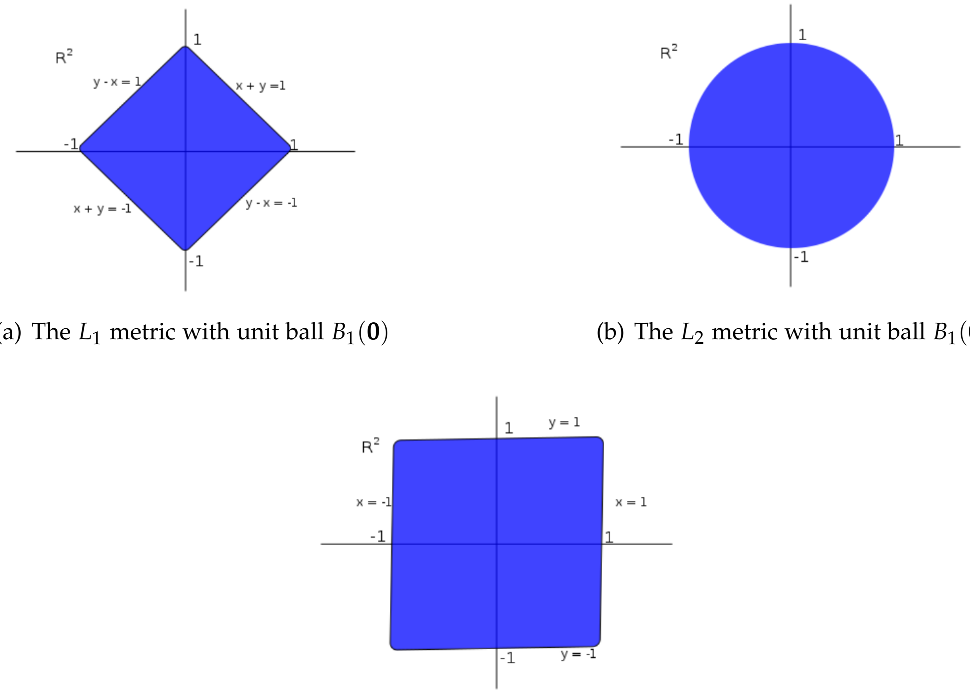
where $\mathcal{H}$ is called the **metric space**, e.g., $\mathbb{R}^m$ is a common metric space.

The reason for defining matric is to describe **convergence** in high dimensions. Let's define the corresponding definitions related to convergence again:

Definition 8.2 [Open ball] The open ball is a set $B_r(\boldsymbol{a})$ such that

$$B_r(\boldsymbol{a}) := \{\boldsymbol{x} \in \mathcal{H} \mid d(\boldsymbol{x},\boldsymbol{a}) < r\} \tag{8.1}$$

Some illustrations for $B_1(\boldsymbol{0})$ in the metric space $\mathbb{R}^2$ is shown as follows:



(a) The L_1 metric with unit ball $B_1(\mathbf{0})$

(b) The L_2 metric with unit ball $B_1(\mathbf{0})$

(c) The L_∞ metric with unit ball $B_1(\mathbf{0})$

Figure 8.1: illustrations for $B_1(\mathbf{0})$ in the metric space $\mathbb{R}^2$

Definition 8.3 [Convergence] The sequence defined on the metric space $(\mathcal{H}, d)$ is convergent to x_0 if $d(x_n, x_0) \to 0$ as $n \to \infty$, which is denoted as $\lim_{n\to\infty} x_n = x_0$, or simply $x_n \to x_0$

Check rudin's book in page 32 for the concepts about Openness, closedness, neighborhood, boundness, **compactness**, limit points, **connectness**. In particular, let's discuss something important.

Definition 8.4 [Compact] A set K is said to be compact ifs every open cover has a finite subcover.

The proof of the proposition that K is compact if it is bounded and closed is left as exercise.

Definition 8.5 A set S is **pathwise** connected if for any two points $\boldsymbol{x}, \boldsymbol{y} \in S$, there exists a "**path** " Γ connecting $\boldsymbol{x}$ and $\boldsymbol{y}$, where Γ is a continuous function $[0,1] \mapsto S$ with the property that $\Gamma(0) = \boldsymbol{x}, \Gamma(1) = \boldsymbol{y}$.

Definition 8.6 [Domain] A domian is an open set which is **path-wise connected**.

R An open set is not necessarily path-wise connected.

Definition 8.7 [Oscillation]

1. The **oscillation** of $f : X \subseteq \mathbb{R}^m \mapsto \mathbb{R}$ on a set $E \subseteq X$ is $\omega(f;E) = d(f(E))$, where d denote the diameter of the set $f(E)$, i.e.,
$$\omega(f;E) = d(f(E)) = \sup_{x,y \in E} d(f(x), f(y)).$$
2. The **oscillation** of f at a point $\boldsymbol{a}$ is
$$w(f;\boldsymbol{a}) = \lim_{r \to 0+} \omega(f; B_r(\boldsymbol{a}))$$

Local Properties.

Proposition 8.1 — **Useful Properties.** Let f be a function mapping a metric space $\mathcal{H}$ to $\mathbb{R}^m$, then

1. f is continuous at the point $\boldsymbol{a}$ iff $\omega(f;\boldsymbol{a}) = 0$
2. If f is continuous at $\boldsymbol{a}$ with $f(\boldsymbol{a}) > 0$, then $f > 0$ in some neighborhood of $\boldsymbol{a}$
3. The linear combinations of continuous functions $(\alpha f + \beta g)$, component-wise products $(f \cdot g)$, or component-wise quotients $(\frac{f}{g}, g_i \neq 0)$ are also continuous functions

Proposition 8.2 — **Global Properties.**

1. Let $f : K \mapsto \mathbb{R}^n$ be a continuous function with K being compact, then we have
   (a) f is **uniformly continuous** on K
   (b) f is **bounded** on K

(c) f assumes its maximum and minimum on K

2. Intermediate Value property: If $f: E \mapsto \mathbb{R}$ is continuous, where E is **path-wise connected**, then $f(a) = A$ and $f(b) = B$ implies for all c between A and B, there exists $c \in E$ such that $f(c) = C$.

■ **Example 8.1** 1.

$$f(x,y) = \begin{cases} \frac{xy}{x^2+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$$

For any (x,y) on the line $y = mx$, we have

$$f(x,mx) = \frac{mx^2}{x^2+m^2x^2} = \frac{m}{1+m^2},$$

thus $\lim_{(x,y)\to(0,0)} f(x,y)$ does not exist sicne different paths toward the origin may lead to a different limit. There is another interesting fact:

$$\begin{cases} \lim_{x\to 0}\lim_{y\to 0} f(x,y) = 0, \\ \lim_{y\to 0}\lim_{x\to 0} f(x,y) = 0 \end{cases}$$

2.

$$f(x,y) = \begin{cases} x + y\sin\frac{1}{x}, x \neq 0 \\ 0, x = 0 \end{cases}$$

Apply ε-δ language to verify that $\lim_{(x,y)\to(0,0)} f(x,y) = 0$, but

$$\begin{cases} \lim_{x\to 0}\lim_{y\to 0} f(x,y) = 0 \\ \lim_{y\to 0}\lim_{x\to 0} f(x,y) \text{ does not exist} \end{cases}$$

R The continuity of a function at $x = a$ does not necessaily imply the interchangeability of limit processes. However, the uniform convergence can enable us to arrive at positive results.

3.

$$f(x,y) = \begin{cases} \dfrac{x^2y}{x^4+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$$

For any (x,y) on the line $y = mx$, we have

$$f(x,mx) = \frac{mx^2 \cdot x}{x^4 + m^2x^2} = \frac{mx}{x^2+m^2} \to 0, \text{ as } x \to 0$$

However, for any (x,y) on the line $y = x^2$, we have

$$f(x,x^2) = \frac{x^2 \cdot x^2}{x^4+x^4} = \frac{1}{2},$$

which means f is not continuous at $(0,0)$.

4.

$$f(x,y) = \begin{cases} \dfrac{x^2y}{x^6+y^2}, & (x,y) \neq (0,0) \\ 0, & (x,y) = (0,0) \end{cases}$$

When $y = mx$, we have $f(x,mx) = \frac{mx}{x^4+m^2} \to 0$.

When $y = x^3$, we have $f(x,x^3) = \frac{1}{2x} \to \infty$, i.e., f is unbounded near origin.

Question. Is it possible that a sphere S^2 and a circle S^1 are situated such that the distance from any point on the sphere to any point on the circle is the same, i.e., is it possible for a function

$$d(x,y) : S^2 \times S^1 \mapsto \mathbb{R}$$

remains constant? When will this fact be possible in $\mathbb{R}^k$?