

**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	ix
Notations	xi
1	Week1 1
1.1	Wednesday 1
1.1.1	Introduction to Set 1
1.2	Quiz 5
1.3	Friday 6
1.3.1	Proof of Schroder-Bernstein Theorem 6
1.3.2	Connectedness of Real Numbers 10
2	Week2 13
2.1	Wednesday 13
2.1.1	Review and Announcement 13
2.1.2	Irrational Number Analysis 13
2.2	Friday 21
2.2.1	Set Analysis 21
2.2.2	Set Analysis Meets Sequence 22
2.2.3	Completeness of Real Numbers 23
3	Week3 27
3.1	Tuesday 27
3.1.1	Application of Heine-Borel Theorem 27
3.1.2	Set Structure Analysis 29
3.1.3	Reviewing 31

3.2	Friday	33
3.2.1	Review	33
3.2.2	Continuity Analysis	34
4	Week4	41
4.1	Wednesday	41
4.1.1	Function Analysis	41
4.1.2	Continuity Analysis	46
4.2	Friday	50
4.2.1	Continuity Analysis	50
4.2.2	Monotone Analysis	53
4.2.3	Cantor Set	55
5	Week5	59
5.1	Wednesdays	59
5.1.1	Differentiation	59
5.1.2	Basic Rules of Differentiation	61
5.1.3	Analysis on Differential Calculus	62
5.2	Friday	67
5.2.1	Analysis on Derivative	67
5.2.2	Analysis on Mean-Value Theorem	68
5.3	Saturday: Comments on Quiz 1	73
5.3.1	First Question	73
5.3.2	Second Question	73
5.3.3	Third Question	74
5.3.4	Fourth Question	75
5.3.5	Fifth Question	75
5.3.6	Grading policy	76

6	Week6	77
6.1	Wednesday	77
6.1.1	Reviewing	77
6.1.2	Convergence Analysis	78
6.2	Friday	83
6.2.1	Recap	83
6.2.2	Riemann Integration	84
7	Week7	89
7.1	Wednesday	89
7.1.1	Integrable Analysis	89
7.1.2	Elementary Calculus Analysis	92
7.2	Friday	97
7.2.1	Improper Intergrals	97
8	Week8	105
8.1	Friday	105
8.1.1	Introduction to metric space	105
9	Week9	111
9.1	Friday	111
9.1.1	Preliminaries	112
9.1.2	Differentiation	114
10	Week10	121
10.1	Wednesday	121
10.1.1	Preliminaries on Notations	121
10.1.2	Analysis on multi-variate differentiation	122

10.2	Friday	128
10.2.1	Multi-variate Taylor's Theorem	128
10.2.2	Application: Optimality Condition	130
11	Week11	135
11.1	Wednesday	135
11.1.1	Recap	135
11.1.2	Introduction to Implicit Function Theorem	137
11.2	Friday	143
11.2.1	Analysis on IFT	143
11.2.2	Applications on IFT	144

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

11.2. Friday

11.2.1. Analysis on IFT

This lecture will talk about the full version of IFT.

Elementary Version. $F : U(x_0, y_0) \rightarrow \mathbb{R}$, which is $\mathcal{C}^p, p \geq 1$.

$$F(x_0, y_0) = 0, \quad F_y(x_0, y_0) \neq 0$$

which implies that there exists $I_x = \{x \mid |x - x_0| < \alpha\}$ and $I_y = \{y \mid |y - y_0| < \beta\}$ and $f \in \mathcal{C}^p(I_x; I_y)$ such that

$$F(x, y) = 0 \quad \forall (x, y) \in I_x \times I_y \iff y = f(x)$$

and

$$f'(x) = -\frac{F_x(x, f(x))}{F_y(x, f(x))}$$

Generalized version. $F : U(\mathbf{x}_0, y_0) (\subseteq \mathbb{R}^m \times \mathbb{R}) \rightarrow \mathbb{R}$, which is $\mathcal{C}^p, p \geq 1$.

$$F(\mathbf{x}_0, y_0) = 0, \quad F_y(\mathbf{x}_0, y_0) \neq 0$$

which implies that there exists $I_{\mathbf{x}} = \{\mathbf{x} \in \mathbb{R}^m \mid |\mathbf{x} - \mathbf{x}_0| < \alpha\}$ and $I_y = \{y \in \mathbb{R} \mid |y - y_0| < \beta\}$ and $f \in \mathcal{C}^p(I_{\mathbf{x}}; I_y)$ such that

$$F(\mathbf{x}, y) = 0 \quad \forall (\mathbf{x}, y) \in I_{\mathbf{x}} \times I_y \iff y = f(\mathbf{x})$$

and

$$Df(\mathbf{x}) = -\frac{1}{F_y(\mathbf{x}, f(\mathbf{x}))} D_{\mathbf{x}} F(\mathbf{x}, f(\mathbf{x}))$$

where $Df(\mathbf{x}) = \nabla^T f(\mathbf{x})$ and $D_{\mathbf{x}} F(\mathbf{x}, f(\mathbf{x})) = \nabla_{\mathbf{x}}^T F(\mathbf{x}, f(\mathbf{x}))$.

Full Version. $F : U(\mathbf{x}_0, \mathbf{y}_0) (\subseteq \mathbb{R}^m \times \mathbb{R}^n) \rightarrow \mathbb{R}^n$, which is $\mathcal{C}^p, p \geq 1$.

$$F(\mathbf{x}_0, \mathbf{y}_0) = \mathbf{0}, \quad D_{\mathbf{y}}F(\mathbf{x}_0, \mathbf{y}_0) \text{ is invertible}$$

which implies that there exists $I_{\mathbf{x}} = \{\mathbf{x} \in \mathbb{R}^m \mid |\mathbf{x} - \mathbf{x}_0| < \alpha\}$ and $I_{\mathbf{y}} = \{\mathbf{y} \in \mathbb{R}^n \mid |\mathbf{y} - \mathbf{y}_0| < \beta\}$ and $f \in \mathcal{C}^p(I_{\mathbf{x}}; I_{\mathbf{y}})$ such that

$$F(\mathbf{x}, \mathbf{y}) = \mathbf{0} \quad \forall (\mathbf{x}, \mathbf{y}) \in I_{\mathbf{x}} \times I_{\mathbf{y}} \iff \mathbf{y} = f(\mathbf{x})$$

and

$$Df(\mathbf{x}) = -[D_{\mathbf{y}}F(\mathbf{x}, f(\mathbf{x}))]^{-1} D_{\mathbf{x}}F(\mathbf{x}, f(\mathbf{x}))$$

where $Df(\mathbf{x}) \in \mathbb{R}^{n \times m}$; $D_{\mathbf{y}}F(\mathbf{x}, f(\mathbf{x})) \in \mathbb{R}^{n \times n}$; and $D_{\mathbf{x}}F(\mathbf{x}, f(\mathbf{x})) \in \mathbb{R}^{n \times m}$.

Proof. Fix m , induction on n .

1. As $n = 1$, it is done.
2. The rest are similar to the proof in elementary version.

■

In this lecture and next upcoming Wednesday, we will talk about the application of IFT. For example, how to apply Chain rule to differentiate; how to compute Jacobian matrix, and the inverse. Pay attention to computational aspect. This will show up in second quiz, as well as the final, too.

11.2.2. Applications on IFT

Inverse Function Theorem.

Theorem 11.4 Given a function $f : E \rightarrow \mathbb{R}^m$, where E is a **domain** (pre-assume it is connected) in \mathbb{R}^m with the property that:

1. $f \in \mathcal{C}^p(E; \mathbb{R}^m), p \geq 1$
2. $Df(\mathbf{x}_0)$ is invertible, where $\mathbf{x}_0 \in E$

which implies that

1. g is invertible near $f(\mathbf{x}_0)(:=\mathbf{y}_0)$, i.e., there exists $U(\mathbf{x}_0) \subseteq E$ and $V(\mathbf{y}_0) \in \mathbb{R}^m$ such that g is a \mathcal{C}^p -diffeomorphism from $U(\mathbf{x}_0)$ to $V(\mathbf{x}_0)$; and

$$Dg(\mathbf{y}) = [Df(g(\mathbf{y}))]^{-1}$$

Note that \mathcal{C}^p diffeomorphism means g is one-to-one onto mapping and $f \in \mathcal{C}^p$.

Proof. Define $F(\mathbf{x}, \mathbf{y}) = f(\mathbf{x}) - \mathbf{y}$, $F: E \times \mathbb{R}^m \rightarrow \mathbb{R}^m$ is \mathcal{C}^p .

$D_{\mathbf{x}}F(\mathbf{x}_0; \mathbf{y}_0) = Df(\mathbf{x}_0)$ is invertible.

$$F(\mathbf{x}_0, \mathbf{y}_0) = 0.$$

Applying IFT, we imply that there exists a neighborhood $I_{\mathbf{x}} = \{\mathbf{x} \in E \mid |\mathbf{x} - \mathbf{x}_0| < \alpha\}$ and $I_{\mathbf{y}} = \{\mathbf{y} \in \mathbb{R}^m \mid |\mathbf{y} - \mathbf{y}_0| < \beta\}$ and $g \in \mathcal{C}^p(I_{\mathbf{y}}; I_{\mathbf{x}})$ such that

$$F(\mathbf{x}, \mathbf{y}) = 0 \quad \forall (\mathbf{x}, \mathbf{y}) \in I_{\mathbf{x}} \times I_{\mathbf{y}} \iff \mathbf{x} = g(\mathbf{y}),$$

i.e., $f(g(\mathbf{y})) = \mathbf{y}$ iff $\mathbf{x} = g(\mathbf{y})$; and

$$Dg(\mathbf{y}) = -[D_{\mathbf{x}}F(g(\mathbf{y}), \mathbf{y})]^{-1} D_{\mathbf{y}}F(g(\mathbf{y}), \mathbf{y}).$$

Notet that $D_{\mathbf{y}}F(g(\mathbf{y}), \mathbf{y}) = -I$, and therefore

$$Dg(\mathbf{y}) = [Df(\mathbf{x})]^{-1}$$

■

Pay attention to the order of derivative when applying full version IFT.

Rank Theorem.

Definition 11.2 [Rank] The **rank** of a vector function $f: U(\subseteq \mathbb{R}^m) \rightarrow \mathbb{R}^n$ at a point $\mathbf{x} \in U$ is defined to be the *rank* of $Df(\mathbf{x})$. ■

Theorem 11.5 — Rank Theorem. Suppose $f \in \mathcal{C}^p(U(\mathbf{x}_0; \mathbb{R}^n))$, where $U(\mathbf{x}_0)$ is a neighborhood of $\mathbf{x}_0 \in \mathbb{R}^m$. If f has the same constant rank k at every point $\mathbf{x} \in U(\mathbf{x}_0)$, then there exists a neighborhood $N(\mathbf{x}_0)$ of \mathbf{x}_0 and a neighborhood $N(\mathbf{y}_0)$ of $\mathbf{y}_0 := f(\mathbf{x}_0)$ and two \mathcal{C}^p -diffeomorphism,

$$u = \phi(\mathbf{x}) \text{ in } N(\mathbf{x}_0) \quad v = \psi(\mathbf{y}) \text{ in } N(\mathbf{y}_0)$$

such that $v = \psi \circ f \circ \phi^{-1}(u)$ takes the form

$$\mathbf{u} := (u_1, \dots, u_k, u_{k+1}, \dots, u_m) \rightarrow v = (v_1, \dots, v_n) := v(u_1, \dots, u_k, 0, \dots, 0)$$

R Given $f: \mathbf{x} \in U(\mathbf{x}_0) \rightarrow \mathbf{y} \in V(\mathbf{y}_0)$, if f has constant rank k , then we have

$$\phi^{-1}: u \in \phi(N(\mathbf{x}_0)) \rightarrow U(\mathbf{x}_0)$$

and

$$\psi: y \in V(\mathbf{y}_0) \rightarrow v \in \psi^{-1}(N(\mathbf{y}_0))$$

and

$$u = (u_1, \dots, u_k, u_{k+1}, \dots, u_m) \rightarrow (u_1, \dots, u_k, 0, \dots, 0).$$

Proof. (I) For $f := (f_1, \dots, f_n)$, we have

$$Df(\mathbf{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1} & \dots & \frac{\partial f_n}{\partial x_m} \end{pmatrix} (\mathbf{x}_0)$$

w.l.o.g., the first k principal minor

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1} & \dots & \frac{\partial f_k}{\partial x_k} \end{pmatrix} (\mathbf{x}_0)$$

is non-singular, which implies that there exists $N(\mathbf{x}_0)$ such that for $\forall x \in N(\mathbf{x}_0)$,

$$\begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_k} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1} & \dots & \frac{\partial f_k}{\partial x_k} \end{pmatrix}(\mathbf{x})$$

is non-singular.

(II) Define

$$\phi(\mathbf{x}) = \begin{pmatrix} f_1(\mathbf{x}) \\ \vdots \\ f_k(\mathbf{x}) \\ x_{k+1} \end{pmatrix}$$

then

$$D\phi(\mathbf{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \dots & \frac{\partial f_1}{\partial x_k} & | & \\ \vdots & & \ddots & & \vdots \\ \frac{\partial f_k}{\partial x_1} & \dots & \frac{\partial f_k}{\partial x_k} & | & \\ & & & & \mathbf{I} \end{pmatrix}$$

which is invertible as well, which implies that ϕ^{-1} exists, and ϕ is a \mathcal{C}^p -differentiaerom.

(III) Let $g = f \circ \phi^{-1}(\mathbf{u})$, which follows that

$$\begin{aligned} y_1 &= f_1 \circ \phi^{-1}(u_1, \dots, u_m) = u_1 \\ &\vdots \\ y_k &= f_k \circ \phi^{-1}(u_1, \dots, u_m) = u_k \\ y_{k+1} &= f_{k+1} \circ \phi^{-1}(u_1, \dots, u_m) = g_{k+1}(u_1, \dots, u_m) \\ &\vdots \\ y_n &= f_n \circ \phi^{-1}(u_1, \dots, u_m) = g_n(u_1, \dots, u_m) \end{aligned}$$

Note that $Dg(u) = Df(\phi^{-1}(u)) \circ D\phi^{-1}(u)$ has rank k at every point, but note that

$$Dg(u) = \begin{pmatrix} \mathbf{I}_{k \times k} & \mathbf{0} \\ * & \frac{\partial g[k+1:n]}{\partial u[k+1:m]} \end{pmatrix}$$

which implies $\frac{\partial g[k+1:n]}{\partial u[k+1:m]}$ is a zero matrix, which implies $g[k+1:n]$ is independent of $u[k+1:m]$, i.e.,

$$g_{k+1}(u_1, \dots, u_m) = g_{k+1}(u_1, \dots, u_k).$$

(IV) Define

$$\psi(y) = v = (v_1, \dots, v_n),$$

where

$$\begin{aligned} \psi_1(y) &= y_1 \\ &\vdots \\ \psi_k(y) &= y_k \\ \psi_{k+1}(y) &= y_{k+1} - g_{k+1}(y_1, \dots, y_k) \\ &\vdots \\ \psi_n(y) &= y_n - g_n(y_1, \dots, y_k) \end{aligned}$$

and

$$D\psi(y_0) = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ * & \mathbf{I} \end{pmatrix},$$

which is invertible.

(V) It follows that $\psi \circ f \circ \phi^{-1} = \psi \circ g$, which maps $(u_1, \dots, u_k, u_{k+1}, u_m)$ to

$$(u_1, \dots, u_k, 0, 0, \dots, 0).$$

■