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 $\mathcal{M}$

11.2. Friday

11.2.1. Analysis on IFT

This lecture will talk about the full verison of IFT.

Elementrary Version. $F : U(x_0, y_0) \to \mathbb{R}$, 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

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

and

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

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

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

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

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

and

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

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

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

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

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

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

and

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

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

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 $f : E \to \mathbb{R}^m$, 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. $Df(\boldsymbol{x}_0)$ is invertible, where $\boldsymbol{x}_0 \in E$

which implies that

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

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

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

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

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

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

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

$$F(\boldsymbol{x},\boldsymbol{y}) = 0 \quad \forall (\boldsymbol{x},y) \in I_{\boldsymbol{x}} \times I_y \Longleftrightarrow \boldsymbol{x} = g(\boldsymbol{y}),$$

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

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

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

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

■

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

Rank Theorem.

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

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

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

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

$$\boldsymbol{u} := (u_1, \ldots, u_k, u_{k+1}, \ldots, u_m) \to v = (v_1, \ldots, v_n) := v(u_1, \ldots, u_k, 0, \ldots, 0)$$

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

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

and

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

and

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

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

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

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

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

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

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

is non-singular.

(II) Define

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

then

$$D\phi(\boldsymbol{x}) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1} & \cdots \frac{\partial f_1}{\partial x_k} \mid & \\ \vdots & \ddots & \vdots \\ \frac{\partial f_k}{\partial x_1} & \cdots \frac{\partial f_k}{\partial x_k} \mid & \\ & & \boldsymbol{I} \end{pmatrix}$$

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

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

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

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

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

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

$$D\psi(\boldsymbol{y}_0) = \begin{pmatrix} \boldsymbol{I} & \mathbf{0} \\ *** & \boldsymbol{I} \end{pmatrix},$$

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

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

■