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 9

Week9

9.1. Friday

We are now in the multi-variate differentiation part.

Comments on question in last lecture. The question left in the last lecture is that

> In dimension $\mathbb{R}^k$ with k to be determined, is it possible to find the smallest k such that a sphere S^2 and a circle S^1 have a way of putting to make each point from S^2 to S^1 have the same distance?

The answer is $k = 5$, Let's give an example. We define the sphere and the circle to be

$$S^2 = \{(x,y,z,0,0) \mid x^2 + y^2 + z^2 = 1\},$$
$$S^1 = \{(0,0,0,u,v) \mid u^2 + v^2 = 1\}$$

Therefore, the distance between any two points on the sphere and the circile, respectively, is

$$d = \sqrt{x^2 + y^2 + z^2 + u^+ v^2} \equiv \sqrt{2}$$

Why $k \leq 4$ is not ok? Now we give a instructive proof:

Proof. For the case $k \leq 4$, let $\boldsymbol{c}_2$ denote the center of sphere S^2 with radius r_2; $\boldsymbol{c}_3$ deote the center of circle S^1 with radius r_1. Any point on S^2 can be written as $\boldsymbol{c}_2 + \boldsymbol{x}$ with $\boldsymbol{x} \in M := \{\boldsymbol{x} \in \mathbb{R}^3 : \|\boldsymbol{x}\| = r_2\}$; and any point on S^1 can be written as $\boldsymbol{c}_1 + \boldsymbol{y}$ with $\boldsymbol{x} \in N := \{\boldsymbol{y} \in \mathbb{R}^2 : \|\boldsymbol{y}\| = r_1\}$. It follows that the distance between any two points can be expressed as:

$$d(\boldsymbol{c}_2+\boldsymbol{x},\boldsymbol{c}_1+\boldsymbol{y}) = \|(\boldsymbol{c}_2+\boldsymbol{x})-(\boldsymbol{c}_1+\boldsymbol{y})\| = \|\boldsymbol{x}\|^2+\|\boldsymbol{c}_2-\boldsymbol{y}-\boldsymbol{c}_1\|^2+2\langle \boldsymbol{x},\boldsymbol{c}_2-\boldsymbol{y}-\boldsymbol{c}_1\rangle$$

Note that the distance between any point in S^2 and any point in S^1 is the same. In particular, $d(\boldsymbol{c}_2+\boldsymbol{x},\boldsymbol{c}_1+\boldsymbol{y}) = d(\boldsymbol{c}_2-\boldsymbol{x},\boldsymbol{c}_1+\boldsymbol{y})$, which implies $\langle \boldsymbol{x},\boldsymbol{c}_2-\boldsymbol{y}-\boldsymbol{c}_1\rangle = 0, \forall \boldsymbol{x}\in M, \boldsymbol{y}\in N$, i.e.,

$$\boldsymbol{c}_2-\boldsymbol{c}_1+N\subseteq M^{\perp},$$

and therefore

$$\begin{aligned}\dim(\boldsymbol{c}_2-\boldsymbol{c}_1+N) = \dim(N) = 2 \\ \le \dim(M^{\perp}) = k-3,\end{aligned}$$

i.e., $k\le 5$ is the sufficient condition of the problem.

■

9.1.1. Preliminaries

Notations. Here we use lower-case and bolded alphabet to denote a vector, e.g., $\boldsymbol{x}$; upper-case and bolded alphabet to denote a matrix, e.g., $\boldsymbol{A}$; unbolded alphabet to denote a scalar, e.g., x_1 or A_1

Define $\boldsymbol{x} := (x_1,\dots,x_m)$ and $\boldsymbol{y} := (y_1,\dots,y_m)$, we define the L_2 norm

$$\begin{aligned}d(\boldsymbol{x},\boldsymbol{y}) :&= \|\boldsymbol{x}-\boldsymbol{y}\| \\ &= \left[(x_1-y_1)^2+\cdots+(x_m-y_m)^2\right]^{1/2}\end{aligned}$$

and the L_1, L_S (sup) norm:

$$\begin{aligned}d_1(\boldsymbol{x},\boldsymbol{y}) :&= \|\boldsymbol{x}-\boldsymbol{y}\| := |x_1-y_1|+\cdots+|x_m-y_m| \\ d_S(\boldsymbol{x},\boldsymbol{y}) :&= \|\boldsymbol{x}-\boldsymbol{y}\|_S := \max_{1\le i\le m}(x_i-y_i)\end{aligned}$$

Definition 9.1 [Vector Norm] A norm $\|\cdot\|$ is a function from a vector space X to $\mathbb{R}$ such that

1. $\|\boldsymbol{x}\| \ge 0$, $\forall x \in X$; and $\|\boldsymbol{x}\| = 0$ iff $\boldsymbol{x} = \boldsymbol{0}$;
2. $\|\lambda \boldsymbol{x}\| = |\lambda| \|\boldsymbol{x}\|$, for $\forall x \in X, \lambda \in \mathbb{R}$;
3. $\|\boldsymbol{x}_1 + \boldsymbol{x}_2\| \le \|\boldsymbol{x}_1\| + \|\boldsymbol{x}_2\|$

R Any norm defines a metric: $d(\boldsymbol{x},\boldsymbol{y}) = \|\boldsymbol{x} - \boldsymbol{y}\|$. From now on, we pre-assume the norm to be L_2 norm and the metric to be L_2 metric, unless specified.

The norm are used to masure the length of a vector, or the distance of two vectors. Correspondingly, the inner product can be used to measure the angle between two angles:

Definition 9.2 [Inner Product] An **inner product** is a binary operation function $\langle\cdot,\cdot\rangle : X \times X \mapsto \mathbb{R}$ such that

1. $\langle x,x\rangle \ge 0$ for $\forall x \in X$ and $\langle x,x\rangle = 0$ iff $\boldsymbol{x} = 0$
2. $\langle \boldsymbol{x},\boldsymbol{y}\rangle = \langle \boldsymbol{y},\boldsymbol{x}\rangle$, for $\forall \boldsymbol{x},\boldsymbol{y} \in X$
3. $\langle \lambda\boldsymbol{x},\boldsymbol{y}\rangle = \lambda\langle \boldsymbol{x},\boldsymbol{y}\rangle$ for $\forall x,y \in X$ and $\forall \lambda \in \mathbb{R}$
4. $\langle x+y,z\rangle = \langle x,z\rangle + \langle y,z\rangle$

R

1. Any inner product always defines a norm, i.e., $\|\boldsymbol{x}\| = \sqrt{\langle \boldsymbol{x},\boldsymbol{x}\rangle}$.
2. The angle θ between vector $\boldsymbol{x},\boldsymbol{y}$ is defined as:

$$\theta = \cos^{-1}\left(\frac{\langle \boldsymbol{x},\boldsymbol{y}\rangle}{\|\boldsymbol{x}\|\|\boldsymbol{y}\|}\right)$$

In particular, $\boldsymbol{x}$ and $\boldsymbol{y}$ are orthogonal, i.e., $\langle \boldsymbol{x},\boldsymbol{y}\rangle = 0$ if $\theta = \pm\frac{\pi}{2}$; $\boldsymbol{x},\boldsymbol{y}$ are parallel if $\theta = 0$, or $\theta = \pm\pi$.

9.1.2. Differentiation

Review for One-dimension. Given a function $f : \mathbb{R} \mapsto \mathbb{R}$, the derivative is defined as

$$f'(x_0) = \lim_{x \to x_0} \frac{f(x) - f(x_0)}{x - x_0}$$

or equivalently, the value $f'(x_0)$ is said to be the derivative of f if it satisfies the equality:

$$\begin{aligned}
0 &= \lim_{x \to x_0} \left[\frac{f(x) - f(x_0)}{x - x_0} - f'(x_0) \right] \\
&= \lim_{x \to x_0} \frac{f(x) - f(x_0) - f'(x_0)(x - x_0)}{x - x_0} \\
&= \lim_{x \to x_0} \left| \frac{f(x) - [f(x_0) + f'(x_0)(x - x_0)]}{x - x_0} \right|
\end{aligned}$$

The interpretation is that the affine (linear function) $f(x_0) + f'(x_0)(x - x_0)$ approximates f near x_0 in at least first order. We can use the similar way to define the high-dimension derivative.

High-Dimension Derivative.

Definition 9.3 [Differentiable] A map $f : U \mapsto \mathbb{R}^n$, where U is open in $\mathbb{R}^m$, is **differentiable** at $\boldsymbol{x}_0 \in U$ if

$$\lim_{\boldsymbol{x} \to \boldsymbol{x}_0} \frac{\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - \boldsymbol{L}(\boldsymbol{x}_0)(\boldsymbol{x} - \boldsymbol{x}_0)\|}{\|\boldsymbol{x} - \boldsymbol{x}_0\|} = 0, \tag{9.1}$$

where $\boldsymbol{L}(\boldsymbol{x}_0)$ is said to be the derivative of $f(\boldsymbol{x})$ at $\boldsymbol{x} = \boldsymbol{x}_0$, which is often denoted as $Df(\boldsymbol{x}_0)$, or $f'(\boldsymbol{x}_0)$.

R Note that $f(\boldsymbol{x}), f(\boldsymbol{x}_0) \in \mathbb{R}^n$, but $(\boldsymbol{x} - \boldsymbol{x}_0) \in \mathbb{R}^m$, thus $L(\boldsymbol{x}_0)$ is a **linear transformation** from $\mathbb{R}^m$ to $\mathbb{R}^n$, i.e., a $n \times m$ matrix.

Interpretation. We re-write $f(\boldsymbol{x}) : \mathbb{R}^m \mapsto \mathbb{R}^n$ (a n-vector valued function of a m-vector argument) as:

$$f(\boldsymbol{x}) = \begin{pmatrix} f_1(\boldsymbol{x}) & \cdots & f_n(\boldsymbol{x}) \end{pmatrix},$$

with each $f_i : \mathbb{R}^m \mapsto \mathbb{R}$ being a scalar-valued function of a m-vector argument. Let's study only one component with 2-argument function first, i.e., $n = 1, m = 2$.

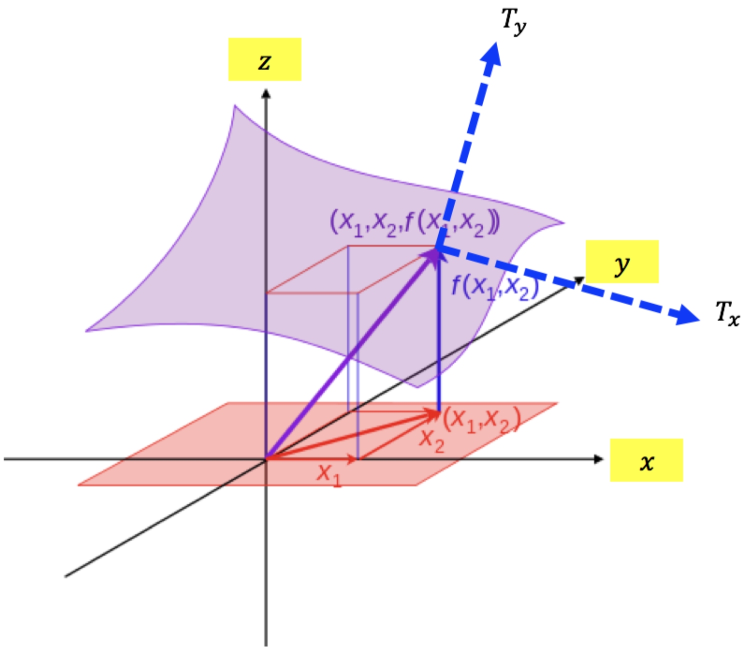

Figure 9.1: Diagram for partial derivatives

Given a surface $S : z = f(x_1, x_2)$, we study a point on the surface S, say $M_0(x_1, x_2, f(x_1, x_2))$. If the surface S intersects the plane $y = x_2$ with the point M_0, then the **partial derivative** $\frac{\partial f}{\partial x_1}(x_1, x_2)$ denotes the slope of the tangent line at M_0; the tangent line T_{x_1} can be denoted as a vector $T_{x_1} := (1, 0, \frac{\partial f}{\partial x_1}(x_1, x_2))$.

Similarly, if the surface S intersects the plane $x = x_1$ with the point M_0, then $\frac{\partial f}{\partial x_2}(x_1, x_2)$ denotes the slope of the tangent line at M_0; the tangent line T_{x_2} can be denoted as a vector $T_{x_2} := (0, 1, \frac{\partial f}{\partial x_2}(x_1, x_2))$.

Furthermore, the plane $\text{span}\{T_{x_1}, T_{x_2}\}$ denotes the tangent plane at point M_0.

Corollary 9.1 f is differentiable at $\boldsymbol{x}_0$ implies all partial derivatives of f at $\boldsymbol{x}_0$ exist.

The converse is not true. Let's raise a counter-example to explain that.

■ **Example 9.1**

$$f(x_1,x_2) = \begin{cases} \dfrac{x_1x_2}{x_1^2+x_2^2}, & (x_1,x_2) \neq (0,0) \\ 0, & (x_1,x_2) = (0,0) \end{cases}$$

When $x_2 = mx_1$, we have

$$f(x_1,mx_1) = \frac{mx_1^2}{x_1^2+mx_1^2} = \frac{m}{1+m^2},$$

i.e., f is not differentiable at the origin as it is not continuos at the origin.

However, we can verify that $\frac{\partial f}{\partial x_1}(0,0) = 0 = \frac{\partial f}{\partial x_2}(0,0)$. ■

What guarntees f to be differntiable if all partial derivatives exist?

$$\frac{\|f(\boldsymbol{x}) - f(\boldsymbol{x}_0) - Df(\boldsymbol{x}_0)(\boldsymbol{x}-\boldsymbol{x}_0)\|}{\|\boldsymbol{x}-\boldsymbol{x}_0\|} \to 0,$$

where $f : \mathbb{R}^m \mapsto \mathbb{R}^n$ and $Df(\boldsymbol{x}_0)$ be a $n \times m$ matrix. We re-write the formula for the derivative of f at $\boldsymbol{x} = \boldsymbol{x}_0$ first. Here we write $f(\boldsymbol{x}_0)$ in column vector form:

$$\begin{pmatrix} f_1(\boldsymbol{x}) \\ f_2(\boldsymbol{x}) \\ \vdots \\ f_n(\boldsymbol{x}) \end{pmatrix} = \begin{pmatrix} f_1(\boldsymbol{x}_0) \\ f_2(\boldsymbol{x}_0) \\ \vdots \\ f_n(\boldsymbol{x}_0) \end{pmatrix} + \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(x_0) & \frac{\partial f_1}{\partial x_2}(x_0) & \cdots & \frac{\partial f_1}{\partial x_m}(x_0) \\ \frac{\partial f_2}{\partial x_1}(x_0) & \frac{\partial f_2}{\partial x_2}(x_0) & \cdots & \frac{\partial f_2}{\partial x_m}(x_0) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial x_1}(x_0) & \frac{\partial f_n}{\partial x_2}(x_0) & \cdots & \frac{\partial f_n}{\partial x_m}(x_0) \end{pmatrix} \begin{pmatrix} x_1 - x_{0,1} \\ x_2 - x_0 \\ \vdots \\ x_m - x_0 \end{pmatrix} + o(\|\boldsymbol{x}-\boldsymbol{x}_0\|), \tag{9.2}$$

where $o(\|\boldsymbol{x}-\boldsymbol{x}_0\|)$ is a $m \times 1$ vector that has order less than $\|\boldsymbol{x}-\boldsymbol{x}_0\|$.

Each entry in LHS of (9.2) can be expressed as:

$$f_i(\boldsymbol{x}) = f_i(\boldsymbol{x}_0) + \nabla^{\mathrm{T}} f_i(\boldsymbol{x}_0) \cdot (\boldsymbol{x}-\boldsymbol{x}_0) + o(\|\boldsymbol{x}-\boldsymbol{x}_0\|),$$

with $\nabla^{\mathrm{T}} f_1(\boldsymbol{x}_0) = \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(\boldsymbol{x}_0) & \cdots & \frac{\partial f_1}{\partial x_m}(\boldsymbol{x}_0) \end{pmatrix}$ to be a row vector, and $(\boldsymbol{x}-\boldsymbol{x}_0)$ to be a column vector.

Sufficient Condition for differentiability. Recall that we have faced a function that is everywhere differentiable but nowhere monotone, which is very counter-inuitive. If adding the condition that such function is continuously differentiable, then such function is monotone. Similarly, the gap for corollary(9.1) is continuous differentiability.

> **Theorem 9.1** Let $f : U \mapsto \mathbb{R}^n$, where $U \subseteq \mathbb{R}^m$ is open. If all partial derivatives of f are **continuous** in U, then f is differentiable in U.

Proof for $n = 2, m = 1$ case. Consider $f : \mathbb{R}^2 \mapsto \mathbb{R}$, for $x_0 \in U$ with small h, k, we have:

$$f(x_0 + h, y_0 + k) - f(x_0, y_0) = [f(x_0 + h, y_0 + k) - f(x_0 + h, y_0)] + [f(x_0 + h, y_0) - f(x_0, y_0)] \tag{9.3a}$$

$$\text{(by Mean Value Theorem, } \exists (c, d) \in [0, h] \times [0, k] \text{ such that)} \tag{9.3b}$$

$$\begin{aligned}
&= k\frac{\partial f}{\partial y}(x_0 + h, y_0 + c) + h\frac{\partial f}{\partial x}(x_0 + d, y_0) && (9.3c)\\
&= k\frac{\partial f}{\partial y}(x_0, y_0 + c) + o(h) + h\frac{\partial f}{\partial x}(x_0, y_0) + o(h) && (9.3d)\\
&= k\frac{\partial f}{\partial y}(x_0, y_0) + h\frac{\partial f}{\partial x}(x_0, y_0) + o(h) + o(k) && (9.3e)
\end{aligned}$$

Or we write it in compact matrix form:

$$f(x_0 + h, y_0 + k) = f(x_0, y_0) + \begin{pmatrix} \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix} \begin{pmatrix} h \\ k \end{pmatrix} + o(\|(h, k)\|),$$

i.e., the derivative of f at $x = x_0$ exists, say $(\frac{\partial f}{\partial x}, \frac{\partial f}{\partial y})$. Thus f is differentiable. ■

The same proof can easily extend to $f : \mathbb{R}^n \mapsto \mathbb{R}$ by adding and subtracting $n - 1$ terms and then apply mean value theorem for n times; the extension to $f : \mathbb{R}^n \mapsto \mathbb{R}^m$ is clear by the same proof for each component.

Proof. Consider in general $f : \mathbb{R}^n \mapsto \mathbb{R}^m$, and it suffices to show (9.2), and moreover, it suffices to show each entry $f_i(\boldsymbol{x})$ in (9.2) can be expressed as

$$f_i(\boldsymbol{x}) = f_i(\boldsymbol{x}_0) + \nabla^{\mathrm{T}} f_i(\boldsymbol{x}_0) \cdot (\boldsymbol{x} - \boldsymbol{x}_0) + o(\|\boldsymbol{x} - \boldsymbol{x}_0\|).$$

Let's abuse notation to refer g as the i-th entry of f_i, i.e., $g : \mathbb{R}^n \mapsto \mathbb{R}$.

Let $\varepsilon > 0$, by continuity of partial derivatives, we can pick $\delta > 0$ such that for any $\|\boldsymbol{h}\| < \delta$, we have

$$\left\| \frac{\partial g}{\partial x_k}(\boldsymbol{x} + h) - \frac{\partial g}{\partial x_k}(\boldsymbol{x}) \right\| < \varepsilon, \qquad k = 1, 2, \dots, n \tag{9.4}$$

Let $\boldsymbol{L} = (\frac{\partial g}{\partial x_1}(\boldsymbol{x}), \dots, \frac{\partial g}{\partial x_n}(\boldsymbol{x}))$; $\phi_k(\boldsymbol{h}) = (h_1, h_2, \dots, h_k, 0, \dots, 0)$, $\phi_0(\boldsymbol{h}) = \boldsymbol{0}$, and suppose $\|\boldsymbol{h}\| < \delta$, , then we have:

$$\begin{aligned} |g(\boldsymbol{x} + \boldsymbol{h}) - g(\boldsymbol{x}) - \boldsymbol{L}\boldsymbol{h}| &= \left| \sum_{k=1}^{n} [g(\boldsymbol{x} + \phi_k(\boldsymbol{h})) - g(\boldsymbol{x} + \phi_{k-1}(\boldsymbol{h})) - \frac{\partial g}{\partial x_k}(\boldsymbol{x}) h_k] \right| && (9.5) \\ &\leq \sum_{k=1}^{n} \left| g(\boldsymbol{x} + \phi_k(\boldsymbol{h})) - g(\boldsymbol{x} + \phi_{k-1}(\boldsymbol{h})) - \frac{\partial g}{\partial x_k}(\boldsymbol{x}) h_k \right| && (9.6) \end{aligned}$$

By the Mean Value Theorem, there exists $\boldsymbol{c}_k \in [\boldsymbol{x} + \phi_{k-1}(\boldsymbol{h}), \boldsymbol{x} + \phi_k(\boldsymbol{h})]$ such that $g(\boldsymbol{x} + \phi_k(\boldsymbol{h})) - g(\boldsymbol{x} + \phi_{k-1}(\boldsymbol{h})) = \frac{\partial g}{\partial x_k}(\boldsymbol{c}_k) h_k$, and therefore

$$\begin{aligned} |g(\boldsymbol{x} + \boldsymbol{h}) - g(\boldsymbol{x}) - \boldsymbol{L}\boldsymbol{h}| &\leq \sum_{k=1}^{n} \left| \frac{\partial g}{\partial x_k}(\boldsymbol{c}_k) h_k - \frac{\partial g}{\partial x_k}(\boldsymbol{x}) h_k \right| && (9.7) \\ &= \sum_{k=1}^{n} \left| \frac{\partial g}{\partial x_k}(\boldsymbol{c}_k) - \frac{\partial g}{\partial x_k}(\boldsymbol{x}) \right| |h_k| && (9.8) \\ &< \varepsilon \sum_{k=1}^{n} |h_k| && (9.9) \end{aligned}$$

where (9.9) is due to the continuty of partial derivatives. To make life easier, we specify the norm to be L_1 norm, which follows that

$$|g(\boldsymbol{x} + \boldsymbol{h}) - g(\boldsymbol{x}) - \boldsymbol{L}\boldsymbol{h}| \leq \varepsilon \|\boldsymbol{h}\|_1,$$

which implies g is differentiable at $\boldsymbol{x}$. In paricular, if every component f_i is differentiable, then f is differentiable.

∎

Planning for following weeks. The next lecture will talk about the chain rule, the inverse, the derivative of inverse, the directional derivative, the gradient. Next Friday will talk about Inverse function theorem and Implicit function theorem, which may last for two weeks.