MAS242 해석학 II
Notes

한승우

November 9, 2023

Contents

Chapter 1

Differentiation

1.1 Higher order partial derivatives

Definition 1.1.1

Given $f : U \to \mathbb{R}$ where U is an open set in $\mathbb{R}^m$, define $\partial_{ij} \triangleq \partial_i(\partial_j f)(x)$ for each $i, j \in [m]$ to be *2nd order partial derivatives*. Any higher order partial derivatives can be defined inductively.

Definition 1.1.2: C^k-regularity

$f : U \to \mathbb{R}$ is C^k-regular if all partial derivatives up to order k and they are continuous.

Theorem 1.1.1

$f : U(\subseteq \mathbb{R}^2) \to \mathbb{R}$ is C^2 at a point $c \in U$, i.e., $\exists \delta > 0$, f is C^2 in $B_\delta(c)$. Then, $\partial_{12} f(c) = \partial_{21} f(c)$.

Proof. Let $|h| < \delta$. Define $A(h) \triangleq f(c_1 + h_1, c_2 + h_2) - f(c_1 + h_1, c_2) - f(c_1, c_2 + h_2) + f(c_1, c_2)$. Define $u(x_1) \triangleq f(x_1, c_2 + h_2) - f(x_1, c_2)$ and $v(x_2) \triangleq f(c_1 + h_1, x_2) - f(c_1, x_2)$. Note that u and v are differentiable.

Then, $A(h) = u(c_1 + h_1) - u(c_1)$ and $A(h) = v(c_2 + h) - v(c_2)$. By MVT, $\exists c_1^* \in (c_1, c_1 + h_1)$ and $c_2^* \in (c_2, c_2 + h_2)$ s.t. $A(h) = u'(c_1^*)h_1 = h_1\big(\partial_1 f(c_1^*, c_2 + h) - \partial_1 f(c_1^*, c_2)\big) = h_1 h_2 \partial_{21} f(c_1^*, c_2^*)$

Similarly, $\exists c_1^{**}, c_2^{**}$ such that $A(h) = h_1 h_2 \partial_{12} f(c_1^{**}, c_2^{**})$. $\partial_{21} f(c_1^*, c_2^*) = \partial_{12} f(c_1^{**}, c_2^{**})$. Hence, as $|h| \to 0$, due to the continuity, $\partial_{21}(c) = \partial_{12}(c)$. □

Corollary 1.1.1

Suppose $f : U(\subseteq \mathbb{R}^m) \to \mathbb{R}$ is C^k at $c \in U$. Then $\partial_{j_1 j_2 \cdots j_k} f(c) = \partial_{j_1' j_2' \cdots j_k'}$ where $j_1' \cdots$ are a permutation of $j_1 \cdots$.

1.2 Extreme Values of differentiable Functions

Definition 1.2.1: Hessian

Let $f : U(\subseteq \mathbb{R}^m) \to \mathbb{R}$ be C_2 in U. Suppose $p \in U$ is a critical point of f, i.e., $\nabla f(p) = 0$. Define

$$\mathcal{H}f(x) \triangleq \begin{pmatrix} \partial_{11}f(x) & \partial_{21}f(x) & \cdots & \partial_{m1}f(x) \\ \partial_{12}f(x) & \partial_{22}f(x) & \cdots & \partial_{m2}f(x) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_{1m}f(x) & \partial_{2m}f(x) & \cdots & \partial_{mm}f(x) \end{pmatrix}.$$

(Sometimes $\mathcal{H}f(x) = D^2 f(x)$.)
Define $D(x) = \det \mathcal{H}f(x)$. (Note that $\mathcal{H}f(x)$ is symmetric when f is C^2 by the theorem above.)

Theorem 1.2.1 2nd-order derivative test for two variable functions.

When $m = 2$ and f is C^2, a critical point p is

- a local maximum if $D(p) > 0$ and $\partial_{11}f(p) > 0$ (or $\partial_{22}f(p) > 0$).
- a local minimum if $D(p) > 0$ and $\partial_{11}f(p) < 0$ (or $\partial_{22}f(p) < 0$).
- a saddle point if $D(p) < 0$.

The test fails when $D(p) = 0$.

Proof. Given a unit vector $\mathbf{u} = (u_1, u_2) \in \mathbb{R}^2$, $D_{\mathbf{u}}f = \nabla f \cdot \mathbf{u} = u_1\partial_1 f + u_2\partial_2 f$, and thus

$$D^2_{\mathbf{u}}f = (u_1\partial_1 + u_2\partial_2)(u_1\partial_1 f + u_2\partial_2 f) = u_1^2\partial_{11}f + u_1u_2(2\partial_{12}f) + u_2^2\partial_{22}f.$$

WLOG, $u_1 \neq 0$. Set $z = u_2/u_1$. Then,

$$D^2_{\mathbf{u}}f(p) = u_1^2\big(\partial_{11}f(p) + 2\partial_{12}f(p)z + \partial_{22}f(p)z^2\big).$$

Note that, if $D(p) > 0$, $D^2_{\mathbf{u}}f(p)$ has no real root.

- If $D(p) > 0$ and $\partial_{11}f(p) < 0$, Then, $D^2\mathbf{u} < 0$ for all unit vector $\mathbf{u}$.
- If $D(p) > 0$ and $\partial_{11}f(p) > 0$, Then, $D^2\mathbf{u} > 0$ for all unit vector $\mathbf{u}$.
- If $D(p) < 0$, $D^2_{\mathbf{u}}f(p)$ has different signs depending on $\mathbf{u}$.

For general m?

$$D_{\mathbf{u}}(D_{\mathbf{u}}f) = D_{\mathbf{u}}\sum_{j=1}^{m}\partial_j f u_j = \sum_{j=1}^{m}\big((\nabla\partial_j f)\cdot\mathbf{u}\big)u_j = \sum_{j=1}^{m}\sum_{k=1}^{m}u_k u_j\partial_{kj}f.$$

Hence,

$$D^2_{\mathbf{u}}f(p) = \mathbf{u}^{\mathrm{T}}\cdot D^2 f(p)\cdot\mathbf{u}$$

Since $D^2 f(p)$ is symmetric, its eigenvalues $\lambda_1, \cdots, \lambda_m$ exists and they are real numbers. Also, there exists an $m \times m$ orthogonal matrix $\mathcal{O}$ such that $D^2 f(p) = \mathcal{O}\Lambda(p)\mathcal{O}^{\mathrm{T}}$ where $\Lambda(p)$ is the diagonal matrix with entries are the eigenvalues.

Then, we can write $D^2_{\mathbf{u}}f(p) = \mathbf{u}\mathcal{O}\Lambda(p)\mathcal{O}^{\mathrm{T}}\mathbf{u}^{\mathrm{T}} = (\mathbf{u}\mathcal{O})\Lambda(p) = (\mathbf{u}\mathcal{O})^{\mathrm{T}}$. Since $\mathcal{O}$ is orthogonal, $\mathbf{u}\mathcal{O}$ is another arbitrary unit vector. □

Theorem 1.2.2 Generalized 2nd order partial derivatives test

When f is C^2, a critical point p is

- a local maximum if all eigenvalues of $D^2 f(p)$ are negative.

- a local minimum if all eigenvalues of $D^2 f(p)$ are positive.
- a saddle point if there are both negative eigenvalues and positive eigenvalues.

The test fails when there are zero eigenvalues.

Chapter 2

Inverse Function Theorem

2.1 Jacobian

Definition 2.1.1: Jacobian

Let $\mathbf{f}\colon U(\subseteq \mathbb{R}^m) \to \mathbb{R}^n$ be differentiable. The function $J_f : U \to \mathbb{R}$ defined by

$$J_{\mathbf{f}}(\mathbf{x}) = \det \begin{bmatrix} \partial_1 f_1(\mathbf{x}) & \cdots & \partial_n f_1(\mathbf{x}) \\ \vdots & \ddots & \vdots \\ \partial_1 f_n(\mathbf{x}) & \cdots & \partial_n f_n(\mathbf{x}) \end{bmatrix}$$

is called the *Jacobian* of $\mathbf{f}$ at $\mathbf{x}$.

Lemma 2.1.1

If $f : V(\subseteq \mathbb{R}^n) - \mathbb{R}$ and $\mathbf{g}\colon U \to V$ are differentiable, then

$$J_{f\circ \mathbf{g}}(\mathbf{x}) = J_f\big(\mathbf{g}(\mathbf{x})\big) \cdot J_{\mathbf{g}}(\mathbf{x}).$$

Note:-

The linear mapping $\mathrm{d}\mathbf{f}(c)$ is invertible if and only if $J_{\mathbf{f}}(c)$ is nonzero.

2.2 The Inverse Function Theorem

Lemma 2.2.1 Contraction Mapping Principle

Let (X, d) be a complete metric space. Let $\varphi\colon X \to X$. Suppose that there exists $M \in [0, 1)$ such that $d\big(\varphi(x_1), \varphi(x_2)\big) \le Md(x_1, x_2)$. (We call it a *contraction mapping*.) Then, there uniquely exists $x_* \in X$ such that $\varphi(x_*) = x_*$.

Proof. Fix any $x_0 \in X$. Since $\{x_j\}_{j\in\mathbb{Z}_+}$, where $x_j = \varphi(x_{j-1})$ for each $j \in \mathbb{Z}_+$, is continuous. It converges to some x_*. As φ is continuous, we have $\varphi(x_*) = x_*$. The uniqueness follows trivially. □

Note:-

- For each $v \in \mathbb{R}^n \setminus \{0\}$, $|Av| = |v| \cdot \left|A\frac{v}{|v|}\right| \le \|A\|_L \cdot |v|$. The result is trivial when $v = 0$.
- For each $u \in \mathbb{R}^n$ with $|u| = 1$, $|ABu| \le \|A\|_L |Bu| \le \|A\|_L \|B\|_L$. Hence, $\|AB\|_L = \|A\|\|B\|$.
- Given invertible $A \in L(\mathbb{R}^n.\mathbb{R}^n)$, $A^{-1} : \mathbb{R}^n \to \mathbb{R}^n$ is linear. Moreover, $\|A\|_L > 0$.

Lemma 2.2.2

Given two linear mappings $A, B : \mathbb{R}^n \to \mathbb{R}^n$ with invertibility of A,

$$\|A - B\|_L \cdot \|A^{-1}\|_L < 1 \implies B \text{ is invertible.}$$

Proof. Let $\|A^{-1}\|_L = 1/\alpha$ and $\|B - A\|_L = \beta$ so that $\beta < \alpha$. Then, for every $\mathbf{x} \in \mathbb{R}^n$,

$$\begin{aligned}\alpha|\mathbf{x}| &= \alpha|A^{-1}A\mathbf{x}| \le \alpha\|A^{-1}\| \cdot |A\mathbf{x}| \\ &= |A\mathbf{x}| \le |(A-B)\mathbf{x}| + |B\mathbf{x}| \le \beta|\mathbf{x}| + |B\mathbf{x}|;\end{aligned}$$

hence $(\alpha-\beta)|\mathbf{x}| \le |B\mathbf{x}|$ where $\mathbf{x} \in \mathbb{R}^n$ is arbitrary. As $\alpha > \beta$, it holds that $B\mathbf{x} = 0 \implies \mathbf{x} = 0$. □

Corollary 2.2.1

The set $\Omega \subseteq L(\mathbb{R}^n, \mathbb{R}^n)$ of invertible linear transformations is open.

Lemma 2.2.3

The mapping from Ω onto Ω defined by $A \mapsto A^{-1}$ is continuous.

Proof. Let A and B be invertible linear transformations from $\mathbb{R}^n$ to $\mathbb{R}^n$. Let $\|A^{-1}\| = 1/\alpha$ and $\|B-A\|_L = \beta$. We have $(\alpha-\beta)|\mathbf{x}| \le |B\mathbf{x}|$ by the same reasoning as in the proof of Lemma 2.2.2. Hence, the following holds.

$$\forall \mathbf{y} \in \mathbb{R}^n,\ (\alpha - \beta)|B^{-1}\mathbf{y}| \le |BB^{-1}\mathbf{y}| = |\mathbf{y}|$$

This shows that $\|B^{-1}\|_L \le (\alpha - \beta)^{-1}$.

Hence, we have

$$\|B^{-1} - A^{-1}\|_L \le \|B^{-1}\|_L \|A - B\|_L \|A^{-1}\|_L \le \frac{\beta}{\alpha(\alpha-\beta)}.$$

This implies that $\|B^{-1} - A^{-1}\|_L \to 0$ as $B \to A$. □

Theorem 2.2.1 Inverse Function Theorem

Let $\mathbf{f}$: $E(\subseteq \mathbb{R}^n) \to \mathbb{R}^n$ be C^1 in E and $\mathbf{c} \in E$. Suppose that $J_{\mathbf{f}}(\mathbf{c}) \ne 0$. Then, the following hold.

(i) There exists a neighborhood U of $\mathbf{a}$ such that $\mathbf{f}\big|_U$ is bijective and $V \triangleq \mathbf{f}(U)$ is open.

(ii) The inverse map of $\mathbf{f}\big|_U$ is C^1 in V.

Proof. Let $A \triangleq \mathrm{d}\mathbf{f}(\mathbf{c})$. Define $\lambda \in \mathbb{R}_+$ by $2\lambda\left\|A^{-1}\right\|_L = 1$. Since $\mathrm{d}\mathbf{f}$ is continuous, there exists a neighborhood U of $\mathbf{c}$ such that $\|\mathrm{d}\mathbf{f}(\mathbf{x}) - A\|_L < \lambda$ for each $\mathbf{x} \in U$.

Given a point $\mathbf{y} \in \mathbb{R}^n$, we define $\varphi(\cdot;\mathbf{y})$ by

$$\begin{aligned}\varphi(\cdot;\mathbf{y}) : B_\delta(\mathbf{c}) &\longrightarrow \mathbb{R}^n \\ \mathbf{x} &\longmapsto \mathbf{x} + A^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{x}))\end{aligned}$$

Note that $\mathbf{x}$ is a fixed point of $\varphi(\cdot;\mathbf{y})$ if and only if $A^{-1}(\mathbf{y} - \mathbf{f}(\mathbf{x})) = 0$, i.e., $\mathbf{y} = \mathbf{f}(\mathbf{x})$. Note also that φ is differentiable and $\mathrm{d}\varphi(\mathbf{x};\mathbf{y}) = \mathrm{Id} - A^{-1}\,\mathrm{d}\mathbf{f}(\mathbf{x}) = A^{-1}\big(A - \mathrm{d}\mathbf{f}(\mathbf{x})\big)$ for each $\mathbf{x} \in U$.

Hence, for all $\mathbf{x} \in U$,

$$\|\mathrm{d}\varphi(\mathbf{x};\mathbf{y})\|_L = \left\|A^{-1}\big(A - \mathrm{d}\mathbf{f}(\mathbf{x})\big)\right\|_L \le \|A^{-1}\|_L \cdot \|A - \mathrm{d}\mathbf{f}(\mathbf{x})\|_L < 1/(2\lambda) \cdot \lambda = 1/2.$$

Thus, MVT gives

$$|\varphi(\mathbf{x}_1;\mathbf{y})-\varphi(\mathbf{x}_2;\mathbf{y})|\le\frac{1}{2}|\mathbf{x}_1-\mathbf{x}_2|$$

whenever $\mathbf{x}_1,\mathbf{x}_2\in U$. Note that this implies there is at most one fixed point of $\varphi(\cdot;\mathbf{y})$ in U, i.e., $\mathbf{f}\big|_U$ is bijective.

Now, we shall show that $V=\mathbf{f}(U)$ is open. Take any $\mathbf{y}_0\in V$. There (uniquely) exists $\mathbf{x}_0\in U$ such that $\mathbf{y}_0=\mathbf{f}(\mathbf{x}_0)$. Fix any $r\in\mathbb{R}_+$ such that $\overline{B}\subseteq U$ where $B=B_r(\mathbf{x}_0)$. Take any $\mathbf{y}\in B_{\lambda r}(\mathbf{y}_0)$. Then,

$$|\varphi(\mathbf{x}_0;\mathbf{y})-\mathbf{x}_0|=|A^{-1}(\mathbf{y}-\mathbf{y}_0)|<\|A^{-1}\|_L\lambda r=\frac{r}{2}.$$

Moreover, for any $\mathbf{x}\in\overline{B}$,

$$|\varphi(\mathbf{x};\mathbf{y})-\mathbf{x}_0|\le|\varphi(\mathbf{x};\mathbf{y})-\varphi(\mathbf{x}_0;\mathbf{y})|+|\varphi(\mathbf{x}_0;\mathbf{y})-\mathbf{x}_0|\le\frac{1}{2}|\mathbf{x}-\mathbf{x}_0|+\frac{r}{2}<r.$$

This directly implies that $\varphi(\overline{B};\mathbf{y})\subseteq B\subseteq\overline{B}$. Hence, $\varphi(\cdot,\mathbf{y})$ is a contraction mapping on a complete metric space $\overline{B}$. By Lemma 2.2.1, there exists a fixed point $\mathbf{x}\in\overline{B}$, which satisfies $\mathbf{y}=\mathbf{f}(\mathbf{x})$. Thus, $\mathbf{y}\in f(\overline{B})\subseteq f(U)=V$. Hence, $B\subseteq V$, V is open. This proves (i).

Now, let $\mathbf{g}\colon V\to U$ be the local inverse map of $\mathbf{f}\big|_U$. Take any $\mathbf{y}\in V$ and $\mathbf{y}+\mathbf{k}\in V$. There are unique $\mathbf{x}\in U$ and $\mathbf{x}+\mathbf{h}\in U$ such that $\mathbf{y}=\mathbf{f}(\mathbf{x})$ and $\mathbf{y}+\mathbf{k}=\mathbf{f}(\mathbf{x}+\mathbf{h})$. Then, we have

$$\varphi(\mathbf{x}+\mathbf{h};\mathbf{y})-\varphi(\mathbf{x};\mathbf{y})=\mathbf{h}+A^{-1}\big(\mathbf{f}(\mathbf{x})-\mathbf{f}(\mathbf{x}+\mathbf{h})\big)=\mathbf{h}-A^{-1}\mathbf{k},$$

which implies $|\mathbf{h}-A^{-1}\mathbf{k}|\le|h|/2$. Hence, $|A^{-1}\mathbf{k}|\ge|h|/2$ is obtained by the triangle inequality; $|\mathbf{h}|\le2\|A^{-1}\|_L|\mathbf{k}|=\lambda^{-1}|\mathbf{k}|$.

Then, since $\|\,d\mathbf{f}(\mathbf{x})-A\|_L\|A^{-1}\|_L<\lambda\cdot1/(2\lambda)=1/2$, Lemma 2.2.2 implies that $d\mathbf{f}(\mathbf{x})$ is invertible. Let $T\triangleq d\mathbf{f}(\mathbf{x})$. Then, we have

$$\mathbf{g}(\mathbf{y}+\mathbf{k})-\mathbf{g}(\mathbf{y})-T^{-1}\mathbf{k}=\mathbf{h}-T^{-1}\mathbf{k}=-T^{-1}\big(\mathbf{f}(\mathbf{x}+\mathbf{h})-\mathbf{f}(\mathbf{x})-T\mathbf{h}\big),$$

and thus

$$\frac{|\mathbf{g}(\mathbf{y}+\mathbf{k})-\mathbf{g}(\mathbf{y})-T^{-1}\mathbf{k}|}{|\mathbf{k}|}\le\frac{\|T^{-1}\|_L}{\lambda}\cdot\frac{|\mathbf{f}(\mathbf{x}+\mathbf{h})-\mathbf{f}(\mathbf{x})-T\mathbf{h}|}{|\mathbf{h}|}.$$

The equation implies that $\mathbf{g}$ is differentiable on V, and that $d\mathbf{g}(\mathbf{y})=T^{-1}=d\mathbf{f}(\mathbf{g}(\mathbf{y}))^{-1}$. Since $d\mathbf{g}$ is a composition of continuous functions, $d\mathbf{g}$ itself is continuous. □

Corollary 2.2.2

Let $\mathbf{f}\colon E(\subseteq\mathbb{R}^n)\to\mathbb{R}^n$ be C^1 in E and $J_\mathbf{f}(\mathbf{x})\ne0$ for all $\mathbf{x}\in E$. Then, for every open set $W\subseteq E$, $\mathbf{f}(W)$ is open.

Proof. This directly follows from (i) of Theorem 2.2.1. □

2.3 Implicit Function Theorem

Definition 2.3.1

- If $\mathbf{x}=(x_1,\cdots,x_n)\in\mathbb{R}^n$ and $\mathbf{y}=(y_1,\cdots,y_m)\in\mathbb{R}^m$, let us write $(\mathbf{x},\mathbf{y})$ for the point $(x_1,\cdots,x_n,y_1,\cdots,y_m)\in\mathbb{R}^{n+m}$.
- Every $A\in L(\mathbb{R}^{n+m},\mathbb{R}^n)$ can be split into $A_x\in L(\mathbb{R}^n)$ and $A_y\in L(\mathbb{R}^m,\mathbb{R}^n)$ where $A(\mathbf{h},\mathbf{k})=A_x\mathbf{h}+A_y\mathbf{k}$ for each $\mathbf{h}\in\mathbb{R}^n$ and $\mathbf{k}\in\mathbb{R}^m$.

Lemma 2.3.1

If $A \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$ and if A_x is invertible, then

$$\forall \mathbf{k} \in \mathbb{R}^m,\ \exists! \mathbf{h} \in \mathbb{R}^n,\ A(\mathbf{h}, \mathbf{k}) = \mathbf{0}.$$

Proof. $A(\mathbf{h}, \mathbf{k}) = A_x \mathbf{h} + A_y \mathbf{k} = \mathbf{0}$ if and only if $\mathbf{h} = -(A_x)^{-1} A_y \mathbf{k}$. □

Theorem 2.3.1 Implicit Function Theorem

Let $\mathbf{f}\colon E \to \mathbb{R}^n$ be a C^1 mapping where E is an open set in $\mathbb{R}^{n+m}$. Let $(\mathbf{a}, \mathbf{b}) \in E$ satisfy $\mathbf{f}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$. Let $A = \mathrm{d}\mathbf{f}(\mathbf{a}, \mathbf{b})$ and suppose A_x is invertible. Then, there exist open sets $U \subseteq \mathbb{R}^{n+m}$ and $W \subseteq \mathbb{R}^m$ that satisfy the following.

(i) $(\mathbf{a}, \mathbf{b}) \in U$ and $\mathbf{b} \in W$.

(ii) $\forall \mathbf{y} \in W,\ \exists! \mathbf{x} \in \mathbb{R}^n,\ (\mathbf{x}, \mathbf{y}) \in U \wedge \mathbf{f}(\mathbf{x}, \mathbf{y}) = 0$.

(iii) If the unique $\mathbf{x}$ in (ii) is denoted by $\mathbf{g}(\mathbf{y})$, then $\mathbf{g}\colon W \to \mathbb{R}^n$ is C^1 on W.

(iv) Moreover, $\mathrm{d}\mathbf{g}(\mathbf{b}) = -(A_x)^{-1} A_y$.

Proof. Define $\mathbf{F}\colon E \to \mathbb{R}^{n+m}$ by $\mathbf{F}(\mathbf{x}, \mathbf{y}) \triangleq (\mathbf{f}(\mathbf{x}, \mathbf{y}), \mathbf{y})$. Then, $\mathbf{F}$ is C^1. Since $\mathbf{f}(\mathbf{a}, \mathbf{b}) = \mathbf{0}$, if $\mathbf{r}(\mathbf{h}, \mathbf{k}) \triangleq \mathbf{f}(\mathbf{a}+\mathbf{h}, \mathbf{b}+\mathbf{k}) - A(\mathbf{h}, \mathbf{k})$, we have $\lim_{\mathbf{h},\mathbf{k}\to\mathbf{0}} |\mathbf{r}(\mathbf{h}, \mathbf{k})|/|(\mathbf{h}, \mathbf{k})| = 0$. Hence, from

$$\mathbf{F}(\mathbf{a}+\mathbf{h}, \mathbf{b}+\mathbf{k}) - \mathbf{F}(\mathbf{a}, \mathbf{b}) = (\mathbf{f}(\mathbf{a}+\mathbf{h}, \mathbf{b}+\mathbf{k}), \mathbf{k}) = (A(\mathbf{h}, \mathbf{k}), \mathbf{k}) + (\mathbf{r}(\mathbf{h}, \mathbf{k}), \mathbf{0}),$$

it is obtained that $\mathrm{d}\mathbf{F}(\mathbf{a}, \mathbf{b})(\mathbf{h}', \mathbf{k}') = (A(\mathbf{h}', \mathbf{k}'), \mathbf{k}')$ for each $(\mathbf{h}', \mathbf{k}') \in \mathbb{R}^{n+m}$. If $\mathrm{d}\mathbf{F}(\mathbf{a}, \mathbf{b})(\mathbf{h}', \mathbf{k}') = \mathbf{0}$, then $\mathbf{k}' = 0$ and $A(\mathbf{h}', \mathbf{0}) = \mathbf{0}$; thus $\mathbf{h}' = \mathbf{0}$ as A_x is invertible. Hence, $\mathrm{d}\mathbf{F}(\mathbf{a}, \mathbf{b})$ is invertible; Theorem 2.2.1 can be applied to $\mathbf{F}$ at $(\mathbf{a}, \mathbf{b})$.

By Theorem 2.2.1, there exists a neighborhood $U \subseteq E$ of $(\mathbf{a}, \mathbf{b})$ such that $\mathbf{F}\big|_U$ is bijective, $\mathbf{F}(U)$ is open, and its inverse is C^1. Let $W \triangleq \{\mathbf{y} \in \mathbb{R}^m \mid (\mathbf{0}, \mathbf{y}) \in \mathbf{F}(U)\}$. W is open as $\mathbf{F}(U)$ is open. Noting that $\mathbf{b} \in W$, we finish the proof for (i).

Take any $\mathbf{y} \in W$. Then, there exists $(\mathbf{x}, \mathbf{y}) \in U$ such that $\mathbf{F}(\mathbf{x}, \mathbf{y}) = (\mathbf{0}, \mathbf{y})$; thus $\mathbf{f}(\mathbf{x}, \mathbf{y}) = \mathbf{0}$. If $\mathbf{x}, \mathbf{x}'$ are two such point corresponding to $\mathbf{y}$, then

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

However, as $\mathbf{F}$ being injective, $\mathbf{x} = \mathbf{x}'$. This proves (ii).

Let $V \triangleq \mathbf{F}(U)$. Let $\mathbf{G}\colon V \to U$ be the inverse of $\mathbf{F}$, which is C^1 by Theorem 2.2.1. Hence, for each $\mathbf{y} \in W$, from $\mathbf{F}(\mathbf{g}(\mathbf{y}), \mathbf{y}) = (\mathbf{0}, \mathbf{y})$, we have $(\mathbf{g}(\mathbf{y}), \mathbf{y}) = \mathbf{G}(\mathbf{0}, \mathbf{y})$. This directly shows that $\mathbf{g}$ is C^1 as well. This proves (iii).

Let $\mathbf{\Phi}\colon W \to U$ be defined by $\mathbf{\Phi}(\mathbf{y}) = \mathbf{G}(\mathbf{0}, \mathbf{y}) = (\mathbf{g}(\mathbf{y}), \mathbf{y})$, which is C^1, indeed. Then, $\mathrm{d}\mathbf{\Psi}(\mathbf{y}) = (\mathrm{d}\mathbf{g}(\mathbf{y}), I_m)$. Differentiating both sides of the equality $\mathbf{f}(\mathbf{\Phi}(\mathbf{y})) = \mathbf{0}$, we get

$$\mathrm{d}\mathbf{f}(\mathbf{\Phi}(\mathbf{y}))\,\mathrm{d}\mathbf{\Phi}(\mathbf{y}) = 0.$$

Putting $\mathbf{y} := \mathbf{b}$, as $\mathbf{\Phi}(\mathbf{b}) = (\mathbf{a}, \mathbf{b})$, we get $A\,\mathrm{d}\mathbf{\Phi}(\mathbf{b}) = 0$, or

$$A_x\,\mathrm{d}\mathbf{g}(\mathbf{b}) + A_y = 0,$$

i.e., $\mathrm{d}\mathbf{g}(\mathbf{b}) = -(A_x)^{-1} A_y$. □

Definition 2.3.2: C^1-norm

Suppose $\varphi : \mathbb{R}^n \to \mathbb{R}$ is C^1. Then,

$$\|\varphi\|_{C^0(\overline{\Omega})} \triangleq \sup_{\mathbf{x}\in\Omega} |\varphi(x)|$$

$$\|\varphi\|_{C^1(\overline{\Omega})} \triangleq \|\varphi\|_{C^0(\overline{\Omega})} + \sum_{j=1}^{n} \|\partial_j \varphi\|_{C^0(\overline{\Omega})}.$$

This is only for Example 2.3.1.

Example 2.3.1 (Level Sets)

Define $\Omega \triangleq \{(x_1,x_2) \in \mathbb{R}^2 \mid |x_2| \le 1\}$. Given two constants, $a, b \in \mathbb{R}$ with $a < b$, define $\overline{\varphi}(x_1,x_2) = ax_1$ and $\overline{\psi}(x_1,x_2) = bx_1$. Then, $\Gamma_0 = \{\mathbf{x} \in \Omega \mid \overline{\varphi}(\mathbf{x}) - \overline{\psi}(\mathbf{x}) = 0\} = \{\mathbf{x} \in \Omega \mid x_1 = 0\}$.

Suppose that $\varphi, \psi : \Omega \to \mathbb{R}$ satisfy

$$\|\varphi - \overline{\varphi}\|_{C^1(\overline{\Omega})} + \|\psi - \overline{\psi}\|_{C^1(\overline{\Omega})} \le \frac{1}{4}|a-b|.$$

Then, what would be the expression for $\Gamma = \{\mathbf{x} \in \Omega \mid \varphi(\mathbf{x}) - \psi(\mathbf{x}) = 0\}$?

Observe that $(\varphi - \psi) = (\varphi - \overline{\varphi}) + (\overline{\varphi} - \overline{\psi}) + (\overline{\psi} - \psi)$ and thus $|(\varphi-\psi)(x_1,x_2) - (a-b)x_1| \le |a-b|/4$. This implies $\lim_{x_1 \to \pm\infty} (\varphi - \psi)(x_1,x_2) = \mp\infty$. Hence, for every $x_2 \in [-1,1]$, there exists $x_1^* \in \mathbb{R}$ such that $(\varphi-\psi)(x_1^*, x_2) = 0$.

Moreover, $\partial_1(\varphi - \psi) = \partial_1(\varphi - \overline{\varphi}) + (a-b) + \partial_1(\overline{\psi} - \psi)$, and thus $|\partial_1(\varphi-\psi)| \ge \frac{3}{4}|a-b| > 0$. Hence, the x_1^* in the previous paragraph is unique. This means that $\Gamma = \{(f(x_2), x_2) \mid x_2 \in \mathbb{R}\}$ for some f.

$(\varphi-\psi)(f(x_2),x_2) - (\overline{\varphi} - \overline{\psi})(f(x_2),x_2) = -(\overline{\varphi} - \overline{\psi})(f(x_2),x_2) = (b-a)f(x_2)$. Hence,

$$f(x_2) = \frac{(\varphi - \overline{\varphi})(f(x_2),x_2) - (\psi - \overline{\psi})(f(x_2),x_2)}{b-a}.$$

This is the implicit representation of f. Moreover, $|f(x_2)| = \dfrac{|b-a|/4}{|b-a|} = 1/4$.

2.4 Applications of IMFT: Lagrange's Method

Theorem 2.4.1 Optimization Under Multiple Constraints

Let $f, g_1, g_2, \cdots, g_k : E \to \mathbb{R}$ be C^1 where E is an open set in $\mathbb{R}^n$ and $n > k$. Let $Z \triangleq \cap_{j=1}^{k} \{\mathbf{z} \in \mathbb{R}^n \mid g_j(\mathbf{z}) = 0\}$. Suppose $\mathbf{z}_0 \in Z$ is a local maximum point with respect to f on Z. Suppose also that

$$\Delta \triangleq \det \begin{bmatrix} \partial_1 g_1(\mathbf{z}_0) & \cdots & \partial_1 g_k(\mathbf{z}_0) \\ \vdots & \ddots & \vdots \\ \partial_k g_1(\mathbf{z}_0) & \cdots & \partial_k g_k(\mathbf{z}_0) \end{bmatrix} \neq 0.$$

Then, there exists $\lambda_1, \lambda_2, \cdots, \lambda_k \in \mathbb{R}$ such that $\nabla f(\mathbf{z}_0) = \sum_{m=1}^{k} \lambda_m \nabla g_m(\mathbf{z}_0)$.

Proof. Since $\Delta \neq 0$, there exists a unique solution $(\lambda_1, \cdots, \lambda_k) \in \mathbb{R}^k$ for the linear system

$$\begin{bmatrix} \partial_1 g_1(\mathbf{z}_0) & \cdots & \partial_1 g_k(\mathbf{z}_0) \\ \vdots & \ddots & \vdots \\ \partial_k g_1(\mathbf{z}_0) & \cdots & \partial_k g_k(\mathbf{z}_0) \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \vdots \\ \lambda_k \end{bmatrix} = \begin{bmatrix} \partial_1 f(\mathbf{z}_0) \\ \vdots \\ \partial_k f(\mathbf{z}_0) \end{bmatrix}.$$

For each point $\mathbf{z} = (z_1, \cdots, z_n) \in \mathbb{R}^n$, let $\mathbf{x} = (z_1, \cdots, z_k)$ and $\mathbf{y} = (z_{k+1}, \cdots, z_n)$. Let $\mathbf{z}_0 = (\mathbf{x}_0, \mathbf{y}_0)$. Let $\mathbf{g} \colon E \to \mathbb{R}^k$ be defined by $\mathbf{g}(\mathbf{z}) = (g_1(\mathbf{z}), \cdots, g_k(\mathbf{z}))$.

Since $\mathbf{g}$ is C^1, $\mathbf{g}(\mathbf{z}_0) = 0$, and $\left(\mathrm{d}\mathbf{g}(\mathbf{z}_0)\right)_x$ is invertible, by Theorem 2.3.1, there exists an open neighborhood $W \subseteq \mathbb{R}^{n-k}$ of $\mathbf{y}_0$ and a C^1 function $\mathbf{s} \colon W \to \mathbb{R}^k$ such that $\mathbf{g}(\mathbf{s}(\mathbf{y}), \mathbf{y}) = \mathbf{0}$ for each $\mathbf{y} \in W$. Note that $\mathbf{s}(\mathbf{y}_0) = \mathbf{x}_0$.

Define $F \colon W \to \mathbb{R}$ by $\mathbf{y} \mapsto f(\mathbf{s}(\mathbf{y}), \mathbf{y})$. As $\mathbf{z}_0$ is a local maximum point, so is $\mathbf{y}_0$. Hence, $\nabla F(\mathbf{y}_0) = \mathbf{0}$. For each $j \in [k]$, define $G_j \colon W \to \mathbb{R}$ by $\mathbf{y} \mapsto g_j(\mathbf{s}(\mathbf{y}), \mathbf{y})$. As $(\mathbf{s}(\mathbf{y}), \mathbf{y}) \in Z$, we have $G_j = 0$ for each $j \in [k]$. Thus, $\nabla G_j(\mathbf{y}) = \mathbf{0}$.

Let $\mathbf{s} = (s_1, s_2, \cdots, s_k)$ where each $s_j \colon W \to \mathbb{R}$. Since

$$\begin{aligned} \nabla F(\mathbf{y}) &= \mathrm{d}f(\mathbf{s}(\mathbf{y}), \mathbf{y})\, \mathrm{d}(\mathbf{s}(\mathbf{y}), \mathbf{y}) \\ &= \begin{bmatrix} \partial_1 f(\mathbf{s}(\mathbf{y}), \mathbf{y}) & \cdots & \partial_n f(\mathbf{s}(\mathbf{y}), \mathbf{y}) \end{bmatrix} \begin{bmatrix} \partial_1 s_1(\mathbf{y}) & \partial_2 s_1(\mathbf{y}) & \cdots & \partial_{n-k} s_1(\mathbf{y}) \\ \partial_1 s_2(\mathbf{y}) & \partial_2 s_2(\mathbf{y}) & \cdots & \partial_{n-k} s_2(\mathbf{y}) \\ \vdots & \vdots & \ddots & \vdots \\ \partial_1 s_k(\mathbf{y}) & \partial_2 s_k(\mathbf{y}) & \cdots & \partial_{n-k} s_k(\mathbf{y}) \\ 1 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & 1 \end{bmatrix}, \end{aligned}$$

$\nabla F(\mathbf{y}_0) = \mathbf{0}$ implies

$$\partial_{k+j} f(\mathbf{z}_0) + \sum_{i=1}^{k} \partial_i f(\mathbf{z}_0) \partial_j s_i(\mathbf{y}_0) = 0$$

for each $j \in [n-k]$. Similarly, $\nabla G_m(\mathbf{y}_0) = \mathbf{0}$ for each $m \in [k]$ implies that

$$-\lambda_m \left[\partial_{k+j} g_m(\mathbf{z}_0) + \sum_{i=1}^{k} \partial_i g_m(\mathbf{z}_0) \partial_j s_i(\mathbf{y}_0) \right] = 0$$

for each $j \in [n-k]$ and $m \in [k]$.

Adding the $k+1$ equations together for each $j \in [n-k]$,

$$0 = \left[\partial_{k+j} f(\mathbf{z}_0) - \sum_{m=1}^{k} \lambda_m \partial_{k+j} g_m(\mathbf{z}_0) \right] + \sum_{i=1}^{k} \left[\partial_i f(\mathbf{z}_0) - \sum_{m=1}^{k} \lambda_m \partial_i g_m(\mathbf{z}_0) \right] \partial_j s_i(\mathbf{y}_0).$$

By the definition of $\lambda_1, \cdots, \lambda_k$, we are left with only

$$\partial_j f(\mathbf{z}_0) = \sum_{m=1}^{k} \lambda_m \partial_j g_m(\mathbf{z}_0)$$

for each $j \in \{k+1, \cdots, n\}$. For $j \in [k]$, the same equation holds by the definition of $\lambda_1, \cdots, \lambda_k$. Hence, we have $\nabla f(\mathbf{z}_0) = \sum_{m=1}^{k} \lambda_m \nabla g_m(\mathbf{z}_0)$. $\square$

Chapter 3

Series of Vectors

3.1 Preliminaries

Definition 3.1.1: Normed Vector Space

Let V be a (real/complex) vector space equipped with a norm $||\cdot||$, i.e., the space $(V, ||\cdot||)$ satisfies the following properties.

(i) $0 \in V$

(ii) $||\mathbf{x}|| \geq 0$ for all $x \in V$ and $||\mathbf{x}|| = 0$ iff $\mathbf{x} = \mathbf{0}$. (*positive definiteness*)

(iii) $||\beta\mathbf{x}|| = |\beta| \cdot ||\mathbf{x}||$ for all $\mathbf{x} \in V$ and $\beta \in \mathbb{R}$. (*absolute homogeneity*)

(iv) $||\mathbf{x}_1 + \mathbf{x}_2|| \leq ||\mathbf{x}_1|| + ||\mathbf{x}_2||$ for all $\mathbf{x}_1, \mathbf{x}_2 \in V$. (*triangle inequality*)

Note:-

Note that $(V, ||\cdot||)$ is naturally a metric space with the metric function $d(\mathbf{x}_1, \mathbf{x}_2) = ||\mathbf{x}_1 - \mathbf{x}_2||$.

Definition 3.1.2: Banach Space

A normed vector space $(V, ||\cdot||)$ is called a *Banach space* if, for every Cauchy sequence $\{x_j\}_{j\in\mathbb{N}}$, there exists a unique $\mathbf{x}_* \in V$ such that $\lim_{n\to\infty} ||\mathbf{x}_n - \mathbf{x}_*|| = 0$.

Example 3.1.1

Let A be a compact subset of $\mathbb{R}^n$. $(V, ||\cdot||)$ where $V = \{f : A \to \mathbb{R} \mid f \text{ is continuous}\}$ and $||f|| = \sup_{x\in A} |f(x)|$ forms a Banach space.

Note:-

A Banach space is a normed vector space whose naturally induced metric space is complete.

Definition 3.1.3: Series

Let $(V,\|\cdot\|)$ be a normed vector space. Given a sequence $\{x_j\}_{j\in\mathbb{N}}\subseteq V$, define $S_k \triangleq \sum_{j=1}^{k} x_j$ for each $k\in\mathbb{N}$. Then, each S_k is called a *partial sum* of $\{x_j\}$. If $\{S_k\}_{k\in\mathbb{N}}$ converges to S_* with respect to $\|\cdot\|$, then we write

$$S_* = \sum_{j=1}^{\infty} x_j.$$

If the limit S_* exists, we symbolically say that "$\sum_{j=1}^{\infty} x_j$ converges."

Lemma 3.1.1

Let $(V,\|\cdot\|)$ be a normed vector space. Let $\{x_j\}_{j\in\mathbb{N}}\subseteq V$ be a sequence. If a series $\sum_{j=1}^{\infty} x_j$ converges, then $\lim_{k\to\infty}\|\mathbf{x}_k\| = 0$.

Proof. $\{S_k\}_{k\in\mathbb{N}}$ is a Cauchy sequence. Hence, $\lim_{k\to\infty}\|\mathbf{x}_k\| = \lim_{k\to\infty}\|S_{k+1}-S_k\| = 0$. □

Lemma 3.1.2

Let $(V,\|\cdot\|)$ be a Banach space. Let $\{x_j\}_{j\in\mathbb{N}}\subseteq V$ be a sequence. A series $\sum_{j=1}^{\infty} x_j$ converges if and only if $\{S_k\}_{k\in\mathbb{N}}$ is Cauchy.

Proof. The definition of Banach spaces. □

3.2 Finite Dimensional Banach Spaces

Theorem 3.2.1 Comparison Test

Given two real sequence $\{a_j\}$ and $\{b_j\}$, suppose $0\le a_j\le b_j$ for all $j\ge k_0$ where $k_0\in\mathbb{N}$ is a fixed constant. Then, if $\sum_{j=1}^{\infty} b_j$ converges, then $\sum_{j=1}^{\infty} a_j$ converges.

Proof. Let $S_k = \sum_{j=k_0}^{k} a_j$ and $T_k = \sum_{j=k_0}^{\infty} b_j$. Then, $0\le S_n - S_m = \sum_{j=m+1}^{n} a_j \le \sum_{j=m+1}^{n} b_j = T_n - T_m$ whenever $n\ge m\ge k_0$. As $\{T_k\}_{k\in\mathbb{N}}$ is Cauchy, $\{S_k\}_{k\in\mathbb{N}}$ is Cauchy as well. As $(\mathbb{R},\|\cdot\|)$ is a Banach space, $\sum a_j$ converges. □

Theorem 3.2.2 Absolute Convergence Test

Let $(V,\|\cdot\|)$ be a Banach space. Let $\{\mathbf{x}_j\}_{j\in\mathbb{N}}\subseteq V$ be a sequence. If $\sum_{j=1}^{\infty}\|\mathbf{x}_j\|$ converges (in $\mathbb{R}$), then $\sum_{j=1}^{\infty}\mathbf{x}_j$ converges.

Proof. Let $S_k = \sum_{j=1}^{k}\mathbf{x}_j \in V$ and $T_k = \sum_{j=1}^{k}\|\mathbf{x}_j\| \in \mathbb{R}$. Then, $\|S_n - S_m\| = \left\|\sum_{j=m+1}^{n}\mathbf{x}_j\right\| \le \sum_{j=m+1}^{n}\|\mathbf{x}_j\| = T_n - T_m$ whenever $n\ge m$. As $\{T_k\}$ is Cauchy, $\{S_k\}$ is Cauchy as well. Hence, $\sum \mathbf{x}_j$ converges. □

Theorem 3.2.3 Summation by Parts

Let $\{a_j\}$ and $\{b_j\}$ be two real sequences. If $\sum a_j$ converges and $\{b_j\}$ is monotonic and convergent, then $\sum_{j=1}^{\infty} a_j b_j$ converges.

Proof. Let $S_k = \sum_{j=1}^k a_j b_j \in V$ and $A_k = \sum_{j=1}^k a_j \in \mathbb{R}$. ($A_0 = 0$.) Then, $S_k = \sum_{j=1}^k (A_j - A_{j-1}) b_j = \sum_{j=1}^k A_j b_j - \sum_{j=0}^k A_0 b_{j+1} + A_k b_{k+1} = A_k b_{k+1} - \sum_{j=1}^k A_j (b_{j+1} - b_j)$.

Let $T_k = \sum_{j=1}^k |A_j (b_{j+1} - b_j)|$. Then, whenever $n < m$, we have

$$0 \leq T_m - T_n \leq M \sum_{j=n+1}^m |b_{j+1} - b_j| = M |b_{m+1} - b_{n+1}| \to 0,$$

$\{T_k\}$ is Cauchy, and thus converges; $\{S_k\}$ converges as well. □

3.3 Conditional Convergence

Definition 3.3.1: Conditional Convergence

Given a real sequence $\{a_j\}_{j\in\mathbb{N}} \subseteq \mathbb{R}$, if $\sum a_j$ converges, and if $\sum |a_j|$ does not converge, then we say that $\sum a_j$ *converges conditionally*.

Theorem 3.3.1 Alternating Series Test

Let $\{a_j\}_{j\in\mathbb{N}} \subseteq \mathbb{R}$ be a real sequence. If $a_j \geq 0$ for all $j \in \mathbb{N}$, and if $\lim_{j\to\infty} a_j = 0$, then $\sum (-1)^j a_j$ converges.

Proof. MAS101. □

Example 3.3.1

$\sum (-1)^j / j$ conditionally converges.

Note:-

Given, a real sequence $\{a_j\}$, we shall use the following definition for now.

For $j \in \mathbb{N}$, define

$$a_j^+ \triangleq \frac{|a_j| + a_j}{2} = \begin{cases} a_j & \text{if } a_j \geq 0 \\ 0 & \text{if } a_j < 0 \end{cases} \quad \text{and} \quad a_j^- \triangleq \frac{|a_j| - a_j}{2} = \begin{cases} 0 & \text{if } a_j \geq 0 \\ -a_j & \text{if } a_j < 0 \end{cases}.$$

Then, $a_j^+, a_j^- \geq 0$, $|a_j| = a_j^+ + a_j^-$, and $a_j = a_j^+ - a_j^-$.

Lemma 3.3.1

Let $\{a_j\}_{j\in\mathbb{N}}$ be a real sequence.

(i) If $\sum a_j$ converges absolutely, then both $\sum a_j^+$ and $\sum a_j^-$ converge. Moreover, $\sum a_j = \sum a_j^+ - \sum a_j^-$.

(ii) If $\sum a_j$ converges conditionally, then both $\sum a_j^+$ and $\sum a_j^-$ diverge.

Proof.

(i) By the definition of a_j^+ and a_j^-.

(ii) If one of $\sum a_j^+$ or $\sum a_j^-$ converges, since $a_j = a_j^+ - a_j^-$, the other converges as well. If they both converge, as $|a_j| = a_j^+ + a_j^-$, $\sum a_j$ converges absolutely.

□

Definition 3.3.2: Rearrangement of Series

Let $\phi : \mathbb{N} \to \mathbb{N}$ be bijective. Given a sequence $\{a_j\}_{j\in\mathbb{N}}$, the series $\sum a_{\phi(j)}$ is called a *rearrangement* of $\sum a_j$.

Theorem 3.3.2 Riemann's Rearrangement Theorem

Let $\{a_j\}_{j\in\mathbb{N}}$ be a conditionally convergent real sequence. Then, for any given $-\infty \le \alpha \le \beta \le \infty$ ($\pm\infty$ is allowed for α and β), there exists a rearrangement $\phi : \mathbb{N} \to \mathbb{N}$ such that $\liminf_{k\to\infty} \sum_{j=1}^{k} a_{\phi(j)} = \alpha$ and $\limsup_{k\to\infty} \sum_{j=1}^{k} a_{\phi(j)} = \beta$.

Proof. Let $\{P_j\}_{j\in\mathbb{N}}$ and $\{Q_j\}_{j\in\mathbb{N}}$ be nonnegative terms and absolute value of negative terms of $\{a_j\}_{j\in\mathbb{N}}$. Then, since they differ from $\{a_j^+\}$ and $\{a_j^-\}$ by zero terms, they are also divergent by Lemma 3.3.1.

Let $\{\alpha_\ell\}_{\ell\in\mathbb{N}}$ and $\{\beta_\ell\}_{\ell\in\mathbb{N}}$ be real sequences such that $\lim_{\ell\to\infty} \alpha_\ell = \alpha$ and $\lim_{\ell\to\infty} \beta_\ell = \beta$. Let $k_1, m_1 \in \mathbb{N}$ be the smallest integers such that

- $S_1 \triangleq P_1 + \cdots + P_{k_1} > \beta_1$ and
- $T_1 \triangleq S_1 - (Q_1 + \cdots + Q_{m_1}) < \alpha_1$.

Inductively, define $\{k_\ell\}_{\ell\in\mathbb{N}}$ and $\{m_\ell\}_{\ell\in\mathbb{N}}$ by

- $k_{\ell+1} \triangleq \min\left\{ k \in \mathbb{N}_{>k_\ell} \,\middle|\, T_\ell + \sum_{j=k_\ell+1}^{k} P_j > \beta_{\ell+1} \right\}$
- $S_{\ell+1} \triangleq T_\ell + \sum_{j=k_\ell+1}^{k_{\ell+1}} P_j$
- $m_{\ell+1} \triangleq \min\left\{ m \in \mathbb{N}_{>m_\ell} \,\middle|\, S_{\ell+1} - \sum_{j=m_\ell+1}^{m} Q_j < \alpha_{\ell+1} \right\}$
- $T_{\ell+1} \triangleq S_{\ell+1} - \sum_{j=m_\ell+1}^{m_{\ell+1}} Q_j$

for each $\ell \in \mathbb{N}$. As $k_\ell \to \infty$ and $m_\ell \to \infty$ as $\ell \to \infty$, this construction gives the natural rearrangement $\phi : \mathbb{N} \to \mathbb{N}$.

By the construction, we have $|S_\ell - \beta_\ell| \le P_{k_\ell}$ and $|T_\ell - \alpha_\ell| \le Q_{m_\ell}$ for each $\ell \in \mathbb{N}$. As $P_j, Q_j \to 0$ as $j \to \infty$, we have $S_\ell \to \beta$ and $T_\ell \to \alpha$ as $\ell \to \infty$; α and β are cluster points of $\left\{ \sum_{j=1}^{k} a_{\phi(j)} \right\}_{k\in\mathbb{N}}$ (as long as they are finite).

Moreover, for every sufficiently large $n \in \mathbb{N}$, we have $k_\ell + m_\ell \le n < k_{\ell+1} + m_{\ell+1}$ for some $\ell \in \mathbb{N}$, and thus $\min\{T_\ell, T_{\ell+1}\} \le \sum_{j=1}^{n} a_{\phi(j)} \le S_{\ell+1}$. This, or some more rigorous explanation using arbitrary $\varepsilon \in \mathbb{R}_+$, implies that there do not exist cluster points smaller than α or greater than β. □

3.4 The Cauchy Product

Definition 3.4.1: Cauchy Product

Given two real sequences $\{a_j\}_{j=0}^{\infty}$ and $\{b_j\}_{j=0}^{\infty}$, define

$$C_k \triangleq \sum_{j=0}^{k} a_j b_{k-j}.$$

The series $\sum_{k=1}^{\infty} C_k$ is called the *Cauchy product* of $\sum_{j=0}^{\infty} a_j$ and $\sum_{j=0}^{\infty} b_j$.

Theorem 3.4.1

Let $\{a_j\}_{j=0}^{\infty}$ and $\{b_j\}_{j=0}^{\infty}$ be two real sequences. Let $\sum_{k=0}^{\infty} C_k$ be the Cauchy product of them.

(i) If $\sum a_j$ converges absolutely, and if $\sum b_j$ converges, then $\sum C_k$ converges to $\left(\sum a_j\right)\left(\sum b_j\right)$.

(ii) If both $\sum a_j$ and $\sum b_j$ converge absolutely, $\sum C_k$ converges absolutely as well.

Proof. (ii) directly follows from the inequality $\sum_{k=0}^{n} |C_k| \leq \left(\sum_{j=0}^{n} |a_j|\right)\left(\sum_{j=0}^{n} |b_j|\right)$ as long as (i) is proven.

Let $S_n \triangleq \sum_{k=0}^{n} C_k$, $A_n \triangleq \sum_{j=0}^{n} a_j$, and $B_n \triangleq \sum_{j=0}^{n} b_j$. Let $B \triangleq \lim_{n\to\infty} B_n$ and $\mu_n \triangleq B_n - B$. Then,

$$\begin{aligned} S_n &= \sum_{k=0}^{n} C_k = \sum_{k=0}^{n}\sum_{j=0}^{k} b_{k-j} = \sum_{j=0}^{n} a_j \sum_{k=j}^{n} b_{k-j} \\ &= \sum_{j=0}^{n} a_j B_{n-j} = \sum_{j=0}^{n} a_j (B + \mu_{n-j}) = B\sum_{j=0}^{n} a_j + \sum_{j=0}^{n} a_j \mu_{n-j}. \end{aligned}$$

Claim. $\lim_{n\to\infty} \sum_{j=0}^{n} a_j \mu_{n-j} = 0$.

Take any $\varepsilon \in \mathbb{R}_+$ so there exists $N \in \mathbb{N}$ such that

- $|\mu_n| < \varepsilon$ for all $n \geq N$ (by $\mu_n \to 0$) and
- $\sum_{j=n+1}^{m} |a_j| < \varepsilon$ for all $m > n \geq N$ (by $\sum_{j=0}^{k} |a_j|$ being Cauchy).

As μ_n converges, there exists $\mu^* \triangleq \sup_{n\in\mathbb{N}} |\mu_n|$. Let $K_n \triangleq \sum_{j=0}^{n} a_j \mu_{n-j}$. Whenever $n > 2N$,

$$\begin{aligned} |K_n| &\leq \sum_{j=0}^{n} |a_j| \cdot |\mu_{n-j}| = \sum_{j=0}^{N-1} |a_j| \cdot |\mu_{n-j}| + \sum_{j=N}^{n} |a_j| \cdot |\mu_{n-j}| \\ &\leq \varepsilon \sum_{j=0}^{N-1} |a_j| + \mu^* \sum_{j=N}^{n} |a_j| \leq \varepsilon \left[\sum |a_j| + \mu^*\right]. \end{aligned}$$

Hence, $\lim_{n\to\infty} K_n = 0$; thus $\lim_{n\to\infty} S_n = \left(\sum a_j\right)\left(\sum b_j\right)$. ☐

3.5 Series on Infinite Dimensional Banach Spaces

Definition 3.5.1: Uniform Convergence of Series

Fix a domain $\Omega \subseteq \mathbb{R}^n$. Given a sequence $\{f_j : \Omega \to \mathbb{R}\}_{j\in\mathbb{N}}$, define $F_n : \Omega \to \mathbb{R}$ by

$$F_n(x) := \sum_{j=1}^{n} f_j(x)$$

for each $x \in \Omega$ and $n \in \mathbb{N}$.

(i) If $\lim_{n\to\infty} F_n(x)$ exists for all $x \in \Omega$, then the series $\sum_{j=1}^{\infty} f_j$ is said to *converge pointwise on* Ω.

(ii) Suppose $\sum_{j=1}^{\infty} f_j(x)$ converges pointwise on Ω and let $F(x) \triangleq \lim_{n\to\infty} F_n(x)$. The series $\sum_{j=1}^{\infty} f_j$ is said to *converge uniformly on* Ω if $\{F_n\}_{n=1}^{\infty}$ uniformly converges to F on Ω.

Theorem 3.5.1

If $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^\Omega$ is a sequence of continuous functions and converges uniformly, then $\lim_{n\to\infty} f_n$ is continuous as well.

Proof. MAS241. □

Definition 3.5.2: Uniform Cauchy

A sequence of function $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^\Omega$ is said to be *uniformly Cauchy on* Ω if

$$\forall \varepsilon \in \mathbb{R}_+,\ \exists N_* \in \mathbb{N},\ \forall n, m \geq N_*,\ \forall x \in \Omega,\ |f_n(x) - f_m(x)| < \varepsilon.$$

Lemma 3.5.1

A sequence of function $\{f_n\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^\Omega$ uniformly converges on Ω if and only if $\{f_n\}_{n\in\mathbb{N}}$ is uniformly Cauchy on Ω.

Proof. ($\Rightarrow$) Let $f(x) = \lim_{n\to\infty} f_n(x)$. Take any $\varepsilon \in \mathbb{R}_+$. Then, there exists $N_* \in \mathbb{N}$ such that, if $n \geq N_*$, then $|f_n(x) - f(x)| < \varepsilon/2$ for all $x \in \Omega$. Consequently, whenever $n, m \geq N_*$, $|f_n(x) - f_m(x)| \leq |f_n(x) - f(x)| + |f(x) - f_m(x)| < \varepsilon$.

($\Leftarrow$) For each $x \in \mathbb{R}$, the sequence $\{f_n(x)\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ is Cauchy. As $(\mathbb{R}, |\cdot|)$ is a Banach space, there uniquely exists the limit $f \triangleq \lim_{n\to\infty} f_n$. Take any $\varepsilon \in \mathbb{R}_+$. Then, there exists $N_* \in \mathbb{N}$ such that $|f_m(x) - f_n(x)| < \varepsilon/2$ for all $n, m \geq N_*$ and $x \in \Omega$. From this, we get $f_n(x) - \varepsilon/2 \leq \lim_{m\to\infty} f_m(x) = f(x) \leq f_n(x) + \varepsilon/2$. Hence, $|f_n(x) - f(x)| \leq \varepsilon/2 < \varepsilon$ holds for all $n \geq N_*$ and $x \in \Omega$.

□

Note:-

Lemma 3.5.1 holds for arbitrary sequence of functions from Ω to any Banach space.

Lemma 3.5.2

Let $\{f_j\}_{j\in\mathbb{N}} \subseteq \mathbb{R}^\Omega$ be a series of continuous functions. If $\sum_{j=1}^{\infty} f_j$ converges uniformly on Ω, then $\sum_{j=1}^{\infty} f_j$ is continuous on Ω.

Proof. Lemma 3.5.1. □

Chapter 4

Analysis for Series Functions

4.1 Calculus of Series Functions

Theorem 4.1.1

Given a sequence $\{f_j\}_{j\in\mathbb{N}} \subseteq \mathbb{R}^{(a,b)}$ of differentiable functions, suppose the following.

(i) $\{f_j(x_0)\}_{j\in\mathbb{N}} \subseteq \mathbb{R}$ converges for some $x_0 \in (a,b)$.

(ii) $\{f_j'\}_{j\in\mathbb{N}} \subseteq \mathbb{R}^{(a,b)}$ uniformly converges on (a,b).

Then, $f_j \rightrightarrows f$ for some $f : (a,b) \to \mathbb{R}$ on (a,b). Furthermore, f is differentiable on (a,b) and $\forall x \in (a,b)$, $f'(x) = \lim_{j\to\infty} f_j'(x)$.

Proof. We shall first show the uniform convergence of $\{f_j\}$. Take any $\varepsilon \in \mathbb{R}_+$. Then, there exists $N \in \mathbb{N}$ such that, for all $j,k \geq N$,

$$\big(|f_j(x_0) - f_k(x_0)| < \varepsilon/2\big) \wedge \big(\forall x \in (a,b),\ |f_j'(x) - f_k'(x)| < \varepsilon/2(b-a)\big).$$

By MVT, for all $x, \tilde{x} \in (a,b)$ with $x \neq \tilde{x}$, there exists $x_* \in (a,b)$ such that

$$(f_j - f_k)(x) - (f_j - f_k)(\tilde{x}) = (f_j - f_k)'(x_*) \cdot (x - \tilde{x})$$

Hence, $|(f_j - f_k)(x) - (f_j - f_k)(\tilde{x})| < \varepsilon/2$. Therefore, $|(f_j - f_k)(x)| < \varepsilon$ by triangle inequality obtained by setting $\tilde{x} = x_0$. This directly implies that $\{f_j\}$ is uniformly Cauchy and thus uniformly converges by Lemma 3.5.1. ✓

Let $f_j \to f$. Fixing $x \in (a,b)$, define

$$\psi_j(t) \triangleq \frac{f_j(t) - f_j(x)}{t - x} \quad \text{and} \quad \psi(t) \triangleq \frac{f(t) - f(x)}{t - x}$$

for $t \in (a,b)$ and $t \neq x$. Now, we claim that $\{\psi_j\}_{j\in\mathbb{N}}$ is uniformly Cauchy. Take any $\varepsilon \in \mathbb{R}_+$. Then, for $j,k \geq N$,

$$|\psi_j(t) - \psi_k(t)| = \left|\frac{(f_j - f_k)(t) - (f_j - f_k)(x)}{t - x}\right| < \frac{\varepsilon}{2(b-a)}.$$

Hence, $\{\psi_j\}$ uniformly converges by Lemma 3.5.1, and $\psi_j \to \psi$ as $f_j \to f$.

Let $A_j \triangleq \lim_{t\to x} \psi_j(t) = f_j'(x)$. By the supposition (ii), we have convergence of $\{A_j\}_{j\in\mathbb{N}}$. Now, we claim that $\lim_{t\to x} \psi(t) = \lim_{j\to\infty} A_j$. Let $A_j \to A$. Take any $\varepsilon \in \mathbb{R}_+$. There exists $N' \in \mathbb{N}$ such that, if $j \geq N'$, we have $|\psi(t) - \psi_j(t)| < \varepsilon/3$ for all $t \in (a,b) \setminus \{x\}$ ana $|A_j - A| < \varepsilon/3$. In addition, from the definition of A_j, there exists $\delta \in \mathbb{R}_+$ such that, whenever $0 < |t-x| < \delta$, we have $|\psi_{N'}(t) - A_{N'}| < \varepsilon/3$. Now, we have

$$|\psi(t) - A| \le |\psi(t) - \psi_{N'}(t)| + |\psi_{N'}(t) - A_{N'}| + |A_{N'} - A| < \varepsilon$$

for $0 < |t-x| < \delta$. Hence, $f'(x) = \lim_{t\to x} \psi(t) = \lim_{j\to\infty} f_j'(x)$. □

Corollary 4.1.1 Term-by-Term Differentiation

Given a sequence $\{f_j\}_{j\in\mathbb{N}} \subseteq \mathbb{R}^{(a,b)}$ of differentiable functions, let $F_n = \sum_{j=1}^{n} f_j$. Suppose the following.

(i) $\{F_n(x_0)\}_{n\in\mathbb{N}} \subseteq \mathbb{R}$ converges for some $x_0 \in (a,b)$.
(ii) $\{F_n'\}_{n\in\mathbb{N}} \subseteq \mathbb{R}^{(a,b)}$ uniformly converges on (a,b).

Then, $\{F_n\}$ converges uniformly to a function $F\colon (a,b) \to \mathbb{R}$ on (a,b). Furthermore, F is differentiable on (a,b) and $\forall x \in (a,b)$, $F'(x) = \sum_{j=1}^{\infty} f_j'(x)$.

Example 4.1.1

Let $f_j(x) = \sin(x/j^2)$ for $-1 < x < 1$ and $F_n = \sum_{j=1}^{n} f_j$.

For $x_0 = 0$, the sequence $\{F_n(x_0)\}_{n\in\mathbb{N}}$ converges (to zero). Now, we have $F_n'(x) = \sum_{j=1}^{n} \cos(x/j^2)/j^2$. Then, for $n,m \in \mathbb{N}$ with $m \ge n$, $|F_m'(x) - F_n'(x)| \le \sum_{j=n+1}^{m} 1/j^2 \to 0$ as $n,m \to \infty$. Hence, $\{F_n'\}$ is uniformly Cauchy; and thus it converges uniformly by Lemma 3.5.1. Hence, Corollary 4.1.1 guarantees the uniform convergence and differentiability of $\sum_{j=1}^{\infty} f_j$.

Theorem 4.1.2

Given a sequence $\{f_j\}_{j\in\mathbb{N}} \subseteq \mathbb{R}^{(a,b)}$ of functions Riemann integrable on (a,b), if $f_j \rightrightarrows f$ on (a,b), then f is Riemann integrable on (a,b). Furthermore, $\int_a^b f(x)\,\mathrm{d}x = \lim_{j\to\infty} \int_a^b f_j(x)\,\mathrm{d}x$.

Proof. □

Corollary 4.1.2 Term-by-Term Integration

Given a sequence $\{f_j\}_{j\in\mathbb{N}} \subseteq \mathbb{R}^{(a,b)}$ of functions Riemann integrable on (a,b), suppose $\sum f_j \rightrightarrows F$ for some $F\colon (a,b) \to \mathbb{R}$. Then, $\int_a^b F(x)\,\mathrm{d}x = \lim_{n\to\infty} \int_a^b \sum_{j=1}^{n} f_j(x)\,\mathrm{d}x$.

End.