## Lecture 12: Infinite Series. Operator Functions. Neumann Series.

**Definition 3.** *Let $X$ be a normed space over $\mathbb{K}$ and let $u_j \in X$ for all $j$. If $\lim_{m\to\infty} \sum_{j=0}^{m} u_j$ exists, denote*

$$X \ni \sum_{j=0}^{\infty} u_j := \lim_{m\to\infty} \sum_{j=0}^{m} u_j,$$

*and the infinite series $\sum_{j=0}^{\infty} u_j$ is called convergent. This infinite series is called absolutely convergent iff*

$$\sum_{j=0}^{\infty} \|u_j\| < \infty.$$

**Proposition 18.** *A normed linear space $X$ is a Banach space if and only if every absolutely convergent infinite series with terms in $X$ is convergent.*

*Proof.* ($\Rightarrow$) Suppose $X$ is a Banach space. Let $\sum_{j=0}^{\infty} u_j$ be an absolutely convergent infinite series in $X$. Then for every $\varepsilon > 0$, there exists $N_\varepsilon > 0$ such that for every $n > N_\varepsilon$, $k \geq 0$, we have

$$\sum_{j=n+1}^{n+k} \|u_j\| < \varepsilon.$$

Denote $s_m = \sum_{j=0}^{m} u_j \in X$, then for every $n > N_\varepsilon$, $k \geq 0$, we have

$$\|s_{n+k} - s_n\| = \|\sum_{j=n+1}^{n+k} u_j\| \leq \sum_{j=n+1}^{n+k} \|u_j\| < \varepsilon.$$

Hence the sequence $\{s_n\}$ is a Cauchy sequence in $X$. Since $X$ is a Banach space, the limit $\lim_{n\to\infty} s_n$ exists. Therefore, the infinite series $\sum_{j=0}^{\infty} u_j$ converges.

($\Leftarrow$) Suppose every absolutely convergent infinite series with terms in $X$ is convergent. We need to prove that $X$ is a Banach space. Here is the sketch of the proof. Let $\{x_n\} \subset X$ be a Cauchy sequence.

- Construct a subsequence $\{x_{n_k}\}$ so that $\|x_{n_k} - x_{n_{k-1}}\| \leq \frac{1}{2^k}$ for all $k \geq 1$.
- Prove that the series $\left(x_{n_0} + \sum_{k=0}^{\infty} (x_{n_{k+1}} - x_{n_k})\right)$ is absolutely convergent, hence it is convergent.
- Therefore $\lim_{m\to\infty} x_{n_{m+1}} = \lim_{m\to\infty} \left(x_{n_1} + \sum_{k=0}^{m} (x_{n_{k+1}} - x_{n_k})\right)$ exists. Denote $x = \lim_{m\to\infty} x_{n_m}$.
- Combining with the assumption that the sequence $\{x_n\} \subset X$ is a Cauchy sequence, prove that $\lim_{j\to\infty} x_j = x$.

□

**Theorem 3** (Theorem and Definition)**.** *Let $X$ be a Banach space over $\mathbb{K}$ (where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$) and a series*

$$F(z) := \sum_{j=0}^{\infty} a_j z^j, \ z \in \mathbb{K}, \quad a_j \in \mathbb{K} \text{ for all } j$$

*such that*

$$\sum_{j=0}^{\infty} |a_j|\,|z|^j < \infty \quad \text{for all } \ z \in \mathbb{C} \text{ with } \ |z| < r \quad \text{and some fixed } r > 0.$$

*Then for each $A \in B(X)$ with $\|A\| < r$, the series $\sum_{j=0}^{\infty} a_j A^j$ is also an element in $B(X)$.*

*Proof.* Let $A \in B(X)$ with $\|A\| < r$. From the assumption on the series, we have $\sum_{j=0}^{\infty} |a_j|\,\|A\|^j < \infty$. For every $j \geq 1$, we have

$$\|a_j A^j\| \leq |a_j|\|A\|^j.$$

Therefore, by the comparison test, the series $\sum_{j=0}^{\infty} a_j A^j$ is absolutely convergent. Since $B(X)$ is a Banach space, the series $\sum_{j=0}^{\infty} a_j A^j$ is a convergent series. That is $\sum_{j=0}^{\infty} a_j A^j \in B(X)$. □

**Definition 4.** *Let $A : X \to Y$ and $B : Y \to X$ be linear operators, where $X$ and $Y$ are linear vector spaces over $\mathbb{K}$. If $AB = I_Y$ and $BA = I_X$, $A$ is said to be bijective and denote $A^{-1} = B$.*

**Example 1.** *Let $X \neq \{0\}$ be a Banach space over $\mathbb{K}$.*

1. *Exponential Function. For each $A \in B(X)$, the infinite series $\sum_{j=0}^{\infty} \frac{1}{j!} A^j$ is also an element in $B(X)$. Denote*

$$e^A := \sum_{j=0}^{\infty} \frac{1}{j!} A^j \in B(X).$$

   *Moreover, for all $t, s \in \mathbb{K}$, we have*

$$e^{tA}\, e^{sA} = e^{(t+s)A}.$$

2. *Neumann Series. Let $A \in B(X)$ with $\|A\| < 1$. Then the following statements hold.*

   *(a) The infinite series $\sum_{j=0}^{\infty} A^j$ is also an element in $B(X)$. The series $\sum_{j=0}^{\infty} A^j$ is called the Neumann series.*

   *(b) The operator $(I - A) \in B(X)$ is bijective and $(I - A)^{-1} = \sum_{j=0}^{\infty} A^j$.*

   *(c) $\|(I - A)^{-1}\| \leq \dfrac{1}{1 - \|A\|}$.*

   *(d) Given $g \in X$, the equation $(I - A)u = g$ with the unknown $u \in X$ has a unique solution*

$$u = (I - A)^{-1} g = \sum_{j=0}^{\infty} A^j g.$$

*Moreover, $u$ can be approximated by*

$$u_n = g + Ag + A^2 g + \ldots + A^{n-1} g$$

*with the error*

$$\|u - u_n\| \le \frac{\|A\|^n}{1 - \|A\|}\|g\|, \quad \textit{for all} \;\; n = 1, 2, \ldots$$

*Proof.* Hint: Using Theorem 3 for $F(z) = \sum_{j=0}^{\infty} \frac{1}{j!} z^j = e^z$ and $F(z) = \sum_{j=0}^{\infty} z^j = \frac{1}{1-z}$, respectively.

For Example 2, let's verify $\sum_{j=0}^{\infty} A^j = (I - A)^{-1}$ and $\|(I-A)^{-1}\| \le \frac{1}{1 - \|A\|}$. Obviously,

$$(I - A)\left(\sum_{j=0}^{\infty} A^j\right) = I \quad \text{and} \quad \left(\sum_{j=0}^{\infty} A^j\right)(I - A) = I.$$

Hence $\sum_{j=0}^{\infty} A^j = (I - A)^{-1}$. From the inequality

$$\left\|\sum_{j=0}^{m} A^j\right\| \le \sum_{j=0}^{m} \|A\|^j \le \sum_{j=0}^{\infty} \|A\|^j = \frac{1}{1 - \|A\|},$$

letting $m \to \infty$, we have

$$\|(I - A)^{-1}\| = \left\|\sum_{j=0}^{\infty} A^j\right\| \le \frac{1}{1 - \|A\|}.$$

(d). Suppose $u$ is a solution for $(I - A)u = g$. Then

$$(I - A)^{-1} g = (I - A)^{-1}(I - A)u = Iu = u.$$

Finally, since

$$u - u_n = A^n g + A^{n+1} g + \cdots = A^n (I + A + \cdots) g = A^n (I - A)^{-1} g,$$

we have

$$\|u - u_n\| \le \|A^n\| \|(I - A)^{-1}\| \|g\| \le \frac{\|A\|^n}{1 - \|A\|}\|g\|.$$

Note: $u_n$ is the iteration generated from the Banach fixed point theorem for $u = Tu$ with $Tu = Au + g$ and the Lipschitz constant is $\|A\|$:

$$\|Tu - Tv\| = \|Au - Av\| = \|A(u - v)\| \le \|A\| \|u - v\|.$$

□

### 2.13 Fréchet Derivative

**Definition 1** (Definition and Theorem)**.** *Let $X$ and $Y$ be normed linear spaces over $\mathbb{K}$ (where $\mathbb{K} = \mathbb{R}$ or $\mathbb{K} = \mathbb{C}$). An operator $F : X \to Y$ is Fréchet differentiable (F-differentiable) at $a \in X$ if and only if there exists a bounded linear operator $DF(a) : X \to Y$ such that*

$$\lim_{h \to 0} \frac{\|F(a+h) - F(a) - DF(a)(h)\|}{\|h\|} = 0. \tag{11}$$

*An operator $DF(a)$ (if exists, i.e., $DF(a) \in B(X,Y)$ and $DF(a)$ satisfies Equation (11)) is unique and is called the Fréchet-derivative of $F$ at $a$.*

*Proof.* Suppose $F : X \to Y$ is Fréchet differentiable at $a \in X$ and there are two bounded linear operators $L_1, L_2 : X \to Y$ such that

$$\lim_{h \to 0} \frac{\|F(a+h) - F(a) - L_1(h)\|}{\|h\|} = 0 \quad \text{and} \quad \lim_{h \to 0} \frac{\|F(a+h) - F(a) - L_2(h)\|}{\|h\|} = 0.$$

Combining with

$$0 \le \frac{\|L_1(h) - L_2(h)\|}{\|h\|} \le \frac{\|F(a+h) - F(a) - L_2(h)\| + \|-(F(a+h) - F(a) - L_1(h))\|}{\|h\|},$$

we have

$$\lim_{h \to 0} \frac{\|L_1(h) - L_2(h)\|}{\|h\|} = 0.$$

Let $L = L_1 - L_2 \in B(X,Y)$, then $\lim_{h \to 0} \frac{\|L(h)\|}{\|h\|} = 0$. We will show that $L(x) = 0$ for all $x \in X$. Since $L$ is linear, $L(0) = 0$. Fix $x \in X, x \neq 0$. For $t \in \mathbb{K}$, if $t \to 0$ then $tx \in X$ and $tx \to 0$. Therefore,

$$0 = \lim_{t \to 0} \frac{\|L(tx)\|}{\|tx\|} = \lim_{t \to 0} \frac{|t|\|L(x)\|}{|t|\|x\|} = \frac{\|L(x)\|}{\|x\|}.$$

The second equality holds because $L$ is linear. Therefore, $\|L(x)\| = 0\|x\| = 0$, so $L(x) = 0$ for all $x \in X \setminus \{0\}$. Hence $L(x) = x$ for all $x \in X$. In other words, an operator $DF(a)$ (if exists) is unique. □

**Example 1.** *1. If $F \in B(X,Y)$ then $F$ is F-differentiable everywhere and $DF(a) = F$ for all $a \in X$.*

*2. Let $F : \mathbb{R}^n \to \mathbb{R}$ and suppose $F \in C^1(\mathbb{R}^n)$ (i.e., $\partial_i f$ exists and is continuous on $\mathbb{R}^n$, $1 \le i \le n$). Then $DF(a) \in B(\mathbb{R}^n, \mathbb{R})$ is defined by*

$$DF(a)(h) := \nabla f(a) \cdot h. \quad \textit{(dot product)}$$

*3. Let $F : \mathbb{R}^n \to \mathbb{R}^m$ and suppose $F \in C^1(\mathbb{R}^n, \mathbb{R}^m)$ (i.e., $\frac{\partial f_j}{x_i}$ exists and is continuous, $1 \le i \le n$, $1 \le j \le m$). Then $F$ is F-differentiable and*

$$DF(a)(h) := \begin{bmatrix} \frac{\partial f_1}{\partial x_1}(a) & \cdots & \frac{\partial f_1}{\partial x_n}(a) \\ \vdots & \vdots & \vdots \\ \frac{\partial f_m}{\partial x_1}(a) & \cdots & \frac{\partial f_m}{\partial x_n}(a) \end{bmatrix} h. \quad \textit{(matrix multiplication)}$$

*The matrix itself is the usual Jacobian matrix.*