Functional Analysis Notes
Linear Operator and Linear Functional

Zheng Xie

June 2, 2018

1 Linear Operator

Definition 1.1 (linear operator)
Let $\mathcal{X}, \mathcal{Y}$ be two vector space, D is a subspace of $\mathcal{X}$, $T : D \to \mathcal{Y}$ is a map, D is called the **domain** of T, sometimes written as $D(T)$, $R(T) = \{Tx|x \in D\}$ is called the **range** of T. If

$$T(\alpha x + \beta y) = \alpha Tx + \beta Ty, \forall x, y \in D, \forall \alpha, \beta \in \mathbb{K}$$

Then T is called a **linear operator**.

Definition 1.2 (linear functional)
Let f be a linear operator, if f is real-valued or complex-valued, then f is called a **linear functional**, written as $f(x)$ or $\langle f, x\rangle$.

Definition 1.3 (continuous)
Let $\mathcal{X}, \mathcal{Y}$ be a normed vector space, $T : D(T) \to \mathcal{Y}$ is a linear operator. T is called **continuous** at $x_0 \in D(T)$, if

$$\{x_n\} \subset D(T), x_n \to x_0 \Rightarrow Tx_n \to Tx_0$$

Proposition 1.1
Let T be a linear operator, then T is continuous in $D(T)$ if and only if T is continuous at $x = 0$.
Proof.
Suppose that T is continuous at 0, then for any $\{x_n\} \subset D(T)$, $x_0 \in D(T)$, $x_n \to x_0$, we have

$$T(x_n - x_0) \to T0 = 0$$

Thus

$$Tx_n \to Tx_0 \quad \square$$

Definition 1.4 (bounded)
Let $\mathcal{X}, \mathcal{Y}$ be normed vector spaces, linear operator $T : \mathcal{X} \to \mathcal{Y}$ is called **bounded** if there exists a constant $M \geq 0$ such that $\forall x \in \mathcal{X}$

$$\|Tx\|_{\mathcal{Y}} \leq M\|x\|_{\mathcal{X}}$$

Proposition 1.2
Let $T : \mathcal{X} \to \mathcal{Y}$ be a linear operator, then T is continuous if and only if T is bounded.
Proof.
$\Rightarrow$:
Suppose not, then we can obtain a sequence in $\mathcal{X}$ such that for each n,

$$\|Tx_n\|_{\mathcal{Y}} > n\|x_n\|_{\mathcal{X}}$$

Let $y_n = \frac{x_n}{n\|x_n\|}$ then $y_n \to 0$ while $\|Ty_n\| > 1$, leading a contradiction.
$\Leftarrow$:
If T is bounded then it's easy to see that T is continuous at $x = 0$ so T is continuous. $\square$

Definition 1.5
Let $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ denote the set of all bounded linear operators from $\mathcal{X}$ to $\mathcal{Y}$. Particularly, we denote $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ by $\mathcal{L}(\mathcal{X})$ simply and denote $\mathcal{L}(\mathcal{X}, \mathbb{K})$ by X^*.

Definition 1.6 (norm)
Let $T : \mathcal{X} \to \mathcal{Y}$ be a linear operator then the norm of T is defined by

$$\|T\| = \sup_{x \in \mathcal{X} \setminus \{0\}} \frac{\|Tx\|}{\|x\|} = \sup_{\|x\|=1} \|Tx\|$$

Theorem 1.1
Let $\mathcal{X}$ be a normed vector space, $\mathcal{Y}$ be a Banach space, then $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ equipped with the linear operations

$$(\alpha_1 T_1 + \alpha_2 T_2)(x) = \alpha_1 T_1 x + \alpha_2 T_2 x \quad \forall x \in \mathcal{X}$$

and the norm $\|\cdot\|$ is a Banach space.
Proof.
It's clear that $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ is a vector space. We first show that $\|\cdot\|$ is a norm.
(1) $\|T\| \geq 0$, $\|T\| = 0$ if and only if $\forall x \in \mathcal{X}$, $Tx = 0$, i.e. $T = 0$.
(2) $\|\alpha T\| = \sup_{\|x\|=1} \|\alpha Tx\| = |\alpha| \sup_{\|x\|=1} \|Tx\| = |\alpha|\|T\|$.
(3)

$$\begin{aligned}\|T_1 + T_2\| &= \sup_{\|x\|=1} \|T_1 x + T_2 x\| \\ &\leq \sup_{\|x\|=1} \|T_1 x\| + \sup_{\|x\|=1} \|T_2 x\| = \|T_1\| + \|T_2\|\end{aligned}$$

Then we show that $\mathcal{X}$ is complete.
Let $\{T_n\}$ be a Cauchy sequence, then for any $\epsilon > 0$, there exists N such that for any $n > N$ and any $p \in \mathbb{N}$ we have

$$\|T_{n+p} - T_n\| < \epsilon$$

i.e. for all $x \in \mathcal{X}$

$$\|(T_{n+p} - T_n)(x)\| \leq \epsilon \|x\|$$

Fix x, then $\{T_n x\}$ is a Cauchy sequence, since $\mathcal{Y}$ is complete, $T_n x \to y \in \mathcal{Y}$, define T by $y = Tx$, then $T_n \to T$.
It remains to show that $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. It's easy to see that T is linear.
For any $x \in \mathcal{X}$ with $\|x\| = 1$, there exists N such that for all $n > N$

$$\|Tx - T_n x\| = \|y - T_n x\| \leq 1$$

So

$$\|Tx\| \leq \|T_n x\| + 1 \leq (\|T_n\| + 1)\|x\|$$

Hence $\|T\| \leq \|T_n\| + 1$, i.e. T is bounded and therefore continuous. $\square$

Proposition 1.3
If $T : \mathcal{X} \to \mathcal{Y}$ is a linear operator where $\mathcal{X}$ is a finite-dimensional normed vector space, then T is continuous.
Proof.
Note that T can be represented by a matrix (t_{ij}), and any two norms of a finite-dimensional vector space are

equivalent. Suppose that $\mathcal{X} = \mathbb{K}^n$, $\mathcal{Y} = \mathbb{K}^m$, then we have

$$\begin{aligned}\|Tx\| &= (\sum_{i=1}^{m}|\sum_{j=1}^{n} t_{ij}x_j|^2)^{\frac{1}{2}} \\ &\leq (\sum_{i=1}^{m}(\sum_{j=1}^{n}|t_{ij}|^2 \cdot \sum_{j=1}^{n}|x_j|^2))^{\frac{1}{2}} \\ &= (\sum_{i=1}^{m}\sum_{j=1}^{n}|t_{ij}|^2)^{\frac{1}{2}}\|x\|\end{aligned}$$

Hence T is bounded and therefore continuous. $\square$

2 Riesz Representation Theorem

Let $\mathcal{X}$ be a Hilbert space, $\forall y \in \mathcal{X}$, if we define

$$f_y : x \mapsto \langle x, y\rangle, \forall x \in \mathcal{X}$$

Then it's easy to see that $f_y \in \mathcal{X}^*$.
Moreover,

$$|f_y(x)| \leq \|y\|\|x\|$$

It follows that $\|f_y\| \leq \|y\|$. Since $|f_y(y)| = \langle y, y\rangle = \|y\|^2$, $\|f_y\| = \|y\|$.
The inverse proposition is also true, that is

Theorem 2.1 (Riesz Representation theorem)
Let $\mathcal{X}$ be a Hilbert space, $f \in X^*$, then there exists a unique $y_f \in \mathcal{X}$ such that

$$f(x) = \langle x, y_f\rangle, \quad \forall x \in \mathcal{X}$$

Proof.
Assume that $f \neq 0$, then $W = ker(f) = \{x \in \mathcal{X} | f(x) = 0\}$ is a proper subspace of $\mathcal{X}$, so we can pick $x_0 \in W^{\perp}$ with $f(x_0) = 1$.
Then for any $x \in \mathcal{X}$,

$$f(x - f(x)x_0) = f(x) - f(x) \cdot f(x_0) = 0$$

So $x - f(x)x_0 \in W$ and it follows that

$$\langle x - f(x)x_0, x_0\rangle = \langle x, x_0\rangle - f(x)\langle x_0, x_0\rangle = 0$$

Hence

$$f(x) = \langle x, \frac{x_0}{\|x_0\|^2}\rangle$$

Take $y_f = \frac{x_0}{\|x_0\|^2}$ and we have

$$f(x) = \langle x, y_f\rangle$$

Assume that there is another y_f' then

$$\langle x, y_f\rangle - \langle x, y_f'\rangle = \langle x, y_f - y_f'\rangle = 0$$

implies that $y_f = y_f'$. $\square$

Theorem 2.2
Let $\mathcal{X}$ be a Hilbert space, $a(x, y)$ be a conjugate bilinear function on $\mathcal{X}$, and there exists $M > 0$ such that

$$|a(x, y)| \leq M\|x\|\|y\| \quad , \forall x, y \in \mathcal{X}$$

Then there exists a unique $A \in \mathcal{L}(\mathcal{X})$ such that

$$a(x,y) = (x, Ay) \quad \forall x, y \in \mathcal{X}$$

and

$$\|A\| = \sup_{(x,y)\in\mathcal{X}\times\mathcal{X},(x,y)\neq 0} \frac{|a(x,y)|}{\|x\|\|y\|}$$

Proof.
For each $y \in \mathcal{X}$, it's easy to verify that $a(x,y)$ is a continuous linear functional, by Riesz representation theorem, there exists $z = z(y)$ such that $a(x,y) = \langle x, z\rangle$, then define

$$A : y \to z(y)$$

and we have $a(x,y) = (x, Ay), \forall x, y \in \mathcal{X}$.
Since

$$\begin{aligned}
\langle x, A(a_1y_1 + a_2y_2)\rangle &= a(x, a_1y_1 + a_2y_2) \\
&= \bar{a}_1 a(x, y_1) + \bar{a}_2 a(x, y_2) \\
&= \bar{a}_1 (x, Ay_1) + \bar{a}_2 (x, Ay_2) \\
&= \langle x, a_1Ay_1\rangle + \langle x, a_2Ay_2\rangle \\
&= \langle x, a_1Ay_1 + a_2Ay_2\rangle \quad \forall x, y_1, y_2 \in \mathcal{X}, \forall a_1, a_2 \in \mathbb{K}
\end{aligned}$$

Thus

$$A(a_1y_1 + a_2y_2) = a_1Ay_1 + a_2Ay_2$$

Moreover,

$$\|Ax\|^2 = \langle Ax, Ax\rangle = a(Ax, x) \leq M\|Ax\|\|x\|$$

So

$$\|Ax\| \leq M\|x\|$$

i.e. $A \in \mathcal{L}(\mathcal{X})$. $\square$

3 Baire Category And Open Mapping Theorem

Definition 3.1 (nowhere dense)
Let $\mathcal{X}$ be a metric space, $E \subset \mathcal{X}$ is called a **nowhere dense** set if $\bar{E}$ has empty interior.

Proposition 3.1
Let $\mathcal{X}$ be a metric space, then $E \subset \mathcal{X}$ is nowhere dense if and only if for any ball $B(x_0, r_0)$, there exists $B(x_1, r_1) \subset B(x_0, r_0)$ such that

$$\bar{E} \cap \overline{B(x_1, r_1)} = \varnothing$$

Proof.
$\Rightarrow$:
Suppose that E is nowhere dense, then $\bar{E}$ cannot contain any ball. Thus for any $B(x_0, r_0)$, there exists some $x_1 \in \overline{E} \setminus B(x_0, r_0)$. Since $\bar{E}$ is closed, x_1 is an interior point thus there exists a ball $B(x_1, r_1) \subset B(x_0, r_0)$ such that $\overline{B(x_1, r_1)} \cap \bar{E} = \varnothing$.
$\Leftarrow$:
Assume that E is not nowhere dense, then there exists a ball $\overline{B(x_0, r_0)} \subset \bar{E}$, thus for any ball $B(x_1, r_1) \subset B(x_0, r_0)$, we have $\overline{B(x_1, r_1)} \cap \bar{E} = \overline{B(x_1, r_1)}$, leading a contradiction. $\square$

Proposition 3.2
A set is nowhere dense if and only if its closure is nowhere dense.
Proof.
$\Rightarrow$:
Let A be a nowhere dense set, then $\bar{A}$ has empty interior. The closure of $\bar{A}$ is itself Hence $\bar{A}$ is nowhere dense.

$\Leftarrow$:
Trivial. $\square$

Proposition 3.3
The complement of a closed nowhere dense set is a dense open subset, and thus the complement of a nowhere dense set is a set with dense interior.
Proof.
Let A be a closed nowhere dense set, then $\bar{A} = A$ has empty interior and A^c is open.
For any open set $O \subset \mathcal{X}$, O cannot be contained in A thus $O \cap A^c \neq \varnothing$. So A^c is dense. $\square$

Definition 3.2 (category)
Let $\mathcal{X}$ be a metric space. E is called a **(Baire) first category set** if

$$E = \bigcup_{n=1}^{\infty} E_n$$

where each E_n is nowhere dense.
Other sets are called **(Baire) second category sets**.

Theorem 3.1 (Baire Category theorem)
A complete metric space is of second category.
Proof.
Let $\mathcal{X}$ be a complete metric space.
Assume that $\mathcal{X}$ is of first category, then

$$\mathcal{X} = \bigcup_{n=1}^{\infty} E_n$$

where each E_n is nowhere dense.
Pick any ball $B(x_0, r_0)$ in $\mathcal{X}$, since E_1 is nowhere dense, there exists a ball $B(x_1, r_1) \subset B(x_0, r_0)$ such that

$$\overline{B(x_1, r_1)} \cap \bar{E}_1 = \varnothing$$

Suppose that we have chosen the n^{th} ball $B(x_n, r_n)$ such that

$$\overline{B(x_n, r_n)} \cap \bar{E}_n = \varnothing$$

Then we can choose $(n+1)^{th}$ ball $B(x_{n+1}, r_{n+1}) \subset B(x_n, r_n)$ such that

$$\overline{B(x_{n+1}, r_{n+1})} \cap \bar{E}_{n+1} = \varnothing$$

since E_{n+1} is nowhere dense. We assume that for each n, $r_n < \frac{1}{2^n}$. Hence we can obtain a sequence $\{x_n\}$ inductively. Note that for each $p \in \mathbb{N}$,

$$d(x_{n+p}, x_n) < r_n = \frac{1}{2^n} \to 0 \text{ as } n \to \infty$$

Thus $\{x_n\}$ is Cauchy and we can assume that $x_n \to x \in \mathcal{X}$ as $\mathcal{X}$ is complete.
Also note that for each $n \in \mathbb{N}$, $d(x, x_n) \leq r_n = \frac{1}{2^n}$ (let $p \to \infty$), so $x \in \overline{B(x_n, r_n)}$, it follows that $x \notin E_n$. Therefore,

$$x \notin \bigcup_{n=1}^{\infty} E_n = \mathcal{X}$$

leading a contradiction. $\square$

Remark.
An equivalent statement of Baire Category theorem:
Let $\mathcal{X}$ be a complete metric space, $\{U_n\}$ is a sequence of open dense sets, then $\bigcap_{n=1}^{\infty} U_n$ is also dense in $\mathcal{X}$.

Proof.
$\Rightarrow$:
Assume that $\{U_n\}$ is a sequence of open dense sets and $\bigcap_{n=1}^{\infty} U_n$ is not dense in complete metric space $\mathcal{X}$, then there exists a ball

$$B(x_0, r_0) \subset \mathcal{X} \setminus \bigcap_{n=1}^{\infty} U_n$$

Note that for each U_n, $E_n = U_n^c$ is closed and nowhere dense, thus

$$\begin{aligned}\mathcal{X} \setminus \bigcap_{n=1}^{\infty} U_n &= \mathcal{X} \cap (\bigcap_{n=1}^{\infty} U_n)^c \\ &= \mathcal{X} \cap \bigcup_{n=1}^{\infty} E_n = \bigcup_{n=1}^{\infty} E_n\end{aligned}$$

Therefore,

$$B(x_0, r_0) \subset \bigcup_{n=1}^{\infty} E_n$$

However, $B(x_0, r_0)$ as a complete metric space can't be covered by countably many nowhere dense sets, leading a contradiction.
$\Leftarrow$:
Assume that complete metric space $\mathcal{X}$ can be covered by countably many nowhere dense sets, i.e.

$$\mathcal{X} = \bigcup_{n=1}^{\infty} E_n$$

where each E_n is nowhere dense.
Naturally we can assume that each E_n is closed. However, this implies

$$\varnothing = \mathcal{X}^c = (\bigcup_{n=1}^{\infty} E_n)^c = \bigcap_{n=1}^{\infty} E_n^c$$

Let $U_n = E_n^c$, then each U_n is open and dense, therefore

$$\bigcap_{n=1}^{\infty} U_n = \varnothing$$

leading a contradiction since $\varnothing$ can't be dense. $\square$

The statement is also true if $\mathcal{X}$ is a locally compact Hausdorff space.

Theorem 3.2
Let $\mathcal{X} = C[0,1]$, $E \subset \mathcal{X}$ is the set of nowhere differentiable functions, then E^c is of first category.
Proof.
Let E_n be the set of all $f \in \mathcal{X}$ for which there exists $x_0 \in [0,1]$ such that

$$|f(x) - f(x_0)| \le n|x - x_0|, \quad \forall x \in [0,1]$$

We first show that for each $n \in \mathbb{N}$, E_n is nowhere dense.
Note that $E_n = \bar{E}_n$. In fact, if there exists a cluster point $f \in \bar{E}_n$ such that for any $x_0 \in [0,1]$ and any $x \in [0,1]$, we have

$$|f(x) - f(x_0)| > n|x - x_0|$$

Fix x and x_0, take

$$\epsilon = \frac{1}{4}(|f(x) - f(x_0)| - n|x - x_0|)$$

There exists $g \in E_n$ such that $\|f-g\|_\infty < \epsilon$, and

$$\begin{aligned}|f(x)-f(x_0)| &= |f(x)-g(x)+g(x)-g(x_0)+g(x_0)-f(x_0)| \\ &\le |f(x)-g(x)|+|g(x)-g(x_0)|+|f(x_0)-g(x_0)| \\ &\le 2\epsilon + |g(x)-g(x_0)|\end{aligned}$$

Hence

$$\begin{aligned}|g(x)-g(x_0)| &\ge |f(x)-f(x_0)| - 2\epsilon \\ &= |f(x)-f(x_0)| - \frac{1}{2}(|f(x)-f(x_0)| - n|x-x_0|) \\ &= \frac{1}{2}(|f(x)-f(x_0)| + n|x-x_0|) \\ &> n|x-x_0|\end{aligned}$$

which implies that $g \notin E_n$, leading a contradiction. So it suffices to show that each point of E_n is not an interior point. Given $\epsilon > 0$, define g_ϵ on $[0,T]$ by

$$g_\epsilon(x) = \begin{cases} \frac{2\epsilon x}{T}, & 0 \le x < \frac{T}{4} \\ \frac{\epsilon}{2} - \frac{2\epsilon(x-\frac{T}{4})}{T}, & \frac{T}{4} \le x < \frac{3T}{4} \\ -\frac{\epsilon}{2} + \frac{2\epsilon(x-\frac{3T}{4})}{T}, & \frac{3T}{4} \le x \le T \end{cases}$$

where $T = \frac{1}{k}$, $k \in \mathbb{N}$. Then we extend g_ϵ to $[0,1]$ periodicly and denote it by $g(x)$, $g(x)$ satisfies that $\|g\|_\infty \le \frac{\epsilon}{2}$
For any $f \in E_n$, by Weierstrass Theorem, there exists a polynomial $p(x)$ on $[0,1]$ such that

$$\|f-p\|_\infty < \frac{\epsilon}{2}$$

Since $p(x)$ is a polynomial on $[0,1]$, we can assume that $|p'(x)| \le M'$ for some constant M' and moreover, we can assume that for any $x_0 \in [0,1]$ and any $x \in [0,1]$, there exists some M such that $|p(x)-p(x_0)| \le M|x-x_0|$.
Now let $h = p+g$, then

$$\|h-f\|_\infty \le \|f-p\|_\infty + \|f-g\|_\infty \le \epsilon$$

And there exists some large k such that $\frac{2\epsilon}{T} > M+n$. Then for any $x_0 \in [0,1]$ there exists some $x \in [0,1]$ such that

$$|g(x)-g(x_0)| \ge M+n$$

And it follows

$$|h(x)-h(x_0)| \ge |g(x)-g(x_0)| - |p(x)-p(x_0)| \ge n$$

Hence $h \notin E_n$, so f is not an interior point.
Since f is arbitrary, we can conclude that E_n has empty interior and therefore it is nowhere dense. Note that $\forall f \in E^c$, there exists some point $x_0 \in (0,1)$ such that f is differentiable at x_0, so $|f'(x_0)|$ exists and it is dominated by some integer. It follows that

$$E^c \subset \bigcup_{n=1}^{\infty} E_n$$

So E^c is of first category. $\square$

Remark.
Since each E_n is closed and nowhere dense, E_n^c is open and dense. Thus

$$\bigcap_{n=1}^{\infty} E_n^c$$

is dense in $C[0,1]$.
Moreover, $\bigcap_{n=1}^{\infty} E_n^c \subset E \subsetneq X$, thus E is also dense in $C[0,1]$.
In general, if A^c is of first category, then A is dense. The converse proposition is not true.

Theorem 3.3 (open mapping theorem)
Let $\mathcal{X}$, $\mathcal{Y}$ be Banach spaces, if $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ and T is onto, then T is an open map.

Corollary 3.1 (Banach inverse mapping theorem)
Let $\mathcal{X}$, $\mathcal{Y}$ be Banach spaces, if $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ and T is bijective, then $T^{-1} \in \mathcal{L}(\mathcal{Y}, \mathcal{X})$.

Definition 3.3 (closed linear operator)
Let $\mathcal{X}$, $\mathcal{Y}$ be normed vector spaces, $T : \mathcal{X} \to \mathcal{Y}$ is a linear operator, T is called **closed** if

$$\begin{cases} x_n \in D(T), x_n \to x \\ Tx_n \to y \end{cases} \Rightarrow \begin{cases} x \in D(T) \\ y = Tx \end{cases}$$

Definition 3.4 (graph)
Let $\mathcal{X}$, $\mathcal{Y}$ be vector spaces, $T : \mathcal{X} \to \mathcal{Y}$, then the graph of T is

$$G(T) = \{(x, Tx) | x \in D(T)\} \subset \mathcal{X} \times \mathcal{Y}$$

Note.
If $\mathcal{X}$ and $\mathcal{Y}$ are normed vector spaces, $T : \mathcal{X} \to \mathcal{Y}$ is a linear operator, then we can define a norm on the product space $\mathcal{X} \times \mathcal{Y}$ by

$$\|x\|_G = \|x\|_{\mathcal{X}} + \|Tx\|_{\mathcal{Y}}$$

Now we can give another equivalent definition of closed linear operator:
Let $\mathcal{X}$, $\mathcal{Y}$ be normed vector spaces, $T : \mathcal{X} \to \mathcal{Y}$ is a linear operator, T is called **closed** if $G(T)$ is closed with respect to $\| \cdot \|_G$.

Theorem 3.4 (bounded linear transformation theorem)
Let X be a normed vector space, Y be a Banach space, $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, then T can be extended to $\overline{D(T)}$ denoted by T_1 satisfying
(1) $T_1|_{D(T)} = T$.
(2) $\|T_1\| = \|T\|$.
Proof.
Define T_1 on $\overline{D(T)}$ by

$$\forall x \in \overline{D(T)}, \exists x_n \in D(T), x_n \to x.$$

We shall verify that T is well-defined, i.e. the limit of Tx_n exists and doesn't depend on the selection of $\{x_n\}$. First note that T is continuous, so there exists $M > 0$ such that

$$\|Tx\| \le M\|x\|, \forall x \in D(T)$$

Then for each $p \in \mathbb{N}$, we have

$$\|Tx_{n+p} - Tx_n\| \le M\|x_{n+p} - x_n\| \quad , \forall n \in \mathbb{N}$$

This implis that $\{Tx_n\}$ is also a Cauchy sequence and therefore, converges to some point $y \in \mathcal{Y}$ as $\mathcal{Y}$ is complete. Suppose $\{x'_n\}$ is another sequence converging to x, then

$$\|Tx'_n - Tx_n\| \le M\|x'_n - x_n\| \le M(\|x'_n - x\| + \|x_n - x\|) \to 0$$

Hence the limit doesn't depend on the selection of $\{x_n\}$. Obviously T_1 is still a linear operator and $T_1|_{D(T)} = T$. Also, $\|T_1 x\| \le \|T\|\|x\|$, thus $\|T_1\| = \|T\|$. $\square$

Corollary 3.2 (equivalent norm theorem)
Let $\mathcal{X}$ be a vector space, $\| \cdot \|_1$ and $\| \cdot \|_2$ are two norms on $\mathcal{X}$, if $(\mathcal{X}, \| \cdot \|_1)$ and $(\mathcal{X}, \| \cdot \|_1)$ are both complete and $\| \cdot \|_2$ is stronger than $\| \cdot \|_1$, then these two norms are equivalent.
Proof.

Consider identity map $I : (\mathcal{X}, \|\cdot\|_2) \to (\mathcal{X}, \|\cdot\|_1)$, since $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$, there exists M such that for each $x \in \mathcal{X}$

$$\|Ix\|_1 = \|x\|_1 \le M\|x\|_2$$

So I is continuous. And since I is a bijection, by Banach theorem, I^{-1} is also continuous and thus there exists M' such that

$$\|x\|_2 = \|I^{-1}x\|_2 \le M'\|x\|_1$$

Therefore $\|\cdot\|_1$ is stronger than $\|\cdot\|_2$. $\square$

Theorem 3.5 (closed graph theorem)
Let $\mathcal{X}$, $\mathcal{Y}$ be Banach spaces. If $T : \mathcal{X} \to \mathcal{Y}$ is a closed linear operator, and $D(T)$ is closed, then T is continuous.
Proof.
We show that $\mathcal{X}$ is also complete with respect to the graph norm

$$\|x\|_G = \|x\|_{\mathcal{X}} + \|Tx\|_{\mathcal{Y}}$$

Let $\{x_n\} \subset D(T)$ be a Cauchy sequence with respect to the graph norm, then for each $p \in \mathbb{N}$

$$\|x_{n+p} - x_n\|_G = \|x_{n+p} - x_n\|_{\mathcal{X}} + \|T(x_{n+p} - x_n)\|_{\mathcal{Y}} \to 0 \text{ as } n \to \infty$$

This implies $\|x_{n+p} - x_n\|_{\mathcal{X}} \to 0$ as $n \to \infty$.
Since $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ is complete, $(D(T), \|\cdot\|_{\mathcal{X}})$ as a closed subspace of $\mathcal{X}$ is also complete, we know that x_n converges to some $x \in D(T) \subset \mathcal{X}$.
For any $p \in \mathbb{N}$ and $\epsilon > 0$, there exists some N such that when $n > N$,

$$\|T(x_{n+p} - x_n)\|_{\mathcal{Y}} < \epsilon$$

Hence $\{Tx_n\}$ is a Cauchy sequence, note that $\mathcal{Y}$ is complete, so there exists $y \in \mathcal{Y}$ such that $Tx_n \to y$.
Also, T is closed implies $(x, y) \in G(T)$, and therefore $y = Tx$,

$$\|x_n - x\|_G = \|x_n - x\|_{\mathcal{X}} + \|Tx_n - Tx\|_{\mathcal{Y}} \to 0 \text{ as } n \to \infty$$

So $\mathcal{X}$ is also complete with respect to the graph norm.
Obviously $\|\cdot\|_G$ is stronger than $\|\cdot\|_{\mathcal{X}}$, by equivalent norm theorem, $\|\cdot\|_{\mathcal{X}}$ is stronger than $\|\cdot\|_G$, hence there exists M such that

$$\|x\|_{\mathcal{X}} + \|Tx\|_{\mathcal{Y}} \le M\|\cdot\|_{\mathcal{X}}$$

Thus $\|Tx\|_{\mathcal{Y}} \le M\|\cdot\|_{\mathcal{X}}$ and therefore T is continuous. $\square$

Example 3.1 A closed linear operator may not be bounded.
Consider $C[0,1]$ and $T : f \mapsto \frac{df}{dt}$, then $D(T) = C^1[0,1]$, we can show that T is closed: Suppose $\{f_n(t)\} \subset D(T)$, $f_n \to f \in D(T)$, $\frac{df_n}{dt} \to y$, note that this is uniformly convergent since the norm is $\|\cdot\|_\infty$, hence

$$\lim_{n\to\infty} \int_0^1 \frac{df_n(t)}{dt} dt$$

Remark.
A continuous linear operator T is a closed linear operator, because by B.L.T, we can always assume that the domain of any continuous linear opeartor is closed, then $x_n \to x$ implies $Tx_n \to Tx$, $x \in D(T)$ so $(x, Tx) \in G(T)$.

Theorem 3.6 (uniform boundedness theorem)
Let $\mathcal{X}$ be a Banach space, $\mathcal{Y}$ be a normed vector space. If $W \subset \mathcal{L}(\mathcal{X}, \mathcal{Y})$ and for each $x \in \mathcal{X}$,

$$M_x = \sup_{T\in W} \|Tx\| < \infty$$

Then there is a finite constant M such that $\|T\| \le M$ for all $T \in W$.
Proof(Version 1).
For any $x \in \mathcal{X}$, define

$$\|x\|_W = \|x\|_{\mathcal{X}} + \sup_{T\in W} \|Tx\|_{\mathcal{Y}}.$$

It's easy to verify that $\|\cdot\|_W$ is a norm on $\mathcal{X}$ stronger than $\|\cdot\|_{\mathcal{X}}$. We want to show that two norms are equivalent, then by equivalent norm theorem, it suffices to show that $(\mathcal{X}, \|\cdot\|_W)$ is complete.
Suppose $\{x_n\}$ is a Cauchy sequence with respect to $\|\cdot\|_W$, i.e.

$$\|x_m - x_n\| + \sup_{T\in W} \|T(x_m - x_n)\|_{\mathcal{Y}} \to 0 \text{ as } m, n \to \infty$$

Since $(\mathcal{X}, \|\cdot\|_{\mathcal{X}})$ is complete, there exists $x \in \mathcal{X}$ such that $\|x_n - x\|_{\mathcal{X}} \to 0$ as $n \to \infty$. Also, $\forall \epsilon > 0$, $\exists N = N(\epsilon)$ such that

$$\sup_{T\in W} \|Tx_m - Tx_n\|_{\mathcal{Y}} < \epsilon, \quad \forall m, n > N$$

Hence for each $T \in W$,

$$\|Tx_m - Tx_n\|_{\mathcal{Y}} < \epsilon, \quad \forall m, n > N$$

Let $n \to \infty$, then we have

$$\|Tx_m - Tx\|_{\mathcal{Y}} \le \epsilon, \quad \forall m > N$$

Now take the supremum over W and we get

$$\sup_{T\in W} \|Tx_n - Tx\|_{\mathcal{Y}} \le \epsilon, \quad \forall n > N$$

Since ϵ is arbitrary,

$$\|x_n - x\| + \sup_{T\in W} \|T(x_n - x)\|_{\mathcal{Y}} \to 0 \text{ as } n \to \infty$$

i.e. $\|x_n - x\|_W \to 0$, hence $\|\cdot\|_W$ and $\|\cdot\|_{\mathcal{X}}$ are equivalent, thus $\exists M$ such that

$$\|x\|_{\mathcal{X}} + \sup_{T\in W} \|Tx\|_{\mathcal{Y}} = \|x\|_W \le M\|x\|_{\mathcal{X}} \quad, \forall x \in \mathcal{X}$$

Therefore

$$\sup_{T\in W} \|Tx\|_{\mathcal{Y}} \le M\|x\|_{\mathcal{X}} \quad, \forall x \in \mathcal{X}$$

So for each $T \in W$, each $x \in \mathcal{X}$

$$\|Tx\|_{\mathcal{Y}} \le M\|x\|_{\mathcal{X}}$$

That is,

$$\|T\| \le M \quad, \forall T \in W \quad \square$$

Proof(Version 2).
Define C_n by

$$C_n = \{x \in \mathcal{X} : \|Tx\| \le n, \forall T \in W\}$$

Since for each $x \in \mathcal{X}$, $\sup_{T\in W} \|Tx\| < M_x < \infty$, so there exists some large n such that $x \in C_n$, therefore

$$\mathcal{X} = \bigcup_{n=1}^{\infty} C_n$$

Also, note that for each n,

$$C_n = \bigcap_{T\in W} \{x \in \mathcal{X} : \|Tx\| \le n\}$$

Obviously, $\{x \in \mathcal{X} : \|Tx\| \le n\}$ is closed hence C_n is closed.
By Baire Category theorem, $\mathcal{X}$ as a complete metric space cannot be covered by countably many nowhere dense sets, therefore there exists n_0 such that C_{n_0} has non-empty interior.
So we can assume that $\overline{B(x_0, \epsilon)} \subset C_{n_0}$, then for any $x \in \mathcal{X}$ with $\|x\| \le \epsilon$ and any $T \in W$

$$\|T(x + x_0)\| \le n_0$$

So

$$\|Tx\| \le n_0 + \|Tx_0\|, \forall T \in W$$

Thus for any $x \in \mathcal{X}$ with $\|x\| \le 1$, $\|\epsilon x\| \le \epsilon$ and

$$\|T\epsilon x\| \le n_0 + \|Tx_0\|, \forall T \in W$$

It follows

$$\|Tx\| \le \frac{n_0 + \|Tx_0\|}{\epsilon}, \forall T \in W$$

Let $M = \frac{n_0+\|Tx_0\|}{\epsilon}$ and take supremum over W, then we have $\|T\| \le M$ □

Theorem 3.7 (Banach-Steinhaus)
Let $\mathcal{X}$ be a Banach space, $\mathcal{Y}$ be a normed vector space, D is a dense subset of $\mathcal{X}$, $A_n (n = 1, 2, \cdots), A \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$. Then $\forall x \in \mathcal{X}$

$$\lim_{n\to\infty} A_n x = Ax$$

if and only if
(1) $\|A_n\| \le M$ for some constant M.
(2) $\forall x \in D$, $\lim_{n\to\infty} A_n x = Ax$.
Proof.
$\Rightarrow$:
Since for each $x \in \mathcal{X}$, $\lim_{n\to\infty} \|A_n x\| = \|Ax\| < \infty$, by uniform boundedness theorem, there exists M such that $\|A_n\| \le M$ for each $n \in \mathbb{N}$.
(2) is trivial.
$\Leftarrow$:
We know that there exists M such that $\|A_n\| \le M'$, take $M = \max\{M', \|A\|\}$.
$\forall x \in \mathcal{X}$ and $\forall \epsilon > 0$, there exists $y \in D$ such that $\|x - y\| < \frac{\epsilon}{3M}$ and for this y, there exists $N \in \mathbb{N}$ such that when $n > N$,

$$\|A_n y - Ay\| < \frac{\epsilon}{3}$$

Then we have

$$\begin{aligned}
\|A_n x - Ax\| &\le \|A_n x - A_n y\| + \|A_n y - Ay\| + \|Ay - Ax\| \\
&\le \|A_n\|\|x - y\| + \|A_n y - Ay\| + \|A\|\|x - y\| \\
&< M \cdot \frac{\epsilon}{3M} + \frac{\epsilon}{3} + M \cdot \frac{\epsilon}{3M} \\
&= \epsilon
\end{aligned}$$

That is, $\|A_n x - Ax\| \to 0$ as $n \to \infty$. □

Theorem 3.8 (Lax-Milgram)
Let $a(x, y)$ be a conjugate bilinear function on Hilbert space $\mathcal{X}$ satisfying
(1) $\exists M > 0$ s.t. $|a(x, y)| \le M\|x\|\|y\|$, $\forall x, y \in \mathcal{X}$.
(2) $\exists \delta > 0$ s.t. $|a(x, x)| \ge \delta\|x\|^2$, $\forall x \in \mathcal{X}$.
Then there exists a unique bounded linear operator $A \in \mathcal{L}(\mathcal{X})$ with continuous inverse satisfying

$$a(x, y) = \langle x, Ay\rangle, \forall x, y \in \mathcal{X}$$

$$\|A^{-1}\| \le \frac{1}{\delta}$$

Proof.
By theorem 2.2 we know that there exists a unique $A \in \mathcal{L}(\mathcal{X})$, then we shall show that A is a bijection.
Suppose $a(x, y_1) = a(x, y_2)$, then $\langle x, Ay_1\rangle = \langle x, Ay_2\rangle$, $\forall x \in \mathcal{X}$. Particularly, $\langle y_1 - y_2, A(y_1 - y_2)\rangle = 0$.
However,

$$0 = \langle y_1 - y_2, A(y_1 - y_2)\rangle = |a(y_1 - y_2, y_1 - y_2)| \ge \delta\|y_1 - y_2\|^2$$

thus $y_1 = y_2$, so A is injective.
To show A is onto, it suffices to show that $R(A)$ is closed and $R(A)^\perp = \{0\}$.
Let $\{y_n\}$ be any sequence in $R(A)$, say, $y_n \to y \in \mathcal{X}$, we know that there is a corresponding sequence $\{x_n\}$

such that $y_n = Ax_n$.
Note that for each $p \in \mathbb{N}$,

$$\begin{aligned}\|x_{n+p} - x_n\|^2 &\leq \frac{1}{\delta}|a(x_{n+p} - x_n, x_{n+p} - x_n)| \\ &= \frac{1}{\delta}\langle x_{n+p} - x_n, A(x_{n+p} - x_n)\rangle \\ &\leq \frac{1}{\delta}\|x_{n+p} - x_n\|\|A(x_{n+p} - x_n)\| \\ &= \frac{1}{\delta}\|x_{n+p} - x_n\|\|y_{n+p} - y_n\|\end{aligned}$$

Thus

$$\|x_{n+p} - x_n\| \leq \frac{1}{\delta}\|y_{n+p} - y_n\| \to 0 \text{ as } n \to \infty$$

Since $\mathcal{X}$ is complete, there exists $x \in \mathcal{X}$ such that $x_n \to x$.
Note that A is continuous, so $y = \lim_{n\to\infty} Ax_n = Ax \in R(A)$, thus $R(A)$ is closed.
Suppose there is a $x_0 \in \mathcal{X}$ such that $\langle x_0, Ay\rangle = 0$ for all $y \in \mathcal{X}$.
Then particularly, $0 = |\langle x_0, Ax_0\rangle| = |a(x_0, x_0)| \geq \delta\|x_0\|^2$, which implies that $x_0 = 0$. Hence A is onto.
By Banach theorem, $A^{-1} \in \mathcal{L}(\mathcal{X})$, and we also have

$$\delta\|x\|^2 \leq |a(x, x)| = |\langle x, Ax\rangle| \leq \|A\|\|x\|^2$$

Hence $\delta\|x\| \leq \|Ax\|$, $\forall x \in \mathcal{X}$ and equivalently,

$$\|A^{-1}y\| \leq \frac{1}{\delta}\|y\|, \quad \forall y \in \mathcal{X}$$

i.e. $\|A^{-1}\| \leq \frac{1}{\delta}$. □

Theorem 3.9 (Lax)
Let $\mathcal{X}$ and $\mathcal{Y}$ be Banach spaces, $T_n \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$ and $\forall y \in \mathcal{Y}$, there exists unique x_n, x such that

$$T_n x_n = y, \quad Tx = y$$

If $\forall x \in \mathcal{X}$,

$$\|Tx - T_n x\| \to 0 \text{ as } n \to \infty \text{ (compatibility)}$$

Then $x_n \to x$ if and only if $\exists C$ such that $\|T_n^{-1}\| \leq C$ for all $n \in \mathbb{N}$.
Proof.
$\Rightarrow$:
For fixed $y \in \mathcal{Y}$, let $x_n = T_n^{-1}y$, $x = T^{-1}y$, suppose $x_n \to x$, then

$$T_n^{-1}y \to T^{-1}y$$

This implies $\sup_n \|T_n^{-1}y\| < \infty$, by uniform boundedness theorem, there exists C such that

$$\|T_n^{-1}\| \leq C$$

$\Leftarrow$:

$$\begin{aligned}\|x_n - x\| &= \|T_n^{-1}y - T_n^{-1}T_n x\| \\ &\leq \|T_n^{-1}\|\|y - T_n x\| \\ &\leq C\|Tx - T_n x\| \to 0 \text{ (by compatibility)} \ \square\end{aligned}$$

4 Hahn-Banach Theorem

Definition 4.1 (sublinear functional)
Let $\mathcal{X}$ be a vector space on $\mathbb{R}$, $f : \mathcal{X} \to \mathbb{R}$ is called a **sublinear functional** if

(1) $\forall \lambda > 0$, $\forall x \in \mathcal{X}$, $f(\lambda x) = \lambda f(x)$. (positive homogeneity)
(2) $\forall x, y \in \mathcal{X}$, $f(x+y) \leq f(x) + f(y)$. (subadditivity).

Theorem 4.1 (real Hahn-Banach theorem)
Let $\mathcal{X}$ be a vector space on $\mathbb{R}$, $\mathcal{Y}$ is a subspace. p is a sublinear functional on $\mathcal{X}$. f is a linear functional on Y dominated by p, then there exists an extension of f on $\mathcal{X}$ denoted by F such that $F|_{\mathcal{Y}} = f$, $F \leq p$.
Proof.
Step 1.
Let S be the set of all tuples (A, g), where $A = D(g)$, g is a linear functional, the extension of f dominated by p. S is equipped with the partial order $\leq$ which is defined by

$$(A_1, g_1) \leq (A_2, g_2) \quad \text{if } A_1 \subset A_2,\ g_2|_{A_1} = g_1$$

Then S is a partially ordered set. We want to find the maximal element of S
Let $C \subset S$ be any chain of S. Suppose that $C = \{(A_1, g_1), (A_2, g_2), \cdots\}$, where

$$(A_k, g_k) \leq (A_{k+1}, g_{k+1}), \quad \forall k \in \mathbb{N}$$

Then we need to find an upper bound for C.
Define A by

$$A = \bigcup_{n=1}^{\infty} A_n$$

Then $\forall x \in A$, there exists $k \in \mathbb{N}$ such that $x \in A_k$, and we can define g by

$$g(x) = g_k(x)$$

So $D(g) = A$. We shall verify that $(A, g) \in S$.

(1) Obviously, $g|_{\mathcal{Y}} = f$.

(2) For any $x_1, x_2 \in A$, $c_1, c_2 \in \mathbb{R}$, there exists k such that $x_1, x_2 \in A_k$, and thus $c_1x_1 + c_2x_2 \in A_k$, so

$$g(c_1x_1 + c_2x_2) = g_k(c_1x_1 + c_2x_2) = c_1 g_k(x_1) + c_2 g_k(x_2) = c_1 g(x_1) + c_2 g(x_2)$$

i.e. $g(x)$ is a linear functional.

(3) For any $x \in A$, there exists k such that $x \in A_k$, so $g(x) = g_k(x) \leq p(x)$. Hence g is dominated by p on A.

Hence every chain of S has an upper bound, by Zorn's lemma, there exists a maximal element $(\mathcal{X}_m, F)$.

Step 2.
It remains to show that $\mathcal{X}_m = \mathcal{X}$.
Assume that $\mathcal{X}_m \subsetneq \mathcal{X}$, then $\exists x_0 \in \mathcal{X} \setminus \mathcal{X}_m$.
Now let $\mathcal{X}'_m$ be the space spanned by $\mathcal{X}_m$ and x_0, i.e.

$$\mathcal{X}'_m = \mathcal{X}_m \oplus \{\lambda x_0 | \lambda \in \mathbb{R}\}$$

So each $x \in \mathcal{X}'_m$ has a unique decomposition $x = x_m + cx_0$, where $x_m \in \mathcal{X}_m$, $c \in \mathbb{R}$
Then define F' on $\mathcal{X}'_m$ by

$$F'(x) = F'(x_m + cx_0) = F(x_m) + ck$$

where k is a constant to be specified. (actually $k = F'(x_0)$)
For any $x_1, x_2 \in \mathcal{X}_m$,

$$\begin{aligned} F(x_1) - F(x_2) = F(x_1 - x_2) = F(x_1 + x_0 - x_2 - x_0) &\leq p(x_1 + x_0 - x_2 - x_0) \\ &\leq p(x_1 + x_0) + p(-x_2 - x_0) \end{aligned}$$

So we have

$$-F(x_2) - p(-x_2 - x_0) \le -F(x_1) + p(x_1 + x_0) \quad ,\forall x_1, x_2 \in \mathcal{X}_m$$

Take the infimum and supremum over $\mathcal{X}_m$ on the right side and left side respectively, we have

$$\sup_{x \in \mathcal{X}_m} (F(-x) - p(-x - x_0)) \le \inf_{x \in \mathcal{X}_m} (-F(x) + p(x + x_0))$$

Let k be any real number in between, then for any $x = x_m + cx_0 \in \mathcal{X}'_m$,

1. if $c = 0$
$$F'(x) = F(x_m) \le p(x_m) = p(x)$$
2. if $c > 0$
$$\begin{aligned} F'(x) &= F(x_m) + ck \\ &\le F(x_m) + c(-F(\frac{x_m}{c}) + p(\frac{x_m}{c} + x_0)) \\ &= F(x_m) - F(x_m) + p(x_m + cx_0) = p(x_m + cx_0) \end{aligned}$$
3. if $c < 0$
$$\begin{aligned} F'(x) &= F(x_m) + ck \\ &= F(x_m) + (-c)(-k) \\ &\le F(x_m) + (-c) \cdot -(F(-\frac{x_m}{c}) - p(-\frac{x_m}{c} - x_0)) \\ &= F(x_m) - F(x_m) + p(x_m + cx_0) = p(x_m + cx_0) \end{aligned}$$

So F' is dominated by p and thus $(\mathcal{X}'_m, F') \in S$, leading a contradiction.
Therefore $\mathcal{X}_m = \mathcal{X}$. $\square$

Theorem 4.2 (complex Hahn-Banach theorem)
Let $\mathcal{X}$ be a complex vector space, p a seminorm on $\mathcal{X}$, $\mathcal{Y}$ a vector subspace of $\mathcal{X}$, f_0 a linear functional on $\mathcal{Y}$ satisfying $|f_0(x)| \le p(x)$, $\forall x \in \mathcal{Y}$, then there exists a linear functional f satisfying
(1) $|f(x)| \le p(x)$, $\forall x \in \mathcal{X}$.
(2) $f(x) = f_0(x)$, $\forall x \in \mathcal{Y}$.
Proof.
skip.

Corollary 4.1
Let $\mathcal{X}$ be a normed vector space, $\mathcal{Y}$ is a vector subspace. Suppose y^* is a continuous linear functional on Y, then there exists a continuous linear functional x^* on $\mathcal{X}$ with $\|x^*\| = \|y^*\|$ such that $x^*(y) = y^*(y)$ for all $y \in \mathcal{Y}$.
Proof.
Define $p(x) = \|y^*\|\|x\|$ on $\mathcal{X}$, then y^* is dominated by p on $\mathcal{Y}$. By Hahn-Banach theorem, there exists a linear functional x^* on $\mathcal{X}$ such that

$$x^*|_{\mathcal{Y}} = y^*, \quad x^* \le p$$

Note that,

$$\|x^*(x)\| \le \|y^*\|\|x\|$$

On the other hand, by definition of operator norm, $\|x^*\| \ge \|y^*\|$.
So $\|x^*\| \le \|y^*\|$, it follows $\|x^*\| = \|y^*\|$. $\square$

Corollary 4.2
Let $\mathcal{X}$ be a normed vector space, $\mathcal{Y}$ be a closed vector subspace, $x_0 \in \mathcal{X} \setminus \mathcal{Y}$, then there exists a $x^* \in \mathcal{X}^*$ with

$$x^*(y) = 0, \quad \forall y \in \mathcal{Y}$$

$$x^*(x_0) = 1$$

$$\|x^*\| = \frac{1}{d}$$

where $d = \inf \|x_0 - y\|, y \in \mathcal{Y}$
Proof.
Consider vector subspace $\mathcal{X}' = \mathcal{Y} \oplus \{\lambda x_0 | \lambda \in \mathbb{R}\}$ and define f on $\mathcal{X}'$ by

$$f(y + \lambda x_0) = \lambda$$

It's easy to verify that f is a continuous linear functional on $\mathcal{X}'$, then by corollary 4.1, there exists a continuous linear functional x^* such that

$$x^*|_{\mathcal{X}'} = f$$

$$x^*(x_0) = 1$$

$$\|x^*\| = \|f\|$$

Now it suffices to show that $\|f\| = \frac{1}{d}$.
Note that $\mathcal{Y}$ is closed, so by definition of d, there exists a point $y' \in \mathcal{Y}$ such that $\|y' - x_0\| = d$, then $|f(y' - x_0)| \leq \|f\| d$, so $\|f\| \geq \frac{1}{d}$.
On the other hand, $\forall x \in \mathcal{X}'$, $x = y + \lambda x_0$, $y \in \mathcal{Y}$,

$$\begin{aligned} \|x\| = \|y + \lambda x_0\| = |\lambda| \|\frac{y}{\lambda} + x_0\| \\ \geq |\lambda| \cdot d = |f(y + \lambda x_0)| \cdot d = |f(x)| \cdot d \end{aligned}$$

So $\|f\| \leq \frac{1}{d}$ and therefore $\|f\| = \frac{1}{d}$. $\square$

Corollary 4.3
Let $\mathcal{X}$ be a normed vector space, $x_0 \in \mathcal{X}$, $x_0 \neq 0$, then there is a $x^* \in \mathcal{X}^*$ such that

$$|x^*(x_0)| = \|x_0\|$$

$$\|x^*\| = 1$$

Proof.
Consider subspace $Y = \{0\}$ then apply Corollary 4.2, there exists a $f \in \mathcal{X}^*$ such that

$$f(x_0) = 1$$

$$\|f\| = \frac{1}{d} = \frac{1}{\|x_0\|}$$

Then let $x^* = \|x_0\| f$. $\square$

Definition 4.2 (maximal vector subspace)
Let $\mathcal{X}$ be a vector space, $M \subset \mathcal{X}$ be a proper subspace, if $\forall M_1 \subset \mathcal{X}$, $M_1 \supset M$, we have $M_1 = \mathcal{X}$, then we call M a **maximal vector subspace**.

Note.
This definition is not true for the infinite-dimensional vector space, consider a infinite orthonormal basis $\{e_n\}$ of $\mathcal{X}$, then $M = span\{e_n\}$ is a proper subspace but it cannot satisfy the following proposition.

Proposition 4.1
Let $\mathcal{X}$ be a vector space, $M \subset \mathcal{X}$ is a maximal vector subspace if and only if M is a proper subspace and there exists $x_0 \in \mathcal{X}$ such that

$$\mathcal{X} = M \oplus \{\lambda x_0 | \lambda \in \mathbb{K}\}$$

Definition 4.3 (hyperplane)
Let $\mathcal{X}$ be a vector space, $M \subset \mathcal{X}$ be a maximal vector subspace, then for any $x_0 \in \mathcal{X}$,

$$L = M + x_0$$

is called a **hyperplane**.

Theorem 4.3
Let $\mathcal{X}$ be a normed vector space, $L \subset \mathcal{X}$ be a hyperplane if and only if there exists a non-zero linear functional f and $r \in \mathbb{R}$, such that

$$L = H_f^r = \{x \in \mathcal{X} | f(x) = r\}$$

Particularly, L is a closed hyperplane if and only if f is a continuous linear functional.
Proof.
$\Rightarrow$:
Suppose L is a hyperplane, and $L = M + x_0$, where M is a maximal subspace of $\mathcal{X}$ and $x_0 \in \mathcal{X}$.
If $x_0 \notin M$, then $\mathcal{X} = \bar{M} \oplus \{\lambda x_0 | \lambda \in \mathbb{R}\}$.
Now we can define a linear functional f on $\mathcal{X}$ by

$$f(m + \lambda x_0) = \lambda, \quad \forall m \in M, \lambda \in \mathbb{R}$$

Also, for each $x \in L = M + x_0$,

$$f(x) = f(m + x_0) = 1$$

So $L = H_f^1$.
If $x_0 \in M$, then pick any $x_1 \notin M$, similarly define f by

$$f(m + \lambda x_1) = \lambda, \quad \forall m \in M, \lambda \in \mathbb{R}$$

Then $L = M = H_f^0$.
$\Leftarrow$:
If $L = H_f^r$ where f is a non-zero linear functional on $\mathcal{X}$, first note that

$$ker(f) = H_f^0$$

is a vector subspace.
Pick any $y \in \mathcal{X} \setminus ker(f)$, then $\forall x \in \mathcal{X}$,

$$x - \frac{f(x)}{f(y)} y \in ker(f)$$

i.e. $\exists x_0 \in ker(f)$ such that $x = x_0 + \frac{f(x)}{f(y)} y$, so $\mathcal{X} = ker(f) \oplus \{\lambda y | \lambda \in \mathbb{R}\}$.
Hence $ker(f) = H_f^0$ is a maximal vector subspace.
Since f is linear, we can assume that $f(y) = r$, then $\forall x \in H_f^r$,

$$f(x - y) = f(x) - f(y) = r - r = 0, \quad x - y \in ker(f)$$

So $H_f^r = ker(f) + y$ is a hyperplane.

Moreover, if L is a closed hyperplane, then $ker(f)$ is closed.
Assume that the corresponding linear functional f is not continuous, then $\forall n \in \mathbb{N}$, we can find $x_n \in \mathcal{X}$ with $\|x_n\| = 1$, $|f(x_n)| > n$.
Pick $y \in \mathcal{X} \setminus ker(f)$ such that $f(y) = 1$, then $y - \frac{y}{f(x_n)} x_n$ is a sequence in $ker(f)$ converging to y and thus $y \in ker(f)$, leading a contradiction.
On the other hand, if f is continuous, then consider any sequence $\{x_n\} \subset ker(f)$, say $x_n \to x$, then $f(x_n) \to f(x)$, since $f(x_n) = 0$ for each n, so $f(x) = 0$ and thus $x \in ker(f)$. Hence $ker(f)$ is closed and therefore L is closed. $\square$

Definition 4.4 (separation)
Let $\mathcal{X}$ be a normed vector space, $L = H_f^r$ be a hyperplane, $E, F \subset \mathcal{X}$, E and F are **separated** by L if

$$\forall x \in E, \quad f(x) \geq r$$

$$\forall x \in F, \quad f(x) \leq r$$

Theorem 4.4 (geometric Hahn-Banach theorem)
Let $\mathcal{X}$ be a normed vector space on $\mathbb{R}$, $C \subset \mathcal{X}$ is a proper convex subset, containing 0 as an interior point, suppose $x_0 \notin C$, then there exists a hyperplane L separating x_0 and C.
Proof.
First note that $p(x_0) \geq 1$. Then consider subspace $\mathcal{Y} = \{\lambda x | \lambda \in \mathbb{R}\}$.
Define f_0 on $\mathcal{Y}$ by

$$f_0(\lambda x_0) = \lambda p(x_0)$$

Then $f_0(\lambda x_0) = \lambda p(x_0) \leq p(\lambda x_0)$.
By Hahn-Banach theorem, there exists a linear functional f on $\mathcal{X}$ such that

$$f|_{\mathcal{Y}} = f_0, \quad f \leq p \text{ for all } x \in \mathcal{X}$$

Consider hyperplane $L = H_f^1$, note that $f(x_0) = f_0(x_0) = p(x_0) \geq 1$ and for each $x \in C$,

$$f(x) \leq p(x) \leq 1 \quad \square$$

Note.
In fact, C can be any proper convex set containing non-empty interior. (By translation)
Moreover, we can show that H_f^r is closed. Note that $p(x)$ is uniformly continuous since C has non-empty interior, so there exists M with

$$|p(x-y)| \leq M\|x-y\|$$

Also,

$$|f(x)| \leq \max\{p(x), p(-x)\}$$

which yields

$$|f(x)| \leq M\|x\|$$

So f is bounded and therefore H_f^r is closed.

Theorem 4.5
Let $\mathcal{X}$ be a normed vector space, $C_1, C_2 \subset \mathcal{X}$ be two convex subset with

$$\mathring{C}_1 \neq \varnothing, \quad \mathring{C}_1 \cap C_2 = \varnothing$$

Then there exists $s \in \mathbb{R}$ and non-zero continuous linear functional f such that H_f^s separates C_1 and C_2, i.e.

$$f(x) \leq s, \quad \forall x \in C_1$$

$$f(x) \geq s, \quad \forall x \in C_2$$

Proof.
Define $C = C_1 - C_2$, then it's easy to verify that C is a non-empty convex set with non-empty interior and $0 \notin \mathring{C}$ since $\mathring{C}_1 \cap C_2 = \varnothing$.
Then by geometric Hahn-Banach theorem, there exists a closed hyperplane $L = H_f^r$ separates $\mathring{C}$ and 0.
Assume that there exists $r \in \mathbb{R}$ such that $f(x) \leq r \leq f(0)$, $\forall x \in \mathring{C}$.
Since f is linear, $r \leq 0$. Moreover, since f is continuous, $\forall x \in C$, we have $f(x) \leq r \leq 0$.
By definition of C, $\forall y \in C_1$, $\forall z \in C_2$,

$$f(y-z) \leq 0, \quad f(y) \leq f(z)$$

Take supremum over C_1 and infimum over C_2 respectively, and pick any s in between, then we have

$$\sup_{x \in C_1} f(x) \leq s \leq \inf_{x \in C_2} f(x)$$

i.e. H_f^s separates C_1 and C_2. $\square$

Corollary 4.4 (Ascoli)
Let $\mathcal{X}$ be a normed vector space on $\mathbb{R}$, $C \subset \mathcal{X}$ is a closed convex set, then $\forall x_0 \in \mathcal{X} \setminus C$, $\exists f \in \mathcal{X}^*$ and $a \in \mathbb{R}$ such that

$$f(x) < a < f(x_0), \quad \forall x \in C$$

Proof.
Since C is closed, we can find an open ball $B = B(x_0, \epsilon)$ with $B \cap C = \varnothing$.
Then apply Theorem 4.5 and we know that there exists $f \in \mathcal{X}^*$ and $s \in \mathbb{R}$ such that

$$\sup_{x \in C} f(x) \le s \le \inf_{x \in B} f(x)$$

There must exists $y \in \mathcal{X}$ with $f(y) = -1$ and we can also find sufficiently small δ with $0 < \delta < \frac{\epsilon}{\|y\|}$ such that $x_0 + \delta y \in B$.
Let $a = f(x_0 + \delta y)$, then

$$\sup_{x \in C} f(x) < a < f(x_0) \quad \square$$

Corollary 4.5 (Mazur)
Let $\mathcal{X}$ be a normed vector space, $C \subset \mathcal{X}$ is a closed convex set with non-empty interior, $F \subset \mathcal{X}$ is a linear manifold, $\mathring{C} \cap F = \varnothing$.
Then there exists a closed hyperplane L containing F such that C is on one side of L.
Proof.
Suppose $F = x_0 + M$, where $x_0 \in \mathcal{X}$, M is a vector subspace.
By theorem 4.5, there exists a hyperplane H_f^r separates C and F, such that

$$f(E) \le r \le f(F) = f(x_0 + M)$$

Let $r_0 = r - f(x_0)$, then we have $f(M) \ge r - f(x_0) = r_0$.
Since f is linear and M is a subspace, it's easy to verify that $f(M) \equiv 0$, i.e. $M \subset ker(f)$. Hence $F = x_0 + M \subset x_0 + ker(f) = H_f^s$, where $s = f(x_0)$.
Note that $f(C) \le r \le f(x_0 + M) = s$, therefore $L = H_f^s$ is what desired. $\square$

5 Dual Space And Weak Convergence

Definition 5.1 (dual space)
Let $\mathcal{X}$ be a normed vector space, the set of all continuous linear functional on $\mathcal{X}$ is a Banach space equipped with the supremum norm

$$\|f\| = \sup_{\|x\|=1} |f(x)|$$

It is called the **dual space** of $\mathcal{X}$, denoted by $\mathcal{X}^*$.

Note.
By theorem 1.1 we know that X^* is complete.

Example 5.1 (L^p spaces)
Let $p > 1$, then $(L^p)^* = L^q$, where q is the conjugate index of p. ($\frac{1}{p} + \frac{1}{q} = 1$)
Particularly, $(L^1)^* = L^\infty$.
Given $p \ge 1$, for each $g \in L^q$, we define a continuous linear functional on L^p by

$$F_g(f) = \int_X f(x)g(x)d\mu, \quad \forall f \in L^p$$

where μ is Lebesgue measure.
It can be shown that

$$\sigma : g \mapsto F_g$$

is surjective and isometric.

Definition 5.2 (double dual space)
Let $\mathcal{X}$ be any normed vector space, then the dual space of X^*, denoted by X^{**}, called the **double dual space** of $\mathcal{X}$.

Theorem 5.1
Let $\mathcal{X}$ be a normed vector space, then there is an isometric embedding $T : \mathcal{X} \to \mathcal{X}^{**}$.
Proof.
For each $x \in \mathcal{X}$, we can define $J_x : X^* \to \mathbb{K}$ by $J_x(f) = f(x)$, $\forall f \in \mathcal{X}^*$.
It's easy to verify that J_x is a linear functional on X^* and moreover

$$|J_x(f)| = |f(x)| \leq \|f\|\|x\|$$

Hence $\|J_x\| \leq \|x\|$, $J_x \in X^{**}$.
Consider canonical map $T : \mathcal{X} \to \mathcal{X}^{**}, x \mapsto J_x$. Still, it's easy to verify that T is linear, injective and

$$\|Tx\| = \|J_x\| \leq \|x\|$$

implies T is continuous, and thus by Banach theorem T is an embedding.
Moreover, by Hahn-Banach, for any given $x \in \mathcal{X}$, we can find an $f_x \in \mathcal{X}^*$ such that $f_x(x) = \|x\|$, $\|f_x\| = 1$, so $\|J_x\| \geq |J_x(f_x)| = |f_x(x)| = \|x\|$ and therefore $\|J_x\| = \|x\|$, i.e. T is an isometry. $\square$

Definition 5.3 (reflexive space)
Let $\mathcal{X}$ be a normed vector space, if canonical map $T : \mathcal{X} \to \mathcal{X}^{**}$ is onto, then $\mathcal{X}$ is called **reflexive**.

Note.
For $p > 1$, L^p is reflexive but for $p = 1, \infty$, L^p is not reflexive.

Definition 5.4 (adjoint operator)
Let $\mathcal{X}$, $\mathcal{Y}$ be normed vector spaces, $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, then the **adjoint operator** of T, denoted by T^*, is defined by

$$T^* : \mathcal{Y}^* \to \mathcal{X}^*$$
$$f(Tx) = (T^*f)(x), \quad \forall f \in \mathcal{Y}^*, \forall x \in \mathcal{X}$$

Theorem 5.2
Let $\mathcal{X}, \mathcal{Y}$ be two normed vector spaces, then $* : T \mapsto T^*$ is an isometric embedding from $\mathcal{L}(\mathcal{X}, \mathcal{Y})$ to $\mathcal{L}(\mathcal{Y}^*, \mathcal{X}^*)$.
Proof.
First we show that * is linear.
For any $x \in \mathcal{X}$, $f \in \mathcal{Y}^*$, $\alpha, \beta \in \mathbb{K}$,

$$\begin{aligned}(\alpha T_1 + \beta T_2)^*(f)(x) &= f((\alpha T_1 + \beta T_2)(x)) \\ &= f(\alpha T_1(x) + \beta T_2(x)) \\ &= \alpha f(T_1(x)) + \beta f(T_2(x)) \\ &= \alpha T_1^*(f)(x) + \beta T_2^*(f)(x) \\ &= (\alpha T_1^* + \beta T_2^*)(f)(x)\end{aligned}$$

So $*(\alpha T_1 + \beta T_2) = (\alpha T_1 + \beta T_2)^* = \alpha T_1^* + \beta T_2^* = \alpha * (T_1) + \beta * (T_2)$. Then we check * is an isometry.
For each $f \in \mathcal{Y}^*$,

$$|T^*(f)| = |f \circ T| \leq \|f\|\|T\|$$

So $\|T^*\| \leq \|T\|$.
For each $x \in \mathcal{X}$ and each $T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, by Hahn-Banach we can find an $f \in \mathcal{Y}^*$ such that

$$f(Tx) = \|Tx\|, \quad \|f\| = 1$$

Hence we have

$$\|T^*\|\|x\| = \|T^*\|\|f\|\|x\| \geq |T^*(f)(x)| = |f(Tx)| = \|Tx\|$$

So $\|T^*\| \geq \|T\|$. Therefore $\|T^*\| = \|T\|$ and T is an isometry. $\square$

Definition 5.5 (weak convergence)
Let $\mathcal{X}$ be a normed vector space, $\{x_n\} \subset \mathcal{X}$, $x \in \mathcal{X}$, say $\{x_n\}$ converges to x weakly if $\forall f \in \mathcal{X}^*$, we have

$$\lim_{n\to\infty} f(x_n) = f(x)$$

denoted by $x_n \rightharpoonup x$. x is called the **weak limit** of $\{x_n\}$.
We can also define the weak convergence of operators.
Let $\mathcal{Y}$ also be a normed vector space and $T_n, T \in \mathcal{L}(\mathcal{X}, \mathcal{Y})$, then $T_n \rightharpoonup T$ if for any $f \in \mathcal{Y}^*$, $x \in \mathcal{X}$

$$\lim_{n\to\infty} f(T_n x) = f(Tx) \text{ (or equivalently, } f \circ T_n \to f \circ T)$$

Proposition 5.1
If weak limit exists, then it's unique. If strong (norm) limit exists then it is also the weak limit.
Proof.
If x and y are both weak limits of $\{x_n\}$, then for each $f \in \mathcal{X}^*$,

$$f(x) = \lim_{n\to\infty} f(x_n) = f(y)$$

Via Hahn-Banach we know that $x = y$.
Also, for each $f \in \mathcal{X}^*$

$$|f(x_n) - f(x)| \le \|f\| \|x_n - x\| \to 0$$

Thus $x_n \rightharpoonup x$. $\square$

Remark.
We have shown that if $x_n \to x$, then $x_n \rightharpoonup x$, but it's not true conversely.
In fact, $x_n \to x$ if and only if $x_n \rightharpoonup x$ and $\|x_n\| \to \|x\|$.

Theorem 5.3 (Mazur)
Let $\mathcal{X}$ be a normed vector space, $x_n \rightharpoonup x_0$, then $\forall \epsilon > 0$, $\exists n \in \mathbb{N}$ and $\lambda_i \ge 0$ $(i = 1, 2, \cdots, n)$ with $\sum_{i=1}^n \lambda_i = 1$ such that

$$\|x_0 - \sum_{i=1}^n \lambda_i x_i\| \le \epsilon$$

Proof.
Let $M = \overline{conv(\{x_n\})}$, then M is a closed convex subset of $\mathcal{X}$. It suffices to show $x_0 \in M$.
Assume that $x_0 \notin M$, then by Ascoli theorem, there exists $f \in \mathcal{X}^*$ and $a \in \mathbb{R}$ such that

$$f(M) < a < f(x_0)$$

Hence for each $n \in \mathbb{N}$, $f(x_n) < a < f(x_0)$, which is contradict to $x_n \rightharpoonup x_0$. $\square$

For any normed vector space $\mathcal{X}$, since X^* is a Banach space, we can also consider weak convergence in X^*.
Suppose $\{f_n\} \subset \mathcal{X}^*$, $f \in \mathcal{X}^*$, then $f_n \rightharpoonup f$ if $\forall x^{**} \in X^{**}$,

$$x^{**}(f_n) \to x^{**}(f)$$

We know $\mathcal{X} \subset \mathcal{X}^{**}$, sometimes it's not necessary to introduce $\mathcal{X}^{**}$ especially for the case that $\mathcal{X}$ is not reflexive.
Hence we shall consider another kind of convergence:

Definition 5.5 (weak-* convergence)
Let $\mathcal{X}$ be a normed vector space, $\{f_n\} \subset \mathcal{X}^*$, $f \in \mathcal{X}^*$, say $\{f_n\}$ is weak-* convergent to f, if $\forall x \in \mathcal{X}$, $f_n(x) \to f(x)$. f is called the **weak-* limit** of $\{f_n\}$, denoted by $f_n \xrightarrow{w*} f$.

Remark.
Since $\mathcal{X} \subset \mathcal{X}^{**}$, the weak convergence on X^* implies the weak-* convergence on $\mathcal{X}^*$, i.e.
if $\{f_n\} \subset \mathcal{X}^*$, $f \in \mathcal{X}^*$, then

$$f_n \rightharpoonup f \Rightarrow f_n \xrightarrow{w*} f$$

Obviously, if $\mathcal{X}$ is reflexive, then they are equivalent.

Theorem 5.4
Let $\mathcal{X}$ be a normed vector space, M^* is a dense subset of $\mathcal{X}^*$, $\{x_n\} \subset \mathcal{X}$, $x \in \mathcal{X}$, then $x_n \rightharpoonup x$ if and only if

(1) $\|x_n\| \leq C$ for some $C \in \mathbb{R}$.
(2) $\forall f \in M^*$, $f(x_n) \to f(x)$.
Proof.
Since $\mathcal{X} \subset \mathcal{X}^{**}$, hence each x_n can be viewed as an operator on $\mathcal{X}^*$, denoted by J_{x_n}.
Then by Banach-Steinhaus it's done. $\square$

Theorem 5.5
Let $\mathcal{X}$ be a Banach space, M is a dense subset of $\mathcal{X}$, $\{f_n\} \subset \mathcal{X}^*$, $f \in \mathcal{X}^*$, then $f_n \xrightarrow{w*} f$ if and only if
(1) $\|f_n\| \leq C$ for some $C \in \mathbb{R}$.
(2) $\forall x \in M$, $f(x_n) \to f(x)$.
Proof.
Apply Banach-Steinhaus. $\square$

Remark.
In general, we have

$$\text{uniform convergence} \Rightarrow \text{strong convergence} \Rightarrow \text{weak convergence}$$

The converse proposition is not true.
Some important counter-examples are as follows.

Example 5.2 (convergent strongly but not uniformly)
Consider l^2 and $T \in l^2$ defined by

$$T : x = (x_1, x_2, \cdots, x_n, \cdots) \mapsto Tx = (x_2, x_3, \cdots, x_n, \cdots)$$

Let $T_n \triangleq T^n$, then

$$T_n x = (x_{n+1}, x_{n+2}, \cdots), \quad \forall x = (x_1, x_2, \cdots) \in l^2$$

Since $T^n(e_{n+1}) = e_1$, $\|e_n\| = 1$, $\forall n \in \mathbb{N}$, so $\|T_n\| \geq \|T_n(e_{n+1})\| = 1$. Hence T_n cannot converge to 0 uniformly, but for each $x \in l^2$,

$$\|T_n x\| = (\sum_{i=1}^{\infty} |x_{n+i}|^2)^{\frac{1}{2}} \to 0 \text{ as } n \to \infty$$

So $T_n \to 0$.

Example 5.3 (convergent weakly but not strongly)
Consider l^2 and $S \in l^2$ defined by

$$S : x = (x_1, x_2, \cdots, x_n, \cdots) \mapsto Tx = (0, x_1, x_2, \cdots, x_n, \cdots)$$

Let $S_n \triangleq S^n$, then $\|S_n x\| = \|x\|$, $\forall x \in l^2$, so $S_n \nrightarrow 0$.
But for each $f = (y_1, y_2, \cdots) \in l^2$,

$$|f(S_n(x))| = |\sum_{i=1}^{\infty} y_{i+n} x_i| \leq (\sum_{i=1}^{\infty} |y_{i+n}|^2)^{\frac{1}{2}} \|x\| \to 0 \text{ as } n \to \infty$$

So $S_n \rightharpoonup 0$. $\square$

Definition 5.6 (weak sequentially compact)
A is called **weak sequentially compact**, if any sequence $\{x_n\} \subset A$ has a weak convergent subsequence.

Definition 5.7 (weak-* sequentially compact)
A is called **weak-* sequentially compact**, if any sequence $\{x_n\} \subset A$ has a weak-* convergent subsequence.

Theorem 5.6
Let $\mathcal{X}$ be a separable normed vector space, then any bounded sequence $\{f_n\} \subset \mathcal{X}^*$ has a weak-* convergent subsequence.

Proof.
Suppose $D = \{x_n\}$ is a countable dense subset of $\mathcal{X}$, $\|f_n\| \le M$.
First consider sequence $\{f_n(x_1)\}$, by Bolzano-Weierstrass there exists a convergent subsequence $\{f_n^{(1)}(x_1)\}$.
Then consider sequence $\{f_n^{(1)}(x_2)\}$, still, we can find a convergent subsequence $\{f_n^{(2)}(x_2)\}$ and so on we obtain countable convergent subsequences like $\{f_n^{(m)}(x_m)\}$, $\forall m \in \mathbb{N}$.
Now define $g_k(x) = f_k^{(k)}(x)$, then for each $x_m \in D$, $g_n(x_m)$ converges.
Since g_n as a subsequence of f_n is bounded, then by Banach-Steinhaus,

$$\forall x \in \mathcal{X}, \lim_{n\to\infty} g_n(x) = g(x)$$

That is,

$$g_n \xrightarrow{w*} g \quad \square$$

Theorem 5.7 (Banach)
Let $\mathcal{X}$ be a normed vector space, if $\mathcal{X}^*$ is separable, then $\mathcal{X}$ is separable.
Proof.
Suppose $\{f_n\}$ is a dense subset of $\mathcal{X}^*$, let

$$g_n = \frac{f_n}{\|f_n\|}$$

Then $\forall g \in S_1 = \{f \in \mathcal{X}^* | \|f\| = 1\}$, we know that there is a subsequence $\{f_{n_k}\}$ such that $f_{n_k} \to g$, and since $\|f_{n_k}\| \to 1$,

$$\begin{aligned}\|g_{n_k} - g\| &\le \|g_{n_k} - f_{n_k}\| + \|f_{n_k} - g\| \\ &= \|\frac{f_{n_k}}{\|f_{n_k}\|} - f_{n_k}\| + \|f_{n_k} - g\| \to 0\end{aligned}$$

Hence $\{g_n\}$ is a countable dense subset of S_1.
Since for each g_n, $\|g_n\| = 1$, so by definition we can find an $x_n \in \mathcal{X}$ with $\|x_n\| = 1$ such that

$$|g_n(x_n)| \ge \frac{1}{2}$$

Let $\mathcal{X}_0 = \overline{span\{x_n\}}$, obviously $\mathcal{X}_0$ is a separable closed subspace of $\mathcal{X}$.
In fact, $\mathcal{X}_0 = \mathcal{X}$, suppose not, then we can pick $y \in \mathcal{X} \setminus \mathcal{X}_0$ with $\|y\| = 1$, by Hahn-Banach, there is an $f \in \mathcal{X}^*$ with $f(\mathcal{X}_0) = 0$, $f(y) = 1$, $\|f\| = 1$.
Since $f \in S_1$, there is a sequence $\{g_{n_k}\}$ such that $g_{n_k} \to f$, then for each $k \in \mathbb{N}$

$$\begin{aligned}\|f - g_{n_k}\| &= \sup_{\|x\|=1} |f(x) - g_{n_k}(x)| \\ &\ge |f(x_{n_k}) - g_{n_k}(x_{n_k})| \\ &= |g_{n_k}(x_{n_k})| \ge \frac{1}{2}\end{aligned}$$

leading a contradiction. $\square$

Theorem 5.8 (Pettis)
Let $\mathcal{X}$ be a reflexive space, then any closed subspace $\mathcal{X}_0 \subset \mathcal{X}$ is also reflexive.
Proof.
We shall show that, if $z_0 \in \mathcal{X}_0^{**}$, then $z_0 \in \mathcal{X}_0$, i.e. $\exists x_0 \in \mathcal{X}_0$ such that

$$z_0(f_0) = f_0(x_0), \quad \forall f_0 \in \mathcal{X}_0^*$$

Define $P : \mathcal{X}^* \to \mathcal{X}_0^*$ by

$$P : f \mapsto f|_{\mathcal{X}_0}$$

Since

$$\|P(f)\| = \|f|_{\mathcal{X}_0}\| \le \|f\|$$

Also it's easy to verify P is linear so $P \in \mathcal{L}(\mathcal{X}^*, \mathcal{X}_0^*)$, $P^* \in \mathcal{L}(\mathcal{X}_0^{**}, \mathcal{X}^{**})$.
Then define $z \triangleq P^*(z_0)$, $z \in \mathcal{X}^{**}$, since $\mathcal{X}$ is reflexive, there is an $x \in \mathcal{X}$ such that

$$z(f) = f(x), \quad \forall f \in \mathcal{X}^*$$

In fact, $x \in \mathcal{X}_0$, suppose not, there exists $g \in \mathcal{X}^*$ with

$$g(x) = 1, \quad g(\mathcal{X}_0) = 0$$

It follows $P(g) = 0$, however

$$0 = \langle z_0, P(g) \rangle = \langle P^*(z_0), g \rangle = z(g) = g(x) = 1$$

leading a contradiction.
So far we have shown that, given $z_0 \in \mathcal{X}_0^{**}$, there is an $x \in \mathcal{X}_0$ such that

$$\langle P^*(z_0), f \rangle = \langle f, x \rangle, \quad \forall f \in \mathcal{X}^*$$

It remains to show that $\langle z_0, f_0 \rangle = \langle f_0, x \rangle$, $\forall f_0 \in \mathcal{X}_0^*$.
By Hahn-Banach, given $f_0 \in \mathcal{X}_0^*$, there is an extension $f \in \mathcal{X}^*$ with

$$f|_{\mathcal{X}_0} = f_0, \quad \|f\| = \|f_0\|$$

Hence $f_0 = P(f)$, and we have

$$\begin{aligned} \langle z_0, f_0 \rangle &= \langle z_0, P(f) \rangle = \langle P^*(z_0), f \rangle \\ &= \langle f, x \rangle = \langle f_0, x \rangle \quad \square \end{aligned}$$

Theorem 5.9 (Eberlein-Smulian)
Let $\mathcal{X}$ be a reflexive space, then the unit ball of $\mathcal{X}$ is weak sequentially compact.
Moreover, the closed unit ball is weak self-sequentially compact.
Proof.
We show that any bounded sequence $\{x_n\} \subset \mathcal{X}$ has a weak convergent subsequence.
Let $\mathcal{X}_0 = \overline{span\{x_n\}}$, obviously $\mathcal{X}_0$ is a closed separable subspace.
By Pettis, since $\mathcal{X}$ is reflexive, $\mathcal{X}_0$ is also reflexive and thus $\mathcal{X}_0^{**}$ is separable, by Banach $\mathcal{X}_0^*$ is also separable.
Now define $J_n = Tx_n$, where T is the canonical map from $\mathcal{X}_0$ to $\mathcal{X}_0^{**}$, then $\{J_n\}$ is also bounded.
Consider separable space $\mathcal{X}_0^*$ and bounded sequence $\{J_n\} \subset X_0^{**}$, apply theorem 5.6, there exists a weak-* convergent subsequence $\{J_{n_k}\}$ and $J_0 \in \mathcal{X}_0^{**}$ with

$$J_{n_k} \xrightarrow{w*} J_0$$

Since X_0^{**} is reflexive, there exists $x_0 = T^{-1}J_0$, so for each $f_0 \in \mathcal{X}_0^*$,

$$f_0(x_{n_k}) = J_{n_k}(f_0) \to J_0(f_0) = f_0(x_0)$$

For any $f \in \mathcal{X}^*$, define $P : f \mapsto f|_{\mathcal{X}_0}$, it's clear that $P \in \mathcal{L}(\mathcal{X}^*, \mathcal{X}_0^*)$, and

$$P^* \in \mathcal{L}(\mathcal{X}_0^{**}, \mathcal{X}^{**}) = \mathcal{L}(\mathcal{X}_0, \mathcal{X})$$

So for each $y_0 \in \mathcal{X}_0$, naturally we have $y_0 = P^*y_0$, moreover

$$\begin{aligned} f(x_{n_k}) &= \langle f, x_{n_k} \rangle = \langle f, P^*(x_{n_k}) \rangle = \langle P(f), x_{n_k} \rangle \\ &= \langle f_0, x_{n_k} \rangle = f_0(x_{n_k}) \to f_0(x_0) = f(x_0) \end{aligned}$$

i.e. $\forall f \in \mathcal{X}^*$, $f(x_{n_k}) \to f(x_0)$, hence $x_{n_k} \rightharpoonup x_0$.
Therefore, any bounded subset of $\mathcal{X}$ is weak sequentially compact, particularly, the unit ball is weak sequentially compact.
Then consider closed unit ball, suppose $x_{n_k} \rightharpoonup x_0$ and $\|x_{n_k}\| \leq 1$, then by Hahn-Banach, there exists $f \in \mathcal{X}^*$ with

$$f(x_0) = \|x_0\|, \quad \|f\| = 1$$

Hence,

$$\|x_0\| = f(x_0) = \lim_{k \to \infty} f(x_{n_k}) \leq \|f\| \sup_k \|x_{n_k}\| \leq 1$$

which implies that x_0 is in the closed unit ball, so the closed unit ball is weak self-sequentially compact. $\square$

6 Spectrum

Definition 6.1 (spectrum)
Let $\mathcal{X}$ be a Banach space, $T : D(T) \subset \mathcal{X} \to \mathcal{X}$ is a closed linear operator.
Suppose $\lambda \in \mathbb{C}$, then if

(1) $\lambda I - T$ is not injective (i.e. $(\lambda I - T)^{-1}$ doesn't exist)
Then λ is an **eigenvalue**, the set of all eigenvalues denoted by $\sigma_p(T)$ is called the **point spectrum** of T.

(2) $\lambda I - T$ is injective, and $R(\lambda I - T) = \mathcal{X}$
Then λ is an **regular value**, the set of all regular values denoted by $\rho(T)$ is called the **resolvent set** of T.
Since T is a closed linear operator, so $\lambda I - T$ is a closed operator hence $(\lambda I - T)^{-1}$ is also closed, by closed graph theorem, $(\lambda I - T)^{-1} \in \mathcal{L}(\mathcal{X})$, therefore $\rho(T)$ can also be defined as

$$\rho(T) = \{\lambda \in \mathbb{C} | (\lambda I - T)^{-1} \in \mathcal{L}(\mathcal{X})\}$$

(3) $\lambda I - T$ is injective, and $R(\lambda I - T) \neq \mathcal{X}$, $\overline{R(\lambda I - T)} = \mathcal{X}$
The set of all these λ's, denoted by $\sigma_c(T)$, is called the **continuous spectrum** of T.

(4) $\lambda I - T$ is injective, and $R(\lambda I - T) \neq \mathcal{X}$, $\overline{R(\lambda I - T)} \neq \mathcal{X}$
The set of all these λ's, denoted by $\sigma_r(T)$, is called the **residual spectrum** of T.

The **spectrum** of T, denoted by $\sigma(T)$, is defined by

$$\sigma(T) = \mathbb{C} \setminus \rho(T)$$

Also, it's clear that

$$\sigma(T) = \sigma_p(T) \cup \sigma_c(T) \cup \sigma_r(T)$$

Note.
For finite-dimensional spaces, $\forall \lambda \in \mathbb{C}$, λ is either eigenvalue or regular value (i.e. $(\lambda I - A)^{-1} \in \mathcal{L}(\mathcal{X})$)

Example 6.1
Let $\mathcal{X} = C[0,1]$, $A : u(t) \mapsto tu(t)$, then A is a bounded linear operator and $\sigma(A) = \sigma_r(A) = [0,1]$
Proof.
$\forall \lambda \notin [0,1]$, then $(\lambda I - A)^{-1} = (\lambda - t)^{-1}$ is linear and bounded since

$$\|\frac{1}{\lambda - t} x(t)\| \leq \sup_{t \in [0,1]} \frac{1}{|\lambda - t|} \|x\|$$

So $\sigma(A) \subset [0,1]$.
$\forall \lambda \in [0,1]$, the unique solution of equation

$$(\lambda - t)u(t) = 0$$

is $u(t) \equiv 0$, so $\lambda I - A$ is injective.
And for each $v \in R(\lambda I - A)$, then $v(\lambda) = 0$, hence $1 \notin \overline{R(\lambda I - A)}$ Hence $[0,1] \subset \sigma_r(A)$.
Since $[0,1] \subset \sigma_r(A) \subset \sigma(A) \subset [0,1]$, we have

$$\sigma_r(A) = \sigma(A) = [0,1] \quad \square$$

Example 6.2
Let $\mathcal{X} = L^2[0,1]$, $A : u(t) \mapsto tu(t)$, then A is a bounded linear operator and $\sigma(A) = \sigma_c(A) = [0,1]$
Proof.
Similarly we have $\sigma(A) \subset [0,1]$.

$\forall\lambda \in [0,1]$, $\lambda I - A$ is injective.
Note that

$$\frac{1}{\lambda - t} \notin \mathcal{X}$$

Thus $1 \notin R(\lambda I - A)$, so $R(\lambda I - A) \neq \mathcal{X}$.
Fix $\lambda \in [0,1]$, $\forall f \in \mathcal{X}$ and $\forall \epsilon > 0$, we can define g by

$$g(x) = \begin{cases} 0, & x \in B(\lambda, \epsilon) \\ f(x), & \text{otherwise} \end{cases}$$

Since

$$\|f - g\|_{L^2}^2 = \int_{B(\lambda,\epsilon)\cap[0,1]} |f|^2 d\mu \to 0 \text{ as } \epsilon \to 0$$

So $f \in \overline{R(\lambda I - A)}$, i.e. $\overline{R(\lambda I - A)} = \mathcal{X}$, hence $[0,1] \subset \sigma_c(A)$ and therefore

$$\sigma_c(A) = \sigma(A) = [0,1] \quad \square$$

Definition 6.2 (resolvent)
Given closed linear operator A, consider operator-value function

$$R_\lambda(A) : \rho(A) \to \mathcal{L}(\mathcal{X}), \lambda \mapsto (\lambda I - A)^{-1}, \quad \forall\lambda \in \rho(A)$$

$R_\lambda(A)$ is called the **resolvent** of A.

Theorem 6.1
Let $T \in \mathcal{L}(\mathcal{X})$, $\|T\| < 1$, then $(I - T)^{-1} \in \mathcal{L}(\mathcal{X})$, and

$$\|(I - T)^{-1}\| \leq \frac{1}{1 - \|T\|}$$

Proof(Version 1).
Suppose $(I - T)(x) = 0$, then $x = Ix = Tx$. $\|T\| < 1$ implies $x = 0$, so $I - T$ is injective and its inverse exists. Obviously $(I - T)^{-1}$ is linear.
Now given any $y \in \mathcal{X}$, consider $S_y : \mathcal{X} \to \mathcal{X}, x \mapsto y + Tx$, then for any $x_1, x_2 \in \mathcal{X}$,

$$\|S_y x_1 - S_y x_2\| = \|Tx_1 - Tx_2\| \leq \|T\|\|x_1 - x_2\| < \|x_1 - x_2\|$$

So S_y is a contraction mapping and thus the fixed point x exists, thus we have $x = S_y x = y + Tx$, it follows

$$\|x\| = \|y + Tx\| \leq \|y\| + \|T\|\|x\|$$

$$\|x\| \leq \frac{\|y\|}{1 - \|T\|}$$

Also,

$$x = (I - T)^{-1}y, \quad \|(I - T)^{-1}y\| = \|x\| \leq \frac{\|y\|}{1 - \|T\|}$$

i.e.

$$\|(I - T)^{-1}\| \leq \frac{1}{1 - \|T\|} \quad \square$$

Proof(Version 2).
Since

$$\sum_{k=0}^{n} T^k(I - T) = I - T^{n+1}$$

And

$$\|T^n - 0\| = \|T^n\| \leq \|T\|^n \to 0$$

We have

$$\sum_{k=0}^{\infty} T^k(I-T) = I$$

which implies

$$(I-T)^{-1} = \sum_{k=0}^{\infty} T^k$$

So

$$\|(I-T)^{-1}\| = \|\sum_{k=0}^{\infty} T^k\| \leq \sum_{k=0}^{\infty} \|T\|^k = \frac{1}{1-\|T\|}$$

And it's easy to verify $\sum_{k=0}^{\infty} T^k$ is linear, hence $(I-T)^{-1} \in \mathcal{L}(\mathcal{X})$. □

Remark.
If $\|T\| < 1$, then

$$(I-T)^{-1} = \sum_{k=0}^{\infty} T^k$$

Corollary 6.1
Let A be a closed linear operator, then $\rho(A)$ is open.
Proof.
Suppose $\lambda_0 \in \rho(A)$, then

$$\begin{aligned}\lambda I - A &= (\lambda-\lambda_0)I + (\lambda_0 I - A) \\ &= (\lambda_0 I - A)(I + (\lambda-\lambda_0)(\lambda_0 I - A)^{-1})\end{aligned}$$

when $|\lambda - \lambda_0| < \|(\lambda_0 I - A)^{-1}\|^{-1}$, define B by

$$B \triangleq [I + (\lambda-\lambda_0)(\lambda_0 I - A)^{-1}]^{-1}$$

Then

$$\|(\lambda-\lambda_0)(\lambda_0 I - A)^{-1}\| \leq |\lambda-\lambda_0|\|(\lambda_0 I - A)^{-1}\| < 1$$

Hence

$$\|B\| \leq \frac{1}{1-\|(\lambda-\lambda_0)(\lambda_0 I - A)^{-1}\|} < \infty$$

i.e. $B \in \mathcal{L}(\mathcal{X})$.
And thus

$$\begin{aligned}(\lambda I - A)^{-1} &= [(\lambda_0 I - A)B^{-1}]^{-1} \\ &= B(\lambda_0 I - A)^{-1} \\ &= BR_{\lambda_0}(A) \in \mathcal{L}(\mathcal{X})\end{aligned}$$

That is, for each $\lambda_0 \in \rho(A)$, we can find a ball

$$B(\lambda_0, \epsilon) \subset \rho(A), \text{ where } \epsilon = \|(\lambda_0 I - A)^{-1}\|^{-1}$$

Therefore $\rho(A)$ is open. □

Theorem 6.2 (first resolvent identity)
Let $\lambda, \mu \in \rho(A)$, then

$$R_\lambda(A) - R_\mu(A) = (\mu-\lambda)R_\lambda(A)R_\mu(A)$$

Proof.

$$\begin{aligned}(\lambda I - A)^{-1} &= (\lambda I - A)^{-1}(\mu I - A)(\mu I - A)^{-1} \\ &= (\lambda I - A)^{-1}((\mu-\lambda)I + \lambda I - A)(\mu I - A)^{-1} \\ &= (\mu-\lambda)(\lambda I - A)^{-1}(\mu I - A)^{-1} + (\mu I - A)^{-1}\end{aligned}$$

i.e.
$$R_\lambda(A) = (\mu - \lambda)R_\lambda(A)R_\mu(A) + R_\mu(A) \quad \square$$

Remark.
Given any $\lambda \in \rho(A)$, $R_\lambda(A) = (\lambda I - A)^{-1}$ exists, and $R_\lambda(A) \neq 0$ since 0 is not invertable.
Hence for any $\lambda, \mu \in \rho(A)$, by first resolvent identity, $\lambda = \mu$ if and only if $R_\lambda(A) = R_\mu(A)$.

Theorem 6.3
$R_\lambda(A)$ is an operator-value holomorphic function on $\rho(A)$.
Proof.
First we show $R_\lambda(A)$ is continuous.
Let $\lambda_0 \in \rho(A)$, suppose
$$|\lambda - \lambda_0| < \frac{1}{2\|(\lambda_0 I - A)^{-1}\|}$$
It follows
$$\|(\lambda - \lambda_0)(\lambda_0 I - A)^{-1}\| \leq |\lambda - \lambda_0|\|(\lambda_0 I - A)^{-1}\| < \frac{1}{2}$$
Then define B by
$$B \triangleq [I + (\lambda - \lambda_0)(\lambda_0 I - A)^{-1}]^{-1}$$
We have
$$\|B\| \leq \frac{1}{1 - \|(\lambda - \lambda_0)(\lambda_0 I - A)^{-1}\|} < 2$$
and
$$\begin{aligned}\|R_\lambda(A)\| &= \|BR_{\lambda_0}(A)\| \\ &\leq \|B\|\|R_{\lambda_0}(A)\| \\ &< 2\|R_{\lambda_0}(A)\|\end{aligned}$$
Then by first resolvent identity,
$$\begin{aligned}\|R_\lambda(A) - R_{\lambda_0}(A)\| &\leq |\lambda - \lambda_0|\|R_\lambda(A)\|\|R_{\lambda_0}(A)\| \\ &\leq 2|\lambda - \lambda_0|\|R_{\lambda_0}(A)\|^2 \to 0\end{aligned}$$
as $\lambda \to \lambda_0$.
Then we show that $R_\lambda(A)$ is differentiable, still, by first resolvent identity,
$$\lim_{\lambda \to \lambda_0} \frac{R_\lambda(A) - R_{\lambda_0}(A)}{\lambda - \lambda_0} = -\lim_{\lambda \to \lambda_0} R_\lambda(A)R_{\lambda_0}(A) = -[R_{\lambda_0}(A)]^2$$
(The last equality holds since $R_\lambda(A)$ is continuous) $\square$

Theorem 6.4
Let A be a bounded linear operator, then $\sigma(A) \neq \varnothing$.
Proof.
Assume that $\rho(A) = \mathbb{C}$, then $R_\lambda(A)$ is holomorphic on $\mathbb{C}$.
When $|\lambda| > 2\|A\|$, we have
$$R_\lambda(A) = \frac{1}{\lambda}(I - \frac{A}{\lambda})^{-1}$$
$$\|R_\lambda(A)\| \leq \frac{1}{|\lambda|} \cdot \frac{1}{1 - \|\frac{A}{\lambda}\|} = \frac{1}{|\lambda| - \|A\|} \leq \frac{1}{\|A\|} < \infty$$
So $R_\lambda(A)$ is bounded on $\mathbb{C}$, say $\|R_\lambda(A)\| \leq M$.
For each $f \in (\mathcal{L}(\mathcal{X}))^*$, consider function
$$u_f(\lambda) \triangleq f(R_\lambda(A))$$
Note that u_f is holomorphic on $\mathbb{C}$ since f as a continuous linear functional is holomorphic, and
$$|u_f(\lambda)| = |f(R_\lambda(A))| \leq \|f\|\|R_\lambda(A)\| \leq \|f\|M < \infty$$

implies u_f is a bounded entire function, by Liouville

$$u_f(\lambda) \equiv C_f$$

where C_f is a constant which doesn't depend on λ.
Since f is arbitrary, by Hahn-Banach, $R_\lambda(A)$ is also a constant which doesn't depend on λ, which is contradict to the first resolvent identity. $\square$

Remark.
We know that for finite-dimensional Banach spaces, each bounded linear operator can be viewed as matrix whose eigenvalue always exists (and thus it has non-empty spectrum). Theorem 6.4 implies any bounded linear operator has non-empty spectrum even for infinite-dimensional spaces.

Definition 6.3 (spectral radius)
Let $A \in \mathcal{L}(\mathcal{X})$, the **spectral radius** of A is defined by

$$r_\sigma(A) \triangleq \sup\{|\lambda| | \lambda \in \sigma(A)\}$$

Note.
If $|\lambda| > \|A\|$, then

$$(\lambda I - A)^{-1} = \sum_{k=0}^{\infty} \frac{A^k}{\lambda^{k+1}} \in \mathcal{L}(\mathcal{X})$$

So $\lambda \notin \sigma(A)$, this implies

$$r_\sigma(A) \leq \|A\|$$

Theorem 6.5 (Gelfand)
Let $\mathcal{X}$ be a Banach space, $A \in \mathcal{L}(\mathcal{X})$, then

$$r_\sigma(A) = \|A^n\|^{\frac{1}{n}}$$