MA4211 Functional Analysis

Thang Pang Ern, Loh Wei Xuan Ryan, and Malcolm Tan Jun Xi

Contents

References

1 Kreyszig, E. (1989). *Introductory Functional Analysis with Applications*. Wiley.
2 Sasane, A. (2017). *A Friendly Approach to Functional Analysis*. World Scientific.

Chapter 1
Metric Spaces

1.1
Metric Spaces

Functional Analysis is essentially the study of infinite-dimensional Linear Algebra.

Example 1.1 (Euclidean metric/distance). Recall the familiar metric in Euclidean space $\mathbb{R}$

$$d(x,y) = |x-y|.$$

We call this the Euclidean metric or Euclidean distance. Naturally, we can extend this to the Euclidean 2-space $\mathbb{R}^2$. Consider $\mathbf{x} = (x_1, x_2)$ and $\mathbf{y} = (y_1, y_2)$ in $\mathbb{R}^2$. Then,

$$d(\mathbf{x},\mathbf{y}) = \sqrt{(x_1-x_2)^2 + (y_1-y_2)^2}.$$

We give the definition of a metric space (Definition 1.1).

Definition 1.1 (metric space). Let X be a set. A metric space is an ordered pair (X,d) equipped with a distance function $d : X \times X \to \mathbb{R}$ such that the following properties are satisifed:

(i) Non-negativity: $d(x,y) \geq 0$

(ii) Positive-definiteness: $x = y$ if and only if $d(x,y) = 0$

(iii) Symmetry: for all $x,y \in X$, we have $d(x,y) = d(y,x)$

(iv) Triangle inequality: for all x,y,z, we have $d(x,z) \leq d(x,y) + d(y,z)$

Example 1.2 (MA4211 AY24/25 Sem 2 Tutorial 1). Given two non-empty subsets A,B of a metric space (X,d), their distance is defined as

$$D(A,B) = \inf_{a\in A, b\in B} d(a,b).$$

Consider the power set of X and the function D. Which of the axioms of a metric space does this pair satisfy?

Solution. We first claim that non-negativity is satisfied. For $A,B \subseteq X$, we have

$$d(a,b) \geq 0 \quad \text{which implies} \quad \inf_{a\in A, b\in B} d(a,b) \geq 0.$$

Next, symmetry is satisfied since

$$D(B,A) = \inf_{a\in A, b\in B} d(b,a) = \inf_{a\in A, b\in B} d(a,b)$$

which follows from the fact that d satisfies symmetry. Next, the triangle inequality is satisfied. To see why, let $A,B,C \subseteq X$. Then,

$$D(A,C) = \inf_{a\in A, c\in C} d(a,c) \leq \inf_{a\in A, b\in B} d(a,b) + \inf_{b\in B, c\in C} d(b,c) = D(A,B) + D(B,C)$$

We claim that the property $D(A,B) = 0$ if and only if $A = B$ is not satisfied. Recall from Definition 1.1 that this is called positive-definiteness. Anyway, to see why, let $A = \{x\}$ and $B = \{x_n\}_{n=1}^{\infty}$, where $x_n \to x$. Then, $D(A,B) = 0$ since x_n can be arbitrarily close to x for large n. However, $A \neq B$. $\square$

Example 1.3 (MA4211 AY24/25 Sem 2 Homework 1)**.** Let (X,d) be a metric space. Define

$$\rho(x,y) = \frac{d(x,y)}{1+d(x,y)}.$$

Show that ρ is a metric on X and that the identity map on X is a homeomorphism (a continuous, bijective map with continuous inverse) between (X,d) and (X,ρ). Thus, every metric space is homeomorphic to a space with a bounded metric.

Solution. We first show that ρ is a metric on X. Note that non-negativity, symmetry and homogeneity are obvious (as usual) so it suffices to only prove that it satisfies the triangle inequality. Let $x,y,z \in X$ such that

$$d(x,z) \le d(x,y) + d(y,z) \quad \text{since} \quad d \text{ satisfies the triangle inequality.}$$

As such,

$$\rho(x,y) + \rho(y,z) - \rho(x,z) = \frac{d(x,y)}{1+d(x,y)} + \frac{d(y,z)}{1+d(y,z)} - \frac{d(x,z)}{1+d(x,z)}$$

Let $a = d(x,y)$, $b = d(y,z)$, and $c = d(x,z)$. Then, the above becomes

$$\begin{aligned}
\frac{a}{1+a} + \frac{b}{1+b} - \frac{c}{1+c} &= \frac{a(1+b)(1+c) + b(1+a)(1+c) - c(1+a)(1+b)}{(1+a)(1+b)(1+c)} \\
&= \frac{a+ab+ac+abc+b+ab+bc+abc-c-ac-bc-abc}{(1+a)(1+b)(1+c)} \\
&= \frac{a+b-c+2ab+abc}{(1+a)(1+b)(1+c)}
\end{aligned}$$

Since $a+b-c \ge 0$ is equivalent to $d(x,z) \le d(x,y) + d(y,z)$, then

$$\frac{a+b-c+2ab+abc}{(1+a)(1+b)(1+c)} \ge \frac{2ab+abc}{(1+a)(1+b)(1+c)} \ge 0$$

and we conclude that ρ is indeed a metric on X.

For the second part, we wish to show that

$$\mathrm{id}_X : (X,d) \to (X,\rho) \quad \text{where} \quad x \mapsto x \quad \text{is a homeomorphism.}$$

Note that a map is continuous if it takes convergent sequences in the domain metric to convergent sequences in the codomain metric. Suppose $x_n \to x$ in (X,d). Then, $d(x_n,x) = 0$. As such, $\rho(x_n,x) = 0$. Similarly, we can consider the inverse map

$$\mathrm{id}_X = (\cdot)^{-1} : (X,\rho) \to (X,d) \quad \text{where} \quad x \mapsto x.$$

Again, it is clear that $\rho(x_n,x) = 0$ implies $d(x_n,x) = 0$, so id_X is indeed a homeomorphism. $\square$

Definition 1.2 ($\mathbb{R}^\infty$)**.** Define $\mathbb{R}^\infty$ to be the space of all infinite sequences of real numbers, i.e. $(x_1,x_2,\ldots)$ where $x_1,x_2,\ldots \in \mathbb{R}$.

Example 1.4 ($\mathbb{R}^\infty$)**.** We have the infinite sequences $(0,0,\ldots,)$ and $(1,2,3,\ldots,100,\ldots)$ in $\mathbb{R}^\infty$.

Example 1.5. For $X = \mathbb{R}$, we can define

$$d(x,y) = \min\{|x-y|, 1\} \quad \text{such that it is a metric.}$$

Example 1.6. For $X = \mathbb{R}^\infty$, let $\mathbf{x} = (x_1, x_2, \ldots)$ and $\mathbf{y} = (y_1, y_2, \ldots)$, where each element is in $\mathbb{R}$. Then, one can check that

$$d(\mathbf{x}, \mathbf{y}) = \sup d(x_i, y_i) \quad \text{is a metric.}$$

Example 1.7 (ℓ^∞). We give an introduction to the sequence space ℓ^∞. This example gives one an impression of how surprisingly general the concept of a metric space is. We can define

$$X = \{\text{bounded sequences of complex numbers}\}.$$

So, every element of X is a complex sequence ξ_j such that for all $j = 1, 2, \ldots$, we have

$$|\xi_j| \leq c_x \quad \text{where} \quad c_x \text{ is a real number which may depend on } x.$$

Then, the following is a metric:

$$d(\mathbf{x}, \mathbf{y}) = \sup_{j \in \mathbb{N}} |\xi_j - \eta_j| \quad \text{where} \quad y = (\eta_1, \eta_2, \ldots) \in X$$

Definition 1.3 (translation invariant and homogeneous metrics). Let X be a normed space and d be a metric. We say that d is translation invariant if

$$\text{for all } x, y, z \in X \quad \text{we have} \quad d(x, y) = d(x + z, y + z).$$

Similarly, d is said to be homogeneous if

$$\text{for all } x, y \in X \text{ and } \alpha \in \mathbb{F} \quad \text{we have} \quad d(\alpha x, \alpha y) = |\alpha| d(x, y).$$

Example 1.8 (MA4211 AY24/25 Sem 2 Tutorial 2). Let X be a normed space and d be a metric. We say that d is translation invariant if

$$\text{for all } x, y, z \in X \quad \text{we have} \quad d(x, y) = d(x + z, y + z).$$

Similarly, d is said to be homogeneous if

$$\text{for all } x, y \in X \text{ and } \alpha \in \mathbb{F} \quad \text{we have} \quad d(\alpha x, \alpha y) = |\alpha| d(x, y).$$

(a) Show that the metric induced on X by its norm is translation invariant and homogeneous.

(b) Show that if d is a translation invariant and homogeneous metric on X, then $\|x\| = d(x, 0)$ defines a norm on X.

Solution.

(a) We have

$$\|x - y\| = \|(x - z) + (z - y)\| = \|x + z\| + \|y + z\|$$

so d is translation invariant.

Next, we prove that d is homogeneous. We have

$$\|\alpha x - \alpha y\| = |\alpha| \|x - y\|.$$

(b) Since d is non-negative, then $\|\cdot\|$ is also non-negative. Positive-definiteness is also clear. Then,

$$d(\alpha x, 0) = \|\alpha x\| = |\alpha| \, \|x\| = |\alpha| \, d(x, 0)$$

so homogenity holds. Lastly, we prove that $\|\cdot\|$ satisfies the triangle inequality. We have

$$\begin{aligned}
\|x - z\| &= d(x - z, 0) \\
&= d(x, z) \quad \text{since } d \text{ is translation invariant} \\
&\le d(x, y) + d(y, z) \quad \text{since } d \text{ satisfies the triangle inequality} \\
&= d(x - y, 0) + d(y - z, 0) \quad \text{since } d \text{ is translation invariant} \\
&= \|x - y\| + \|y - z\|
\end{aligned}$$

so $\|\cdot\|$ defines a norm on X. □

Definition 1.4 (function space). A function space is a set of functions holding the properties of a vector space structure, norm, or inner product. In particular, it has either of the following properties:

- **Vector space structure:**
  - **Closure under addition:** for any $f, g \in \mathcal{C}[a, b]$, we have $f + g \in \mathcal{C}[a, b]$
  - **Closure under scalar multiplication:** for any $k \in \mathbb{R}$, $kf \in \mathcal{C}[a, b]$ for $f \in \mathcal{C}[a, b]$
- **Norm:** A norm $\|\cdot\|$ is a function $\|\cdot\| : \mathcal{C}[a, b] \to \mathbb{R}$ that satisfies the following:
  - **Non-negativity:** $\|f\| \ge 0$ for all $f \in \mathcal{C}[a, b]$, and $\|f\| = 0$ if and only if $f = 0$
  - **Scalar multiplication:** $\|kf\| = |k|\|f\|$ for any $k \in \mathbb{R}$ and $f \in \mathcal{C}[a, b]$
  - **Triangle inequality:** $\|f + g\| \le \|f\| + \|g\|$ for all $f, g \in \mathcal{C}[a, b]$
- **Inner product:** An inner product $\langle \cdot, \cdot \rangle$ is a function $\langle \cdot, \cdot \rangle : \mathcal{C}[a, b] \times \mathcal{C}[a, b] \to \mathbb{R}$ that satisfies the following:
  - **Conjugate symmetry:** $\langle f, g \rangle = \langle g, f \rangle$
  - **Linearity in the first argument:** $\langle kf + g, h \rangle = k\langle f, h \rangle + \langle g, h \rangle$ for any $k \in \mathbb{R}$
  - **Positive-definiteness:** $\langle f, f \rangle \ge 0$ for all $f \in \mathcal{C}[a, b]$, and $\langle f, f \rangle = 0$ if and only if $f = 0$

Example 1.9. Let

$$\mathcal{C}[a, b] \quad \text{denote} \quad \text{the set of continuous functions on } [a, b].$$

Example 1.10 (function space). Let $X = \mathcal{C}[a, b]$, for which we recall that this refers to the set of continuous functions on $[a, b]$. Let $f, g \in \mathcal{C}[a, b]$. Then,

$$d(f, g) = \max_{x \in [a, b]} |f(x) - g(x)| \quad \text{and} \quad d(f, g) = \sqrt{\int_0^L |f(x) - g(x)|^2 \, dx} \quad \text{are metrics.}$$

Example 1.11 (Hamming distance). Consider the two English words 'word' and 'wind' of the same length for which the second and third letters differ. Since two letters differ, we say that their Hamming distance is 2. We write

$$d(\text{wind}, \text{word}) = 2.$$

In this case, d is a metric. The reader can read Kreyszig p. 9 Question 10 to prove that the Hamming distance is indeed a metric.

Example 1.12 (Hamming metric; Kreyszig p. 9 Question 10). Let X be the set of all ordered triples of zeros and ones. Show that X consists of eight elements and a metric d on X defined by

$$d(x,y) = \text{number of places where } x \text{ and } y \text{ have different entries.}$$

Solution. First, non-negativity clearly holds as it is impossible for two words of the same length to differ by a negative number of letters. So, $d(x,y) \geq 0$. Symmetry is also obvious.

For positive-definiteness, suppose $x = y$. Then, x and y are the same word. Hence, they are two words of the same length which differ by 0 letters. By definition of d, we have $d(x,y) = 0$. Similarly, if $d(x,y) = 0$, then x and y are two words of the same length that are of Hamming distance 0, which implies that they differ by 0 letters. As such, $x = y$.

Lastly, we prove that d satisfies the triangle inequality. Let x, y and z be three words of length n. We can explicitly define d as follows:

$$d(x,y) = \sum_{i=1}^{n} \mathbf{1}_{\{x_i \neq y_i\}} \quad \text{where} \quad \mathbf{1}_{\{x_i \neq y_i\}} = \begin{cases} 1 & \text{if } x_i \neq y_i; \\ 0 & \text{if } x_i = y_i \end{cases}$$

We note that for each position $1 \leq i \leq n$, the inequality

$$\mathbf{1}_{\{x_i \neq z_i\}} \leq \mathbf{1}_{\{x_i \neq y_i\}} + \mathbf{1}_{\{y_i \neq z_i\}}.$$

To see why, if $x_i = z_i$, then $\mathbf{1}_{\{x_i \neq z_i\}} = 0$ and the inequality holds trivially since $\mathbf{1}_{\{x_i \neq y_i\}}$ and $\mathbf{1}_{\{y_i \neq z_i\}}$ are non-negative. If $x_i \neq z_i$, then either $x_i \neq y_i$, or $y_i \neq z_i$, or both. Hence, at least one of $\mathbf{1}_{\{x_i \neq y_i\}}$ or $\mathbf{1}_{\{y_i \neq z_i\}}$ is 1, and the inequality holds. Hence,

$$\sum_{i=1}^{n} \mathbf{1}_{\{x_i \neq z_i\}} \leq \sum_{i=1}^{n} \mathbf{1}_{\{x_i \neq y_i\}} + \sum_{i=1}^{n} \mathbf{1}_{\{y_i \neq z_i\}} \quad \text{or equivalently} \quad d(x,z) \leq d(x,y) + d(y,z).$$

so the triangle inequality is satisfied. $\square$

Definition 1.5 (ℓ^p-space). Let $p \geq 1$ be a fixed real number. Each element in the space ℓ^p is a sequence $(x_1, \ldots)$ such that $|x_1|^p + \ldots$ converges. So,

$$\ell^p = \left\{ \mathbf{x} \in \mathbb{R}^\infty : \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p} < \infty \right\}.$$

Definition 1.6 (p-norm). Every element in ℓ^p-space is equipped with a norm, known as the p-norm. We define it as follows (will not be strict with the use of either x or $\mathbf{x}$):

$$\text{if } x \in \ell^p \quad \text{then} \quad \|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p}$$

Example 1.13 (MA4211 AY24/25 Sem 2 Tutorial 1). Give an example of a sequence x_n such that

$$x_n \to 0 \quad \text{but} \quad x_n \notin \ell^p \text{ for any } 1 \leq p < \infty.$$

Solution. Consider

$$x_n = \frac{1}{\log(n+1)} \quad \text{for which} \quad \lim_{n\to\infty} x_n = 0.$$

Then, we have

$$\|x_n\|_p = \frac{1}{[\log(n+1)]^p} \quad \text{so} \quad \sum_{n=1}^{\infty} \|x_n\|_p = \sum_{n=1}^{\infty} \frac{1}{[\log(n+1)]^p} \text{ diverges.}$$

□

Example 1.14 (MA4211 AY24/25 Sem 2 Homework 1)**.** Recall that for $1 \le p < \infty$, we defined

$$\ell^p = \left\{ x_i \in \mathbb{R}^\infty \mid \sum_{i=1}^{\infty} |x_i|^p < \infty \right\} \quad \text{with norm} \quad \|x\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p}.$$

Show that if $p \le q$, then $\|x\|_q \le \|x\|_p$ and hence $\ell^p \subseteq \ell^q$.

Solution. We have

$$\|x\|_q = \left(\sum_{i=1}^{\infty} |x_i|^q \right)^{1/q} = \left[\left(\sum_{i=1}^{\infty} |x_i|^q \right)^{p/q} \right]^{1/p} \le \left(\sum_{i=1}^{\infty} |x_i|^{q(p/q)} \right)^{1/p} = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p}$$

The inequality holds because

$$\left[\left(\sum_{i=1}^{\infty} |x_i|^q \right)^{p/q} \right]^{1/p} \le \left(\sum_{i=1}^{\infty} |x_i|^{q(p/q)} \right)^{1/p} \quad \text{is equivalent to} \quad \left(\sum_{i=1}^{\infty} |x_i|^q \right)^{p/q} \le \sum_{i=1}^{\infty} |x_i|^{q(p/q)}$$

Let

$$a_i = |x_i|^q.$$

Then, it suffices to prove that

$$\left(\sum_{i=1}^{\infty} a_i \right)^r \le \sum_{i=1}^{\infty} a_i^r \quad \text{where} \quad r = \frac{p}{q} \le 1.$$

To see why this holds, note that for all $a, b \ge 0$ and $0 < r \le 1$, we have

$$(a+b)^r \le a^r + b^r.$$

Let $f(a) = a^r + b^r - (a+b)^r$. Then, $f'(a) = ra^{r-1} - r(a+b)^{r-1} \ge 0$. Since $f(0) = 0$, then $f(a) \ge 0$. As such, the inequality $(a+b)^r \le a^r + b^r$ holds. We can extend it to any finite sum by induction; for an infinite sum, use a limiting argument. As such, for any sequence in ℓ^p, it must be contained in ℓ^q. □

Definition 1.7 (L^p)**.** Define the L^p space as follows:

$$L^p[a,b] = \left\{ f : [a,b] \to \mathbb{R} \quad \text{such that} \quad \int_a^b |f(x)|^p \, dx < \infty \right\}$$

Theorem 1.1 (Young's inequality)**.** Suppose $\alpha, \beta > 0$. Then,

$$\alpha\beta \le \frac{\alpha^p}{p} + \frac{\beta^q}{q} \quad \text{where} \quad \frac{1}{p} + \frac{1}{q} = 1.$$

Example 1.15. Let $t = 1/p$. Then,

$$\begin{aligned}\ln\left(t\alpha^p + (1-t)\beta^q\right) &\geq t\ln\left(\alpha^p\right) + (1-t)\ln\left(\beta^q\right) \quad \text{since } \ln \text{ is concave down}\\ &= \frac{1}{p}\ln\left(\alpha^p\right) + \frac{1}{q}\ln\left(\beta^q\right)\\ &= \ln\alpha + \ln\beta\\ &= \ln\alpha\beta\end{aligned}$$

Taking exponentials on both sides yields the desired result.

Theorem 1.2 (Hölder's inequality). For $1 \leq p, q < \infty$ such that $1/p + 1/q = 1$, we have

$$\sum_{i=1}^{\infty} |x_i y_i| \leq \|x\|_p \|y\|_q .$$

Theorem 1.3 (Minkowski's inequality). For $1 \leq p < \infty$, we have

$$\|x+y\|_p \leq \|x\|_p + \|y\|_p .$$

Please refer to my MA4262 notes for proofs of Theorems 1.2 and 1.3. Minkowski's inequality (Theorem 1.3) looks like a generalisation of the triangle inequality to some ℓ^p space, where $1 \leq p < \infty$. We will see in Example 1.16 that the inequality does not hold if $0 < p < 1$ as it would violate the triangle inequality.

Example 1.16 (MA4211 AY24/25 Sem 2 Tutorial 2). Consider the space ℓ^p for $0 < p < 1$, with $\|\cdot\|_p$ defined as usual. Which of the properties of a norm does $\|\cdot\|_p$ satisfy?

Solution. We claim that $\|\cdot\|_p$ satisfies all properties of a norm in Definition 2.6 except the triangle inequality. Verifying the other properties is easy. We have

$$\|(1,0)+(0,1)\|_p = \|(1,1)\|_p = 2^{1/p} \geq 2 = 1+1 = \|(1,0)\|_p + \|(0,1)\|_p$$

so the triangle inequality is not satisfied. ☐

Example 1.17 (MA4211 AY24/25 Sem 2 Tutorial 1). Prove the reverse triangle inequality

$$|d(x,z) - d(z,y)| \leq d(x,y) \quad \text{for all } x,y,z \text{ in a metric space } (X,d).$$

Solution. Recall the triangle inequality which states that

$$d(x,z) \leq d(x,y) + d(y,z) \quad \text{for } x,y,z \text{ in a metric space } (X,d).$$

So,

$$d(x,z) - d(z,y) \leq d(x,y) \quad \text{where we used the fact that the metric } d \text{ is symmetric.}$$

If $d(x,z) - d(z,y) \geq 0$, then the result follows by taking absolute value; otherwise, we now consider the case where $d(x,z) < d(z,y)$. So,

$$d(x,y) \geq d(y,z) - d(x,z) > 0.$$

Taking absolute value again yields the desired result. ☐

1.2 Convergence, Completeness, and Topology

Definition 1.8 (convergence of sequence)**.** Let x_n be a sequence in $\mathbb{R}$. We say that

$$x_n \to x \quad \text{if} \quad \text{there exists } x \in \mathbb{R} \text{ such that } \lim_{n\to\infty} d(x_n, x) = 0.$$

One should recall from MA2108 that this is equivalent to saying that

$$\text{for all } \varepsilon > 0 \text{ there exists } N \in \mathbb{N} \quad \text{such that for all } n \geq N \text{ we have } d(x_n, x) < \varepsilon.$$

Definition 1.9 (continuous function)**.** Let (X, d_X) and (Y, d_Y) be metric spaces. We say that f is continuous at $x_0 \in X$ if

$$\text{for all } \varepsilon > 0 \text{ there exists } \delta > 0 \quad \text{such that } d_X(x, x_0) < \delta \text{ implies } d_Y(f(x), f(x_0)) < \varepsilon.$$

Theorem 1.4. Let (X, d_X) and (Y, d_Y) be metric spaces and $T : X \to Y$ be a map. Then,

$$T \text{ is continuous} \quad \text{if and only if} \quad x_n \to x \text{ implies } T(x_n) \to T(x).$$

Definition 1.10 (isometry)**.** Let (X, d_X) and (Y, d_Y) be metric spaces and $T : X \to Y$. We say that

$$T \text{ is an isometry} \quad \text{if} \quad d_X(x_1, x_2) = d_Y(T(x_1), T(x_2)).$$

Definition 1.11 (isometric spaces)**.** If there exists a bijective isometry T between two metric spaces X and Y, then we say that X and Y are isometric.

Definition 1.12 (open and closed intervals)**.** Let

$$(a, b) \text{ and } [a, b] \quad \text{denote} \quad \text{the open interval and closed interval in } \mathbb{R} \text{ respectively.}$$

Definition 1.13 (open and closed balls)**.** Define

$$B(x, r) = \{y \in \mathbb{R}^n : d(x, y) < r\} \quad \text{denote} \quad \text{the open ball in } \mathbb{R}^n$$
$$\overline{B}(x, r) = \{y \in \mathbb{R}^n : d(x, y) \leq r\} \quad \text{denote} \quad \text{the closed ball in } \mathbb{R}^n$$

Here, each ball is centred at x and is of radius r.

Definition 1.14 (open set)**.** A subset $A \subseteq X$ is open if

$$\text{for all } x \in A \text{ there exists } r > 0 \quad \text{such that} \quad B(x, r) \subseteq A.$$

Definition 1.15 (closed set)**.** A subset of a metric space $S \subseteq X$ is closed if

$$\text{its complement } S^c \quad \text{is open.}$$

Example 1.18 (MA4211 AY24/25 Sem 2 Tutorial 2)**.** Show that the closed unit ball of a normed linear space is closed.

Solution. Let $(X, \|\cdot\|)$ be a normed space. Define the closed unit ball as follows:

$$\overline{B(0,1)} = \{x \in X : \|x - y\| \leq 1 \text{ where } y \in X\}$$

It suffices to show that $\overline{B(0,1)}$ contains all of its limit points. Let $x_n \in \overline{B(0,1)}$ be such that $x_n \to x$. Since x_n converges, then for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$\text{for all } n \geq N \quad \text{we have} \quad \|x - x_n\| < \varepsilon$$

By the triangle inequality,

$$\|x - y\| \leq \|x - x_n\| + \|x_n - y\| < 1 + \varepsilon.$$

Since $\varepsilon > 0$ can be made sufficiently small, we see that $\|x - y\| \leq \|x_n - y\| \leq 1$. We conclude that $\|x\|$ is contained in $\overline{B(0,1)}$, or to be more explicit, the limit point of x_n is contained in the closed unit ball, making $\overline{B(0,1)}$ a closed set. □

Definition 1.16 (topology)**.** Given a set X, a topology $\mathcal{T}$ on X is a collection of subsets of X satisfying the following properties:

(i) $\emptyset, X \in \mathcal{T}$

(ii) $\mathcal{T}$ is closed under arbitrary unions

(iii) $\mathcal{T}$ is closed under finite intersection

Definition 1.17 (limit point)**.** In a topological space, a point x is a limit point of a sequence x_n if for every neighbourhood of x, there exists $N \in \mathbb{N}$ such that for all $n \geq n$, x_n belongs to that neighbourhood.

Definition 1.18 (continuous function)**.** If $f : X \to Y$ is a function between two topological spaces X and Y, then f is continuous if the pre-image of every open set in Y is open in X.

Example 1.19 (MA4211 AY24/25 Sem 2 Homework 1)**.** Let X and Y be normed linear spaces over the same field $\mathbb{F}$ and having norms $\|\cdot\|_1$ and $\|\cdot\|_2$ respectively. For $(x,y) \in X \times Y$, define

$$\|(x,y)\| = \|x\|_1 + \|y\|_2 .$$

Show that this defines a norm on $X \times Y$ and that the projection maps

$$\pi_1 : (x,y) \mapsto x \quad \text{and} \quad \pi_2 : (x,y) \mapsto y \quad \text{are continuous.}$$

Solution. Note that

$$\|(0,0)\| = \|0\|_1 + \|0\|_2 = 0$$

and conversely, if $\|(x,y)\| = 0$, then $\|x\|_1 + \|y\|_2 = 0$. Since the 1-norm and 2-norm are non-negative, then $\|x\|_1 = \|y\|_2 = 0$, so $\|\cdot\|$ is positive-definite.

Non-negativity and homogeneity of $\|\cdot\|$ is clear. Lastly,

$$\begin{aligned}
\|(x_1 + x_2, y_1 + y_2)\| &= \|x_1 + x_2\|_1 + \|y_1 + y_2\|_2 \\
&\leq \|x\|_1 + \|x\|_2 + \|y\|_1 + \|y_2\|_2 \quad \text{since } \|\cdot\|_1 \text{ and } \|\cdot\|_2 \text{ satisfy triangle inequality} \\
&= \|x_1\| + \|y_1\| + \|x_2\|_2 + \|y_2\|_2 \\
&= \|(x_1, y_1)\| + \|(x_2, y_2)\|
\end{aligned}$$

so $\|\cdot\|$ satisfies the triangle inequality. As such, $\|\cdot\|$ defines a norm on $X \times Y$.

We wish to prove for every $\varepsilon > 0$, there exists $\delta > 0$ such that

$$\|(x,y)-(x_0,y_0)\| < \delta \quad \text{implies} \quad \|\pi_1(x,y)-\pi_1(x_0,y_0)\| < \varepsilon.$$

We have

$$\|\pi_1(x,y)-\pi_1(x_0,y_0)\| = \|x-x_0\|_1 \le \|x-x_0\| + \|y-y_0\|_2 < \delta$$

so π_1 is a continuous map. Similarly, π_2 is a continuous map. □

Definition 1.19 (closure). Let X be a topological space. For $A \subseteq X$, the closure of A, denoted by $\mathrm{cl}(A)$ or $\overline{A}$, is defined as follows:

$$\overline{A} = A \cup \{\text{limit points of } A\}$$

Definition 1.20 (dense set). Let X be a topological space. If $D \subseteq X$ such that

$$\overline{D} = X \quad \text{then} \quad D \text{ is dense in } X.$$

Example 1.20 (MA4211 AY24/25 Sem 2 Homework 1). Suppose that (X,ρ) and (Y,d) are metric spaces and A is a dense subset of X. Show that if

$$F: X \to Y \text{ and } G: X \to Y \text{ are two continuous functions} \quad \text{such that} \quad F = G \text{ on } A \quad \text{then} \quad F = G \text{ on } X.$$

Solution. Since A is a dense subset of X, we take some sequence x_n in A such that $x_n \to x$ in (X,ρ). As F and G are continuous, we have

$$x_n \xrightarrow{\rho} x \quad \text{implies} \quad F(x_n) \xrightarrow{d} F(x) \text{ and } G(x_n) \xrightarrow{d} G(x).$$

Since $x_n \in A$ for all $n \in \mathbb{N}$, then $F(x_n) = G(x_n)$, i.e. these sequences are the same. These two sequences converge in (Y,d), i.e.

$$F(x) = \lim_{n\to\infty} F(x_n) \quad \text{and} \quad G(x) = \lim_{n\to\infty} G(x_n).$$

By the unqiueness of limits, we must have $F(x) = G(x)$. As this holds for all $x \in X$, then $F = G$ on X. □

Definition 1.21 (separable space). A topological space is separable if it has a countable dense subset.

Theorem 1.5. Let $1 \le p < \infty$. Then, ℓ^p is separable.

Proof. Define $X \subseteq \ell^p$ to be the collection of sequences of the form

$$(x_1, x_2, \ldots, x_n, 0, 0, \ldots) \quad \text{where } x_i \in \mathbb{Q}.$$

As X is a countable union of countable sets, X is countable. Let $y \in \ell^p$ and $\varepsilon > 0$ be arbitrary. That is,

$$\sum_{i=1}^{\infty} |y_i|^p < \infty.$$

Also, there exists $n \in \mathbb{N}$ such that

$$\sum_{i=n+1}^{\infty} \frac{\varepsilon^p}{2}.$$

Since $\mathbb{Q}$ is dense in $\mathbb{R}$, we can choose $x \in X$ such that

$$\sum_{i=1}^{n} |y_i - x_i|^p < \frac{\varepsilon^p}{2}.$$

Then,

$$\|y-x\|_p^p = \sum_{i=1}^{n} |y_i - x_i|^p + \sum_{i=n+1}^{\infty} |y_i|^p < \frac{\varepsilon^p}{2} + \frac{\varepsilon^p}{2} = \varepsilon^p.$$

Taking the p^{th} root, we obtain $d(x,y) = \|x-y\|_p < \varepsilon$. □

Theorem 1.6. ℓ^∞ is not separable.

Proof. Let $y \in \mathbb{R}^\infty$ be a sequence of 0s and 1s. Define $z \in \mathbb{R}$ to be as follows:

$$z = \frac{y_1}{2} + \frac{y_2}{2^2} + \frac{y_3}{2^3} + \ldots \quad \text{so} \quad \text{we infer that } 0 \le z \le 1.$$

Because there are uncountably many real numbers in $[0,1]$, it follows there are uncountably many distinct sequences $y \in \{0,1\}^{\mathbb{N}}$. Denote this uncountable family by $\mathcal{Y} \subset \ell^\infty$. Note that for any two distinct sequences

$$y = (y_1, y_2, y_3, \ldots) \quad y' = (y_1', y_2', y_3', \ldots) \quad \in \mathcal{Y},$$

there is at least one index i such that $y_i \neq y_i'$. Because each coordinate is either 0 or 1, at index i, we have$|y_i - y_i'| = 1$. Hence,

$$\|y - y'\|_\infty = \sup_{n \in \mathbb{N}} |y_n - y_n'| \ge 1.$$

In fact, it is exactly 1 if the two sequences differ in at least one place (and cannot exceed 1 because each coordinate difference is 0 or 1). Around each $y \in \mathcal{Y}$, consider the open ball $B(y, 1/3)$ of radius $1/3$. Since any two distinct y, y' are at distance $\|y - y'\|_\infty = 1$, their balls $B(y,1/3)$ and $B(y',1/3)$ cannot overlap. In other words, these balls are pairwise disjoint.

Suppose on the contrary that $D \subseteq \ell^\infty$ is a countable dense set. Then, for each $y \in \mathcal{Y}$, $B(y,1/3)$ must contain at least one point of D. However, there are uncountably many such disjoint balls $B(y,1/3)$ since $\mathcal{Y}$ is uncountable. A single countable set D cannot meet each of these uncountably many disjoint balls in a distinct point. This leads to a contradiction. □

Definition 1.22 (Cauchy sequence). A sequence x_n in a metric space X is Cauchy if

$$\text{for all } \varepsilon > 0 \text{ there exists } N \in \mathbb{N} \quad \text{such that for all } m, n \ge N \text{ we have } d(x_m, x_n) < \varepsilon.$$

Proposition 1.1. Every convergent sequence is Cauchy.

Remark 1.1. The converse of Proposition 1.1 is not true, i.e. not every Cauchy sequence is convergent.

Definition 1.23 (complete metric space). A metric space is called complete if every Cauchy sequence in the space converges.

Example 1.21. ℓ^p and ℓ^∞ are complete.

Example 1.22. The space of continuous functions on $[a,b]$

$$\mathcal{C}[a,b] \quad \text{equipped with the norm} \quad d_\infty(f,g) = \sup_{x\in[a,b]} |f(x)-g(x)| \quad \text{is complete.}$$

However, for any $1 \le p < \infty$,

$$\mathcal{C}[a,b] \quad \text{equipped with the norm} \quad d_p(f,g) = \left(\int_a^b |f(x)-g(x)|^p \, dx\right)^{1/p} \quad \text{is not complete.}$$

Having said that, the completion of this is denoted by $L^p[a,b]$.

Example 1.23 (MA4211 AY24/25 Sem 2 Tutorial 1). Show that the limits of Cauchy sequences in a complete metric space are unique.

Solution. Let x_n be a Cauchy sequence in a complete metric space. Suppose x_n has two limits x and y. Then, for all $\varepsilon > 0$, there exist $N_1, N_2 \in \mathbb{N}$ such that

$$\text{for all } n \ge N_1 \text{ we have } |x_n - x| < \frac{\varepsilon}{2} \quad \text{and} \quad \text{for all } n \ge N_2 \text{ we have } |x_n - y| < \frac{\varepsilon}{2}.$$

By the triangle inequality,

$$|x-y| \le |x_n - x| + |x_n - y| < \varepsilon.$$

Since ε can be made arbitrarily small, then $x = y$, i.e. the limits of the Cauchy sequence are the same. □

Example 1.24 (MA4211 AY24/25 Sem 2 Tutorial 1). Let

$$(X,d) \text{ and } (Y,\rho) \text{ be metric spaces} \quad \text{and} \quad \text{suppose } f : X \to Y \text{ is uniformly continuous.}$$

Prove that under f, the image of every Cauchy sequence is a Cauchy sequence.

Solution. Since $f : X \to Y$ is uniformly continuous, for all $\varepsilon > 0$, there exists $\delta > 0$ such that for all $x, y \in X$ such that

$$d(x,y) < \delta \quad \text{we have} \quad \rho(f(x), f(y)) < \varepsilon.$$

Let x_n be a Cauchy sequence in the metric space (X,d). Then, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that for all $m, n \ge N$, we have $d(x_m, x_n) < \delta$. Hence, $\rho(f(x_m), f(x_n)) < \varepsilon$. □

Theorem 1.7 (completion of metric space). Every metric space can be completed, and the completion is unique up to isometry.

We first state a rough proof sketch of Theorem 1.7 before formally proving it. Given a metric space (X,d), we can construct a new metric space $\left(\widetilde{X}, \widetilde{d}\right)$. Then, there exists $W \subseteq \widetilde{X}$ such that X is isometric to W and $\overline{W} = \widetilde{X}$. So, $\widetilde{X}$ is complete and $\widetilde{X}$ is unique up to isometry.

Proof. Let X be a metric space. Suppose x_n and y_n be Cauchy sequences of X. We say that

$$x_n \sim y_n \quad \text{if and only if} \quad \lim_{n\to\infty} d(x_n, y_n) = 0.$$

Define $\widetilde{X}$ to be the set of equivalence clases determined by $\sim$. We write $x_n \in \widetilde{X}$. Then,

$$\widetilde{d}(\widetilde{x}, \widetilde{y}) = \lim_{n\to\infty} d(x_n, y_n) \quad \text{where} \quad x_n, y_n \in X.$$

By applying the triangle inequality twice, we obtain

$$\begin{aligned} d(x_n, y_n) &\le d\left(x_n, x_n'\right) + d\left(x_n', y_n\right) \\ &\le d\left(x_n, x_n'\right) + d\left(x_n', y_n'\right) + d\left(y_n', y_n\right) \end{aligned}$$

Also, by applying the triangle inequality, we have

$$d\left(x_n', y_n'\right) \le d\left(x_n', x_n\right) + d(x_n, y_n) + d\left(y_n, y_n'\right).$$

We claim that $d(x_n, y_n)$ is a Cauchy sequence. To see why, let $\varepsilon > 0$ be arbitrary. Recall that x_n and y_n are both Cauchy sequences. So, there exists $N \in \mathbb{N}$ such that for all $m, n \ge N$, we have

$$d(x_n, x_m) < \varepsilon \quad \text{and} \quad d(y_n, y_m) < \varepsilon.$$

Then,

$$\begin{aligned} d(x_n, y_n) &\le d(x_n, x_m) + d(x_m, y_n) \\ &\le d(x_n, x_m) + d(x_m, y_m) + d(y_m, y_n) \end{aligned}$$

Let $x \in X$ and consider the infinite sequence $(x, x, \ldots)$. This sequence belongs to some element of $\widetilde{X}$, so we call it $\widetilde{x}$. Define $\varphi : X \to \widetilde{X}$ via $\varphi(x) = \widetilde{x}$. We also define $W = \varphi(X)$. Then,

$$\widetilde{d}(\varphi(x), \varphi(y)) = \widetilde{d}(\widetilde{x}, \widetilde{y}) = \lim_{n\to\infty} d(x_n, y_n) = \lim_{n\to\infty} d(x, y) = d(x, y).$$

Let $\widetilde{x} \in \widetilde{X}$ and consider $x_n \in \widetilde{x}$. Then, for all $\varepsilon > 0$, there exists $N \in \mathbb{N}$ such that

$$d(x_n, x_N) < \frac{\varepsilon}{2} \quad \text{for all } n \ge N.$$

Now, consider the sequence $(x_N, x_N, \ldots)$, so $\widetilde{x}_N \in W$. As such,

$$\widetilde{d}(\widetilde{x}, \widetilde{x}_N) = \lim_{n\to\infty} d(x_n, x_N) \le \frac{\varepsilon}{2} < \varepsilon.$$

This shows that W is dense in $\widetilde{X}$.

Next, let $\widetilde{x}_n$ be a Cauchy sequence of elements in $\widetilde{X}$. Since W is dense in $\widetilde{X}$, for each $n \in \mathbb{N}$, there exists $\widetilde{w}_n \in W$ such that $\widetilde{d}(\widetilde{x}_n, \widetilde{w}_n) < 1/n$. Hence,

$$\begin{aligned} \widetilde{d}(\widetilde{w}_m, \widetilde{w}_n) &\le \widetilde{d}(\widetilde{w}_m, \widetilde{x}_m) + \widetilde{d}(\widetilde{x}_m, \widetilde{x}_n) + \widetilde{d}(\widetilde{x}_n, \widetilde{w}_n) \\ &< \frac{1}{m} + \widetilde{d}(\widetilde{x}_m, \widetilde{x}_n) + \frac{1}{n} \end{aligned}$$

Hence, $\widetilde{w}_n$ is Cauchy. For each $\widetilde{w}_n$, define $w_n \in X$ by $w_n = \varphi^{-1}(\widetilde{w}_n)$. Since φ is an isometry, then w_n is Cauchy. Let $w_n \in \widetilde{X}$. Then,

$$\begin{aligned} \widetilde{d}(\widetilde{x}_n, \widetilde{x}) &\le \widetilde{d}(\widetilde{x}_n, \widetilde{w}_m) + d(\widetilde{w}_m, \widetilde{x}) \\ &< \frac{1}{n} + d(\widetilde{w}_m, \widetilde{x}) \end{aligned}$$

So,

$$\widetilde{d}(\widetilde{w}_m, \widetilde{x}) = \lim_{n\to\infty} d(w_m, w_n).$$

As such, the sequence w_n is Cauchy, which implies

$$\lim_{n\to\infty} \widetilde{d}(\widetilde{x}_n, \widetilde{x}) = 0 \quad \text{or equivalently} \quad \widetilde{x}_n \to \widetilde{x}.$$

Finally, we need to argue that Y is another complete metric space and there is an isometric embedding $f : X \to Y$ with $f(X)$ dense in Y, then Y is isometric to $\widetilde{X}$. Concretely, each Cauchy sequence x_n in X gives a Cauchy sequence $f(x_n)$ in Y. Since Y is complete, then $f(x_n)$ converges to some point in Y. Define a map $\psi : \widetilde{X} \to Y$ by sending $\widetilde{x}$ to the limit of $f(x_n)$ in Y. ψ is well-defined, isometric and surjective by the density of $f(X)$ in Y. We conclude that $\widetilde{X}$ is unique up to isometry. □

Example 1.25 (continuous extension theorem; MA4211 AY24/25 Sem 2 Tutorial 1). [†] Suppose W is a dense subset of a metric space X and f is a uniformly continuous function from W into some complete metric space Y. Prove that f has a unique continuous extension to X. That is,

$$\text{there exists a unique continuous function } F : X \to Y \quad \text{such that} \quad F(w) = f(w) \text{ for all } w \in W.$$

Solution. Since W is dense in X, there exists a sequence $\{w_n\}_{n=1}^{\infty} \subseteq W$ such that $w_n \to x$ in X. We first show that $\{f(w_n)\}$ is a Cauchy sequence in Y. Since $f : W \to Y$ is uniformly continuous, then for every $\varepsilon > 0$, there exists $\delta > 0$ such that for all $u, v \in W$, we have

$$d_X(u, v) < \delta \quad \text{implies} \quad d_Y(f(u), f(v)) < \varepsilon.$$

For m, n large enough, suppose w_m and w_n are within $\delta/2$ of x, so

$$d_X(w_m, w_n) \le d_X(w_m, x) + d_X(w_n, x) < \frac{\delta}{2} + \frac{\delta}{2} = \delta \quad \text{by the triangle inequality.}$$

Hence,

$$d_Y(f(w_m), f(w_n)) < \varepsilon.$$

As such, $\{f(w_n)\}$ is a Cauchy sequence. As Y is a complete metric space, then the limit of the sequence $\{f(w_n)\}$ exists in Y. Now, we define

$$F : X \to Y \quad \text{such that} \quad F(x) = \lim_{n\to\infty} f(w_n).$$

Note that F extends f. To see why, if $x \in W$, we can take the constant sequence $w_n = x$, so $w_n \to x$, which implies

$$F(x) = \lim_{n\to\infty} f(w_n) = \lim_{n\to\infty} f(x) = f(x).$$

We then show that F is continuous at x. Suppose $x, x' \in X$ such that $d_X(x, x') < \delta/3$. Then, choose sequences $w_n \to x$ and $w'_n \to x'_n$. For large n, we have

$$d_X(w_n, x) < \frac{\delta}{3} \quad \text{and} \quad d_X\left(w'_n, x'\right) < \frac{\delta}{3}.$$

[†] As remarked by Tianle, this was taken from Rudin's Principles of Mathematical Analysis p. 99 Question 11.

By the triangle inequality, $d_X(w_n, w_n') < \delta$. As $f : W \to Y$ is uniformly continuous, then $d_Y(f(w_n), f(w_n')) < \varepsilon$ for large n. As F extends f, it follows that $d_Y(F(x), F(x')) < \varepsilon$[‡].

Lastly, we prove that the extension is unique. Suppose there exists another continuous map $\widetilde{F} : X \to Y$ such that $\widetilde{F}(w) = f(w)$ for all $w \in W$. Since $\widetilde{F}$ is continuous, then

$$\lim_{n\to\infty} \widetilde{F}(w_n) = \widetilde{F}\left(\lim_{n\to\infty} w_n\right) = \widetilde{F}(x) \quad \text{but also} \quad \lim_{n\to\infty} \widetilde{F}(w_n) = \lim_{n\to\infty} f(w_n) = F(x)$$

so $\widetilde{F} = F$. The result follows. □

[‡]Here, we are proving a stronger result, which is that f has a unique uniformly continuous extension to X.

Chapter 2
Normed Spaces and Banach Spaces

2.1
Vector Spaces

Definition 2.1 (vector space). A vector space is a set V together with a field $\mathbb{F}$ equipped with two operations (addition $+$ and multiplication $\cdot$)

$$+ : V \times V \to V \quad \text{and} \quad \cdot : \mathbb{F} \times V \to V$$

satisfying the following properties 2

(i) $+$ is commutative

(ii) $+$ is associative

(iii) There exists an additive identity

(iv) There exists an additive inverse

(v) $\cdot$ is associative

(vi) There exists a multiplicative identity

(vii) The distributivity properties hold

In this course, we are interested in infinite-dimensional vector spaces. Next, we are interested in continuous functions over $\mathbb{R}$. Recall that the set of continuous functions on $[a,b]$ over $\mathbb{R}$ is denoted by $\mathcal{C}[a,b]$. We define the addition and scalar multiplication functions by the obvious way — that is for any $f,g \in \mathcal{C}[a,b]$, we have

$$(f+g)(x) = f(x) + g(x) \quad \text{and} \quad (\alpha f)(x) = \alpha \cdot f(x),$$

where $\alpha \in \mathbb{R}$.

Example 2.1. ℓ^p for $1 \le p \le \infty$ is a vector space.

Example 2.2. $\mathbb{R}^\infty$ is a vector space.

Definition 2.2 (linear combination and span). A linear combination is a vector of the form

$$\mathbf{v} = \alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n \quad \text{for all } \alpha_i \in \mathbb{F} \text{ and } \mathbf{v}_i \in V.$$

Given a set of vectors S, its span, denoted by $\operatorname{span}(S)$, is the set of all linear combinations of its elements.

Definition 2.3 (linear independence). A set of vectors $\mathbf{v}_1, \ldots, \mathbf{v}_n$ is linearly independent if for any $\alpha_1, \ldots, \alpha_n \in \mathbb{F}$, we have

$$\alpha_1 \mathbf{v}_1 + \ldots + \alpha_n \mathbf{v}_n = \mathbf{0} \quad \text{implies} \quad \alpha_1 = \ldots = \alpha_n = 0.$$

Definition 2.4 (dimension). Let V be a vector space. If there exists $S \subseteq V$ such that $|S| < \infty$ and $\operatorname{span}(S) = V$, then V is finite-dimensional. Otherwise, V is infinite-dimensional.

Example 2.3 (infinite-dimensional vector spaces). We give some examples of infinite-dimensional vector spaces. For $1 \leq p < \infty$, $\ell^p(\mathbb{R})$, which denotes the space of real sequences a_n such that

$$\sum_{n=1}^{\infty} |a_n|^p \text{ converges} \quad \text{is infinite-dimensional.}$$

$\ell^\infty(\mathbb{R})$, which denotes the space of bounded real sequences equipped with the supremum norm, is also infinite-dimensional. Lastly, the space of polynomials with real coefficients, denoted by $\mathbb{R}[x]$, is infinite-dimensional.

Definition 2.5 (basis). If $\mathcal{B}$ is an independent spanning set for a finite-dimensional vector space V, then we say that $\dim(V) = |\mathcal{B}|$. Such a set $\mathcal{B}$ is a basis. More formally, we call it a Hamel basis.

In Definition 1.4 on the definition of a function space, we briefly discussed the definition of a norm. We also talked about the p-norm in Definition 1.6. Now, we will formally introduce p-norms.

Definition 2.6 (norm). Given a vector space V, a norm on V is a function $\|\cdot\| : V \to R$ satisfying the following properties for any $\mathbf{u}, \mathbf{v} \in V$:

(i) **Non-negativity:** $\|\mathbf{v}\| \geq 0$ for all $\mathbf{v} \in V$

(ii) **Positive-definiteness:** $\|\mathbf{v}\| = 0$ if and only if $\mathbf{v} = 0$

(iii) **Homogeneity:** For any $\alpha \in \mathbb{F}$, we have $\|\alpha\mathbf{v}\| = |\alpha| \|\mathbf{v}\|$

(iv) **Triangle inequality:** $\|\mathbf{u}+\mathbf{v}\| \leq \|\mathbf{u}\| + \|\mathbf{v}\|$

It is important to note that a metric space is a space where we wish to measure distance, but a normed space is a space where we wish to measure the Euclidean distance between two vectors. We then discuss some properties of normed spaces.

Definition 2.7. Let $\mathbf{u}, \mathbf{v}, \mathbf{x}, \mathbf{y}, \mathbf{z}$ be vectors in a normed space and α be a real-valued scalar. Then, we have the following:

(i) $d(\mathbf{u}, \mathbf{v}) = \|\mathbf{u} - \mathbf{v}\|$, which defines the metric induced by the norm, also known as the norm-induced metric

(ii) We have the familiar p-norm

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{\infty} |x_i|^p \right)^{1/p}$$

(iii) $d(\mathbf{x}+\mathbf{z}, \mathbf{y}+\mathbf{z}) = d(\mathbf{x}, \mathbf{y})$, which denotes translation invariance

(iv) $d(\alpha\mathbf{x}, \alpha\mathbf{y}) = |\alpha| d(\mathbf{x}, \mathbf{y})$, which denotes the homogeneity of the metric

Definition 2.8 (Banach space). A normed space where the space is complete with respect to the induced metric is called a Banach space.

Example 2.4 (MA4211 AY24/25 Sem 2 Tutorial 2). Consider the space of continuously differentiable functions $\mathcal{C}^1([a,b])$ with the $\mathcal{C}^1$-norm

$$\|f\| = \sup_{a \leq x \leq b} |f(x)| + \sup_{a \leq x \leq b} |f'(x)|.$$

Prove that $\mathcal{C}^1([a,b])$ is a Banach space with respect to the given norm.

Solution. We first show that $\|\cdot\|$ is indeed a norm. As usual, non-negativity, positive-definiteness and homogeneity are obvious. It suffices to prove that it satisfies the triangle inequality. We have

$$\begin{aligned}\|f+g\| &= \sup_{a\le x\le b}|f(x)+g(x)|+\sup_{a\le x\le b}\left|f'(x)+g'(x)\right| \\ &\le \sup_{a\le x\le b}|f(x)|+\sup_{a\le x\le b}|g(x)|+\sup_{a\le x\le b}\left|f'(x)\right|+\sup_{a\le x\le b}\left|g'(x)\right| \\ &= \sup_{a\le x\le b}|f(x)|+\sup_{a\le x\le b}\left|f'(x)\right|+\sup_{a\le x\le b}|g(x)|+\sup_{a\le x\le b}\left|g'(x)\right| \\ &= \|f\|+\|g\|\end{aligned}$$

Next, let f_n be a Cauchy sequence in $\mathcal{C}^1([a,b])$. Then, for every $\varepsilon>0$, there exists $N\in\mathbb{N}$ such that for all $m,n\ge N$, we have

$$\|f_m-f_n\| = \sup_{a\le x\le b}|f_m(x)-f_n(x)|+\sup_{a\le x\le b}\left|f_m'(x)-f_n'(x)\right|<\varepsilon.$$

As $\mathcal{C}^1([a,b])$ equipped with the supremum norm is a Banach space, then f_n and f_n' converge uniformly to some continuous functions f and g respectively. By the Fundamental Theorem of Calculus, we have

$$f_n(x)-f_n(a)=\int_a^x f_n'(t)\ dt.$$

Taking the limit as $n\to\infty$, we have

$$f(x)-f(a)=\int_a^x g(t)\ dt.$$

Since $f_n\to f$ uniformly and $f_n'\to g$ uniformly, then $\|f_n-f\|\to 0$, which implies every Cauchy sequence in $\mathcal{C}^1([a,b])$ converges to a limit in $\mathcal{C}^1([a,b])$. □

Example 2.5 (MA4211 AY24/25 Sem 2 Tutorial 3). For $f:[0,1]\to\mathbb{R}$, let

$$\|f\|=\sup_{x\in[0,1]}\left\{x^2|f(x)|\right\}.$$

(a) Prove that $\|\cdot\|$ defines a norm on the linear space $\mathcal{C}([0,1])$ of continuous functions $f:[0,1]\to\mathbb{R}$.

(b) Prove that $\mathcal{C}([0,1])$ is not complete with respect to this norm.

Solution.

(a) We will only prove that $\|\cdot\|$ satisfies the triangle inequality. Note that

$$\begin{aligned}\|f+g\| &= \sup_{x\in[0,1]}\left\{x^2|f(x)+g(x)|\right\} \\ &\le \sup_{x\in[0,1]}\left\{x^2|f(x)|+x^2|g(x)|\right\} && \text{by the triangle inequality} \\ &\le \sup_{x\in[0,1]}\left\{x^2|f(x)|\right\}+\sup_{x\in[0,1]}\left\{x^2|g(x)|\right\} && \text{by the identity } \sup(A+B)\le\sup(A)+\sup(B) \\ &= \|f\|+\|g\|\end{aligned}$$

Hence, $\|\cdot\|$ defines a norm.

(b) Let $\{f_n\}_{n\in\mathbb{N}}$ be a Cauchy sequence in $\mathcal{C}([0,1])$. It suffices to show that $\{f_n\}_{n\in\mathbb{N}}$ does not converge to some value in this space. Let

$$f_n(x)=\begin{cases}0 & \text{if } 0\le x<1/n;\\ 1 & \text{if } x\ge 1/n.\end{cases}$$

Then, $\{f_n\}_{n\in\mathbb{N}}$ is Cauchy. To see why, let $\varepsilon > 0$ be arbitrary. In particular, we can set $\varepsilon = 1$. Then, there exists $N \in \mathbb{N}$ such that for all $m,n \geq N$, we have

$$\|f_m(x) - f_n(x)\| = \sup_{x\in[0,1]} \left\{x^2 |f_m(x) - f_n(x)|\right\} \leq 1.$$

However,

$$f_n \to f = \begin{cases} 0 & \text{if } x = 0; \\ 1 & \text{if } x > 0 \end{cases} \quad \text{pointwise.}$$

Clearly, $f(x)$ is not continuous at $x = 0$. As such, there is no limit of $\{f_n\}_{n\in\mathbb{N}}$ in $\mathcal{C}[0,1]$, showing that the space is not complete. $\square$

Example 2.6 (MA4211 AY24/25 Sem 2 Tutorial 3)**.** Let $\|\cdot\|_1$ and $\|\cdot\|_2$ be two norms on a linear space X. We say that

$$\|\cdot\|_2 \text{ is } \textit{stronger} \text{ than } \|\cdot\|_1 \quad \text{if} \quad \text{for any sequence } \{x_n\}_{n\in\mathbb{N}} \subseteq X \text{ we have } \|x_n\|_2 \to 0 \text{ implies } \|x_n\|_1 \to 0.$$

Show that $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$ if and only if there exists a constant $C > 0$ such that $\|x\|_1 \leq C\|x\|_2$ for all $x \in X$.

Solution. For the forward direction, suppose no such C exists. Then, for all $n \in \mathbb{N}$, we can always find a sequence $\{x_n\}_{n\in\mathbb{N}} \subseteq X$ such that

$$\|x_k\|_2 = 1 \quad \text{but} \quad \|x_k\|_1 > k.$$

Define $y_k = x_k/k$, so

$$\|y_k\|_2 = \frac{\|x_k\|_2}{k} = \frac{1}{k} \quad \text{but} \quad \|y_k\|_1 = \frac{\|x_k\|_1}{k} > \frac{1}{k}.$$

Since $\|\cdot\|_2$ is stronger than $\|\cdot\|_1$, then the result follows.

For the reverse direction, since $\|x\|_1 \leq C\|x\|_2$, where $\|x_n\|_2 \to 0$, by the squeeze theorem, the result follows. $\square$

Example 2.7 (MA4211 AY24/25 Sem 2 Tutorial 3)**.** Suppose X is a normed linear space with norm $\|\cdot\|$. Define

$$\rho = \frac{\|x\|}{1 + \|x\|}.$$

(a) Define $r : X \times X \to \mathbb{R}$ by

$$r(x,y) = \rho(x-y).$$

Prove that r is a metric on X.

(b) Define the diameter of X with respect to a metric d by

$$\operatorname{diam}(X) = \sup_{x,y\in X} d(x,y).$$

What is the diameter of X with respect to the metric $r(x,y) = \rho(x-y)$?

(c) Prove that $\|x_n - x\| \to 0$ as $n \to \infty$ if and only if $(x_n, x) \to 0$ as $n \to \infty$.

Solution.

(a) Again, we will only prove that r satisfies the triangle inequality as the other properties are deemed trivial. Let $x,y,z \in X$. Then,

$$\begin{aligned}
\rho(x,y)+\rho(y,z) &= \frac{\|x-y\|}{1+\|x-y\|}+\frac{\|y-z\|}{1+\|y-z\|} \\
&\geq \frac{\|x-y\|}{1+\|x-y\|+\|y-z\|}+\frac{\|y-z\|}{1+\|x-y\|+\|y-z\|} && \text{by the triangle inequality} \\
&= \frac{\|x-y\|+\|y-z\|}{1+\|x-y\|+\|y-z\|} \\
&\geq \frac{\|x-z\|}{1+\|x-y\|+\|y-z\|}
\end{aligned}$$

(b) We have

$$\operatorname{diam}(X)=\sup_{x,y\in X}\rho(x,y)=\sup_{x,y\in X}\frac{\|x-y\|}{1+\|x-y\|}=\sup_{x,y\in X}\left(1-\frac{1}{1+\|x-y\|}\right)=1.$$

(c) The forward direction is trivial. For the reverse direction, if $r(x_n,x)\to 0$, then

$$\frac{\|x_n-x\|}{1+\|x_n-x\|}\to 0$$

so $\|x_n-x\|\to 0$. □

Theorem 2.1. A subspace Y of a Banach space X is complete if and only if it is closed in X.

Theorem 2.1 can be reformulated as saying that a subspace of a Banach space X is a Banach space if and only if X is closed. We now prove it.

Proof. We first prove the forward direction. Suppose Y is complete. Let $y \in \overline{Y}$. Then, there exists a subsequence $y_n \subseteq Y$ such that

$$\lim_{n\to\infty} y_n = y.$$

So, y_n is Cauchy and in fact, y_n converges in Y. So, Y is closed in X.

For the reverse direction, suppose Y is closed in X. Let y_n be Cauchy in Y. Since $Y\subseteq X$ and X is complete, then

$$\lim_{n\to\infty} y_n = y.$$

So, $y\in\overline{Y}=Y$. □

Definition 2.9 (convergence). If X is a normed space and x_n is a sequence of elements in X, then we can define

$$s_n=\sum_{k=1}^{n} x_k.$$

If there exists $s\in X$ such that $\lim_{n\to\infty}\|s_n-s\|=0$, then we write

$$s=\sum_{k=1}^{\infty} x_k.$$

Note that if $\sum_{k=1}^{\infty}\|x_k\|<\infty$, the series is said to be absolutely convergent.

Example 2.8 (MA4211 AY24/25 Sem 2 Tutorial 2)**.** Let Y be a closed subspace of a normed linear space $(X, \|\cdot\|)$. Let X/Y denote the quotient space (elements of X/Y are additive cosets). For $x+Y \in X/Y$, define the quotient norm $\|\cdot\|_q$ by

$$\|x+Y\|_q = \inf_{y\in Y} \|x-y\|.$$

Show that $\|\cdot\|_q$ is a norm on X/Y. Also, if X is a Banach space, show that X/Y is a Banach space under the quotient norm.

Solution. Clearly, the norm satisfies positive-definiteness, non-negativity and homogeneity. We prove that it satisfies the triangle inequality. We have

$$\begin{aligned}
\|x+Z\|_q &= \inf_{z\in Z} \|x-z\| \\
&\leq \inf_{z\in Z} \|(x-y)+(y-z)\| \\
&= \|x-y\| + \inf_{z\in Z} \|y-z\| \\
&= \inf_{y\in Y} \|x-y\| + \inf_{z\in Z} \|y-z\| \\
&\leq \|x+Y\|_q + \|y+Z\|_q
\end{aligned}$$

so $\|\cdot\|_q$ is a norm on X/Y.

Next, suppose X is a Banach space. We wish to prove that X/Y is also a Banach space, i.e. every Cauchy sequence in X/Y converges to a limit in X/Y (norm on X/Y was established earlier). For each $x \in X$, let $\widehat{x} = x+Y \in X/Y$. Consider $\widehat{x}_n$ such that

$$\sum_{n=1}^{\infty} \|\widehat{x}_n\| \quad \text{converges.}$$

We have for every $n \in \mathbb{N}$, there exists $x_n \in \widehat{x}_n$ such that $\|x_n\| \leq 2\|\widehat{x}_n\|$. Since the mentioned sum converges, then

$$\sum_{n=1}^{\infty} \|x_n\| \quad \text{converges.}$$

As X is Banach, then

$$\sum_{n=1}^{\infty} x_n \quad \text{converges in } X \quad \text{to some } x \in X.$$

From definition of the norm in X/Y, it follows that

$$\sum_{n=1}^{\infty} \widehat{x}_n \text{ converges to } \widehat{x} \quad \text{in } X/Y.$$

We conclude that X/Y is Banach as well. ☐

Definition 2.10 (c_0)**.** Let $\mathbb{F}$ be a field. The space c_0 is the set of all real or complex sequences that converge to zero, formally defined as follows:

$$c_0 = \left\{ \{x_n\}_{n\in\mathbb{N}} \in \mathbb{F}^{\mathbb{N}} : \lim_{n\to\infty} x_n = 0 \right\} \quad \text{where} \quad \mathbb{F} \text{ is either } \mathbb{R} \text{ or } \mathbb{C}$$

The space c_0 is equipped with the supremum norm, i.e.

$$\left\| \{x_n\}_{n\in\mathbb{N}} \right\|_\infty = \sup_{n\in\mathbb{N}} |x_n|.$$

With this norm, c_0 is a Banach space because it is a normed space and it is complete. For the first property, it is clear that the supremum norm satisfies the properties of a norm. As for completeness, if $\left\{x_n^{(k)}\right\}_{n\in\mathbb{N}}$ in c_0 is a Cauchy sequence in the supremum norm, then it must converge uniformly to a sequence $\{x_n\}_{n\in\mathbb{N}}$. Since each term of a Cauchy sequence of elements in c_0 must go to zero, the limit sequence also belongs to c_0, ensuring completeness.

Example 2.9 (MA4211 AY24/25 Sem 2 Homework 1)**.** Much of our motivation comes from wanting to do Linear Algebra in infinite dimensions. In order to compute $\mathbf{Ax}$, where $\mathbf{A}$ is an infinite matrix and $\mathbf{x}$ is an infinite sequence, we need expressions like

$$\sum_{k=1}^{\infty} a_k x_k \quad \text{to converge.}$$

Even better, we would like these to converge absolutely, i.e.

$$\sum_{k=1}^{\infty} |a_k x_k| < \infty.$$

In this problem, we investigate what conditions are needed to guarantee convergence.

(a) Give an example of $\mathbf{x},\mathbf{y} \in \mathbb{R}^\infty$ such that

$$\sum_{k=1}^{\infty} |x_k y_k| \quad \text{does not converge.}$$

(b) Give an example of $\mathbf{x},\mathbf{y} \in \mathbb{R}^\infty \setminus \ell^\infty$ such that

$$\sum_{k=1}^{\infty} |x_k y_k| \quad \text{does not converge.}$$

(c) Prove or disprove: If $\mathbf{x} \in \mathbb{R}^\infty \setminus \ell^\infty$, then there exists $\mathbf{y} \in \ell^1$ such that

$$\sum_{k=1}^{\infty} |x_k y_k| \quad \text{does not converge.}$$

(d) Recall Hölder's inequality (Theorem 1.2): if $1 \le p,q < \infty$ such that $1/p + 1/q = 1$, then

$$\sum_{k=1}^{\infty} |x_k y_k| \le \|\mathbf{x}\|_p \|\mathbf{y}\|_q.$$

Prove that this inequality still holds if we take $q = \infty$ and use the convention that $\frac{1}{\infty} = 0$.

(e) Prove or disprove: If $p,q \in [1,\infty]$ such that $1/p + 1/q \ge 1$ and $\mathbf{x} \in \ell^p$, $\mathbf{y} \in \ell^q$, then

$$\sum_{k=1}^{\infty} |x_k y_k| \quad \text{converges.}$$

(f) Prove or disprove: If $p,q \in [1,\infty]$ such that $1/p + 1/q < 1$, then there exist $\mathbf{x} \in \ell^p, \mathbf{y} \in \ell^q$ such that

$$\sum_{k=1}^{\infty} |x_k y_k| \quad \text{does not converge.}$$

Solution.

(a) Let $\mathbf{x} = \mathbf{y} = (1,\ldots)$ be two sequences of ones, so

$$\sum_{k=1}^{\infty} |x_k y_k| = \sum_{k=1}^{\infty} 1 \quad \text{which is divergent.}$$

(b) Let $x_k = y_k = k$ for all $k \in \mathbb{N}$. Then,

$$\sum_{k=1}^{\infty} |x_k y_k| = \sum_{k=1}^{\infty} k^2 \quad \text{which does not converge.}$$

(c) Yes. Recall that $\mathbb{R}^\infty \setminus \ell^\infty$ denotes the set of unbounded sequences. Let $x_k = 2^k$ and $y_k = 1/2^k$, so

$$\sum_{k=1}^{\infty} |x_k y_k| = \sum_{k=1}^{\infty} 1 \quad \text{which does not converge.}$$

(d) If $q = \infty$, then $p = 1$, so we wish to prove that

$$\sum_{k=1}^{\infty} |x_k y_k| \le \left(\sum_{k=1}^{\infty} |x_k| \right) \cdot \sup_{k \in \mathbb{N}} |y_k| .$$

This is obvious because

$$\sum_{k=1}^{\infty} |x_k y_k| = |x_1 y_1| + |x_2 y_2| + \ldots \le \left| x_1 \cdot \sup_{k\in\mathbb{N}} |y_k| \right| + \left| x_2 \cdot \sup_{k\in\mathbb{N}} |y_k| \right| + \ldots = \sup_{k\in\mathbb{N}} |y_k| \sum_{k=1}^{\infty} |x_k|$$

(e) Recall Example 1.14, which mentioned that if $1 \le p \le q$, then $\ell^p \subseteq \ell^q$. Hence, $\ell^1 \subseteq \ell^q$. It suffices to prove that if $\mathbf{x}, \mathbf{y} \in \ell^1$, then

$$\sum_{k=1}^{\infty} |x_k y_k| \quad \text{converges.}$$

This holds because

$$\sum_{k=1}^{\infty} |x_k y_k| \le \left(\sum_{k=1}^{\infty} |x_k| \right) \left(\sum_{k=1}^{\infty} |y_k| \right)$$

As such, the statement is true.

(f) The statement is true[†]. We shall consider three cases.

- **Case 1:** Suppose $p = q = \infty$. Then, we can take $\mathbf{x} = \mathbf{y}$ to be the constant sequence 1. As such,

$$\sum_{k=1}^{\infty} |x_k y_k| = \sum_{k=1}^{\infty} 1 \quad \text{which diverges.}$$

- **Case 2:** Suppose exactly one of p or q is ∞. Say $p = \infty$, then $1/q < 1$. Let $\mathbf{x}$ be the constant sequence 1 and $\mathbf{y}$ be the sequence $1/k$. Then, $\mathbf{x} \in \ell^\infty$ and

$$\sum_{k=1}^{\infty} \left| \frac{1}{k} \right|^p = \sum_{k=1}^{\infty} \frac{1}{k^p}.$$

  Since $1/p < 1$, then $p > 1$ so the series converges. However,

$$\sum_{k=1}^{\infty} |x_k y_k| = \sum_{k=1}^{\infty} 1 \quad \text{which does not converge.}$$

- **Case 3:** Suppose $1 \le p, q < \infty$. Since $1/p + 1/q < 1$, let

$$\varepsilon = \frac{1}{2} \left(1 - \frac{1}{p} - \frac{1}{q} \right).$$

  Define

$$x_k = \frac{1}{k^{\varepsilon + 1/p}} \quad \text{and} \quad y_k = \frac{1}{k^{\varepsilon + 1/q}}.$$

  Hence,

$$\sum_{k=1}^{\infty} |x_k|^p = \sum_{k=1}^{\infty} \frac{1}{k^{\varepsilon + 1/p}}.$$

  As $1 + \varepsilon p > 1$, then $\mathbf{x} \in \ell^p$. A similar argument shows that $\mathbf{y} \in \ell^q$. However,

$$\sum_{k=1}^{\infty} |x_k y_k| = \frac{1}{k^{2\varepsilon + 1/p + 1/q}} = \sum_{k=1}^{\infty} \frac{1}{k}.$$

  This is precisely the harmonic series, which does not converge!. □

[†]We credit the solution to Lou Yi.

Definition 2.11 (Schauder basis)**.** If X is a normed space and e_k is a sequence of elements of x such that for all $x \in X$, there exists a unique sequence of scalars a_n such that

$$\lim_{n\to\infty} \left\| \sum_{k=1}^{n} a_k e_k - x \right\| = 0 \quad \text{then} \quad \{e_1, \dots, e_n\} \text{ is a Schauder basis.}$$

In a *slightly more general notion*, a Schauder basis is also called a countable basis.

Definition 2.12 (partial order)**.** A partial order $\leq$ on a set S is a binary relation that satisfies the following properties:

(i) Reflexivity: $a \leq a$

(ii) Transitivity: If $a \leq b$ and $b \leq b$, then $a \leq c$

(iii) Antisymmetry: If $a \leq b$ and $b \leq a$, then $a = b$

Definition 2.13 (total order)**.** A total order relation is a partial order in which every element of the set is comparable with every other element of the set, i.e. if $\leq$ is a partial order on a set S, then

$$\text{for any } x, y \in S \text{ such that either } x \leq y \text{ or } y \leq x \quad \text{then} \quad \leq \text{ is a total order on } S.$$

Note that every total order is a partial order, but the converse does not hold.

Theorem 2.2 (Zorn's lemma)**.** If S is a partially ordered set such that every totally ordered subset of S has an upper bound in S, then

$$\text{there exists an element } m \in S \quad \text{such that} \quad \text{for all } a \in S \text{ we have } m \geq a.$$

Theorem 2.3. Every vector space has a Hamel basis.

Proof. If $V = \{\mathbf{0}\}$, then $\emptyset$ is its basis. Suppose $V \neq \emptyset$. Let P be the set of linearly independent subsets of V ordered by inclusion.

Let S_α be a totally ordered subset of P and define

$$M = \bigcup_\alpha S_\alpha.$$

Suppose $\mathbf{v}_1, \dots, \mathbf{v}_n \in M$ such that $\alpha_1 \mathbf{v}_1 + \alpha_n \mathbf{v}_n = \mathbf{0}$ with not all $\alpha_i = 0$. Then, there exists an α such that S_α contains all $\mathbf{v}_1, \dots, \mathbf{v}_n$. By Zorn's lemma (Theorem 2.2), S_α has a maximal element, say $\mathcal{B}$.

Suppose on the contrary that $\mathcal{B}$ does not span V. Then, there exists $\mathbf{v} \in V$ such that $\mathbf{v} \notin \operatorname{span}(\mathcal{B})$. So, $\mathcal{B} \cup \{\mathbf{v}\}$ is linearly independent but $\mathcal{B} \subseteq \mathcal{B} \cup \{\mathbf{v}\}$. This is a contradiction. Hence, $\mathcal{B}$ spans V and is a Hamel basis. Since V was an arbitrary vector space, we conclude that every vector space has a Hamel basis. $\square$

Lemma 2.1 (frame condition)**.** Let $\mathbf{x}_1, \dots, \mathbf{x}_n$ be linearly vectors independent in X. The set $\{\mathbf{x}_1, \dots, \mathbf{x}_n\}$ is a frame in X if there exist constants $0 < A \leq B$ such that for any scalars $\alpha_1, \dots, \alpha_n$, we have

$$A \sum_{i=1}^{n} |\alpha_i| \leq \left\| \sum_{i=1}^{n} \alpha_i \mathbf{x}_i \right\| \leq B \sum_{i=1}^{n} |\alpha_i|.$$

Note that one part of the frame condition in Lemma 2.1 is obvious — by the triangle inequality ,we have

$$\left\|\sum_{i=1}^{n} \alpha_i \mathbf{x}_i\right\| = \sum_{i=1}^{n} |\alpha_i| \, \|\mathbf{x}_i\| \leq \max_{1\leq i\leq n} \|\mathbf{x}_i\| \sum_{i=1}^{n} |\alpha_i|$$

As such,

$$\text{setting } B = \max_{1\leq i\leq n} \|\mathbf{x}_i\| \quad \text{yields part of the desired result.}$$

2.2 Compactness

We first define sequential compactness (Definition 2.14). Note that in metric spaces, compactness and sequential compactness are equivalent. This is a key property of metric spaces, where the structure of the metric allows for this equivalence to hold.

Definition 2.14 (sequential compactness)**.** Let X be a metric space. A subset K of X is said to be sequentially compact if every sequence of elements in K contains a subsequence that converges to an element in K.

Theorem 2.4 (Heine-Borel theorem)**.** A subset of $\mathbb{R}^n$ is

$$\text{compact} \quad \text{if and only if} \quad \text{it is closed and bounded.}$$

Remark 2.1. In any metric space, a compact set is always closed and bounded.

Theorem 2.5 (Bolzano-Weierstrass theorem)**.** Every bounded sequence in $\mathbb{R}^n$ has a convergent subsequence.

Proof. Let x_n be a bounded sequence in $\mathbb{R}^n$. Then, there exists a positive constant $M > 0$ such that the sequence is contained in $[-M,M]^n = C_0$ Then, partition C_0 into 2^n smaller cubes of side length M. By the pigeonhole principle, at least one of these subcubes, call it C_1, contains infinitely many points of the sequence.

From C_1 (which has side length M), divide it again into 2^n subcubes of side length $M/2$. At least one of these subcubes contains infinitely many points of the sequence; call it C_2. Next, from C_2 (side length $M/2$), divide it into 2^n subcubes of side length $M/4$. Again, at least one of these new subcubes contains infinitely many points; call it C_3.

Continue the above process inductively. At the k^{th} step, we will have a cube C_k of side length $M/2^{k-1}$. Subdivide C_k into 2^n smaller cubes of side length $M/2^k$. At least one of these smaller cubes, denoted by C_{k+1}, contains infinitely many points of $\{x_n\}$.

Hence, we obtain a nested sequence of closed cubes as follows:

$$C_0 \supseteq C_1 \supseteq C_2 \supseteq \ldots \supseteq C_k \supseteq \ldots \quad \text{with} \quad \text{side lengths decreasing to } 0.$$

As each cube C_k contains infinitely many terms of the sequence, we can choose at least one index n_k such that $x_{n_k} \in C_k$. We wish to construct a subsequence $\{x_{n_k}\}$. To ensure it is well-defined (i.e. so that $n_1 < n_2 < n_3 < \ldots$), choose each index n_{k+1} to be strictly larger than n_k. This gives us a subsequence $\{x_{n_k}\}$.

Since

$$C_0 \supseteq C_1 \supseteq C_2 \supseteq \ldots \supseteq C_k \supseteq \ldots,$$

each C_k is a closed cube whose diameter goes to 0 as $k \to \infty$. More precisely,

$$\text{the diameter of } C_k \text{ is} \quad \sqrt{n} \cdot \frac{M}{2^{k-1}} \quad \text{which} \quad \text{tends to 0 as } k \text{ goes to infinity.}$$

As $x_{n_k} \in C_k$ for each k, any two points x_{n_k}, x_{n_ℓ} eventually lie in the same small cube for sufficiently large k, ℓ. Hence,

$$\lim_{k,\ell \to \infty} \|x_{n_k} - x_{n_\ell}\| \to 0 \quad \text{which shows that} \quad x_{n_k} \text{ is a Cauchy subsequence in } \mathbb{R}^n.$$

$$\|x_{n_k} - x_{n_\ell}\| \to 0 \quad \text{as } k, \ell \to \infty,$$

As $\mathbb{R}^n$ is complete, then every Cauchy sequence converges to some limit point in $\mathbb{R}^n$. Therefore, the subsequence $\{x_{n_k}\}$ converges and the proof is complete. □

Example 2.10 (MA4211 AY24/25 Sem 2 Tutorial 3)**.** Prove that $K \subseteq \mathbb{R}^n$ is sequentially compact if and only if it is closed and bounded.

Solution. Suppose K is sequentially compact. Then, every sequence in K has a convergent subsequence. So, K is bounded, otherwise there exists some unbounded sequence in K which clearly does not have a convergent subsequence, contradicting sequential compactness. Also, if a sequence in K converges, then its limit must be in K so K must be closed. By the Heine-Borel theorem (Theorem 2.4), K is compact.

For the reverse direction, suppose K is compact. Then, by the Heine-Borel theorem (Theorem 2.4), K is closed and bounded. By the Bolzano-Weierstrass theorem (Theorem 2.5), K has a convergent subsequence. Since K is closed, the limit of every sequence is in K, so K is sequentially compact. □

Theorem 2.6. If $f : K \to Y$ is continuous, then $f(K)$ is compact.

Theorem 2.7 (Heine-Cantor theorem)**.** If $f : K \to Y$ is continuous, where K is compact, then f is uniformly continuous.

We now turn our attention to the extreme value theorem (Theorem 2.8), which extends the familiar result from Calculus and Real Analysis (MA2002 and MA2108, respectively). This generalisation, originally attributed to Weierstrass, provides a broader framework for understanding the attainment of extrema.

Theorem 2.8 (extreme value theorem)**.** If $f : K \to \mathbb{R}$ is continuous, where K is compact, then f attains its maximum and minimum on K.

Lemma 2.2. Let X be a finite-dimensional normed space and let $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ be a basis for X. Then, there exist $0 < A \leq B$ such that

$$\text{for all } \mathbf{x} \in X \text{ where } \mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i \quad \text{we have} \quad A \sum_{i=1}^{n} |\alpha_i| \leq \|\mathbf{x}\| \leq B \sum_{i=1}^{n} |\alpha_i|.$$

Proof. Define

$$S = \left\{ (\alpha_1, \ldots, \alpha_n) \in \mathbb{R}^n : \sum_{i=1}^{n} |\alpha_i| = 1 \right\} \quad \text{which is is closed and bounded, hence compact.}$$

Note that S denotes the unit sphere in $\mathbb{R}^n$, i.e. the boundary of the unit ball. Define

$$f : S \to X \quad \text{where} \quad (\alpha_1, \ldots, \alpha_n) \mapsto \sum_{i=1}^{n} \alpha_i \mathbf{e}_i.$$

Let $(\alpha_1, \ldots, \alpha_n), (\beta_1, \ldots, \beta_n) \in S$. Then,

$$\begin{aligned} \|f(\alpha_1, \ldots, \alpha_n) - f(\beta_1, \ldots, \beta_n)\| &= \left\| \sum_{i=1}^{n} \alpha_i \mathbf{e}_i - \sum_{i=1}^{n} \beta_i \mathbf{e}_i \right\| \\ &= \sum_{i=1}^{n} |\alpha_i - \beta_i| \, \|\mathbf{e}_i\| \\ &\le \max_{1 \le i \le n} \|\mathbf{e}_i\| \sum_{i=1}^{n} |\alpha_i - \beta_i| \end{aligned}$$

Recall that $\|\cdot\| : X \to \mathbb{R}$ is continuous. Since S is compact and f is continuous, the composition $\|\cdot\| \circ f$ is a continuous real-valued function on the compact set S. By the extreme value theorem (Theorem 2.8), it achieves its minimum and maximum on S. Let $A = \min \|f\|$ and $B = \max \|f\|$.

Then, let $(\beta_1, \ldots, \beta_n)$ be a point in S where $\|f\|$ attains a minimum. Then,

$$\left\| \sum_{i=1}^{n} \beta_i \mathbf{e}_i \right\| = A.$$

Since we cannot have all $\beta_i = 0$ and $\{\mathbf{e}_1, \ldots, \mathbf{e}_n\}$ is linearly independent, then $A > 0$. The proof pretty much follows from here. $\square$

Theorem 2.9. Every finite-dimensional normed space is a Banach space.

Theorem 2.10 (equivalence of norms)**.** If X is a finite-dimensional vector space and $\|\cdot\|_a$ and $\|\cdot\|_b$ are norms on X, then there exist $0 < A \le B$ such that for all $\mathbf{x} \in X$, we have

$$A \|\mathbf{x}\|_a \le \|\mathbf{x}\|_b \le B \|\mathbf{x}\|_a$$

Proof. Let $\mathbf{x}_k$ be a Cauchy sequence in X, where X is a finite-dimensional vector space. Let $\mathbf{e}_1, \ldots, \mathbf{e}_n$ be a basis for X. So, we can write

$$\mathbf{x}_k = \sum_{i=1}^{n} \alpha_{ik} \mathbf{e}_i.$$

Fix i and consider the sequence α_{ik}, so $k \in \mathbb{N}$. Now,

$$\left\| \mathbf{x}_k - \mathbf{x}_j \right\| \ge A \sum_{i=1}^{n} \left| \alpha_{ij} - \alpha_{ik} \right| \ge A \left| \alpha_{ij} - \alpha_{ik} \right| \quad \text{which implies} \quad \left| \alpha_{ij} - \alpha_{ik} \right| \le \frac{1}{A} \left\| \mathbf{x}_k - \mathbf{x}_j \right\|.$$

As such, α_{ik} is a Cauchy sequence, hence convergent. Call its limit α_i, i.e.

$$\lim_{k \to \infty} \alpha_{ik} = \alpha_i.$$

Define

$$\mathbf{x} = \sum_{i=1}^{n} \alpha_i \mathbf{e}_i.$$

Then,

$$\|\mathbf{x}_k - \mathbf{x}\| \leq B \sum_{i=1}^{n} |\alpha_{ik} - \alpha_i|.$$

Since α_{ik} is a Cauchy sequence, then $\mathbf{x}_k \to \mathbf{x}$ and the result follows. □

Theorem 2.11 (equivalence of norms)**.** Let X be a finite-dimensional vector space and let $\|\cdot\|_1$ and $\|\cdot\|_2$ be norms on X. Then, there exist constants $0 \leq A < B$ such that

$$A\|\mathbf{x}\|_2 \leq \|\mathbf{x}\|_1 \leq B\|\mathbf{x}\|_2 \quad \text{for all } \mathbf{x} \in X.$$

2.3 Linear Operators

In the case of vector spaces and in particular, normed spaces, a mapping is called an operator. Of special interest are operators which preserve the two algebraic operations of a vector space, so we have the following definition of a linear operator (Definition 2.15):

Definition 2.15 (linear operator)**.** A linear operator T is an operator such that the domain $\mathcal{D}(T)$ is a vector space and the range $\mathcal{R}(T)$ lies in a vector space over the same field such that for all $x, y \in \mathcal{D}(T)$ and $\alpha \in \mathbb{F}$, we have

$$T(x+y) = T(x) + T(y) \quad \text{and} \quad T(\alpha x) = \alpha T(x).$$

Note that some authors may write Tx in place of $T(x)$ — this simplification is standard in Functional Analysis.

Definition 2.15 expresses the fact that a linear operator T is a homomorphism of a vector space (its domain) into another vector space (its codomain), i.e. T preserves the two operations of a vector space in the following sense: on the left of each equation, we first apply a vector space operation (addition or multiplication by a scalar respectively) and then map the resulting vector into Y, whereas on the right, we first map x and y into Y and then perform the vector space operations in Y, with the outcome being the same. This property makes linear operators important.

The reader should vividly recall from MA2001 the concepts of range and nullspace. For a linear operator T, we denote its range space or range using $\mathcal{R}(T)$; its nullspace is denoted by $\mathcal{N}(T)$.

Example 2.11 (identity operator)**.** For all $x \in X$, the identity operator

$$I_X : X \to X \quad \text{is defined by} \quad x \mapsto x.$$

We also simply write I in place of I_X, so $I(x) = x$.

Example 2.12 (zero operator)**.** For all $x \in X$, the zero operator

$$0 : X \to Y \quad \text{is defined by} \quad x \mapsto 0 \text{ (or equivalently } 0x = 0).$$

Example 2.13 (differentiation). Let X denote the vector space of all polynomials on $[a,b]$. We may define a linear operator T on X by setting

$$T(x(t)) = x'(t) \quad \text{for all } x \in X.$$

Here, the prime symbol denotes differentiation with respect to t.

Example 2.14 (integration). A linear operator T from $\mathcal{C}[a,b]$ to itself can be defined by

$$T(x(t)) = \int_a^t x(u)\ du \quad \text{for all } t \in [a,b].$$

Example 2.15 (multiplication operator). Let $X = \mathcal{C}[a,b]$. Another linear operator T from X to itself is defined by

$$T(x(t)) = t(x(t)).$$

Multiplication operators appear in Quantum Mechanics in various contexts, particularly in the representation of position observables. The most relevant connection is in the position operator. In the Schrödinger representation of quantum mechanics, the position operator $\widehat{x}$ acts on wave functions $\psi(x)$ as

$$(\widehat{x}\psi)(x) = x\psi(x).$$

Example 2.16 (dot product and cross product). The dot product $\cdot$ with one fixed factor defines a linear operator $T_1 : \mathbb{R}^3 \to \mathbb{R}$, say

$$T_1(\mathbf{x}) = \mathbf{x} \cdot \mathbf{a} = \xi_1\alpha_1 + \xi_2\alpha_2 + \xi_3\alpha_3 \quad \text{where} \quad \mathbf{a} = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3 \text{ is fixed.}$$

Also, the cross product with one product kept fixed defines a linear operator $T_2 : \mathbb{R}^3 \to \mathbb{R}^3$.

Example 2.17 (matrices). A real matrix $\mathbf{A} = (\alpha_{ij})$ with r rows and n columns defines an operator $T : \mathbb{R}^n \to \mathbb{R}^r$ by means of

$$\mathbf{y} = \mathbf{A}\mathbf{x} \quad \text{or equivalently} \quad \begin{bmatrix} \eta_1 \\ \eta_2 \\ \vdots \\ \eta_r \end{bmatrix} = \begin{bmatrix} \alpha_{11} & a_{12} & \cdots & \alpha_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ a_{r1} & a_{r2} & \cdots & a_{rn} \end{bmatrix} \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ x_n \end{bmatrix}.$$

Here, $\mathbf{x}$ has n rows and $\mathbf{y}$ has r components and both vectors are written as column vectors due to the usual convention of matrix multiplication.

Note that T is linear (we say that T is a linear transformation) because matrix multiplication is a linear operation. If on the other hand $\mathbf{A}$ were complex, then T would be a linear operator from $\mathbb{C}^n$ to $\mathbb{C}^r$.

Theorem 2.12. Let T be a linear operator. Then, the following hold:

(i) $\mathcal{R}(T)$ is a vector space

(ii) If $\dim(\mathcal{D}(T)) = n < \infty$, then $\dim(\mathcal{R}(T)) \leq n$

(iii) $\mathcal{N}(T)$ is a vector space

Corollary 2.1. Linear operators preserve linear independence.

Proof. This follows from **(b)** of Theorem 2.12. □

Theorem 2.13 (linear operator). Let X and Y be arbitrary vector spaces. Let $T : \mathcal{D}(T) \to Y$ be a linear operator such that $\mathcal{D}(T) \subseteq X$ and $\mathcal{R}(T) \subseteq Y$. Then, the following hold:

(i) The inverse $T^{-1} : \mathcal{R}(T) \to \mathcal{D}(T)$ exists if and only if $T(x) = 0$ implies $x = 0$

(ii) If T^{-1} exists, then it is a linear operator

(iii) If $\dim(\mathcal{D}(T)) = n < \infty$ and T^{-1} exists, then $\dim(\mathcal{R}(T)) = \dim(\mathcal{D}(T))$

Lemma 2.3 (inverse of composition). Let X, Y, Z be vector spaces and

$$T : X \to Y \quad \text{and} \quad S : Y \to Z \quad \text{be bijective linear operators.}$$

Then, the inverse $(S \circ T)^{-1} : Z \to X$ of the product ST exists, and $(S \circ T)^{-1} = T^{-1} \circ S^{-1}$.

2.4 Bounded and Continuous Linear Operators

When we were discussing linear operators previously, we did not make use of any norms. We now take norms into account, in the following basic definition:

Definition 2.16 (bounded linear operator). Let X and Y be normed spaces and $T : \mathcal{D}(T) \to Y$ be a linear operator, where $\mathcal{D}(T) \subseteq X$. The operator T is said to be bounded if there exists $c \in \mathbb{R}$ such that for all $x \in \mathcal{D}(T)$, we have

$$\|T(x)\|_Y \leq c\,\|x\|_X .$$

In Definition 2.16, on the LHS of the equation, the norm is that on Y, whereas the norm on the RHS is on X. Actually, we can denote both norms by the same symbol $\|\cdot\|$, i.e. the same subscript, without danger of confusion. Moreover, Definition 2.16 shows that a bounded linear operator maps bounded sets in $\mathcal{D}(T)$ onto bounded sets in Y.

However, our present use of the word 'bounded' is different from that in Calculus — for the latter, a bounded function is one whose range is a bounded set.

As we have

$$\frac{\|T(x)\|}{\|x\|} \leq c,$$

then it shows that c must be at least as big as the supremum of the expression on the left taken over $\mathcal{D}(T) \setminus \{0\}$. So, the smallest possible c in Definition 2.16 is that supremum. We denote this quantity by $\|T\|$, called the norm of T. We write

$$\|T\| = \sup_{x \in \mathcal{D}(T) \setminus \{0\}} \frac{\|T(x)\|}{\|x\|}.$$

On the other hand, if $\mathcal{D}(T) = \{0\}$, then we define $\|T\| = 0$.

Lemma 2.4. Let T be a bounded linear operator. Then, an alternative formula for the norm of T is

$$\|T\| = \sup_{\substack{x \in \mathcal{D}(T) \\ \|x\| = 1}} \|T(x)\| .$$

We then take a look at some typical examples of bounded linear operators so that we can better understand the concept of a bounded linear operator.

Example 2.18 (identity operator). The identity operator $I : X \to X$ on a normed space $X \neq \{0\}$ is bounded and has norm $\|I\| = 1$.

Example 2.19 (zero operator). The zero operator $0 : X \to Y$ on a normed space X is bounded and has norm $\|0\| = 0$.

Example 2.20 (differentiation operator). Let X be the normed space of all polynomials on $J = [0,1]$ with norm given by $\|x\| = \max |x(t)|$ for all $t \in J$. A differentiation operator T is defined on X by

$$T(x(t)) = x'(t).$$

Again, the prime denotes differentiation with respect to t. Note that this operator is linear but not bounded. For example, let $x_n(t) = t^n$, where $n \in \mathbb{N}$. Then, $\|x_n\| = 1$ and

$$T(x_n(t)) = x_n'(t) = nt^{n-1}$$

so

$$\|T(x_n)\| = n \quad \text{and} \quad \|T\| = \frac{\|T(x_n)\|}{\|x_n\|} = n.$$

Since $n \in \mathbb{N}$ is arbitrary, then there does not exist any $c \in \mathbb{R}$ such that $\|T\| \leq c$. As such, T is not a bounded operator.

> **Remark 2.2.** Since differentiation is an important operation, Example 2.20 seems to imply that unbounded operators are also of importance.

Example 2.21 (integral operator). We can define an integral operator $T : \mathcal{C}[0,1] \to \mathcal{C}[0,1]$ by

$$y = T(x) \quad \text{where} \quad y(t) = \int_0^1 k(t,u)\, x(u)\ du.$$

Here, k is a given function, which is called the kernel of T (think of the kernel as the weight, and obviously do not confuse this with the null space of the matrix of a linear transformation), and it is assumed to be continuous on the closed square $G = [0,1] \times [0,1]$ in the tu-plane. This operator is linear.

Moreover, the good news is tht T is bounded. To see why, we note that the continuity of k on the closed square implies k is bounded, i.e.

$$\text{for all } (t,u) \in G \quad \text{there exists } k_0 \in \mathbb{R} \quad \text{such that} \quad |k(t,u)| \leq k_0.$$

Moreover,

$$|x(t)| \leq \max_{t \in J} |x(t)| = \|x\|.$$

Hence,

$$\begin{aligned} \|y\| &= \|T(x)\| \quad \text{by definition of } y \\ &= \max_{t \in J} \left| \int_0^1 k(t,u)\, x(u)\ du \right| \\ &\leq \max_{t \in J} \int_0^1 |k(t,u)|\, |x(u)|\ du \quad \text{by the triangle inequality} \\ &\leq k_0 \|x\| \end{aligned}$$

We conclude that $\|T(x)\| \leq k_0 \|x\|$. This is precisely Definition 2.16 with $c = k_0$, so T is bounded.

Example 2.22 (matrices). Let $\mathbf{A} = \left(\alpha_{jk}\right)$ be a real matrix with r rows and n columns. Then, $\mathbf{A}$ defines an operator $T : \mathbb{R}^n \to \mathbb{R}^r$ by the equation

$$\mathbf{y} = \mathbf{A}\mathbf{x}.$$

Here, $\mathbf{x} = (\xi_1, \ldots, \xi_n)$ and $\mathbf{y} = (\eta_1, \ldots, \eta_r)$ are column vectors with n and r rows respectively. In terms of the components, $\mathbf{y} = \mathbf{A}\mathbf{x}$ becomes

$$\eta_j = \sum_{k=1}^{n} \alpha_{jk}\xi_k.$$

Clearly, T is linear because matrix multiplication is a linear operation.

We claim that T is bounded. Recall that the norm on $\mathbb{R}^n$ is given by

$$\|\mathbf{x}\| = \left(\sum_{m=1}^{n} \xi_m^2\right)^{1/2}.$$

The same idea holds for $y \in \mathbb{R}^r$. As such,

$$\begin{aligned}
\|T(\mathbf{x})\|^2 &= \sum_{j=1}^{r} \eta_j^2 \\
&= \sum_{j=1}^{r} \left(\sum_{k=1}^{n} \alpha_{jk}\xi_k\right)^2 \\
&\leq \sum_{j=1}^{r} \left[\left(\sum_{k=1}^{n} \alpha_{jk}^2\right)^{1/2} \left(\sum_{m=1}^{n} \xi_m^2\right)^{1/2}\right]^2 && \text{by the Cauchy-Shcwarz inequality} \\
&= \|\mathbf{x}\|^2 \sum_{j=1}^{r}\sum_{k=1}^{n} \alpha_{jk}^2
\end{aligned}$$

Since the double sum in the last line does not depend on x, we can express the inequality as

$$\|T(\mathbf{x})\|^2 \leq c^2 \|\mathbf{x}\|^2 \quad \text{where} \quad c^2 = \sum_{j=1}^{r}\sum_{k=1}^{n} \alpha_{jk}^2.$$

Hence, T is bounded.

Theorem 2.14. If a normed space X is finite-dimensional, then every linear operator on X is bounded.

Proof. Let $\dim X = n$ and $\{e_1, \ldots, e_n\}$ be a basis for X. We take any linear combination

$$x = \sum \xi_i e_i$$

and consider any linear operator T on X. Hence,

$$\begin{aligned}
\|T(x)\| &= \left\|\sum_{i=1}^{n} \xi_i T(e_i)\right\| && \text{since } T \text{ is a linear operator} \\
&\leq \sum_{i=1}^{n} |\xi_i| \, \|T(e_i)\| \\
&\leq \max_{1\leq k\leq n} \|T(e_k)\| \sum_{i=1}^{n} |\xi_i|
\end{aligned}$$

Hence,

$$\sum_{i=1}^{n} |\xi_i| \le \frac{1}{c} \left\| \sum_{i=1}^{n} \xi_i e_i \right\| = \frac{1}{c} \|x\| .$$

As such,

$$\|T(x)\| \le \frac{1}{c} \max_{1\le k\le n} \|T(e_k)\| \cdot \|x\|$$

and we conclude that T is bounded. □

Theorem 2.15 (continuity and boundedness). Let $T : X \to Y$ be a linear operator, where X, Y are normed spaces. Then, the following hold:

(i) T is continuous if and only if T is bounded

(ii) If T is continuous at a single point, it is continuous

Proof. We only prove **(i)** as **(ii)** is trivial. We start off with the reverse direction of **(i)**. Suppose T is a bounded linear operator and let $x \in X$. Let $\varepsilon > 0$ be arbitrary and define $\delta = \varepsilon / \|T\|$. Here, division by $\|T\|$ is under the assumption that T is a non-zero linear operator.

For any $y \in X$ such that $\|x - y\| < \delta$, we have

$$\|T(x) - T(y)\| = \|T(x-y)\| \le \|T\| \|x-y\| < \|T\| \delta = \varepsilon.$$

So, T is continuous.

We then prove the forward direction. Suppose T is continuous at some point $y \in X$. Then, for any $\varepsilon > 0$, there exists $\delta > 0$ such that $\|x - y\| < \delta$. Say $\|T(x) - T(y)\| < \varepsilon$. Let $z \in X$, where $z \neq 0$. The remaining part of the proof relies on the clever choice of y —

$$\text{choose } y = x + \frac{\delta}{\|z\|} z \quad \text{or equivalently} \quad x = y - \frac{\delta}{\|z\|} z$$

so

$$y - x = \frac{\delta}{\|z\|} z.$$

Since $\|x - y\| < \delta$, then

$$\|T(x) - T(y)\| = \|T(x-y)\| = \left\| T\left(\frac{\delta}{\|z\|} z\right) \right\| = \frac{\delta}{\|z\|} \cdot \|T(z)\| .$$

This implies

$$\|T(z)\| < \frac{\varepsilon}{\delta} \|z\| .$$

Thus, we can write this as $\|T(z)\| \le c\|z\|$, where $c = \varepsilon/\delta$. By Definition 2.16, we conclude that T is bounded. □

Corollary 2.2 (continuity). Let T be a bounded linear operator. Suppose $x_n, x \in \mathcal{D}(T)$. Then, the following hold:

(i) $x_n \to x$ implies $T(x_n) \to T(x)$

(ii) $\mathcal{N}(T)$ is closed

Definition 2.17 (equal operators)**.** Two operators T_1 and T_2 are equal, i.e. $T_1 = T_2$, if they have the same domain $\mathcal{D}(T_1) = \mathcal{D}(T_2)$ and if $T_1(x) = T_2(x)$ for all $x \in \mathcal{D}(T_1) = \mathcal{D}(T_2)$.

Definition 2.18 (restriction and extension)**.** The restriction of an operator $T : \mathcal{D}(T) \to Y$ to a subset $B \subseteq \mathcal{D}(T)$ is denoted by $T|_B$ and is the following operator:

$$T|_B : B \to Y \quad \text{where} \quad x \mapsto T(x) \text{ for all } x \in B.$$

An extension of T to a set $M \supseteq \mathcal{D}(T)$ is an operator

$$\widetilde{T} : M \to Y \quad \text{such that} \quad \widetilde{T}|_{\mathcal{D}(T)} = T.$$

That is, $\widetilde{T}(x) = T(x)$ for all $x \in \mathcal{D}(T)$.

Definition 2.19 (linear maps and bounded linear maps)**.** We let

$$\begin{array}{lll} \mathcal{L}(X,Y) & \text{denote} & \text{the set of all linear maps from } X \text{ to } Y \quad \text{and} \\ B(X,Y) & \text{denote} & \text{the set of all bounded linear maps from } X \text{ to } Y \end{array}$$

Theorem 2.16 (bounded linear extension)**.** Let $T : \mathcal{D}(T) \to Y$ be a bounded linear operator, where $\mathcal{D}(T)$ lies in a normed space X and Y is a Banach space. Then, T has an extension

$$\widetilde{T} : \overline{\mathcal{D}(T)} \to Y,$$

where $\widetilde{T}$ is a bounded linear operator of the form $\left\|\widetilde{T}\right\| = \|T\|$.

Theorem 2.17. If Y is complete, then $B(X,Y)$ is a Banach space.

2.5 Linear Functionals

Definition 2.20 (functional)**.** We define a functional to be an operator whose range lies on the real line $\mathbb{R}$ or in the complex plane $\mathbb{C}$.

We will denote functionals by lowercase letters $f, g, h, \ldots$, the domain of f by $\mathcal{D}(f)$, the range by $\mathcal{R}(f)$, and the value of f at an $x \in \mathcal{D}(f)$ by $f(x)$. Functionals are operators.

Definition 2.21 (linear functional)**.** A linear functional f is a linear operator with domain in a vector space X and range in the scalar field K of X, thus

$$f : \mathcal{D}(f) \to K,$$

where $K = \mathbb{R}$ if X is real and $K = \mathbb{C}$ if X is complex.

Naturally, we have the following definition of a bounded linear functional (Definition 2.22):

Definition 2.22 (bounded linear functional)**.** A bounded linear functional f is a bounded linear operator with range in the scalar field of the normed space X in which $\mathcal{D}(f)$ lies, i.e. there exists $c \in \mathbb{R}$ such that for all $x \in \mathcal{D}(f)$, the inequality

$$|f(x)| \leq c\,\|x\| \quad \text{holds.}$$

Furthermore, the norm of f is

$$\|f\| = \sup_{x \in \mathcal{D}(f) \setminus \{0\}} \frac{|f(x)|}{\|x\|} \quad \text{or} \quad \|f\| = \sup_{\substack{x \in \mathcal{D}(f) \\ \|x\|=1}} |f(x)|.$$

Theorem 2.18 (continuity and boundedness)**.** A linear functional f with domain $\mathcal{D}(f)$ in a normed space is continuous if and only if f is bounded.

Example 2.23 (norm)**.** Let $(X, \|\cdot\|)$ be a normed space. Then the norm $\|\cdot\| : X \to \mathbb{R}$ is a non-linear functional on X.

Example 2.24 (dot product)**.** The dot product with one factor kept fixed defines a functional $f : \mathbb{R}^3 \to \mathbb{R}$ by means of

$$f(x) = \mathbf{x} \cdot \mathbf{a} = \xi_1 \alpha_1 + \xi_2 \alpha_2 + \xi_3 \alpha_3 \quad \text{where} \quad \mathbf{a} = (\alpha_1, \alpha_2, \alpha_3) \in \mathbb{R}^3 \text{ is fixed.}$$

We see that f is linear and bounded. Moreover, we shall justify that the norm of f is $\|f\| = \|\mathbf{a}\|$. To see why, we have

$$|f(x)| = |\mathbf{x} \cdot \mathbf{a}| \leq \|\mathbf{x}\|\,\|\mathbf{a}\|.$$

Taking the supremum over all $\mathbf{x}$ of norm 1, it follows that $\|f\| \leq \|\mathbf{a}\|$. On the other hand, letting $\mathbf{x} = \mathbf{a}$, we obtain

$$\|f\| \geq \frac{|f(\mathbf{a})|}{\|\mathbf{a}\|} = \frac{\|\mathbf{a}\|^2}{\|\mathbf{a}\|} = \|\mathbf{a}\|.$$

This proves that $\|f\| = \|\mathbf{a}\|$.

Example 2.25 (definite integral)**.** The definite integral is a number if we consider it for a single function — as we do in Calculus. However, the situation changes completely if we consider that integral for all functions in a certain function space. As such, the integral becomes what is called a functional on that space. Call this functional f. As a space, we can choose the set of continuous functions on $[a,b]$, denoted by $\mathcal{C}[a,b]$. Thus, f is defined by

$$f(x) = \int_a^b x(t)\, dt.$$

Note that f is linear. We shall prove that f is bounded and has norm $\|f\| = b - a$. In fact, if we recall the norm on $\mathcal{C}[a,b]$, we obtain the following result which resembles the estimation lemma in Complex Analysis:

$$|f(x)| = \left| \int_a^b x(t)\, dt \right| \leq (b-a) \max_{t \in [a,b]} |x(t)| = (b-a)\,\|x\|.$$

Hence, taking the supremum over all x of norm 1, we obtain $\|f\| \leq b - a$. To obtain the inequality $\|f\| \geq b - a$, we choose $x = x_0 = 1$ and note that $\|x_0\| = 1$. As such,

$$\|f\| \geq \frac{|f(x_0)|}{\|x_0\|} = |f(x_0)| = \int_a^b dt = b - a.$$

It is of basic importance that the set of all linear functionals defined on a vector space X can itself be made into a vector space. This space is denoted by X^* and is called the algebraic dual space of X. Its algebraic operations of vector spaces are defined in a natural way as follows. The sum $f_1 + f_2$ of two functionals f_1 and f_2 is the functional s whose value at every $x \in X$ is

$$s(x) = (f_1 + f_2)(x) = f_1(x) + f_2(x).$$

Also, the product αf of a scalar α and a functional f is the functional p whose value at $x \in X$ is

$$p(x) = (\alpha f)(x) = \alpha f(x).$$

Of course, this is like the usual way of adding functions and multiplying them by constants.

We can further consider the algebraic dual $(X^*)^*$ of X^*, whose elements are linear functionals defined on X^*. We denote $(X^*)^*$ by X^{**} and call it the second algebraic dual space of X. It turns out that there exist some interesting relationships between X and X^{**} as shown in the following table:

Space	**General element**	**Value at a point**
X	x	$-$
X^*	f	$f(x)$
X^{**}	g	$g(f)$

We can obtain a $g \in X^{**}$ which is a linear functional defined on X^* by choosing a fixed $x \in X$ and then setting $g(f) = g_x(f) = f(x)$. The subscript x serves as a reminder that g is obtained by the use of a certain $x \in X$. Here, f is the variable, whereas x is fixed. Hence, g is clearly linear! For any $\alpha, \beta \in \mathbb{F}$, we have

$$g_x(\alpha f_1 + \beta f_2) = (\alpha f_1 + \beta f_2)(x) = \alpha f_1(x) + \beta f_2(x) = \alpha g_x(f_1) + \beta g_x(f_2).$$

As such, g_x is an element of X^{**}.

For each $x \in X$, we associate a functional $g_x \in X^{**}$. This defines a mapping

$$C : X \to X^{**}, \quad \text{where} \quad x \mapsto g_x.$$

The map C is called the canonical mapping of X into X^{**}. It is a linear map because X is a vector space, and

$$C(\alpha x + \beta y)(f) = g_{\alpha x + \beta y}(f) = f(\alpha x + \beta y) = \alpha f(x) + \beta f(y) = \alpha g_x(f) + \beta g_y(f) = \alpha (Cx)(f) + \beta (Cy)(f).$$

Thus, C is a linear map. The map C is also called the canonical embedding of X into X^{**}. To understand this concept, let us first introduce the notion of isomorphism, which is widely applicable.

In Mathematics, we often deal with spaces that share a common structure. Let X denote a set, and let a specific structure be defined on X. For instance, in a metric space, this structure is the metric; in a vector space, it consists of the two algebraic operations, namely addition and scalar multiplication; in a normed space, it includes the two algebraic operations as well as the norm.

Given two spaces X and $\widetilde{X}$ of the same type (for example, two vector spaces), we are interested in whether X and $\widetilde{X}$ are essentially identical, meaning they differ only in the nature of their elements, not their structural properties. If this is the case, we can treat X and $\widetilde{X}$ as two representations of the same abstract space.

This leads to the concept of an isomorphism, which is defined as a bijective mapping T between X and $\widetilde{X}$ that preserves the given structure. For a metric space $X = (X, d)$ and $\widetilde{X} = \left(\widetilde{X}, \widetilde{d}\right)$, an isomorphism T satisfies

$$\widetilde{d}(Tx, Ty) = d(x, y) \quad \text{for all } x, y \in X.$$

In general, we have the following definition:

Definition 2.23 (isomorphism)**.** An isomorphism T of a vector space X onto a vector space $\widetilde{X}$ over a field $\mathbb{F}$ is a bijective map that preserves the two algebraic operations of vector spaces. Specifically, for all $x, y \in X$ and scalars α,

$$T(x+y) = T(x) + T(y) \quad \text{and} \quad T(\alpha x) = \alpha T(x),$$

that is, $T : X \to \widetilde{X}$ is a bijective linear operator. The space $\widetilde{X}$ is then *isomorphic* to X, and X and $\widetilde{X}$ are called *isomorphic vector spaces*.

Isomorphisms for normed spaces are vector space isomorphisms that also preserve norms. We will discuss this in due course.

It can be shown that the canonical mapping C is injective. Since C is linear (as shown previously), it is an isomorphism of X onto the range $\mathcal{R}(C) \subset X^{**}$.

Definition 2.24 (embedding)**.** If X is isomorphic to a subspace of a vector space Y, we say that X is *embeddable* in Y. Hence, X is embeddable in X^{**}, and C is the *canonical embedding* of X into X^{**}.

Chapter 3
Hilbert Spaces

3.1
Inner Products and Hilbert Spaces

Definition 3.1 (inner product)**.** Let X be as vector space. An inner product is a function

$$\langle \cdot, \cdot \rangle : X \times X \to \mathbb{F}$$

satisfying the following properties for any $x, y, z \in X$:

(i) Non-negativity: $\langle x, x \rangle \geq 0$

(ii) Positive-definiteness: $\langle x, x \rangle = 0$ if and only if $x = 0$

(iii) Linearity in the first argument: $\langle x + y, z \rangle = \langle x, z \rangle + \langle y, z \rangle$

(iv) Homogeneity: for any $\alpha \in \mathbb{F}$, $\langle \alpha x, y \rangle = \alpha \langle x, y \rangle$

(v) Conjugate symmetry: $\langle x, y \rangle = \overline{\langle y, x \rangle}$

Example 3.1 (inner product on continuous functions)**.** Recall that $\mathcal{C}[a,b]$ denotes the set of continuous functions on $[a,b]$. Let $f, g \in \mathcal{C}[a,b]$. Then, we define their inner product to be

$$\langle f, g \rangle = \int_a^b f(x) \overline{g(x)}\, dx.$$

Theorem 3.1 (Cauchy-Schwarz inequality)**.** For any $x, y \in X$, we have

$$|\langle x, y \rangle| \leq \|x\| \|y\|.$$

Proof. If $y = 0$, then we are done. Suppose $y \neq 0$. Let $\alpha \in \mathbb{F}$ and we shall consider

$$\begin{aligned} 0 \leq \|x - \alpha y\|^2 &= \langle x - \alpha y, x - \alpha y \rangle \\ &= \langle x, x \rangle - \langle \alpha y, x \rangle - \langle x, \alpha y \rangle + \langle \alpha y, \alpha y \rangle \quad \text{by } \textbf{(ii)} \text{ of Definition 3.1} \\ &= \langle x, x \rangle - \alpha \langle y, x \rangle - \overline{\alpha} \langle x, y \rangle + \alpha \overline{\alpha} \langle y, y \rangle \quad \text{by } \textbf{(iii)} \text{ of Definition 3.1} \end{aligned}$$

Here, $\overline{\alpha} \in \mathbb{F}$ denotes the complex conjugate of α. Note that it satisfies the property $\alpha\overline{\alpha} = |\alpha|^2$. As such,

$$\alpha \langle y, x \rangle + \overline{\alpha \langle y, x \rangle} \leq \langle x, x \rangle + |\alpha|^2 \langle y, y \rangle.$$

The trick is to define

$$\alpha = \frac{\langle x, y \rangle}{\|y\|^2}$$

so the above inequality becomes

$$\frac{\langle x, y \rangle \langle y, x \rangle}{\|y\|^2} + \frac{\overline{\langle x, y \rangle \langle y, x \rangle}}{\|y\|^2} \leq \|x\|^2 + \frac{|\langle x, y \rangle|^2}{\|y\|^4} \cdot \|y\|^2.$$

As such,

$$\frac{2|\langle x, y \rangle|^2}{\|y\|^2} \leq \|x\|^2 + \frac{|\langle x, y \rangle|^2}{\|y\|^2} \quad \text{which implies} \quad \frac{|\langle x, y \rangle|^2}{\|y\|^2} \leq \|x\|^2.$$

The result follows from here. ☐

Theorem 3.2 (triangle inequality). For any $x, y \in X$, we have

$$\|x+y\| \leq \|x\| + \|y\|.$$

Proof. We have

$$\left|\|x+y\|^2\right| = |\langle x+y, x+y\rangle| = \left|\|x\|^2 + \|y\|^2 + \langle x,y\rangle + \langle y,x\rangle\right| \leq \|x\|^2 + 2\langle x,y\rangle + \|y\|^2 = (\|x\| + \|y\|)^2.$$

Hence, the result follows. □

Proposition 3.1 (parallelogram law). Let X be an inner product space. For any $x, y \in X$, we have

$$\|x+y\|^2 + \|x-y\|^2 = 2\left(\|x\|^2 + \|y\|^2\right).$$

Proof. Check homework! □

Proposition 3.2 (polarisation identity). Let X be an inner product space and $x, y \in X$. We have

$$\langle x,y\rangle = \frac{1}{4}\left(\|x+y\|^2 + \|x-y\|^2\right) \quad \text{if } \mathbb{F} = \mathbb{R}$$

and

$$\mathrm{Re}\left(\langle x,y\rangle\right) = \frac{1}{4}\left(\|x+y\|^2 + \|x-y\|^2\right) \quad \text{if } \mathbb{F} = \mathbb{C} \quad \text{and}$$
$$\mathrm{Im}\left(\langle x,y\rangle\right) = \frac{1}{4}\left(\|x+iy\|^2 + \|x-iy\|^2\right) \quad \text{if } \mathbb{F} = \mathbb{C}$$

Definition 3.2 (Hilbert space). A Hilbert space is a complete inner product space. More precisely, a Hilbert space is a vector space over a field $\mathbb{F}$ equipped with an inner product $\langle\cdot,\cdot\rangle H \times H \to \mathbb{F}$ and the inner product induces a norm given by

$$\|x\| = \sqrt{\langle x,x\rangle}.$$

A Hilbert space is a vector space that is complete with respect to this norm.

Definition 3.3 (convex set). A set Y is said to be convex if for all $y_1, y_2 \in Y$ and any $0 < \lambda < 1$, we have

$$z = \lambda y_1 + (1-\lambda) y_2 \in Y.$$

Theorem 3.3. Let X be an inner product space and $\emptyset \neq Y \subseteq X$ such that Y is convex and complete. Then,

$$\text{for any } x \in X \text{ there exists a unique } y_0 \in Y \quad \text{such that} \quad \inf_{y \in Y} \|x-y\| = \|x-y_0\|.$$

Proof. Let

$$\delta = \inf_{y \in Y} \|x-y\|.$$

Then, there exists a sequence $\{y_n\}_{n\in\mathbb{N}}$ in Y such that

$$\|x_n - y_n\| = \delta_n \quad \text{and} \quad \lim_{n\to\infty} \delta_n = \delta.$$

Define $z_n = y_n - x$ so $\|z_n\| = \delta_n$. For any $m,n \in \mathbb{N}$, we he

$$\|z_m + z_n\| = \|y_m + y_n - 2x\| = 2\left\|\frac{1}{2}(y_m + y_n) - x\right\| \geq 2\delta.$$

Next,

$$\begin{aligned}\|y_n - y_m\|^2 &= \|z_n - z_m\|^2 \\ &= -\|z_n + z_m\|^2 + 2\left(\|z_n\|^2 + \|z_m\|^2\right) \\ &\leq (2\delta)^2 + 2\left(\delta_n^2 + \delta_n^2\right)\end{aligned}$$

which is bounded above by $-2\delta^2$. For large $m,n \in \mathbb{N}$, note that δ_m, δ_n both tend to δ. As such, $y_n \to y_0$ and $\|x - y_0\| \geq \delta$. By the triangle inequality, we have

$$\|x - y_0\| \leq \|x - y_n\| + \|y_n - y_0\| = \delta_n + \|y_n - y_0\|.$$

As such, $\|x - y_0\| \leq \delta$. Combining both inequalities yields $\|x - y_0\| = \delta$. □

Definition 3.4 (orthogonal vectors)**.** Let X be an inner product space. For $x, y \in X$, if $\langle x, y\rangle = 0$, then x and y are said to be orthogonal.

Theorem 3.4. Let $Y \subseteq X$, where Y is a complete subspace. Let $y_0 \in Y$ such that

$$\inf_{y \in Y} \|x - y\| = \|x - y_0\|.$$

Then, for all $y \in Y$, we have $\langle x - y_0, y\rangle = 0$.

3.2 Orthogonal Complements and Direct Sums

Definition 3.5 (orthogonal complement)**.** Let X be an inner product space. For $Y \subseteq X$, define

$$Y^\perp = \{x \in X : \langle x, y\rangle = 0 \text{ for all } y \in Y\}.$$

Theorem 3.5 (orthogonal decomposition)**.** Let X be an inner product space. If Y is a closed subspace of X, then

$$X = Y \oplus Y^\perp.$$

That is, for all $x \in X$, there exists a unique $y \in Y$ and $z \in Y^\perp$ such that $x = y + z$.