Normed Vector Spaces

Vector Spaces

Throughout, $\mathbb{F} = \mathbb{R}$ or $\mathbb{C}$. A **vector space** over $\mathbb{F}$ is a nonempty set V equipped with two operations: vector addition and scalar multiplication.

$$\begin{aligned} V \times V &\xrightarrow{+} V \\ (v, w) &\mapsto v + w && \text{Vector Addition} \\ F \times V &\to V \\ (\alpha, v) &\mapsto \alpha v && \text{Scalar Multiplication} \end{aligned}$$

The vector space is an Abelian group, where $u, v, w \in V$ and $\alpha, \beta \in \mathbb{F}$, we have:

(i) $u + (v + w) = (u + v) + w$

(ii) $\exists 0_v \in V$ with $\forall v \in V,\ 0_v + v = v + 0_v = v$

(iii) $(\forall v \in V)(\exists w \in V)$ with $v + w = 0_v$

(iv) $\forall v, w \in V,\ v + w = w + v$

(v) $\alpha(v + w) = \alpha v + \alpha w,\ (\alpha + \beta)v = \alpha v + \beta v$

(vi) $\alpha(\beta w) = (\alpha\beta)w$

(vii) $1 \cdot v = v$

Remarks:

(a) 0_v is unique and known as the zero vector.

(b) The vector w in (iii) is unique, and denoted $-v$.

(c) $0 \cdot v = 0_v$

(d) $(-1) \cdot v = -v$

(e) Property (iv) follows from all the other axioms.

(f) For $n \in \mathbb{N}$, $n \cdot v = \underbrace{v + v + \cdots + v}_{n \text{ times}}$

Subspaces

Let V be a vector space over $\mathbb{F}$. A **subspace** is a nonempty subset $W \subseteq V$ satisfying the following:

(i) $w \in W, \alpha \in \mathbb{F} \to \alpha w \in W$.

(ii) $w_1, w_2 \in W \Rightarrow w_1 + w_2 \in W$.

Remark: 0_v is always a member of any subspace; a subspace is also a vector space.

Proposition: Intersection of Subspaces

If $\{W_i\}_{i \in I}$ is a family of subspaces of V, then, $\bigcap W_i$ is a subspace of V.

Proposition: Union of Subspaces

It is not the case that the union of subspaces of V also a subspace. For example, consider $\mathbb{R}^2$ with the traditional vector space operations:

$$\begin{pmatrix} x \\ y \end{pmatrix} + \begin{pmatrix} x' \\ y' \end{pmatrix} = \begin{pmatrix} x + x' \\ y + y' \end{pmatrix}$$
$$\alpha \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} \alpha x \\ \alpha y \end{pmatrix}$$

If $W_1, W_2 \in V$ are subspaces such that $W_1 \cup W_2$ is a subspace, then $W_1 \subseteq W_2$ or $W_2 \subseteq W_1$.

Generated Subspaces

Let $S \subseteq V$ be any subset of a vector space V. Then,

$$\text{span}(S) = \left\{ \sum_{j=1}^{n} \alpha_j v_j \mid \alpha_1, \dots, \alpha_n \in \mathbb{F}, v_1, \dots, v_n \in S \right\}$$

Remarks:

- $\text{span}(S) \subseteq V$ is a subspace.
- $\text{span}(S) = \bigcap W$, where $S \subseteq W$ and $W \subseteq V$ is a subspace. Thus, $\text{span}(S)$ is the "smallest" subspace containing S, or the subspace generated by S.

Proposition: Quotient Group on Vector Space

Let V be a vector space, and let $W \subseteq V$ is a subspace. Define $u \sim_W v \leftrightarrow u - v \in W$.

(1) $\sim_W$ is an equivalence relation.

(2) If $[v]_W$ denotes the equivalence class of v, then $[v]_W = v + W = \{v + w | w \in W\}$.

(3) $V/W := \{[v]_W | v \in V\}$ is a vector space with $[v_1]_W + [v_2]_W = [v_1 + v_2]_W$ and $\alpha[v]_W = [\alpha v]_W$.

Proof of (1):

- Reflexive: $u \sim_W u$, since $u - u = 0 \in W$.
- Transitive: Suppose $u \sim_W v$, and $v \sim_W z$. Then, $u - v \in W$, and $v - z \in W$. So, $(u - v) + (v - z) \in W$, so $u - z \in W$. Whence, $u \sim_W z$.
- Symmetric: If $u \sim_W v$, then $u - v \in W$, so $-1 \cdot (u - v) \in W$, so $v - u \in W$. Whence, $v \sim_W u$.

Proof of (2):

$$\begin{aligned} [v]_W &= \{u \in V \mid u \sim_W v\} \\ &= \{u \in V \mid u - v \in W\} \\ &= \{u \in V \mid u = v + w \text{ some } w \in W\} \\ &= \{v + w \mid w \in W\} \\ &= v + W \end{aligned}$$

Proof of (3): Prove that the operation is well-defined.

Bases

Let V be a vector space and $S \subseteq V$ be a subset.

(1) S is said to be spanning for V if $\text{span}(S) = V$.

(2) S is linearly independent if, for $\sum_{j=1}^{n} \alpha_j v_j = 0_v$ with $\alpha_1, \dots, \alpha_n \in \mathbb{F}$, $v_1, \dots, v_n \in S$, then $\alpha_1 = \alpha_2 = \cdots = \alpha_n = 0$.

(3) S is a basis for V if S is linearly independent and spanning for V.

Proposition: Existence of Basis

Every vector space admits a basis. If $B_0 \subseteq V$ is linearly independent, $\exists B \subseteq V$ such that B is a basis and $B \supseteq B_0$.

Background: A relation on a set X is a subset $R \subseteq X \times X$. If R is reflexive ($x \sim x$), transitive ($x \sim y, y \sim z \to x \sim z$), and antisymmetric ($x \sim y, y \sim x \to x = y$), then R is an ordering, and we write $x \leq y$.

If $\leq$ is an ordering of X such that $\forall x, y \in X,\ x \leq y$ or $y \leq x$, then $\leq$ is a total (or linear) ordering.

Let $\leq$ be an ordering of X, let $Y \subseteq X$. An upper bound for Y is an element $u \in X$ such that $y \leq u\ \forall y \in Y$. A maximal element in X is an element $m \in X$ such that $x \in X,\ x \geq m \to x = m$.

Example: $\mathbb{N}$ under the division ordering defines $a \leq b \Leftrightarrow a|b$. If we want to find the maximal elements of $A = \{2, 6, 9, 12\}$, we would see that they are 9 and 12 (since no element of A can be divided by 9 and 12). Meanwhile, $\mathbb{N}$ itself has no maximal elements.

This leads us to ask: given an ordered set, $(X, \leq)$, does X admit maximal elements.

Zorn's Lemma (or Axiom): Let $(X, \leq)$ be an ordered set. Suppose that every totally ordered subset, $Y \subseteq X$ has an upper bound in X. Then, X admits at least one maximal element.

The proof of Zorn's Lemma relies on the Axiom of Choice (and Zorn's Lemma is equivalent to the Axiom of Choice).

Proof: Let $X = \{D \mid B_0 \subseteq D \subseteq V\}$ with D linearly independent. Since $B_0 \subseteq X$, $X \neq \emptyset$. Define $D, E \in X$, $D \leq E \Leftrightarrow D \subseteq E$. We will show that X has a maximal element.

Consider any totally ordered subset, $Y = \{D_i\}_{i\in I}$. Consider $D = \bigcup D_i$. Clearly, $B_0 \subseteq D \subseteq V$. Suppose $\sum \alpha_k v_k = 0_v$ with $v_1, \ldots, v_n \in D$. Therefore, $\exists D_j$ with $v_1, \ldots, v_n \in D_j$ because Y is totally ordered. However, by definition, D_j is a linearly independent set — therefore, $\alpha_k = 0$. Thus, D is linearly independent.

Since D is linearly independent, and $B_0 \subseteq D$, it must be the case that $D \in X$. D is also an upper bound for Y. So, by Zorn's Lemma, X has a maximal element, B.

So, $B_0 \subseteq B \subseteq V$, B is independent, and B is maximal in X. We claim that B is a basis for V. Suppose toward contradiction that $\exists v \in V$ such that $v \notin \text{span}(B)$. Consider $B' = B \cup \{v\}$.

Then, $B_0 \subseteq B'$, and B' is linearly independent — if $\sum \alpha_k v_k + \alpha v = 0$, where $v_1, \ldots, v_n \in B$, then either:

- If $\alpha = 0$, then $\alpha_k v_k = 0 \Rightarrow \alpha_k = 0$.
- If $\alpha \neq 0$, then $\sum \alpha_k v_k = -\alpha v$, which means $v \in \text{span}(B)$. $\perp$

Thus, we have a linearly independent set, B', with $B \subseteq B'$, and $B_0 \subseteq B'$. Therefore, $B' \in X$. However, this contradicts the maximality of B. Therefore, $\text{span}(B) = V$, and B is a basis for V.

Examples: Vector Spaces

(1) n-Dimensional Vectors:

$$\mathbb{F}^n = \left\{ \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} \mid x_j \in \mathbb{F} \right\}$$
$$\begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} + \begin{pmatrix} y_1 \\ \vdots \\ y+n \end{pmatrix} = \begin{pmatrix} x_1 + y + 1 \\ \vdots \\ x_n + y + n \end{pmatrix}$$
$$\alpha \begin{pmatrix} x_1 \\ \vdots \\ x_n \end{pmatrix} = \begin{pmatrix} \alpha x_1 \\ \vdots \\ \alpha x_n \end{pmatrix}$$
$$B = \{e_1, \ldots, e_n\}$$

where e_i denotes the unit vector at position i.

(2) $m \times n$ Matrices:

$$\mathbb{M}_{m,n}(\mathbb{F}) = \left\{ \begin{pmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{m1} & \cdots & a_{mn} \end{pmatrix} \mid a_{ij} \in \mathbb{F} \right\}$$
$$(a_{ij}) + (b_{ij}) = (a_{ij} + b_{ij})$$
$$\alpha(a_{ij}) = (\alpha a_{ij})$$
$$B = \{e_{ij}\}$$

where e_{ij} denotes a matrix of 0 everywhere except column i and row j.

(3) Functions with domain Ω:

$$\begin{aligned}\mathcal{F}(\Omega, \mathbb{F}) &= \{f \mid f : \Omega \to \mathbb{F}\}\\ (f+g)(x) &= f(x) + g(x)\\ (\alpha f)(x) &= \alpha f(x)\end{aligned}$$

(4) Bounded functions with domain Ω:

$$\begin{aligned}\ell_\infty(\Omega, \mathbb{F}) &= \{f \in \mathcal{F}(\Omega, \mathbb{F}) \mid \|f\|_u \leq \infty\}\\ \|f\|_u &= \sup_{x\in\Omega} |f(x)|\end{aligned}$$

Exercises:

- Triangle Inequality: $\|f + g\|_u \leq \|f\|_u + \|g\|_u$
- Scalar Multiplication/Absolute Homogeneity: $\|\alpha f\|_u = |\alpha|\|f\|_u$
- Positive Definite: $\|f\|_u = 0 \Rightarrow f = \mathbb{0}$

Proof of Triangle Inequality: Given $x \in \Omega$,

$$\begin{aligned}|(f+g)(x)| &= |f(x) + g(x)|\\ &\leq |f(x)| + |g(x)|\\ &\leq \|f\|_u + \|g\|_u\end{aligned}$$

Therefore,

$$\begin{aligned}\sup |(f+g)(x)| &\leq \|f\|_u + \|g\|_u\\ \|f+g\|_u &\leq \|f\|_u + \|g\|_u\end{aligned}$$

(5) Continuous functions on closed and bounded intervals:

$$C([a, b], \mathbb{F}) = \{f : [a, b] \to \mathbb{F} \mid f \text{ continuous}\}$$

Check that $C([a, b], \mathbb{F}) \subseteq \ell_\infty([a, b], \mathbb{F})$ is a subspace.

(6) Let $f : [a, b] \to \mathbb{R}$ be any function. Let $\mathcal{P} : a = x_0 < x_1 < x_2 < \cdots < x_n = b$.

$$\begin{aligned}\mathrm{var}(f; \mathcal{P}) &:= \sum_{k=1}^{n} |f(x_k) - f(x_{k-1})|\\ \mathrm{var}(f) &= \sup_{\mathcal{P}} \mathrm{var}(f; \mathcal{P})\\ \mathrm{BV}([a, b]) &= \{f : [a, b] \to \mathbb{R} \mid \mathrm{var}(f) < \infty\}\\ \|f\|_{\mathrm{BV}} &= |f(a)| + \mathrm{var}(f)\end{aligned}$$

$\mathrm{BV}([a, b])$ is a vector space.

Question: Is $\mathbb{1}_{\mathbb{Q}} \in \mathrm{BV}([0, 1])$?

(7) Suppose $K \subseteq V$ is a *convex* subset of a vector space: $v, w \in K, t \in [0, 1] \Rightarrow (1 - t)v + tw \in K$. Let $\mathrm{Aff}(K) = \{f : K \to \mathbb{R} \mid f \text{ is affine}\}$, where f is affine if $\forall v, w \in K, t \in [0, 1], f((1 - t)v + tw) = (1 - t)f(v) + tf(w)$.

Exercise: Show that $\mathrm{Aff}(K) \subseteq \mathcal{F}(K, \mathbb{R})$ is a subspace.

(8) Let S be defined as

$$S = \{(a_k)_{k=1}^{\infty} \mid a_k \in \mathbb{F}\}.$$

Under pointwise operations, S is a vector space.

$$\begin{aligned}(a_k)_k + (b_k)_k &= (a_k + b_k)_k\\ \alpha(a_k)_k &= (\alpha a_k)_k\end{aligned}$$

Note 1: $S = \mathcal{F}(\mathbb{N}, \mathbb{F})$.

Note 2: $c_{00} \subseteq \ell_1 \subseteq c_0 \subseteq c \subseteq \ell_\infty \subseteq S$.

- $c_{00} = \{(a_k)_k \mid \text{finitely many } a_k \neq 0\}$
- $c_0 = \{(a_k)k \mid (a_k)_k \to 0\}$

- $c = \{(a_k)_k \mid (a_k)_k \to a < \infty\}$
- $\ell_\infty = \{(a_k)_k \mid \|(a_k)_k\|_u < \infty\}$
- $\ell_1 = \left\{(a_k)_k \mid \sum_{k=1}^{\infty} |a_k| = a < \infty\right\}$

(9) $C_C(\mathbb{R}) \subseteq C_0(\mathbb{R}) \subseteq \ell_\infty(\mathbb{R})$ are all subspaces.

- $C_C(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{F} \mid f \text{ compactly supported}\}$: $f : \mathbb{R} \to \mathbb{F}$ is compactly supported if $\exists [a, b]$ such that $x \notin [a, b] \Rightarrow f(x) = 0$.
- $C_0(\mathbb{R}) = \{f : \mathbb{R} \to \mathbb{F} \mid f \text{ continuous, } \lim_{x\to\pm\infty} f(x) = 0\}$

(10) Let S be any non-empty set.

$$\mathbb{F}(S) := \{f : S \to \mathbb{F} \mid f \text{ finitely supported}\}$$
$$\operatorname{supp}(f) = \{x \in S \mid f(x) \neq 0\}$$

We claim that $\mathbb{F}(S) \subseteq \mathcal{F}(S, \mathbb{F})$ is a subspace. Consider $e_t : S \to \mathbb{F}$ defined as follows:

$$e_t(s) = \begin{cases} 1 & s = t \\ 0 & s \neq t \end{cases}.$$

We claim that $\xi = \{e_t\}_{t\in S}$ is a basis for $\mathbb{F}(S)$.

Indeed, given $f \in \mathbb{F}(S)$, we know that $\operatorname{supp}(f) = \{t_1, \dots, t_n\} \subseteq S$. Therefore, $f = \sum_{k=1}^{n} f(t_k)e_{t_k} \in \operatorname{span}(\xi)$. Therefore, ξ is spanning for $\mathbb{F}(S)$. Suppose $\sum_{k=1}^{n} \alpha_{t_k} e_{t_k} = \mathbb{0}$ for some $\alpha_k \in \mathbb{F}$, $t_k \in S$.

$$\left(\sum_{k=1}^{\alpha_{t_k}} e_{t_k}\right) = \mathbb{0}(t_1)$$
$$\alpha_{t_1} = 0.$$

Similarly, $\alpha_{t_j} = 0$ for $j = 1, \dots, n$. Therefore, ξ is linearly independent. Since ξ is linearly independent and spanning, ξ forms a basis for $\mathbb{F}(S)$.

Note: The free vector space, $\mathbb{F}(S)$, displays the universal property.

> There are functions $\iota : S \to \mathbb{F}(S)$, where $\iota(t) = e_t$, and given any map $\varphi : S \to V$ for V a vector space over $\mathbb{F}$, $\exists!$ linear map $T_\varphi : \mathbb{F}(S) \to V$ such that $\iota \circ T_\varphi = \varphi$.

$$\begin{tikzcd} S \arrow[r, "\iota"] \arrow[dr, "\varphi"'] & \mathbb{F}(S) \arrow[d, "T_\varphi"] \\ & V \end{tikzcd}$$

Proof: Every $f \in \mathbb{F}(S)$ has a unique expression $f = \sum_{k=1}^{n} f(t_k)e_{t_k}$, where $\operatorname{supp}(f) = \{t_1, \dots, t_n\}$. Therefore,

$$T_\varphi(f) := \sum_{k=1}^{n} f(t_k)\varphi(t_k)$$

Exercise: Show T_φ is linear and unique.

Exercise 2: Suppose V is a vector space over $\mathbb{F}$ with basis B. Show that $\mathbb{F}(B) \cong V$. Remember that $V \cong W$ if $\exists\, T : V \to W$ such that T is bijective and linear.

Normed Spaces

To every vector $v \in V$, we want to assign a length to v, $\|v\|$.

A **norm** on a vector space V is a map

$$\|\cdot\| : V \to \mathbb{R}^+$$
$$v \mapsto \|v\| \geq 0$$

such that

(i) Homogeneity: $\|\alpha v\| = |\alpha|\|v\|$

(ii) Triangle Inequality: $\|v + w\| \leq \|v\| + \|w\|$

(iii) Positive definiteness: $\|v\| = 0 \Rightarrow v = \mathbb{0}_V$.

If $p : V \to \mathbb{R}^+$ satisfies (i) and (ii), then p is a **seminorm**.

The pair $(V, \|\cdot\|)$ is called a normed space.

Two norms, $\|\cdot\|$ and $\|\cdot\|'$ are called **equivalent** if $\exists c_1, c_2 \geq 0$ with, $\forall v \in V$,

$$\begin{aligned} \|v\| &\leq c_1\|v\|' \\ \|v\|' &\leq c_2\|v\| \end{aligned}$$

Note: On $\mathbb{R}^n$, all norms are equivalent.

Exercise: If p is any seminorm on V, then $|p(v) - p(w)| \leq p(v - w)$.

Notation: If V is a normed space, then $B_V = \{v \in V \mid \|v\| \leq 1\}$, and $U_V = \{v \in V \mid \|v\| < 1\}$ are the closed and open unit ball respectively.

Examples of Normed Spaces

(1) Given $V = \mathbb{F}^n$ and $x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$, we have different norms:

$$\begin{aligned} \|x\|_1 &= \sum_{j=1}^{n} |x_j| \\ \|x\|_\infty &= \max_{1\leq j\leq n} |x_j| \\ \|x\|_2 &= \left(\sum_{j=1}^{n} |x_j|^2\right)^{1/2}. \end{aligned}$$

In general, for $1 \leq p < \infty$,

$$\|x\|_p = \left(\sum_{j=1}^{\infty} |x_j|^p\right)^{1/p}.$$

Exercise: Show that $\|\cdot\|_1$ and $\|\cdot\|_\infty$ are norms. Show that $\lim_{p\to\infty} \|x\|_p = \|x\|_\infty$.

We want to show that $\|\cdot\|_p$ defines a norm for $1 \leq p < \infty$. If $1 \leq p < \infty$, its conjugate index $q \in [1, \infty]$ whereby $\frac{1}{p} + \frac{1}{q} = 1$. For example, if $p = 1$, then $q = \infty$, and if $p = \infty$, then $q = 1$.

Lemma 1: For $1 < p < \infty$, $p^{-1} + q^{-1} = 1$, $f : [0, \infty) \to \mathbb{R}$, $f(t) = \frac{1}{p}t^p - t + \frac{1}{q}$. Then, $f(t) \geq 0$ for all $t \geq 0$.

Proof 1: We can see that $f'(t) = t^{p-1} - 1$. Then, $f'(t) = 0$ at $t = 1$; $f'(t) > 0$ for $t > 1$ and $f'(t) < 0$ for $t \in [0, 1)$.

So, since $f(t) \geq f(1)$ for all $t \geq 0$, and $f(1) = 0$, $f(t) \geq 0$ for all $t \geq 0$.

Lemma 2: For $1 < p < \infty$, $p^{-1} + q^{-1} = 1$, $z, y \geq 0$, $xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q$.

Proof 2: We know from Lemma 1, $t \leq \frac{1}{p}t^p + \frac{1}{q}$. Multiply by y^q to get

$$ty^q \leq \frac{1}{p}t^p y^q + \frac{1}{q}y^q.$$

Set $t = xy^{1-q}$. Then,

$$xy^{1-q}y^q \leq \frac{1}{p}x^p y^{p-pq}y^q + \frac{1}{q}y^q.$$

Since $\frac{1}{p} + \frac{1}{q} = 1$, $p - pq = -q$, so

$$xy \leq \frac{1}{p}x^p + \frac{1}{q}y^q.$$

With these two lemmas in mind, we get two important inequalities.

Hölder's Inequality: For $1 \leq p \leq \infty$, $p^{-1} + q^{-1} = 1$. Then, for $x, y \in \mathbb{F}^n$,

$$\left|\sum_{j=1}^{n} x_j y_j\right| \leq \|x\|_p \|y\|_q.$$

Proof of Hölder's Inequality: For $p = 1$, the solution is as follows:

$$\begin{aligned}
\left|\sum_{j=1}^{n} x_j y_j\right| &\leq \sum_{j=1}^{n} |x_j||y_j| \\
&\leq \sum_{j=1}^{n} |x_j| \|y\|_\infty \\
&= \|x\|_q \|y\|_\infty,
\end{aligned}$$

and similarly for $p = \infty, q = 1$.

For $1 < p < \infty$, assume $\|x\|_p = \|y\|_q = 1$.

$$\begin{aligned}
\left|\sum_{j=1}^{n} x_j y_j\right| &\leq \sum_{j=1}^{\infty} |x_j||y_j| \\
&\leq \sum_{j=1}^{n} \left(\frac{1}{p}|x_j|^p + \frac{1}{q}|y_j|^q\right) \\
&= \frac{1}{p}\left(\sum_{j=1}^{n} |x_j|^p\right) + \frac{1}{q}\left(\sum_{j=1}^{n} |y_j|^q\right) \\
&= \frac{1}{p} + \frac{1}{q} \\
&= 1.
\end{aligned}$$

If $\|x\|_p = 0$ or $\|y\|_q = 0$, then $x = \mathbb{0}_{\mathbb{F}}$ or $y = \mathbb{0}_{\mathbb{F}}$, the inequality still holds.

Assume $\|x\|_p \neq 0$, $\|y\|_p \neq 0$. Set

$$\begin{aligned}
x' &= \frac{x}{\|x\|_p} \\
y' &= \frac{y}{\|y\|_p}.
\end{aligned}$$

It can be verified that $\|x'\|_p = 1 = \|y'\|_q$. Therefore,

$$\begin{aligned}
\left|\sum_{j=1}^{n} x_j' y_j'\right| &\leq 1 \\
\left|\sum_{j=1}^{n} \frac{x_j}{\|x\|_p} \frac{y_j}{\|y\|_q}\right| &\leq 1 \\
\left|\sum_{j=1}^{n} x_j y_j\right| &\leq \|x\|_p \|y\|_q
\end{aligned}$$

Minkowski's Inequality: Given $x, y \in \mathbb{F}^n$, $1 \leq p \leq \infty$, $\frac{1}{p} = \frac{1}{q} = 1$,

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

Proof of Minkowski's Inequality: We can verify for $p = 1, q = \infty$, and vice versa.

Assume $1 < p < \infty$. Then,

$$\begin{aligned}
\|x+y\|_p^p &= \sum_{j=1}^{n} |x_j + y_j|^p \\
&= \sum_{j=1}^{\infty} |x_j + y_j||x_j + y_j|^{p-1} \\
&\leq \sum_{j=1}^{\infty} |x_j||x_j + y_j|^{p-1} + \sum_{j=1}^{n} |y_j||x_j + y_j|^{p-1} \\
&\leq \left(\sum_{j=1}^{n} |x_j|^p\right)^{1/p} \left(\sum_{j=1}^{n} |x_j + y_j|^{pq-q}\right)^{1/q} + \left(\sum_{j=1}^{n} |y_j|^p\right)^{1/p} \left(\sum_{j=1}^{n} |x_j + y_j|^{pq-q}\right)^{1/q} && \text{Hölder's Inequality} \\
&= \|x\|_p \|x+y\|_p^{p/q} + \|y\|_p \|x+y\|_p^{p/q} \\
&= \left(\|x\|_p + \|y\|_p\right) \|x+y\|_p^{p-1}
\end{aligned}$$

Divide by $\|x+y\|_p^{p-1}$ to get desired inequality.

(2) $\ell_\infty(\Omega, \mathbb{F})$ with $\|\cdot\|_u$. This includes subspaces that inherit the norm, such as

$$\begin{aligned}
C([a,b]) &\subseteq \ell_\infty(\Omega) \\
\ell_\infty(\mathbb{R}) &\supseteq C_0(\mathbb{R}) \supseteq C_C(\mathbb{R})
\end{aligned}$$

Exercise: Show that $C_0(\mathbb{R}) \subseteq \ell_\infty(\mathbb{R})$ is a subspace.

(3) $\Omega = \mathbb{N}$, $\ell_\infty = \ell_\infty(\mathbb{N})$ with $\|\cdot\|_\infty$. Subspaces that inherit the norm are

$$c_{00} \subseteq c_0 \leq \ell_\infty.$$

(4) ℓ_1 with $\|\cdot\|_1$,

$$\|(a_k)_k\|_1 = \sum_{k=1}^{n} |a_k|.$$

(5) $C([a,b])$ with

$$\|f\|_1 = \int_a^b |f(x)| dx.$$

(6) Let $1 \leq p < \infty$.

$$\ell_p = \left\{ (a_k)_{k=1}^{\infty} \mid \sum_{k=1}^{\infty} |a_k|^p < \infty \right\}$$

is a normed space with

$$\|(a_k)_k\|_p = \left(\sum_{k=1}^{\infty} |a_k|^p\right)^{1/p}$$

We will show that the triangle inequality holds for this norm.

$$\begin{aligned}
\left(\sum_{k=1}^{n} |a_k + b_k|^p\right)^{1/p} &= \left\| \begin{bmatrix} a_1 + b_1 \\ \vdots \\ a_n + b_n \end{bmatrix} \right\|_{\ell_p^n} \\
&= \left\| \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} + \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \right\|_{\ell_p^n} \\
&\leq \left\| \begin{bmatrix} a_1 \\ \vdots \\ a_n \end{bmatrix} \right\| + \left\| \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix} \right\|_{\ell_p^\infty} \\
&\leq \|(a_k)_k\|_p + \|(b_k)_k\|_p.
\end{aligned}$$

Taking the limit as $n \to \infty$ (by the definition of an infinite series), we find that $\|(a_k)_k + (b_k)_k\|_p \leq \|(a_k)_k\|_p + \|(b_k)_k\|_p$.

(7) $\text{BV}([a,b]) = \{f : [a,b] \to \mathbb{R} \mid \text{Var}(f) < \infty\}$ with the norm $\|f\|_{\text{BV}} = |f(a)| + \text{Var}(f)$ is a normed space:

$$\begin{aligned} \|f\|_{\text{BV}} &= 0 \\ |f(a)| &= 0 \\ \text{Var}(f) &= 0 \end{aligned}$$

given $t \in (a,b]$, look at the partition $a < t \leq b$. Then,

$$\begin{aligned} \text{Var}(f) &\geq |f(t) - f(a)| + |f(b) - f(t)| \\ f(t) &= 0 \\ f &= \mathbb{0}_f. \end{aligned}$$

(8) $\mathbb{M}_{m,n}(\mathbb{F})$ with

$$\|a\|_{\text{op}} = \sup_{\|\xi\|_{\ell_2^n} \leq 1} \|a\xi\|_{\ell_2^m}$$

is a normed vector space. If $\|a\|_{\text{op}} = 0$, then

$$ae_j = 0 \qquad \forall j \in \{1, \dots, n\}.$$

take the dot product with $i \neq j$

$$\begin{aligned} ae_j \cdot e_i &= a_{ij} \\ &= 0 \end{aligned}$$

so $a_{ij} = 0$ for all a_{ij}, so a is the $\mathbb{0}$ matrix.

(9) Let V, W be vector spaces over $\mathbb{F}$. Then, $\mathcal{L}(V, W) = \{T \mid T : V \to W \text{ linear}\}$, where $T(\alpha v_1 + \beta v_2) = \alpha T(v_1) + \beta T(v_2)$.

$\mathcal{L}(V, W)$ is a vector space with operations

$$\begin{aligned} (T + S)(v) &= T(v) + S(v) \\ (\alpha T)(v) &= \alpha T(v). \end{aligned}$$

Notation: $\mathcal{L}(V) := \mathcal{L}(V, V)$ is all linear operators on V. $\mathcal{L}(V, \mathbb{F}) = V'$ is all linear functionals.

Suppose V and W are normed vector spaces. If $T : V \to W$, set

$$\begin{aligned} \|T\|_{\text{op}} &:= \sup_{\|v\|_v \leq 1} \|T(v)\|_W, \\ \mathbb{B}(V, W) &= \{T \in \mathcal{L}(V, W) \mid \|T\|_{op} \leq \infty\}, \end{aligned}$$

where $\mathbb{B}(V, W)$ is referred to as the set of all bounded linear maps from V to W. $\mathbb{B}(V, W)$ with $\|\cdot\|_{\text{op}}$ is a normed space.

- Homogeneity:

$$\begin{aligned} \|\alpha T\|_{[op]} &= \sup_{\|v\|_V \leq 1} \|\alpha T(v)\|_W \\ &= \sup_{\|v\|_V \leq 1} |\alpha| \|T(v)\|_W \\ &= |\alpha| \sup_{\|v\|_V \leq 1} \|T(v)\|_W \\ &= |\alpha| \|T\|_{\text{op}}. \end{aligned}$$

- Triangle Inequality: for $\|v\|_V \leq 1$,

$$\begin{aligned} \|\left(T + S\right)(v)\|_W &= \|T(v) + S(v)\|_W \\ &\leq \|T(v)\|_W + \|S(v)\|_W \\ &\leq \|T\|_{\text{op}} + \|S\|_{\text{op}} \end{aligned}$$

so

$$\begin{aligned} \|T + S\|_{\text{op}} &= \sup_{\|v\| \leq 1} \|T + S(v)\| \\ &\leq \|T\|_{\text{op}} + \|S\|_{\text{op}} \end{aligned}$$

- Positive Definite: If $\|T\|_{\text{op}} = 0$, then $T(v) = 0$ for all $v \in V$, $\|v\| \leq 1$.

  Let $v \in V$, $v \neq 0$. Then, $\frac{v}{\|v\|} \in B_V$.

$$\begin{aligned} T\left(\frac{v}{\|v\|}\right) &= 0 \\ \frac{1}{\|v\|}T(v) &= 0 \\ T(v) &= 0 \end{aligned}$$

Special Cases: $\mathbb{B}(V) = \mathbb{B}(V, V)$, $V^* = \mathbb{B}(V, \mathbb{F})$.

Exercise: $\mathcal{L}(\mathbb{F}^n, \mathbb{F}^m) = \mathbb{B}(\ell_2^n, \ell_2^m)$.

(10) Inner Product Spaces (expanded upon below).

Inner Product Spaces

An inner product on a vector space V is a pairing

$$V \times V \xrightarrow{\langle \cdot, \cdot \rangle} \mathbb{F}$$

that satisfies

(i) $\langle v_1 + v_2, w\rangle = \langle v_1, w\rangle + \langle v_2, w\rangle$, $\langle \alpha v, w\rangle = \alpha\langle v, w\rangle$.

(ii) $\langle v, w\rangle = \overline{\langle w, v\rangle}$

(iii) $\langle v, v\rangle \geq 0$.

(iv) If $\langle v, v\rangle = 0$, then $v = 0$.

The pair $(V, \langle \cdot, \cdot\rangle)$ is known as an inner product space.

Remarks: $\langle v, w_1 + w_2\rangle = \langle v, w_1\rangle + \langle v, w_2\rangle$, $\langle v, \alpha w\rangle = \overline{\alpha}\langle v, w\rangle$.

If $\langle \cdot, \cdot\rangle$ is an inner product on a linear space V, then set

$$\|v\|_2 := \langle v, v\rangle^{1/2}.$$

Exercise: $\|\alpha v\|_2 = |\alpha|\|v_2\|$, $\|v\|_2 = 0 \Rightarrow v = 0$.

$v, w \in (V, \langle, \cdot, \cdot\rangle)$ are *orthogonal* if $\langle v, w\rangle = 0$.

The Pythagoran theorem states that for $v_1, \dots, v_n \in V$ mutually orthogonal, then

$$\left\|\sum_{i=1}^{n} v_i\right\|^2 = \sum_{j=1}^{n} \|v_j\|^2.$$

For two vectors $v, w \in V$, $P_w(v) = \frac{\langle v, w\rangle}{\langle w, w\rangle}w$.

Exercise: Check that $\langle P_w(v), v - P_w(v)\rangle$, meaning

$$\|v\|^2 = \|P_w(v)\|^2 + \|v - P_w(v)\|^2$$

Cauchy-Schwarz Inequality: In any inner product space,

$$|\langle v, w\rangle| \leq \|v\| \cdot \|w\|.$$

Proof of Cauchy-Schwarz: From the exercise,

$$\begin{aligned} \|v\| &\geq \|P_w(v)\| \\ \|v\| &\geq \left\|\frac{\langle v, w\rangle}{\langle w, w\rangle}w\right\| \\ &= \frac{|\langle v, w\rangle|}{\|w\|^2}\|w\| \end{aligned}$$

therefore,

$$\|v\|\|w\| \geq |\langle v, w\rangle|$$

The triangle inequality follows from the Cauchy-Schwarz inequality.

Proof of Triangle Inequality:

$$\begin{aligned}
\|v+w\|_2^2 &= \langle v+w, v+w\rangle \\
&= \langle v,v\rangle + \langle v,w\rangle + \langle w,v\rangle + \langle w,w\rangle \\
&= \|v\|^2 + \|w\|^2 + \langle v,w\rangle + \overline{\langle v,w\rangle} \\
&= \|v\|^2 + \|w\|^2 + 2\mathrm{Re}\,\langle v,w\rangle \\
&\leq \|v\|^2 + \|w\|^2 + 2|\langle v,w\rangle| \\
&\leq \|v\|^2 + \|w\|^2 + 2\|v\|\|w\| && \text{Cauchy-Schwarz Inequality} \\
&= (\|v\| + \|w\|)^2.
\end{aligned}$$

Take square roots on both sides.

(1) $\ell_2^n = \mathbb{F}^n$ with

$$\left\langle \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}, \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} \right\rangle = \sum_{i=1}^{n} x_i\overline{y_i}.$$

Cauchy-Schwarz is found as

$$\left|\sum_{j=1}^{n} x_j\overline{y_j}\right| \leq \left(\sum_{j=1}^{n} |x_j|^2\right)^{1/2} \left(\sum_{j=1}^{n} |y_j|^2\right)^{1/2}.$$

(2) ℓ_2 with

$$\left\langle (a_j)_j, (b_j)_j \right\rangle = \sum_{j=1}^{\infty} a_j\overline{b}_j.$$

We can see that for any finite n, the Cauchy-Schwarz inequality in ℓ_2^n states

$$\begin{aligned}
\left|\sum_{j=1}^{n} a_j\overline{b_j}\right| &\leq \left(\sum_{j=1}^{n} |a_j|^2\right)^{1/2} \left(\sum_{j=1}^{n} |b_j|^2\right)^{1/2} \\
&\leq \left(\sum_{j=1}^{\infty} |a_j|^2\right)^{1/2} \left(\sum_{j=1}^{\infty} |b_j|^2\right)^{1/2}.
\end{aligned}$$

Taking the limit as $n \to \infty$, we see that $\left\langle (a_j)_j, (b_j)_j \right\rangle$ is convergent.

(3) $C([a,b])$ with

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

(4) Let $V = \mathbb{M}_n(\mathbb{C})$.

Recall that if

$$a = (a_{ij})_{i,j},$$

then

$$a^* = (\overline{a_{ji}})_{i,j}.$$

Let $\mathrm{Tr} : \mathbb{M}_n(\mathbb{C}) \to \mathbb{C}$, $\mathrm{Tr}((a_{ij})) = \sum_{i=1}^{n} a_{ii}$.

- $\mathrm{Tr}(I_n) = n$
- $\mathrm{Tr}(a + \alpha b) = \mathrm{Tr}(a) + \alpha\mathrm{Tr}(b)$
- $\mathrm{Tr}(ab) = \mathrm{Tr}(ba)$

Then, if $\operatorname{Tr}(a^*a) = 0$, then $a = \mathbb{0}_{\mathbb{M}_n}$.

$$\begin{aligned} a^*a &= (\overline{a_{ji}})_{i,j}(a_{ij})_{i,j}\\ &= \left(\sum_{k=1}^{n} \overline{ki}a_{kj}\right)_{i,j}\\ \operatorname{Tr}(a^*a) &= \sum_{i=1}^{n}\sum_{k=1}^{n} \overline{a_{ki}}a_{ki}\\ &= \sum_{i,k=1}^{n} |a_{ki}|^2\\ &= \sum_{i,j=1}^{n} |a_{ij}|^2. \end{aligned}$$

If $\operatorname{Tr}(a^*a) = 0$, then $a_{ij} = 0$ for all i, j.

We define

$$\langle a, b\rangle_{\text{HS}} = \operatorname{Tr}(b^*a).$$

(i) $(b_1 + b_2)^* = b_1^* + b_2^*$
(ii) $(\alpha b)^* = \overline{\alpha}b^*$
(iii) $(b_1b_2)^* = b_2^*b_1^*$
(iv) $b^{**} = b$

The norm is defined as

$$\begin{aligned} \|a\|_{\text{HS}} &= \langle a, a\rangle^{1/2}\\ &= \operatorname{Tr}(a^*a)^{1/2}\\ &= \left(\sum_{i,j=1}^{n} |a_{ij}|^2\right)^{1/2} \end{aligned}$$

Metric Spaces

We looked at normed spaces, where we attach a length $\|v\|$ to very vector v. We can also speak of the distance between two vectors, defined as $d(v, w) = \|v - w\|$.

Notice that the following hold:

- $d(v, w) \geq 0$
- $$\begin{aligned} d(v, w) &= \|v - w\|\\ &= \|(-1)(w - v)\|\\ &= |-1|\,\|w - v\|\\ &= \|w - v\| \end{aligned}$$
- $$\begin{aligned} d(u, w) &= \|u - w\|\\ &= \|u - v + v - w\|\\ &\leq \|u - v\| + \|v - w\|\\ &= d(u, v) + d(v, w). \end{aligned}$$
- $d(v, v) = \|v - v\| = 0$. If $d(v, w) = 0$, then $\|v - w\| = 0$, so $v - w = \mathbb{0}$, so $v = w$.

In Real Analysis I, we studied the properties (such as convergence, limits, and continuity) of a particular normed vector space, namely $(\mathbb{R}, |\cdot|)$. We will expand these concepts to all metric spaces.

Definition of a Metric Space

Let X be a non-empty set. A **metric** on X is a map

$$\begin{aligned} d &: X \times X \to \mathbb{R}^+ \\ (x, y) &\mapsto d(x, y) \geq 0 \end{aligned}$$

such that

(i) Symmetry: $d(x, y) = d(y, x)$ for all $x, y \in X$.

(ii) Triangle Inequality: $d(x, z) \leq d(x, y) + d(y, z)$ for all $x, y, z \in X$.

(iii) Zero Distance: $d(x, x) = 0$

(iv) Definite: $d(x, y) = 0 \Rightarrow x = y$

If d satisfies (i), (ii), and (iii), then d is called a semi-metric. If d satisfies (iv) as well, then d is a metric.

If d is a (semi-)metric on X, the pair (X, d) is called a (semi-)metric space.

Two metrics, d and ρ, on X, are equivalent if $\exists c_1, c_2 \geq 0$ such that $d(x, y) \leq c_1\rho(x, y)$ and $\rho(x, y) \leq c_2 d(x, y)$ for all x, y.

Examples of Metric Spaces

(1) Discrete Metric:

$$d(x, y) = \begin{cases} 1 & x \neq y \\ 0 & x = y \end{cases}$$

for X any set.

(2) Hamming distance: between two bit strings of equal length. Let

$$\begin{aligned} X &= \{0, 1\}^n \\ &= \{0, 1\} \underbrace{\times \cdots \times}_{n \text{ times}} \{0, 1\} \\ d_H((x_j)_1^n, (y_j)_1^n) &= \left|\{j \mid x_j \neq y_j\}\right|. \end{aligned}$$

(3) Any normed space $(V, \|\cdot\|)$ is a metric space.

$$d(v, w) = \|v - w\|.$$

Exercise: Show that if two norms are equivalent, their induced metrics are equivalent.

(4) Subset of Metric Space: If (X, d) is a metric space, and $Y \subseteq X$ is non-empty. Then, (Y, d) is a metric space.

(5) Paris metric: let (X, ρ) be a metric space. Let $p \in X$ be a fixed point.

$$\rho(x, y) := \begin{cases} 0 & x = y \\ \rho(x, p) + \rho(p, y) & x \neq y \end{cases}$$

(6) Bounded metric: Let ρ be a (semi-)metric on X. Set

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

We claim that d is a (semi-)metric. Notice that $0 \leq d(x, y) \leq 1$.

Proof: Clearly, $d(x, y) = d(y, x)$. Additionally, $d(x, x) = 0$. If $d(x, y) = 0$ and ρ is a metric, then $\rho(x, y) = 0$, so $x = y$.

To show the triangle inequality, we examine the function

$$\begin{aligned} f(t) &= \frac{t}{1 + t} \\ f'(t) &= \frac{1}{(1 + t)^2} > 0. \end{aligned}$$

Since ρ satisfies the triangle inequality, $\rho(x,z) \leq \rho(x,y) + \rho(y,z)$. Apply f on both sides. Then,

$$\begin{aligned}\underbrace{\frac{\rho(x,z)}{1+\rho(x,z)}}_{d(x,z)} &\leq \frac{\rho(x,y)+\rho(y,z)}{1+(\rho(x,y)+\rho(y,z))}\\ &= \frac{\rho(x,y)}{1+\rho(x,y)+\rho(y,z)} + \frac{\rho(y,z)}{1+\rho(x,y)+\rho(y,z)}\\ &\leq \underbrace{\frac{\rho(x,y)}{1+\rho(x,y)}}_{d(x,y)} + \underbrace{\frac{\rho(y,z)}{1+\rho(y,z)}}_{d(y,z)}.\end{aligned}$$

(7) If $d_1, \dots, d_n$ are metrics on X, $c_1, \dots, c_n \geq 0$. Then,

$$d(x,y) = \sum_{k=1}^{n} c_k d_k(x,y)$$

is a metric.

(8) Let $\{\rho_k\}_{k=1}^{\infty}$ be a family of semi-metrics. Assume the family is separating — for all $x \neq y$, there exists k such that $\rho_k(x,y) \neq 0$.

Let d_k be defined as

$$d_k(x,y) = \frac{\rho_k(x,y)}{1+\rho_k(x,y)}.$$

Note that $\{d_k\}_{k=1}^{\infty}$ is also separating.

Then,

$$d(x,y) = \sum_{k=1}^{\infty} 2^{-k} d_k(x,y)$$

is a metric.

We will now define the Frechet Metric using this method. Let $X = C(\mathbb{R})$. For each $k = 1,2,3,\dots$, set $p_k(f) = \sup_{x\in[-k,k]} |f(x)|$.

We can verify that p_k defines a seminorm. We can then check $\rho_k(f,g) = p_k(f-g)$ is a semi-metric.

We claim that $\{\rho_k\}$ is separating: if $f \neq g$, then there exists $x_0 \in \mathbb{R}$ with $f(x_0) \neq g(x_0)$. Since f and g are continuous, there is a neighborhood $[x_0-\delta, x_0+\delta]$ such that $f(x) \neq g(x)$ for all $x \in [x_0-\delta, x_0+\delta]$. Find k such that $[x_0-\delta, x_0+\delta] \subseteq [-k,k]$. Then, $\rho_k(f-g) > 0$.

Construct d_k as above, and then d as follows:

$$d_{\text{F}} = \sum \frac{2^{-k} p_k(f-g)}{1+p_k(f-g)}$$

(9) Product of metric spaces: let $(X_k, \rho_k)_{k=1}^{\infty}$ be a countable family of metric spaces. For each k, let

$$d_k(x,y) = \frac{\rho_k(x,y)}{1+\rho_k(x,y)}.$$

Remark: If the ρ_k are already uniformly bounded, let $d_k = \rho_k$.

Let

$$\begin{aligned}X &= \prod_{k=1}^{\infty} X_k\\ &= \{(x_k)_k \mid x_k \in X_k\}\\ &= \left\{ f : \mathbb{N} \to \bigsqcup_{k=1}^{\infty} X_k \mid f(k) \in X_k \right\}.\end{aligned}$$

Define $D : X \times X \to [0, \infty)$ as

$$\begin{aligned} D(x, y) &= \sum_{k=1}^{\infty} 2^{-k}\rho_k(x_k, y_k), \\ D(f, g) &= \sum_{k=1}^{\infty} 2^{-k}\rho(f(k), g(k)). \end{aligned}$$

For example, for each k, let $X_k = \{0, 1\}$ with the discrete metric. Let

$$\begin{aligned} \Delta &= \prod_{k\in\mathbb{N}} \{0, 1\} \\ &= \{(x_k)_k \mid x_k \in \{0, 1\}\} \\ D(x, y) &= \sum_{k=1}^{\infty} 2^{-k}|x_k - y_k| && (x_k)_k, (y_k)_k \in \Delta. \end{aligned}$$

Δ is known as the abstract Cantor set; every compact metric space is a surjective image of the abstract Cantor set.

(10) Geodesic Distance: let $\langle \cdot, \cdot \rangle$ be the standard dot product on $\mathbb{R}^3(\mathbb{R}^n)$, then

$$\begin{aligned} S^2 &= \left\{x \in \mathbb{R}^3 \mid \|x\|_2 = 1\right\} \\ S^{n-1} &= \{x \in \mathbb{R}^n \mid \|x\|_2 = 1\}. \end{aligned}$$

To find the geodesic distance, we take $d(x, y) = \arccos(\langle x, y\rangle)$. We claim d is a metric.

- Symmetry: self-evident.
- $d(x, x) = \arccos(1) = 0$. Suppose $d(x, y) = 0$. Then, $\langle x, y\rangle = 1$, meaning $\|x - y\|^2 = 0$, so $x = y$.
- Let $\theta = \arccos(\langle x, y\rangle)$, $\varphi = \arccos(\langle y, z\rangle)$, where $\theta, \varphi \in [0, \pi]$.

$$\begin{aligned} p_x &= \frac{\langle x, y\rangle}{\langle y, y\rangle} y \\ &= \cos(\theta)y \\ x &= \cos(\theta)y + \sin(\theta)u \end{aligned}$$

  where

$$u = \frac{x - p_x}{\|x - p_x\|}.$$

  Similarly, we can take

$$z = \cos(\varphi)y + \sin(\varphi)v$$

  where

$$v = \frac{z - p_z}{\|z - p_z\|}.$$

  So,

$$\begin{aligned} \langle x, z\rangle &= \cos(\theta)\cos(\varphi) + \sin(\theta)\sin(\varphi)\langle u, v\rangle \\ &\geq \cos(\theta)\cos(\varphi) - \sin(\theta)\sin(\varphi) && \langle u, v\rangle \geq -1 \\ &= \cos(\theta + \varphi). \end{aligned}$$

  Since arccos is decreasing,

$$\begin{aligned} \arccos(\langle x, z\rangle) &\leq \arccos(\cos(\theta + \varphi)) \\ &= \theta + \varphi \\ &= \arccos(\langle x, y\rangle) + \arccos(\langle y, z\rangle). \end{aligned}$$

  Therefore, $d(x, y) \leq d(x, y) + d(y, z)$.
- Let $\Gamma = (V, E)$ be a simple connected graph. We define $d : V \times V \to [0, \infty)$ to be the length of the shortest path between vertices u and v.

  Exercise: Show this is a metric.

(11) Let (X, d) be any metric space. If $E \subseteq X$, define $\operatorname{diam}(E) = \sup_{x,y\in E} d(x, y)$. E is bounded if $\operatorname{diam}(E) < \infty$.

Exercise: If $(V, \|\cdot\|)$ is a normed space, $E \subseteq V$ is a subset, show the following are equivalent:

(i) E is bounded (in the metric sense)

(ii) $\sup_{v \in E} \|v\| < \infty$

(iii) $\exists r > 0$ such that $E \subseteq rB_V$.

Let Ω be any set. The function $f : \Omega \to X$ is bounded if $f(\Omega) \subseteq X$ is bounded. We let.

$$\text{Bd}(\Omega, X) = \{f : \Omega \to X \mid f \text{ is bounded}\}.$$

Remark: $\text{Bd}(\Omega, \mathbb{F}) = \ell_\infty(\Omega, \mathbb{F})$.

(12) $\text{Bd}(\Omega, X)$ with

$$D_u(f, g) = \sup_{x \in \Omega} d(f(x), g(x)).$$

Exercise: Show that D_u defines a metric.

Consider $\text{Bd}(\Omega, \mathbb{F}) = \ell_\infty$. Look at the subset

$$E = \{f \in \text{Bd}(\Omega, \mathbb{F}) \mid f(x) \in \{0, 1\}\}.$$

Then,

$$\begin{aligned} D_u(f, g) &= \sup_{x \in \Omega} |f(x) - g(x)|. \\ &= \begin{cases} 1 & f \neq g \\ 0 & f = g \end{cases}. \end{aligned}$$

When we take a particular subset of $D_u(f, g)$, we find that we get the discrete metric.

Taking an overview of the concepts we have learned so far, we see

$$\text{Inner Product Spaces} \subseteq \text{Normed Vector Spaces} \subseteq \text{Metric Spaces}$$

Topology of Metric Spaces

Throughout this section, let (X, d) be a metric space.

(1) Let $x_0 \in X$, $\delta > 0$.

(i) We say

$$U(x_0, \delta) = \{x \in X \mid d(x, x_0) < \delta\}$$

is the open ball centered at x_0 with radius δ.

(ii) We say

$$B(x_0, \delta) = \{x \in X \mid d(x, x_0) \leq \delta\}$$

is the closed ball.

(iii) We say

$$S(x_0, \delta) = \{x \in X \mid d(x, x_0) = \delta\}$$

is the sphere.

(2) $U \subseteq X$ is open if

$$(\forall x \in U)(\exists \delta > 0) \text{ such that } U(x, \delta) \subseteq U.$$

Let

$$\begin{aligned} \tau_X &= \{U \subseteq X \mid U \text{ open}\} \\ &\subseteq \mathcal{P}(X). \end{aligned}$$

(3) $D \subseteq X$ is closed if D^c is open.

(4) If $x \in U \in \tau_X$, then U is called an open neighborhood of x. If $x \in U \subseteq N$, where $U \in \tau_X$, then N is a neighborhood of x.

$$\mathcal{N}_x = \{N \mid N \text{ is a neighborhood of } x\}$$

(5) Let $A \subseteq X$. The interior of A is

$$A^\circ = \bigcup \{V \mid V \subseteq A, V \text{ open}\}.$$

The closure of A is

$$\overline{A} = \bigcap \{D \mid A \subseteq D, D \text{ closed}\}.$$

The boundary of A is

$$\partial A = \overline{A} \setminus A^0.$$

Exercise: $\overline{A^c} = (A^\circ)^c$, $(\overline{A})^c = (A^c)^\circ$.

Remarks: A° is the largest open set contained in A. So, if V is open and $V \subseteq A$, then $V \subseteq A^\circ$. Similarly, $\overline{D}$ is the smallest closed set containing D. If C is closed and $D \subseteq C$, then $\overline{D} \subseteq C$.

- For example, $(a, b]^\circ = (a, b)$. This is because (a, b) is open and contained in $(a, b]$, so $(a, b) \subseteq (a, b]^\circ$.
- We will show that $\overline{A^c} \subseteq (A^\circ)^c$.

$$A^\circ \subseteq A$$
$$(A^\circ)^c \supseteq A^c$$

The union of open sets is open, so A° is open, so $(A^\circ)^c$ is closed by definition. Therefore,

$$(A^\circ)^c \supseteq \overline{A^c}.$$

Topology of Open Sets in a Metric Space

The open sets τ_X form a topology:

(i) $\emptyset, X \in \tau_X$.

(ii) If $\{V_i\}_{i \in I} \subseteq \tau_X$, then

$$\bigcup_{i \in I} V_i \in \tau_X.$$

(iii) If $V_i, \dots, V_n \in \tau_X$, then

$$\bigcap_{i=1}^{n} V_i \in \tau_X.$$

Remark: This is only true of finite intersections. For a counterexample, if $V_n = (-1/n, 1/n) \subseteq \mathbb{R}$ with the Euclidean metric, then the infinite intersection yields $\{0\}$, which is closed in $\mathbb{R}$ with the Euclidean metric.

Proof:

(1) Clearly, $\emptyset$ (by vacuous truth) and X are open.

(2) Let $x \in \bigcup_{i \in I} V_i$. Then, $\exists i_0 \in I$ with $x \in V_{i_0}$. Since V_{i_0} is open, $\exists \varepsilon > 0$ such that $U(x, \varepsilon) \subseteq V_{i_0} \subseteq \bigcup V_i$.

(3) Let $x \in \bigcap_{i=1}^{n} V_i$. Then, $x \in V_i$ for all $i \in 1, \dots, n$. Since each V_i is open, $\exists \varepsilon_1, \dots, \varepsilon_n$ with $U(x, \varepsilon_i) \subseteq V_i$ for each $i = 1, \dots, n$. Set $\varepsilon = \min\{\varepsilon_i\}_{i=1}^{n}$. Then, $U(x, \varepsilon) \subseteq U(x, \varepsilon_i) \subseteq V_i$ for all i. Therefore, $U(x, \varepsilon) \subseteq \bigcap_{i=1}^{n} V_i$.

Exercise: Show all open balls are open. In particular, show all open intervals are open.

Exercise: Show the following:

(1) $X, \emptyset$ are closed.

(2) If $\{C_i\}_{i \in I}$ is a family of closed sets, then $\bigcap_{i \in I} C_i$ is closed.

(3) For $C_1, \dots, C_n$ closed, then $\bigcup_{i=1}^{n} C_i$ is closed.

(4) Closed balls are closed. Spheres are closed.

Let $x \in X$. Recall that $\mathcal{N}_x$ is the set of all neighborhoods of x.

(i) $N \in \mathcal{N}_x \Leftrightarrow \exists \delta > 0 : U(x, \delta) \in N$

(ii) $N \in \mathcal{N}_x, N \subseteq M \Rightarrow M \in \mathcal{N}_x$

(iii) $N_1, N_2 \in \mathcal{N}_x \Rightarrow N_1 \cap N_2 \in \mathcal{N}_x$

In this sense, $\mathcal{N}_x$ is a directed set with reverse inclusion.

Pointwise Characterization of Subsets

Let $A \subseteq X$.

(i) $x \in A^\circ \Leftrightarrow \exists \delta > 0 : U(x,\delta) \subseteq A$.

(ii) $x \in \overline{A} \Leftrightarrow \forall \delta > 0 : U(x,\delta) \cap A \neq \emptyset$.

(iii) $x \in \partial A \Leftrightarrow \forall \delta > 0 : U(x,\delta) \cap A \neq \emptyset$ and $U(x,\delta) \cap A^c \neq \emptyset$.

Proof: Let $A \subseteq X$

(i)

$$\begin{aligned} x \in A^\circ &\Leftrightarrow x \in \bigcup_{\substack{V \in \tau_X \\ V \subseteq A}} V \\ &\Leftrightarrow \exists V \in \tau_X, V \subseteq A, x \in V \\ &\Leftrightarrow \exists \delta > 0 : U(x,\delta) \subseteq A. \end{aligned}$$

(ii)

$$\begin{aligned} x \notin \overline{A} &\Leftrightarrow x \in (\overline{A})^c \\ &\Leftrightarrow x \in (A^c)^\circ \\ &\Leftrightarrow \exists \delta > 0 : U(x,\delta) \subseteq A^c \\ &\Leftrightarrow \exists \delta > 0 : U(x,\delta) \cap A = \emptyset. \end{aligned}$$

We negate both sides.

(iii)

$$\begin{aligned} x \in \partial A &\Leftrightarrow x \in \overline{A} \setminus A^\circ \\ &\Leftrightarrow x \in \overline{A} \cap (A^0)^c \\ &\Leftrightarrow x \in \overline{A} \cap \overline{A}^c \\ &\Leftrightarrow x \in \overline{A} \text{ and } x \in \overline{A}^c \\ &\Leftrightarrow \forall \delta > 0 : U(x,\delta) \cap A \neq \emptyset, U(x,\delta) \cap A^c \neq \emptyset \end{aligned}$$

Remark: $\overline{U(v,\delta)} = B(v,\delta)$ in a normed space. $\partial U(v,\delta) = \partial B(v,\delta) = S(v,\delta)$ in a normed space. Also, $B(v,\delta)^\circ = U(v,\delta)$.

Proof: We show that $\overline{U}(v,\delta) = B(v,\delta)$. Since $B(v,\delta)$ is closed, and $U(v,\delta) \subseteq B(v,\delta)$, we know $\overline{U(v,\delta)} \subseteq B(v,\delta)$.

Let $w \in B(v,\delta)$. If $\|w - v\| < \delta$, then $w \in U(v,\delta)$. Assume $\|w - v\| = \delta$. Let $u_t = (1-t)v + tw$, where $t \in [0,1]$.

$$\begin{aligned} \|w - u_t\| &= \|w - (1-t)v - tw\| \\ &= \|(1-t)(w - v)\| \\ &= (1-t)\|w - v\| \\ &= (1-t)\delta. \end{aligned}$$

Let $\varepsilon > 0$. Let $t \in (0,1)$ such that $(1-t)\delta < \varepsilon$. Then, $u_t \in U(w,\varepsilon) \cap U(v,\delta)$. Therefore, $w \in \overline{U(v,\delta)}$.

Unions and Intersections of Closure/Interior

Let (X,d) be a metric space.

(i)

$$\left(\bigcup_{i \in I} A_i\right)^\circ \supseteq \bigcup_{i \in I} A_i^\circ \qquad \text{may be strict}$$

(ii)

$$\overline{\bigcap_{i \in I} A_i} \subseteq \bigcap_{i \in I} \overline{A_i}$$

(iii)

$$\bigcap_{k=1}^{n} A_k^\circ = \left(\bigcap_{k=1}^{n} A_k\right)^0$$

(iv)

$$\overline{\bigcup_{k=1}^{n} D_k} = \bigcup_{k=1}^{n} \overline{D_k}$$

Proof:

(i)

$$\begin{aligned} A_i^\circ &\subseteq A_i \\ \bigcup_{i\in I} A_i^\circ &\subseteq \bigcup_{i\in I} A_i \\ \bigcup_{i\in I} A_i^\circ &\subseteq \left(\bigcup_{i\in I} A_i\right)^\circ \end{aligned}$$

Remark: We claim $\overline{\mathbb{Q}} = \mathbb{R}$ under the absolute value metric. We know that $\mathbb{Q} \subseteq \mathbb{R}$, $\mathbb{R}$ is closed, meaning $\overline{\mathbb{Q}} \subseteq \mathbb{R}$. Let $t \in \mathbb{R}$, $\delta > 0$. We know that $(t-\delta, t+\delta) \cap \mathbb{Q} \neq \emptyset$. Therefore, $t \in \overline{\mathbb{Q}}$. Thus, $\overline{\mathbb{Q}} = \mathbb{R}$.

Properties of Boundary

Let $A \subseteq X$.

(1) ∂A is closed.

(2) $\partial A = \partial A^c$

(3) $\overline{A} = A \cup \partial A$

(4) $A \setminus \partial A = A^\circ$

Proof:

(1)

$$\begin{aligned} \partial A &= \overline{A} \setminus A^\circ \\ &= \overline{A} \cap (A^\circ)^c. \end{aligned}$$

(2) Follows from pointwise characterization.

(3) Clearly, $A \cup \partial A \subseteq \overline{A}$. Let $x \in \overline{A}$. If $x \in A$, we're done. Otherwise, $x \in \overline{A} \setminus A \subseteq \overline{A} \setminus A^\circ = \partial A$.

(4)

$$\begin{aligned} A \setminus \partial A &= A \cap (\partial A)^c \\ &= A \cap (\overline{A} \setminus A^\circ)^c \\ &= A \cap \left(\overline{A} \cap (A^\circ)^c\right)^c \\ &= A \cap \left(\overline{A}^c \cup A^\circ\right) \\ &= \left(A \cap \overline{A}^c\right) \cup (A \cap A^\circ) \\ &= A^\circ \end{aligned}$$

Density and Separability

Let (X, d) be a metric space.

(1) $A \subseteq X$ is d-dense if $\overline{A} = X$.

(2) $N \subseteq X$ is nowhere dense if $(\overline{N})^\circ = \emptyset$.

(3) (X, d) is separable if there is a countable dense subset.

Exercise: If $N \subseteq X$ is closed, then N is nowhere dense if and only if N^c is dense.

Exercise: The following are equivalent.

(1) $A \subseteq X$ is dense.
(2) $\forall \emptyset \neq U \in \tau_X$, $U \cap A \neq \emptyset$.
(3) $\forall x \in X, \forall \varepsilon > 0$, $U(x, \varepsilon) \cap A \neq \emptyset$.
(4) $\forall x \in X, \forall \varepsilon > 0, \exists a \in A$ such that $d(x, a) < \varepsilon$.

Let X be a metric space.

(1) A base for τ_X is a family of open subsets $\mathcal{B}$ such that:

$$(\forall U \in \tau_X)\,(\forall x \in U)\exists B \in \mathcal{B} \text{ such that } x \in B \subseteq U.$$

Equivalently,

$$\forall U \in \tau_X, U = \bigcup_{i \in I} B_i. \qquad B_i \in \mathcal{B}$$

(2) We say that (X, d) is second countable if τ_X admits a countable base.

- For any (X, d) a metric space, $\mathcal{B} = \{U(x, \varepsilon) \mid x \in X, \varepsilon > 0\}$ is a base. Indeed, given any $x \in U \subseteq \tau_X$, by definition, $\exists \varepsilon > 0$ such that $U(x, \varepsilon) \subseteq U$. Alternatively, $\mathcal{B}' = \{U(x, 1/n) \mid x \in X, n \geq 1\}$ is a topological base.
- Let $X = \mathbb{R}^d$ with the Euclidean metric. Then, for $\mathcal{B} = \{U(q, 1/n) \mid n \geq 1, q \in \mathbb{Q}^d\}$, we claim this is a base.

  Let $V \subseteq \mathbb{R}^d$ be open, $r \in V$. Since V is open, $\exists \delta > 0$ with $U(r, \delta) \subseteq V$. Find n large such that $1/n < \delta$. Find $q \in \mathbb{Q}^d$ with $\|r - q\| < 1/2n$. This is always possible as $\mathbb{Q}^d$ is dense in $\mathbb{R}^d$.

  Consider $U(q, 1/2n)$. Then, $r \subseteq U(q, 1/2n) \subseteq U(r, \delta) \subseteq V$ because $\|r - q\| < 1/2n$, and if $t \in U(q, 1/2n)$, then

$$\begin{aligned} \|t - r\| &\leq \|t - q\| + \|q - r\| \\ &< 1/2n + 1/2n \\ &= 1/n \\ &< \delta. \end{aligned}$$

Separable, Non-Separable, Dense, and Non-Dense Sets

(1) $\left(R^d, \|\cdot\|_p\right)$ is separable for any $p \in [1, \infty]$. Indeed, $\mathbb{Q}^d \subseteq \mathbb{R}^d$ is the countable dense subset of $\mathbb{R}^d$.

Let $r = \begin{bmatrix} r_1 \\ \vdots \\ r_d \end{bmatrix} \in \mathbb{R}^d$. Find $q = \begin{bmatrix} q_1 \\ \vdots \\ q_d \end{bmatrix} \in \mathbb{Q}^d$ with $|r_j - q_j| < \varepsilon/d$. Then,

$$\begin{aligned} \|r - q\|_1 &= \sum_{j=1}^{d} |r_j - q_j| \\ &< d. \end{aligned}$$

We know that for any vector $r \in \mathbb{R}^d$, we can find a vector q such that

$$\|q - r\|_p \leq c\,\|q - r\|_1,$$

so for arbitrary p, find q such that $\|q - r\|_1 < \varepsilon/c$.

(2) Similarly, $\mathbb{C}_\mathbb{Q} = \{a + bi \mid a, b \in \mathbb{Q}\}$ is also countable, meaning $\mathbb{C}_\mathbb{Q}^d \subseteq \mathbb{C}^d$ is dense and $\mathbb{C}^d$ is dense.

Proposition: Separable Subsets

If (X, d) is separable, and $Y \subseteq X$, then (Y, d) is also separable.

Let $\{a_k\}$ be a countable dense subset in X. Let $N = \{(m, n) \mid U(a_m, 1/n) \cap Y \neq \emptyset\}$. Clearly, N is nonempty. For each $(m, n) \in N$, choose $b_{(m,n)} \in Y \cap U(a_m, 1/n)$. We claim $\{b_{(m,n)} \mid m, n \geq 1\}$ is dense in Y.

Let $y \in Y$, $\varepsilon > 0$. Find N large so that $\frac{1}{n} < \varepsilon/2$. Since $A \subseteq X$ is dense, find $U(y, 1/n) \cap A \neq \emptyset$. Suppose $d(a_m, y) < 1/n$. Then,

$$\begin{aligned} d(b_{(m,n)}, y) &\leq d(b_{(m,n)}, a_m) + d(a_m, y) \\ &< \frac{1}{n} + \frac{1}{n} \\ &= \frac{2}{n} \\ &< \varepsilon. \end{aligned}$$

(1) ℓ_p^n is separable.

(2) $c_{00} = \{(a_k)_{k=1}^n \mid \text{finitely many } a_k \neq 0\}$ with $\|\cdot\|_u$ is separable.

Recall that $e_k = (0, 0, \dots, 1, 0, 0, \dots)$ where 1 is at position k. Consider $E = \mathbb{Q}\text{-span}\{e_k \mid k \geq 1\}$,

$$E = \left\{ \sum_{k=1}^{n} \alpha_k e_k \mid \alpha_k \in \mathbb{Q}, n \geq 1 \right\}.$$

The set E is countable. If we fix $n \geq 1$, we have

$$E_n = \left\{ \sum_{k=1}^{n} \alpha_k e_k \mid \alpha_k \in \mathbb{Q} \right\}.$$

Then, $E = \bigcup E_n$. Note

$$\begin{aligned} \underbrace{\mathbb{Q} \times \mathbb{Q} \times \cdots \times \mathbb{Q}}_{n} &\to E_n \\ (\alpha_1, \dots, \alpha_n) &\mapsto \sum_{k=1}^{n} \alpha_k e_k. \end{aligned}$$

Thus, E_n is countable, and E is a countable union of countable sets.

We claim that E is dense. Given $z \in c_{00}$, $\varepsilon > 0$, we know that $z = \sum_{k=1}^{n} a_k e_k$ for some n and $a_k \in \mathbb{R}$. Find $\alpha_k \in \mathbb{Q}$ such that $|\alpha_k - a_k| < \varepsilon$. Set $w = \sum_{k=1}^{n} \alpha_k e_k$. Then, $\|z - w\|_u = \sup |\alpha_k - a_k| < \varepsilon$.

(3) c_0 with $\|\cdot\|_u$ is separable.

(4) ℓ_∞ is not separable.

Suppose ℓ_∞ were separable. Consider $E = \{(a_k)_k \in \ell_\infty \mid a_k \in \{0, 1\}\}$. Then, E is separable. Recall that $(E, \|\cdot\|_u)$ has the discrete metric.

In the discrete metric, every subset is open, meaning every subset is closed. Therefore, if X is separable and discrete, then X is countable.

However, E is not countable by Cantor's theorem. $\text{card}(E) = 2^{\aleph_0}$.

Alternatively, we can show that

$$(a_k)_k \mapsto \sum_{k=1}^{\infty} 2^{-k} a_k$$

is onto.

Exercise: ℓ_p is separable for $1 \leq p < \infty$.

(5) We will show that

$$\mathbb{P}[0,1]\left\{\sum_{k=1}^{n} a_k x^k \mid a_k \in \mathbb{R}, n \geq 1\right\}$$

is $\|\cdot\|_u$-dense in $C([0,1])$. Using this, we can show that $(C([0,1]), \|\cdot\|_u)$ is separable.

The Cantor Set

$$\begin{aligned}
C_0 &= [0,1]\\
C_1 &= [0,1/3] \cup [2/3,1]\\
C_2 &= [0,1/9] \cup [2/9,1/3] \cup [2/3,7/9] \cup [8/9,1]\\
C_3 &= [0,1/27] \cup [2/27,1/9] \cup \cdots \cup [26/27,1]\\
&\vdots
\end{aligned}$$

In each step, we delete the middle third of each interval. This process repeated ad infinitum yields the Cantor set.

$$\mathcal{C} = \bigcap_{n=1}^{\infty} \bigcup_{k=0}^{3^{n-1}-1} \left(\left[\frac{3k+0}{3^n}, \frac{3k+1}{3^n}\right] \cup \left[\frac{3k+2}{3^n}, \frac{3k+3}{3^n}\right]\right).$$

(i) $\mathcal{C}$ is closed as it is the intersection of closed sets.

(ii) $\text{length}(\mathcal{C}) = 0$. Look at the total length of the removed intervals,

$$\begin{aligned}
l &= \frac{1}{3} + \frac{2}{9} + \frac{4}{27} + \frac{8}{81} + \cdots\\
&= \sum_{k=1}^{\infty} \left(\frac{2^{k-1}}{3^k}\right)\\
&= \frac{1}{2}\sum_{k=1}^{n} \left(\frac{2}{3}\right)^k\\
&= 1.
\end{aligned}$$

Thus, $\text{length}(\mathcal{C}) = 0$.

(iii) $\mathcal{C}$ is nowhere dense — $(\overline{\mathcal{C}})^\circ = \emptyset$. Since $\mathcal{C}$ is closed, $\mathcal{C}^\circ = \emptyset$.

Suppose $\mathcal{C}^\circ \neq \emptyset$. Then, $\exists x \in \mathcal{C}, \varepsilon > 0$ such that $(x-\varepsilon, x+\varepsilon) \subseteq \mathcal{C}$. So, $(x-\varepsilon, x+\varepsilon) \subseteq \mathcal{C}_n$ for all n.

Note $\mathcal{C}_n$ is the disjoint union of 2^n subintervals, each with length $1/3^n$. Find m so large such that $3^{-m} < \varepsilon$. We know that $(x-\varepsilon, x+\varepsilon) \subseteq C_m$.

However, $(x-\varepsilon, x+\varepsilon)$ has length $2\varepsilon > \frac{2}{3^m}$. Each subinterval in C_m has length $1/3^m$. This implies C_m contains an interval of length greater than $\frac{2}{3^m}$. $\perp$

(iv) $\text{card}(\mathcal{C}) = \text{card}(\mathbb{R})$

Claim 1: Given $n \geq 1$,

$$E_n = \left\{\sum_{k=1}^{n} \frac{w_k}{3^k} \mid w_k \in \{0,2\}\right\}$$

is precisely the set of *left* endpoints of the subintervals of C_n.

For $n = 1$, if $w_1 = 0$, then we get 0, and $w_1 = 2$ yields $2/3$. Meanwhile, if $n = 2$, then we have

$$\begin{aligned}
w_1 = 0, w_2 = 0 &\mapsto 0\\
w_1 = 0, w_2 = 2 &\mapsto 2/9\\
w_1 = 2, w_2 = 0 &\mapsto 2/3\\
w_1 = 2, w_2 = 2 &\mapsto 8/9.
\end{aligned}$$

By induction, we have shown for $n = 1, 2$. Assume this is true for n.

$$\sum_{k=1}^{n+1} w_k 3^{-k} = \underbrace{\sum_{k=1}^{n} w_k 3^{-k}}_{(1)} + \underbrace{w_{n+1} 3^{-(n+1)}}_{(2)}$$

Part (1) denotes one of the left endpoints of C_n, called $C_{n,k}$ for some $1 \leq k \leq 2^n$. Then, if $w_{n+1} = 0$, we get the left endpoint of $C_{n+1,2k-1}$, and if $w_n = 2$, we get the left endpoint of $C_{n+1,2k}$.

Claim 2:

$$C = \left\{\sum_{k=1}^{\infty} w_k 3^{-k} \mid w_k \in \{0, 2\}\right\}$$

is precisely the Cantor set.

Let $x = \sum_{k=1}^{\infty} w_k 3^{-k}$. We will show that $x \in C_n$ for all n. Fix $n \geq 1$. Then,

$$x = \underbrace{\sum_{k=1}^{n} w_k 3^{-k}}_{y} + \underbrace{\sum_{k>n} w_k 3^{-k}}_{z}.$$

From our previous claim, y is the left endpoint of some subinterval of C_n. Additionally,

$$\begin{aligned}
z &= \sum_{k>n} w_k 3^{-k} \\
&\leq 2\sum_{k>n} 3^{-k} \\
&= \frac{2}{3^{n+1}}\left(1 + \frac{1}{3} + \frac{1}{9} + \cdots\right) \\
&= \frac{1}{3^n}.
\end{aligned}$$

Since the length of a subinterval in C_n is exactly 3^{-n}, it is the case that $x = y + z$ remains an element of $C_{n,k}$.

Let $x \in \mathcal{C}$. Then, $x \in C_n$ for all n. Then, $x \in C_1$, so let x_1 be the left endpoint of the interval $C_{1,j}$ that contains x. Then, $|x - x_1| < \frac{1}{3}$, and $x_1 = w_1 3^{-1}$ for some $w_1 \in \{0, 2\}$.

Let x_2 be the left endpoint of the subinterval $C_{2,j}$ that contains x. Then, $|x - x_2| < \frac{1}{3^2}$. Therefore,

$$\begin{aligned}
x_2 &= x_1 + w_2 3^{-2} \\
&= w_1 3^{-1} + w_2 3^{-2}.
\end{aligned}$$

Iterating, we have x_n, the left endpoint of the subinterval $C_{n,j}$ that contains x.

$$x_n = \sum_{k=1}^{n} w_k 3^{-k}.$$

We have that $|x - x_n| < 3^{-n}$.

Therefore, $(x_n)_n \to x$. Also,

$$\begin{aligned}
x_n &= \sum_{k=1}^{n} w_k 3^{-k} \\
&\to \sum_{k=1}^{n} w_k 3^{-k}.
\end{aligned}$$

Thus,

$$x = \sum_{k=1}^{\infty} w_k 3^{-k}.$$

To prove $\text{card}(\mathcal{C}) = \text{card}(\mathbb{R})$, we will show that $\text{card}\left(\{0, 1\}^{\mathbb{N}}\right) = \text{card}(\mathcal{C})$.

$$(a_k)_k \mapsto \sum_{k=1}^{\infty} 2a_k 3^{-k}.$$

Relative (or Subspace) Topology

We know that if (X, d) is a metric space, and $Y \subseteq X$ is any subset, then (Y, d) is a metric space. The question now is: what are the open sets of Y?

For example, let $X = \mathbb{R}$, $Y = [0, 1]$. Consider $U = [0, 1/2)$. U is not open in $\mathbb{R}$, as if $x = 0$, then there is no open ball completely contained in U. However, in Y, U is open.

Let (X, d) be a metric space, $Y \subseteq X$ any subset. $V \subseteq Y$ is open if and only if $\exists U \subseteq X$ open such that $V = U \cap Y$. That is, $\tau_Y = \{U \cap Y \mid U \in \tau_X\}$.

Let V be open in Y. Then, $\forall x \in V$, $\exists \delta_x > 0$ such that $U_Y(x, \delta_x) \subseteq V$. We have $U_Y(x, \delta_x) = \{y \in Y \mid d(y, x) < \delta_x\}$. Let

$$\begin{aligned} U &= \bigcup_{x \in V} U_X(x, \delta_x) \\ U \cap Y &= \left(\bigcup_{x \in V} U_X(x, \delta_x)\right) \cap Y \\ &= \bigcup_{x \in V} U_X(x, \delta_x) \cap Y \\ &= \bigcup_{x \in V} U_Y(x, \delta_x). \end{aligned}$$

Let U be open in X. Then, for $x \in U \cap Y$, $\exists \delta_x$ such that $U(x, \delta_x) \subseteq U$.

(1) ℓ_∞ is not a discrete metric space. However, $E = \{(a_k)_k \mid a_k \in \{0, 1\}\}$ with the induced metric. Then, E is a discrete metric space.

Convergent Sequences

Fix a metric space (X, d). A sequence in X is a map $x : \mathbb{N} \to X$, $n \mapsto x(n) = x_n$.

A natural sequence $(n_k)_k$ is a sequence in $\mathbb{N}$ with $n_k \geq k$ for all k. A subsequence of $(x_n)_n$ is a sequence $(x_{n_k})_k$, where $(n_k)_k$ is a natural sequence.

A sequence $(x_n)_n$ converges to $x \in X$ if $\forall \varepsilon > 0$, $\exists N_\varepsilon \in \mathbb{N}$ such that $n \geq N_\varepsilon$ implies $d(x_n, x) < \varepsilon$. We write $(x_n)_n \xrightarrow{d} x$.

Exercise: A sequence can have at most one limit, as metric spaces are Hausdorff.

Proposition: Equivalent Definitions of Convergence

Given $(x_n)_n \in X$, $x \in X$, the following are equivalent.

(i) $(x_n)_n \to x$ in X

(ii) $(d(x_n, x))_n \to 0$ in $\mathbb{R}$

(iii) $\forall V \in \mathcal{N}_x$, $\exists N \in \mathbb{N}$ with $n \geq N \Rightarrow x_n \in V$.

Exercise: Let (X, ρ) be a metric space, let $d(x, y) = \frac{\rho(x,y)}{1+\rho(x,y)}$. A sequence $(x_n)_n \xrightarrow{d} x$ if and only if $(x_n)_n \xrightarrow{\rho} x$.

Proposition: Convergent Sequences are Bounded

Let $(x_n)_n \to x$ in (X, d). Let $\varepsilon = 1$. Then, $\exists N \in \mathbb{N}$ large such that for $n \geq N$, $d(x_n, x) < 1$.

If $m, n \geq N$, then $d(x_n, x_m) \leq d(x_n, x) + d(x, x_m) < 2$. Let $c = \max_{1 \leq n, m \leq N} d(x_n, x_m)$. Then,

$$\begin{aligned} d(x_n, x_m) &\leq d(x_n, x_N) + d(x_n, x_m) \\ &\leq 1 + c. \end{aligned}$$

Let $k = \max\{1 + c, 2\}$. Then, $\operatorname{diam}(\{x_n\}) \leq k$.

Convergence in Different Metric Spaces

Convergence for Bounded Functions: Recall that for (Y, d) a metric space is

$$\begin{aligned} \text{Bd}(\Omega, Y) &= \{f : \Omega \to Y \mid f \text{ bounded}\} \\ D_u(f, g) &= \sup_{x\in\Omega} d(f(x), g(x)). \end{aligned}$$

Then, $(f_n)_n \to f$ in $\text{Bd}(\Omega, Y)$ if and only if $D_u(f_n, f) \to 0$ in $\mathbb{R}$.

$$\begin{gathered} (\forall\varepsilon > 0)(\exists N \in \mathbb{N}) \text{ such that } n \geq N \Rightarrow D_u(f_n, f) < \varepsilon \\ \Leftrightarrow \\ (\forall\varepsilon > 0)(\exists N \in \mathbb{N}) \text{ such that } n \geq N \Rightarrow \sup_{x\in\Omega} d(f_n(x), f(x)) < \varepsilon \\ \Leftrightarrow \\ (\forall\varepsilon > 0)(\exists N \in \mathbb{N}) \text{ such that } n \geq N \Rightarrow \forall x,\ d(f_n(x), f(x)) < \varepsilon. \end{gathered}$$

This is exactly the definition of uniform convergence.

Since $\ell_\infty(\Omega) = \text{Bd}(\Omega, \mathbb{F})$, convergence in $\ell_\infty(\Omega)$ is uniform convergence. This is also the case for subspaces, such as c, c_0, and c_{00}.

Convergence in the Frechet Metric: Consider a separating family of semimetrics ρ_k on a set X. Set $d_k = \frac{\rho_k}{1+\rho_k}$. We saw that

$$d(x, y) = \sum_{k=1}^{\infty} 2^{-k} d_k(x, y)$$

is a metric on X.

We claim that $(x_n)_n \to x$ in (X, d) if and only if for all $k \geq 1$, $\rho_k(x_n, x) \to 0$.

In the forward direction, we know that $(x_n)_n \to x$ with respect to d if and only if $d(x_n, x) \to 0$ in $\mathbb{R}$. Since $0 \leq 2^{-k} d_k(x_n, x) \leq d(x_n)$ for fixed k, we have that

$$0 \leq d_k(x_n, x) \leq 2^k d(x_n, x),$$

and as $n \to \infty$, $d(x_n, x) \to 0$, meaning $d_k(x_n, x) \to 0$. Therefore, $\rho_k(x_n, x) \to 0$.

In the reverse direction, suppose $\rho_k(x_n, x) \to 0$ in $\mathbb{R}$ as $n \to \infty$ for all $k \geq 1$. Thus, $d_k(x_n, x) \to 0$ as $n \to \infty$ for all $k \geq 1$.

Let $\varepsilon > 0$. Let K be so large such that

$$\sum_{k\geq K} 2^{-k} < \varepsilon/2.$$

Therefore, $d_k(x_n, x) \to 0$ for all $k = 1, \ldots, K$. Therefore, $\exists N_1, \ldots, N_K$ such that for $n \geq N_k$,

$$d_k(x_n, x) < \frac{\varepsilon}{2}.$$

Let $N = \max\{N_1, \ldots, N_K\}$. Therefore, for $n \geq N$,

$$d_k(x_n, x) < \frac{\varepsilon}{2}$$

for all $k = 1, \ldots, K$.

Thus, for all $n \geq N$,

$$\begin{aligned} d(x_n, x) &= \sum_{k=1}^{\infty} 2^{-k} d_k(x_n, x) \\ &= \sum_{k=1}^{K} 2^{-k} d_k(x_n, x) + \sum_{k=K+1}^{\infty} 2^{-k} d_k(x_n, x) \\ &\leq \frac{\varepsilon}{2} \sum_{k=1}^{K} 2^{-k} + \frac{\varepsilon}{2} \\ &< \varepsilon \end{aligned}$$

Therefore, $(x_n)_n \to x$.

Recall that, for the Frechet metric, our set was $X = C(\mathbb{R})$. For $k = 1, 2, 3, \dots$, we had

$$p_k(f) = \sup_{[-k,k]} |f(x)|$$

as our seminorm, and our semimetric was

$$\rho_k(f, g) = p_k(f - g).$$

We also showed that the ρ_k family is separating. We make $d_k(f, g) = \frac{\rho_k(f,g)}{1+\rho_k(f,g)}$ as the bounded family of separating metrics, and

$$d_F(f, g) = \sum_{k=1}^{\infty} \frac{2^{-k}\rho_k(f - g)}{1 + \rho_k(f - g)}.$$

In $(C(\mathbb{R}), d_F)$, $(f_n)_n \to f$ if and only if $\rho_k(f_n, f) \to 0$ for all k, meaning $(f_n)_n \to f$ uniformly on $[-k, k]$ for all k.

This is known as convergence on compact subsets.

Convergence in a Product Space: Let (X, d) and (Y, ρ) be metric spaces. Then,

$$\begin{aligned} X \times Y &= \{(x, y) \mid x \in X, y \in Y\}, \\ D_1((x, y), (x', y')) &= d(x, x') + \rho(y, y') \\ D_\infty((x, y), (x', y')) &= \max\{d(x, x'), \rho(y, y')\}. \end{aligned}$$

Both D_1 and D_∞ are equivalent metrics.

Exercise: $((x_n, y_n))_n \to (x, y)$ if and only if $(x_n)_n \xrightarrow{d} x$ and $(y_n)_n \xrightarrow{\rho} y$.

Series in a Normed Vector Space

Let $(V, \|\cdot\|)$ be a normed vector space. Consider a sequence $(v_k)_k$ of vectors.

$$\begin{aligned} s_1 &= v_1 \\ s_2 &= v_1 + v_2 \\ &\vdots \\ s_n &= \sum_{k=1}^{n} v_k. \end{aligned}$$

If $s_n \to s$ in $(V, \|\cdot\|)$, meaning $\|s_n - s\| \to 0$, then we say the series $\sum_{k=1}^{\infty} v_k$ converges to s. We write

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

The series converges absolutely if

$$\sum_{k=1}^{\infty} \|v_k\|$$

converges in $\mathbb{R}$.

Proposition: Sequential Characterization of Closure

Let (X, d) be a metric space with $A \subseteq X$. $x \in \overline{A}$ if and only if $\exists (a_n)_n$ in A with $(a_n)_n \to X$.

In the forward direction, recall that $x \in \overline{A}$ if and only if $\forall \delta > 0$, $U(x, \delta) \cap A \neq \emptyset$. If $x \in \overline{A}$, then set $\varepsilon_n = 1/n$, and since $U(x, 1/n) \cap A \neq \emptyset$. Let $a_n \in U(x, 1/n) \cap A$. Then, $d(a_n, x) < 1/n \to 0$, meaning $a_n \to x$ and $a_n \in A$.

In the reverse direction, if $(a_n)_n \to x$ and $\varepsilon > 0$, $\exists N$ with $n \geq N \Rightarrow a_n \in U(x, \varepsilon) \cap A$. Thus, $x \in \overline{A}$.

Proposition: Sequential Characterization of Closed Sets

If (X, d) is a metric space, $A \subseteq X$, then the following are equivalent:

(i) A is closed.

(ii) Whenever $(a_n)_n$ in A with $(a_n)_n \xrightarrow{d} x$ in X, then $x \in A$.

Continuous Bounded Functions: $C([a, b]) \subseteq \ell_\infty([a, b])$ is closed under $\|\cdot\|_u$, since if $(f_n)_n \to f$ uniformly, and f_n is continuous, then f is continuous.

Sequence Closure: $c_0 \subseteq \ell_\infty$ is closed under $\|\cdot\|_u$. Let $(f_n)_n$ be a sequence

$$f_1 = (f_1(1), f_1(2), \dots)$$
$$f_2 = (f_2(1), f_2(2), \dots)$$
$$\lim_{k\to\infty} f_n(k) = 0 \qquad \forall n$$

Suppose $(f_n)_n \xrightarrow{\|\cdot\|_\infty} f \in \ell_\infty$.

Let $\varepsilon > 0$. Then, $\exists n \in \mathbb{N}$ such that for $n \geq N$, $\|f - f_n\|_\infty < \varepsilon/2$. Also, $\lim_{k\to\infty} f_N(k) = 0$. Then, $\exists K \in \mathbb{N}$ such that for $k \geq K$, $|f_N(k)| < \varepsilon/2$. Thus, for $k \geq K$,

$$\begin{aligned}|f(k)| &= |f(k) - f_N(k) + f_N(k)|\\ &\leq |f(k) - f_N(k)| + |f_N(k)|\\ &\leq \|f - f_N\|_\infty + |f_N(k)|\\ &< \frac{\varepsilon}{2} + \frac{\varepsilon}{2}\\ &= \varepsilon.\end{aligned}$$

Thus, $f \in c_0$.

Distance to a Set

Let (X, d) be a metric space, $A \subseteq X$. Then, $\text{dist}_A : X \to [0, \infty)$ is defined as

$$\text{dist}_A(x) = \inf_{a\in A} d(x, a).$$

(1) $\overline{A} = \{x \mid \text{dist}_A = 0\}$

(2) $\text{dist}_A(\cdot) = \text{dist}_{\overline{A}}(\cdot)$

(3) $|\text{dist}_A(x) - \text{dist}_A(y)| \leq d(x, y)$

Proof of (1): Let $x \in \overline{A}$. Then, $\exists (a_n)_n$ such that $(a_n)_n \to x$. Then, $d(a_n, x) \to 0$. Since $0 \leq \text{dist}_A(x) \leq d(x, a_n)$, $\text{dist}_A(x) = 0$.

Let x be such that $\text{dist}_A(x) = 0$. By the definition of inf, we construct a_n by finding $a_n \in U(x, 1/n) \cap A$. Thus, $d(a_n, x) \to 0$, meaning $(a_n)_n \to x$, so $x \in \overline{A}$.

Proof of (2): Exercise; use (1).

Proof of (3): For all $a \in A$,

$$\begin{aligned}\text{dist}_A(x) &\leq d(x, a)\\ &\leq d(x, y) + d(y, a).\end{aligned}$$

Therefore,

$$\begin{aligned}\text{dist}_A(x) - d(x, y) &\leq d(y, a)\\ \text{dist}_A(x) - d(x, y) &\leq \inf_{a\in A} d(y, a)\\ &= \text{dist}_A(y)\\ \text{dist}_A(x) - \text{dist}_A(y) &\leq d(x, y).\end{aligned}$$

Similarly,

$$\text{dist}_A(y) - \text{dist}_A(x) \leq d(y, x) = d(x, y)$$

meaning

$$|\text{dist}_A(y) - \text{dist}_A(x)| \leq d(x, y).$$

Continuity

Let (X, d) and (Y, ρ) be metric spaces. A map $f : X \to Y$

(1) is continuous at $x_0 \in X$ if

$$(\forall \varepsilon > 0)(\exists \delta > 0) \text{ such that } d(x, x_0) < \delta \Rightarrow \rho(f(x), f(x_0)) < \varepsilon$$
$$(\forall \varepsilon > 0)(\exists \delta > 0) \text{ such that } x \in U_X(x_0, \delta) \Rightarrow f(x) \in U_Y(f(x_0), \varepsilon)$$
$$(\forall \varepsilon > 0)(\exists \delta > 0) \text{ such that } f(U_X(x_0, \delta)) \subseteq U_Y(f(x_0), \varepsilon).$$

(2) is continuous if f is continuous at every $x_0 \in X$.

Proposition: Equivalent Continuity Criteria

Let $f : (X, d) \to (Y, \rho)$, $x_0 \in X$. The following are equivalent:

(1) f is continuous at x_0;

(2) $(\forall V \in \mathcal{N}_{f(x_0)})(U \in \mathcal{N}_{x_0})$ such that $f(U) \subseteq V$.

(3) $\forall (x_n)_n \to x_0$, $(f(x_n))_n \to f(x_0)$.

(1) $\Leftrightarrow$ (2): Clearly follows from definitions.

(1) $\Rightarrow$ (3): Let $(x_n)_n \to x_0$. Let $\varepsilon > 0$. Then, $\exists \delta > 0$ such that $d(x, x_0) < \delta$ implies $\rho(f(x), f(x_0)) < \varepsilon$.

Thus, $\exists N \in \mathbb{N}$ such that $n \geq N$ implies $d(x_n, x_0) < \delta$. So, if $n \geq N$, $d(x_n, x_0) < \delta$, implying $\rho(f(x_n), f(x_0)) < \varepsilon$. So, $(f(x_n))_n \to f(x_0)$.

(3) $\Rightarrow$ (1): Suppose toward contradiction that $\exists \varepsilon_0 > 0$ such that for $\delta = 1/n$ where $n \in \mathbb{N}$, $\exists (x_n)_n : d(x_n, x_0) < \delta$ and $\rho(f(x_n), f(x_0)) \geq \varepsilon_0$. Then, $(x_n)_n \to x_0$, but $f(x_n)_n \nrightarrow f(x_0)$. $\perp$

Proposition: Topological Criterion for Continuity

Let $f : (X, d) \to (Y, \rho)$. The following are equivalent:

(1) f is continuous.

(2) $\forall V \in \tau_Y$, $f^{-1}(V) \in \tau_X$.

(3) $\forall x \in X, \forall (x_n)_n \to x$, we have $(f(x_n))_n \to f(x)$.

Proof: Exercise.

Proposition: Composition of Functions

Let $(X, d) \xrightarrow{f} (Y, \rho) \xrightarrow{g} (Z, p)$. If f and g are continuous, then $g \circ f$ is continuous.

Proof: Exercise.

Uniform Continuity

Let $f : (X, d) \to (Y, \rho)$.

(1) f is uniformly continuous if

$$(\forall \varepsilon > 0)(\exists \delta > 0) \text{ such that } \forall x, x' \in X, d(x, x') < \delta \Rightarrow d(f(x), f(x')) < \varepsilon$$

(2) f is Lipschitz if $\exists c > 0$ with

$$\rho(f(x), f(x')) \leq c d(x, x')$$

for all $x, x' \in X$.

(3) If $\rho(f(x), f(x')) = d(x, x')$, then f is an isometry. Isometries are always injective.

Exercise:

$$\text{Isometry} \Rightarrow \text{Lipschitz} \Rightarrow \text{Uniformly Continuous} \Rightarrow \text{Continuous}.$$

For example, $f(x) = x^2$ on $[0, \infty)$ is continuous but not uniformly continuous, and $\sqrt{x}$ on $[0, 1]$ is uniformly continuous but not Lipschitz.