Section 13: metric spaces

Definition (Definition 13.1 – metric)

Suppose S is a set. A function $d : S \times S \to [0, \infty)$ is called a metric (or distance) if the following hold:

- **(D1)** Non-degeneracy: $d(x, y) = 0$ iff $x = y$ (hence $d(x, y) > 0$ when $x \neq y$).
- **(D2)** Symmetry: for $x, y \in S$, $d(x, y) = d(y, x)$.
- **(D3)** Triangle inequality: for $x, y, z \in S$, $d(x, y) + d(y, z) \geqslant d(x, z)$.

(S, d) is called a metric space.

Examples. 1. Let $S = \mathbb{R}$. Then $d(x, y) = |x - y|$ is a metric.

2. The n-dimensional Euclidean space $\mathbb{R}^n$ (p. 84): set of all n-tuples $\vec{x} = (x_1, \ldots, x_n)$ ($x_i \in \mathbb{R}$). $d(\vec{x}, \vec{y}) = \left(\sum_i (x_i - y_i)^2 \right)^{1/2}$. (D1), (D2) are easy to check; (D3) is harder.

Convergence in metric spaces

Definition (Definition 13.2)

Suppose (S, d) is a metric space.
A sequence $(s_n) \subset S$ converges to $s \in S$ if $\lim_n d(s_n, s) = 0$ – that is, $\forall\, \varepsilon > 0\ \exists N$ s.t. $d(s_n, s) < \varepsilon$ for $n > N$.
A sequence $(s_n) \subset S$ is Cauchy if $\forall\, \varepsilon > 0\ \exists N$ s.t. $d(s_n, s_m) < \varepsilon$ for $m, n > N$.

Note: for $(\mathbb{R}, d)$, we recover the usual definitions of "converges" and "Cauchy."

Complete metric spaces

Proposition

If (s_n) converges, then it is Cauchy.

Proof. Suppose (s_n) converges to s. Want to show (s_n) is Cauchy. Fix $\varepsilon > 0$, and find N s.t. $d(s_n, s_m) < \varepsilon$ when $n, m > N$. Find N s.t. $d(s_n, s) < \frac{\varepsilon}{2}$ if $n > N$. This N works: if $n, m > N$, then $d(s_n, s_m) \leqslant d(s_n, s) + d(s, s_m) < 2 \cdot \frac{\varepsilon}{2} = \varepsilon$. ∎

Does a Cauchy sequence always converge?

Definition (Definition 13.2 continued)

A metric space (S, d) is called complete if any Cauchy sequence in S converges.

Examples. $\mathbb{R}$ is complete. Will prove: $\mathbb{R}^n$ is complete.

Example: discrete metric

Discrete metric. S is a set. For $x, y \in S$, $d(x,y) = \begin{cases} 0 & x = y \\ 1 & x \neq y \end{cases}$. Is it a metric? **Yes!** Need to check the following:

(D1) $d(x,y) = 0$ iff $x = y$, $d(x,y) > 0$ if $x \neq y$.

(D2) For $x, y \in S$, $d(x,y) = d(y,x)$.

(D3) For $x, y, z \in S$, $d(x,y) + d(y,z) \geqslant d(x,z)$.

Verification of (D3), for x, y, z distinct:
$d(x,y) + d(y,z) = 1 + 1 \geqslant d(x,z) = 1$.

Is this metric space complete? **Yes!** Suppose a sequence (s_n) is Cauchy, and show it converges.
Find N s.t. $d(s_n, s_m) < 1$ for $n, m > N$. Then $s_n = s_{N+1}$ fr $n > N$. In particular, (s_n) converges to s_{N+1}.

A metric space which is not complete

Consider $\mathbb{Q}$, with metric $d(x, y) = |x - y|$ (metric inherited from $\mathbb{R}$). To show that this metric space is not complete, exhibit a Cauchy sequence which doesn't converge.

For $n \in \mathbb{N}$ find $r_n \in (\sqrt{2} - \frac{1}{n}, \sqrt{2})$. $r_n \to \sqrt{2}$ is $\mathbb{R}$, hence (r_n) is Cauchy (in $\mathbb{R}$, and also in $\mathbb{Q}$).

In $\mathbb{R}$, $\lim r_n = \sqrt{2}$. If (r_n) were convergent in $\mathbb{Q}$, it would have converged to the same limit. But $\sqrt{2} \notin \mathbb{Q}$, hence (r_n) is not convergent in $\mathbb{Q}$. ■

Example: Manhattan (taxicab) metric d_1

Consider $\mathbb{R}^n$; for $\vec{x} = (x_1, \ldots, x_n)$ and $\vec{y} = (y_1, \ldots, y_n)$,
$d_1(\vec{x}, \vec{y}) = \sum_{i=1}^n |x_i - y_i|$.
The notation d_1 is different from the textbook (Exercise 13.1), but in line with common mathematical usage.

Proposition

$(\mathbb{R}^n, d_1)$ is a complete metric space.

Proof: $(\mathbb{R}^n, d_1)$ is a metric space.

- Symmetry: $d_1(\vec{x}, \vec{y}) = d_1(\vec{y}, \vec{x})$ (straightforward).

- Non-degeneracy: $d_1(\vec{x}, \vec{y}) = 0$ iff $\vec{x} = \vec{y}$.
$d_1(\vec{x}, \vec{y}) = 0$ iff $\sum_{i=1}^n |x_i - y_i| = 0$ iff $x_i = y_i$ for any i iff $\vec{x} = \vec{y}$.

- Triangle inequality: $d_1(\vec{x}, \vec{z}) \leqslant d_1(\vec{x}, \vec{y}) + d_1(\vec{y}, \vec{z})$.
$\mathrm{RHS} = d_1(\vec{x}, \vec{y}) + d_1(\vec{y}, \vec{z}) = \sum_{i=1}^n |x_i - y_i| + \sum_{i=1}^n |y_i - z_i|$
$= \sum_{i=1}^n \big(|x_i - y_i| + |y_i - z_i|\big) \geqslant \sum_{i=1}^n \big|(x_i - y_i) + (y_i - z_i)\big|$
$= \sum_{i=1}^n |x_i - z_i| = d_1(\vec{x}, \vec{z}) = \mathrm{LHS}$. ■

Proof: $(\mathbb{R}^n, d_1)$ is complete

Lemma (similar to 13.3)

Consider a sequence $(\vec{x}^{(k)})_k$ in $\mathbb{R}^n$, with $\vec{x}^{(k)} = (x_i^{(k)})_{i=1}^n$. Then:

1. $(\vec{x}^{(k)})_k$ *is Cauchy iff* $(x_i^{(k)})_k$ *is Cauchy for* $1 \leqslant i \leqslant n$.
2. $(\vec{x}^{(k)})_k$ *converges to* $\vec{x} = (x_i)_{i=1}^n$ *iff* $\lim_k x_i^{(k)} = x_i$ *for* $1 \leqslant i \leqslant n$.

Lemma $\Rightarrow$ completeness of $(\mathbb{R}^n, d_1)$.

Need to show: if $(\vec{x}^{(k)})_k$ is Cauchy, then it converges.

$\vec{x}^{(k)} = (x_i^{(k)})_{i=1}^n$. $(x_i^{(k)})_k$ is Cauchy for $1 \leqslant i \leqslant n$.

$\mathbb{R}$ is complete $\Rightarrow (x_i^{(k)})_k$ converges to some $x_i \in \mathbb{R}$

$\Rightarrow (\vec{x}^{(k)})_k$ converges to $\vec{x} = (x_i)_{i=1}^n$. ∎

Proof of Lemma, part (1) (part (2) is similar)

$(\vec{x}^{(k)})_k$ is Cauchy iff $(x_i^{(k)})_k$ is Cauchy for $1 \leqslant i \leqslant n$.

$\Rightarrow$: Suppose $(\vec{x}^{(k)})_k$ is Cauchy. Fix i. Want: $(x_i^{(k)})_k$ is Cauchy.
For $\varepsilon > 0$ need to find N s.t. $|x_i^{(k)} - x_i^{(m)}| < \varepsilon$ for $k, m > N$.
Find N s.t. $d_1(\vec{x}^{(k)}, \vec{x}^{(m)}) < \varepsilon$ for $k, m > N$.
$\varepsilon > d_1(\vec{x}^{(k)}, \vec{x}^{(m)}) = \sum_{j=1}^{n} |x_j^{(k)} - x_j^{(m)}| \geqslant |x_i^{(k)} - x_i^{(m)}|$, so this N works for us! $\square$

$\Leftarrow$: Suppose $(x_i^{(k)})_k$ is Cauchy $\forall\, i$. Want: $(\vec{x}^{(k)})_k$ is Cauchy.
Fix $\varepsilon > 0$. Need to find N s.t. $d_1(\vec{x}^{(k)}, \vec{x}^{(m)}) < \varepsilon$ for $k, m > N$.
For $1 \leqslant i \leqslant n$ find $N_i \in \mathbb{N}$ s.t. $|x_i^{(k)} - x_i^{(m)}| < \frac{\varepsilon}{n}$ for $k, m > N_i$.
For $k, m > N = \max\{N_1, \ldots, N_n\}$,
$d_1(\vec{x}^{(k)}, \vec{x}^{(m)}) = \sum_{j=1}^{n} |x_j^{(k)} - x_j^{(m)}| < n \cdot \frac{\varepsilon}{n} = \varepsilon$. $\blacksquare$

The n-dimensional Euclidean space revisited

The n-dimensional Euclidean space $\mathbb{R}^n$ (p. 84): set of all n-tuples $\vec{x} = (x_1, \ldots, x_n)$ $(x_i \in \mathbb{R})$. $d(\vec{x}, \vec{y}) = \left(\sum_i (x_i - y_i)^2 \right)^{1/2}$.

Proposition (Only partially proved in textbook)

$(\mathbb{R}^n, d)$ *is a complete metric space.*

- <u>Symmetry</u>: $d(\vec{x}, \vec{y}) = d(\vec{y}, \vec{x})$ (straightforward).
- <u>Non-degeneracy</u>: $d(\vec{x}, \vec{y}) = 0$ iff $\vec{x} = \vec{y}$.
$d(\vec{x}, \vec{y})^2 = 0$ iff $\sum_{i=1}^{n} (x_i - y_i)^2 = 0$ iff $x_i = y_i$ for any i iff $\vec{x} = \vec{y}$.
- <u>Triangle inequality</u>: requires extra work. We shall define inner product and prove some useful inequalities.

The inner product (not in textbook)

Definition (Inner product and magnitude)

For $\vec{x}, \vec{y} \in \mathbb{R}^n$ define the inner (scalar) product: $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i y_i$. The magnitude of $\vec{x}$ is $\|\vec{x}\| = \left(\sum_{i=1}^n x_i^2 \right)^{1/2} = \langle \vec{x}, \vec{x} \rangle^{1/2} = d(\vec{x}, 0)$.

Observation: $d(\vec{x}, \vec{y}) = \|\vec{x} - \vec{y}\|$.

Properties of inner product.
Symmetry. For $\vec{x}, \vec{y} \in \mathbb{R}^n$, $\langle \vec{x}, \vec{y} \rangle = \langle \vec{y}, \vec{x} \rangle$.
Linearity. For $\vec{x}, \vec{y}, \vec{z} \in \mathbb{R}^n$ and $t \in \mathbb{R}$, $\langle t\vec{x}, \vec{y} \rangle = t \langle \vec{x}, \vec{y} \rangle$, $\langle \vec{x}, \vec{y} + \vec{z} \rangle = \langle \vec{x}, \vec{y} \rangle + \langle \vec{x}, \vec{z} \rangle$

Why do we need inner product? Orthogonality! Without inner products, cannot talk about orthogonal projections.

Triangle inequality and BCS inequality

Lemma (Triangle Inequality lite; not in textbook)

For $\vec{x}, \vec{y} \in \mathbb{R}^n$, $\|\vec{x} + \vec{y}\| \leqslant \|\vec{x}\| + \|\vec{y}\|$.

Lemma $\Rightarrow$ triangle inequality: For $\vec{a}, \vec{b}, \vec{c} \in \mathbb{R}^n$,
$d(\vec{a}, \vec{c}) = \|\vec{a} - \vec{c}\| = \|(\vec{a} - \vec{b}) + (\vec{b} - \vec{c})\|$ (now apply Lemma)
$\leqslant \|\vec{a} - \vec{b}\| + \|\vec{b} - \vec{c}\| = d(\vec{a}, \vec{b}) + d(\vec{b}, \vec{c})$. ∎

Theorem (Bunyakovsky-Cauchy-Schwarz, or BCS)

For $\vec{x}, \vec{y} \in \mathbb{R}^n$, $\left|\langle \vec{x}, \vec{y} \rangle\right| \leqslant \|\vec{x}\| \, \|\vec{y}\|$.

Inner product

Definition (Inner product and magnitude)

For $\vec{x}, \vec{y} \in \mathbb{R}^n$ define the inner (scalar) product: $\langle \vec{x}, \vec{y} \rangle = \sum_{i=1}^n x_i y_i$. The magnitude of $\vec{x}$ is $\|\vec{x}\| = \left(\sum_{i=1}^n x_i^2 \right)^{1/2} = \langle \vec{x}, \vec{x} \rangle^{1/2} = d(\vec{x}, 0)$.

Theorem (Bunyakovsky-Cauchy-Schwarz, or BCS)

For $\vec{x}, \vec{y} \in \mathbb{R}^n$, $\left| \langle \vec{x}, \vec{y} \rangle \right| \leqslant \|\vec{x}\| \, \|\vec{y}\|$.

Proof. Nothing to prove if $\vec{x} = 0$ or $\vec{y} = 0$.

If $\vec{x}, \vec{y} \neq 0$, let $f(t) = \langle \vec{x} + t\vec{y}, \vec{x} + t\vec{y} \rangle = \|\vec{x} + t\vec{y}\|^2 \geqslant 0$. OTOH
$f(t) = \langle \vec{x}, \vec{x} \rangle + \langle \vec{x}, t\vec{y} \rangle + \langle t\vec{y}, \vec{x} \rangle + \langle t\vec{y}, t\vec{y} \rangle = \|\vec{x}\|^2 + 2t\langle \vec{x}, \vec{y} \rangle + t^2\|\vec{y}\|^2$
$= \left(t\|\vec{y}\| + \frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{y}\|} \right)^2 + \left(\|\vec{x}\|^2 - \frac{\langle \vec{x}, \vec{y} \rangle^2}{\|\vec{y}\|^2} \right)$.
$f(t) \geqslant \|\vec{x}\|^2 - \frac{\langle \vec{x}, \vec{y} \rangle^2}{\|\vec{y}\|^2}$, with equality for $t = -\frac{\langle \vec{x}, \vec{y} \rangle}{\|\vec{y}\|^2}$. Thus,
$\|\vec{x}\|^2 - \frac{\langle \vec{x}, \vec{y} \rangle^2}{\|\vec{y}\|^2} \geqslant 0$, or equivalently, $\langle \vec{x}, \vec{y} \rangle^2 \leqslant \|\vec{y}\|^2 \|\vec{x}\|^2$.
Take $\sqrt{\ }$ of both sides. ■