I. Linear Algebra

I.3. Normed Spaces

Lecture based on

https://github.com/gwthomas/math4ml (Garrett Thomas, 2018)

https://mml-book.github.io/ (Deisenroth et al., 2020, Mathematics for Machine Learning)

Prof. Dr. Christoph Lippert

Digital Health & Machine Learning

Metrics generalize the notion of distance from Euclidean space

A **metric** on a set S is a function $d : S \times S \to \mathbb{R}$ that satisfies

i. $d(x,y) \geq 0$, with equality if and only if $x = y$
ii. $d(x,y) = d(y,x)$
iii. $d(x,z) \leq d(x,y) + d(y,z)$ (**triangle inequality**)

for all $x, y, z \in S$.

A **norm** on a real vector space V is a function $\|\cdot\|: V \to \mathbb{R}$ that satisfies

(i) $\|\mathbf{x}\| \geq 0$, with equality if and only if $\mathbf{x} = \mathbf{0}$

(ii) $\|\alpha\mathbf{x}\| = |\alpha|\|\mathbf{x}\|$

(iii) $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$ (**triangle inequality**)

for all $\mathbf{x}, \mathbf{y} \in V$ and all $\alpha \in \mathbb{R}$.

Lemma (Reverse triangle inequality)

$$\|x - y\| \geq \left|\|x\| - \|y\|\right|$$

- A vector space endowed with a norm is a **normed vector space**, or simply a **normed space**.
- Any norm on V induces a distance metric on V:

$$d(\mathbf{x}, \mathbf{y}) = \|\mathbf{x} - \mathbf{y}\|$$

Important norms on $\mathbb{R}^n$:

$$\|\mathbf{x}\|_1 = \sum_{i=1}^{n} |x_i|$$

$$\|\mathbf{x}\|_2 = \sqrt{\sum_{i=1}^{n} x_i^2}$$

$$\|\mathbf{x}\|_p = \left(\sum_{i=1}^{n} |x_i|^p\right)^{\frac{1}{p}} \qquad (p \geq 1)$$

$$\|\mathbf{x}\|_\infty = \max_{1 \leq i \leq n} |x_i|$$

Note that we require $p \geq 1$ for the general definition of the p-norm because the triangle inequality fails to hold if $p < 1$.

Consider a subset $E \subset \mathbb{R}$

1. E is bounded above if there exists $\beta \in \mathbb{R}$ s. t. $x \leq \beta$
2. If E is bounded above:

   Let the **supremum** over E $(\sup E)$ denote the smallest such β.
3. If E is not bounded above, $\sup E = \infty$

Consider a function $f : A \to \mathbb{R}$

Let $E = \{f(\mathbf{x}) | \mathbf{x} \in A\}$

$$\sup E = \sup\{f(\mathbf{x}) | \mathbf{x} \in A\}$$

Sometimes there exists $\mathbf{x}_0 \in A$ such that $f(\mathbf{x}_0) = \sup\{f(\mathbf{x}) | \mathbf{x} \in A\}$

In this case, we speak about the **maximum** rather than the supremum.

We say "The supremum is attained."

Example

Consider

$$f(x) = -x^2 + 2x = x(2 - x)$$

with $x \in [0, 3[$.

$$\max\{f(x) | x \in [0, 3[\} = f(1) = 1$$

The supremum is attained.

Example

Consider $f(x) = x^2 - 2x$, $x \in [0, 3[$
The supremum is not attained.

Let $[a, b]$ be a closed, bounded interval.

Then any continuous function

$$f[a, b] \to \mathbb{R}$$

attains its supremum.

i.e., $\exists x_0 \in [a, b]$: $f(x_0) = \sup\{f(x) | x \in [a, b]\}$

Let $C[a,b] := \{f : [a,b] \to \mathbb{R} \mid f \text{ continuous}\}$

Example

$C[a,b]$ is a vector space, and it has the norm

$$\|f\|_\infty = \max_{x \in [a,b]} |f(x)|$$

A vector space V must satisfy

- (i) There exists a $\mathbf{0} \in V$ such that $\mathbf{x} + \mathbf{0} = \mathbf{x}$ for all $\mathbf{x} \in V$
- (ii) For each $\mathbf{x} \in V$, there exists a $-\mathbf{x}$ such that $\mathbf{x} + (-\mathbf{x}) = \mathbf{0}$
- (iii) There exists a $1 \in \mathbb{R}$ such that $1\mathbf{x} = \mathbf{x}$ for all $\mathbf{x} \in V$
- (iv) **Commutativity**: $\mathbf{x} + \mathbf{y} = \mathbf{y} + \mathbf{x}$ for all $\mathbf{x}, \mathbf{y} \in V$
- (v) **Associativity**: $(\mathbf{x} + \mathbf{y}) + \mathbf{z} = \mathbf{x} + (\mathbf{y} + \mathbf{z})$ and $\alpha(\beta\mathbf{x}) = (\alpha\beta)\mathbf{x}$ for all $\mathbf{x}, \mathbf{y}, \mathbf{z} \in V$ and $\alpha, \beta \in \mathbb{R}$
- (vi) **Distributivity**: $\alpha(\mathbf{x} + \mathbf{y}) = \alpha\mathbf{x} + \alpha\mathbf{y}$ and $(\alpha + \beta)\mathbf{x} = \alpha\mathbf{x} + \beta\mathbf{x}$ for all $\mathbf{x}, \mathbf{y} \in V$ and $\alpha, \beta \in \mathbb{R}$

Let $C[a,b] := \{f : [a,b] \to \mathbb{R} | f \text{continuous}\}$

Example

$C[a,b]$ is a vector space, and it has the norm

$$\|f\|_\infty = \max_{x \in [a,b]} |f(x)|$$

A **norm** on a real vector space V is a function $\|\cdot\| : V \to \mathbb{R}$ that satisfies

i $\|\mathbf{x}\| \geq 0$,
with $\|\mathbf{x}\| = 0 \Leftrightarrow \mathbf{x} = \mathbf{0}$

ii $\|\alpha \mathbf{x}\| = |\alpha| \|\mathbf{x}\|$

iii $\|\mathbf{x} + \mathbf{y}\| \leq \|\mathbf{x}\| + \|\mathbf{y}\|$

Consider a sequence $\{\mathbf{x}_k^\infty\} \subseteq V$.

$$\{\mathbf{x}_k^\infty\} = \{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots\}$$

Then $\{\mathbf{x}_k^\infty\}$ **converges** towards $\mathbf{x} \in V$
if $\|\mathbf{x} - \mathbf{x}_k\| \to 0$ as $k \to \infty$.

This can also be written as
$\mathbf{x}_k \to \mathbf{x}$ as $k \to \infty$
or $\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}$

Exact definition:
For all $\epsilon > 0$, there exists $N \in \mathbb{N}$, such that

$$\|\mathbf{x} - \mathbf{x}_k\| \leq \epsilon$$

for $k \geq N$.

Note that metrics allow limits to be defined for mathematical objects other than real numbers.

- A sequence $\{x_k\} \subseteq S \to x$ if for any $\epsilon > 0$, there exists $N \in \mathbb{N}$ such that $d(x_k, x) < \epsilon$ for all $k \geq N$.
- The definition above is a special case when using the metric $d(x, y) = |x - y|$.

Example

$$f_k(x) = \sum_{n=0}^{k} x^n$$
$$= \frac{1 - x^{k+1}}{1 - x}$$

with $x \in]-1, 1[$

$$f(x) = \sum_{k=0}^{\infty} x^k$$
$$= \frac{1}{1 - x}$$

with $x \in]-1, 1[$

Note:

$$f, f_k \in C\left[0, \frac{1}{2}\right]$$

So this has the norm $\|f\|_\infty = \max_{x \in [0,1/2]} |f(x)|$

Question:
Will

$$f_k \to f$$

in $C[0, 1/2]$?
i.e.,

$$\|f - f_k\|_\infty \to 0$$

for $k \to \infty$?

Here's a fun fact:

For any given finite-dimensional vector space V, all norms on V are equivalent in the sense that for two norms $\|\cdot\|_A, \|\cdot\|_B$, there exist constants $\alpha, \beta > 0$ such that

$$\alpha\|\mathbf{x}\|_A \leq \|\mathbf{x}\|_B \leq \beta\|\mathbf{x}\|_A$$

for all $\mathbf{x} \in V$.

Therefore convergence in one norm implies convergence in any other norm.

This rule may not apply in infinite-dimensional vector spaces such as function spaces, though.

Summary

- Metrics generalize the notion of distance from Euclidean space
- Any norm on V induces a distance metric
- Using norms we can define convergence on real vector spaces
- Using metrics we can define convergence on other types of objects