Computational Statistics & Machine Learning
Lecture 7
Vector, Metric, Banach spaces

Mark Girolami
mag92@cam.ac.uk

Department of Engineering
University of Cambridge

November 3, 2021

Overview

- Vector Space
- Cauchy Sequences
- Complete Vector and Metric spaces
- Banach Space

Vector Space

- Vector space is a set of objects to which addition and multiplication applies
- A vector space V is a collection of elements x - *vectors* with following axioms
  - $x + y = y + x \in V$, addition is commutative and admits closure
  - $x + (y + z) = (x + y) + z \in V$, associative
  - $x + 0 = x$ zero vector exists $\forall x \in V$
  - $x + (-x) = 0$ additive inverse exists $\forall x \in V$
  - For $\alpha \in \mathbb{F}$ ($\mathbb{R}$ or $\mathbb{C}$) scalar multiplication is closed $\alpha x \in V$ for any $x \in V$
  - Scalar multiplication distributive $\alpha(x + y) = \alpha x + \alpha y \in V$
  - Scalar multiplication associative $\alpha(\beta x) = \beta(\alpha x)$
  - Zero and Unit scalar $0 \times x = 0 \in V$, $1 \times x = x$
- Vector space is a *Group* under addition

Inner product space

- $\mathbb{R}^3$ is a vector space - can you suggest some others?
- The *inner product* is a mapping from a pair of vectors x and y to a scalar satisfying the following
  - $\langle x, y\rangle = \langle y, x\rangle$ for $\mathbb{F} = \mathbb{R}$ what if $\mathbb{F} = \mathbb{C}$?
  - $\langle \alpha x + \beta y, z\rangle = \alpha\langle x, z\rangle + \beta\langle y, z\rangle$ where $\alpha, \beta \in \mathbb{R}$
  - $\langle x, x\rangle \geq 0$ for any x
  - $\langle x, x\rangle = 0$ if and only if $x = 0$
- Vector space endowed with *inner product* are termed *inner product spaces*
- $\mathrm{z} = \{z_1, z_2, \cdots, z_n\} \in \mathbb{C}^n$ for a and b $\in \mathbb{C}^n$the inner product defined by $\langle \mathrm{a}, \mathrm{b}\rangle = \sum_{i=1}^{n} a_i^* b_i$
- Functions defined by a variable on interval $x \in [a, b]$ of $\mathbb{R}$ the inner product between functions $f(\cdot)$ and $g(\cdot)$ is defined by

$$\int_a^b f(x)g(x)dx$$

  does this satisfy requirements of an inner product?

The Norm in a vector space space

- As $\langle x, x\rangle \geq 0$ it is always positive (and real) this is referred to as the *norm* of the vector usually represented by $\|x\| = \sqrt{\langle x, x\rangle}$ denoting the *length* of the vector
- In this case the *norm* has been induced by the inner product $\langle \cdot, \cdot \rangle$
- The L_p and l_p norms (we will meet these shortly) are defined with no reference to an inner product , the norm is more general than the inner product
- Now the vector space is endowed with a norm and inner product, two important capabilities emerge
  - The definition of the length of an element of the vector space (norm)
  - The assessment of convergence of a sequence of elements in the vector space

Cauchy-Schwarz

- Given V for any two elements $x, y \in V$
- Schwarz Inequality

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

- Triangle Inequality

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

- Parallelogram Law

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

- Schwarz defines the length of the projection of y onto x and cannot be greater than the norm of y

$$\|y\| \geq \frac{|\langle x, y \rangle|}{\|x\|}$$

- Derive (prove) each of these three statements

Orthogonality and Orthonormality

- The inner product enable the definition of vectors which are orthogonal to each other
- Two elements x and y in an inner product space are orthogonal if and only if $\langle x, y \rangle = 0$
- The function $f(x) = \sin(x)$ and $g(x) = \cos(x)$ defined on closed interval $[-\pi, \pi]$ are orthogonal as

$$(f, g) = \int_{-\pi}^{\pi} \sin(x)\cos(x)dx = 0$$

- Normalising each vector i.e. $\frac{x}{||x||}$ then $\langle x_i, x_j \rangle = \delta_{i,j}$, orthogonal unit length vectors
- The unit vectors define Euclidean , Cartesian space i.e. e_1, e_2, e_3 where $e_1 = (1, 0, 0), e_2 = (0, 1, 0), e_3 = (0, 0, 1)$

Basis in finite dimension

- An orthonormal finite set of vectors $e_1, e_2, \cdots, e_n$ is linearly independent
- This holds if and only if $\sum_{i=1}^{n} c_i e_i = 0 \implies c_i = 0 \;\forall\, i$
- Try and prove this statement
- Any orthonormal set of vectors serves as a *basis* for a finite dimensional inner product space
- A basis for V is a set of linearly independent vectors $\{e_i\}$ of V and *every* vector in V can be expressed as

$$x = \sum_{i=1}^{n} \alpha_i e_i$$

- The $\alpha_1, \alpha_2, \cdots, \alpha_n$ are the coordinates of x
- What happens when $n = \infty$?

Complete vector spaces

- A basis is only defined for finite dimensional spaces, the space is complete
- In infinite dimensions things are a little more interesting as defining an infinite set of orthonormal vectors may not yield a complete space - there may be holes in it
- Think of a Fourier series expansion of the notch function which is well defined, but it does not lie in the space of continuous functions i.e. $x = \sum_{i=1}^{\infty} \alpha_i e_i$ but $x \notin V$
- Assess completeness of an infinite dimensional space using the *Cauchy criterion*

Cauchy Sequences

- Baron Augustin-Louis Cauchy 1789 - 1857

Cauchy Sequences

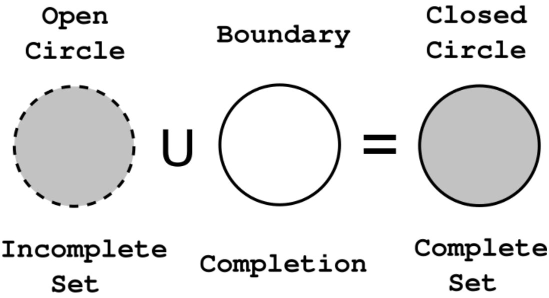
- A sequence of vectors $\{x_1, x_2, x_3, \cdots\}$ is a *Cauchy Sequence* of vectors if for any positive $\epsilon > 0$, there exists a number N that

$$\|x_m - x_n\| < \epsilon$$

  for all $m, n > N$
- What this means is that for a Cauchy sequence the sequence of terms x_m and x_n get closer and closer in an unlimited manner as $m, n \to \infty$
- An infinite sequence of vectors $\{x_1, x_2, \cdots\}$ is said to be convergent if there exists and element $x \in V$ such that $\|x_n - x\| \to 0$
- If every Cauchy sequence in a vector space V converges to a limit that is also in V, then the space V is *complete*.

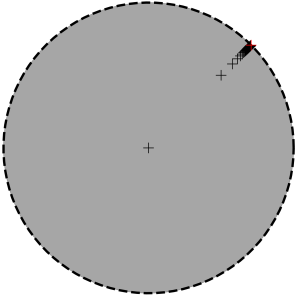
Completeness

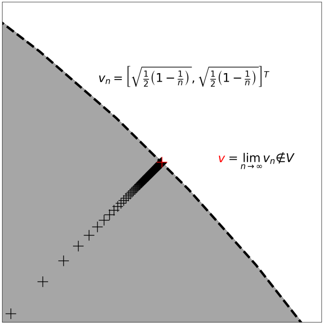
- The diagram below illustrates the notion of completeness.



Completeness

- See the notebook `Lecture_07.ipynb` for example showing the incompleteness of the open circle.





Metric Vector Spaces

- Consider two elements x and y in a vector space V map them through a function ρ to a non-negative real number a
- The function ρ is a *distance* function if
  - $\rho(x, y) = 0$ if and only if $x = y$
  - $\rho(x, y) = \rho(y, x)$
  - $\rho(x, y) \leq \rho(x, z) + \rho(z, y)$ for any $z \in V$
- A vector space having a distance function is a *metric* vector space
- $\mathbb{R}$ with distance function $\rho(x, y) = |x - y|$ forms a metric vector space prove it

Normed Spaces

- A metric space is normed if to each element $x \in V$ there is a non-negative number $\|x\|$ which is the norm of x
- A normed space is also a metric space under the definition of distance $\rho(x, y) = \|x - y\|$
- Every norm induces a distance function the reverse is not neccessarily the case
- $x = (x_1, x_2, \cdots, x_n)$ can define the l^p norm

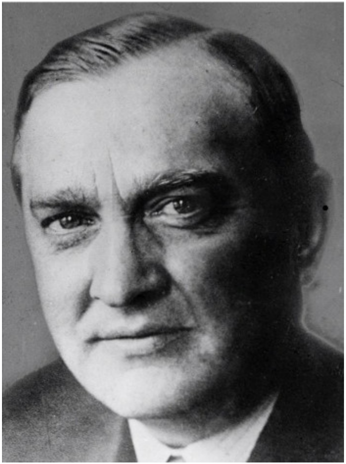
$$\|x\| = \left(\sum_{i-=1}^{n} x_i^p \right)^{\frac{1}{p}}$$

- An alternative space emerges if defining the norm as

$$\|x\| = \max\{|x_k|; 1 \leq k \leq n\}$$

Banach Space

- Stefan Banach 1892 - 1945



Banach Spaces

- If a normed space is *complete* it is called a Banach space
- All finite dimensional normed spaces are complete, they are all Banach spaces
- Infinite dimensional normed spaces that are complete are called Banach spaces
- The space V with elements $x = (x_1, x_2, \cdots,)$ satisfies

$$\sum_{i=1}^{\infty} |x_i|^p < \infty, (p \geq 1)$$

- This set is then a Banach space, the l^p-space, under the defined norm $\|x\| = (\sum_{i=1}^{\infty} |x_i|^p)^{\frac{1}{p}}$

Banach Spaces

- We met the Lebesgue integral in previous lectures
- The Lebesgue integral is used to define the L^p space of functions
- The set of functions $f(x)$ expressed by

$$\int_a^b |f(x)|^p dx < \infty$$

  where dx denotes Lebesgue measure, forms a Banach space, the L^p space under the norm

$$\|f\|_p = \left(\int_a^b |f(x)|^p dx\right)^{\frac{1}{p}}$$

- Completeness of the L^p space can be proved by showing every Cauchy sequence in L^p has a limit in L^p. want to try?
- What applications to Machine Learning can you see for normed Banach spaces?