



Machine Learning

MECH70025

Manab Shrestha

Contents

Chapter 1	Bayesian Decision Theory and Maximum Likelihood _____	Page 2 _____
Chapter 2	Regression Approaches _____	Page 3 _____
Chapter 3	Linear Discriminant Functions _____	Page 4 _____
Chapter 4	Support Vector Machines _____	Page 5 _____
Chapter 5	Multilayer Neural Networks _____	Page 6 _____
Chapter 6	Nonparametric Techniques _____	Page 7 _____
Chapter 7	Nonmetric Methods _____	Page 8 _____
Chapter 8	Unsupervised Learning and Clustering _____	Page 9 _____

Chapter 1

Bayesian Decision Theory and Maximum Likelihood

Chapter 2

Regression Approaches

Chapter 3

Linear Discriminant Functions

Chapter 4

Support Vector Machines

Chapter 5

Multilayer Neural Networks

Chapter 6

Nonparametric Techniques

Chapter 7

Nonmetric Methods

Chapter 8

Unsupervised Learning and Clustering

Random Examples

Definition 8.0.1: Limit of Sequence in \mathbb{R}

Let $\{s_n\}$ be a sequence in \mathbb{R} . We say

$$\lim_{n \rightarrow \infty} s_n = s$$

where $s \in \mathbb{R}$ if \forall real numbers $\epsilon > 0 \exists$ natural number N such that for $n > N$

$$s - \epsilon < s_n < s + \epsilon \text{ i.e. } |s - s_n| < \epsilon$$

Question 1

Is the set $x\text{-axis} \setminus \{\text{Origin}\}$ a closed set

Solution: We have to take its complement and check whether that set is a open set i.e. if it is a union of open balls

Note:-

We will do topology in Normed Linear Space (Mainly \mathbb{R}^n and occasionally \mathbb{C}^n) using the language of Metric Space

Claim 8.0.1 Topology

Topology is cool

Example 8.0.1 (Open Set and Close Set)

Open Set:

- ϕ
- $\bigcup_{x \in X} B_r(x)$ (Any $r > 0$ will do)

Closed Set:

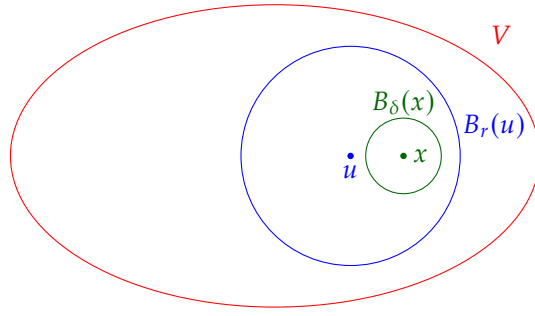
- $B_r(x)$ is open
- X, ϕ
- $\overline{B_r(x)}$

$x\text{-axis} \cup y\text{-axis}$

Theorem 8.0.1

If $x \in$ open set V then $\exists \delta > 0$ such that $B_\delta(x) \subset V$

Proof: By openness of V , $x \in B_r(u) \subset V$



Given $x \in B_r(u) \subset V$, we want $\delta > 0$ such that $x \in B_\delta(x) \subset B_r(u) \subset V$. Let $d = d(u, x)$. Choose δ such that $d + \delta < r$ (e.g. $\delta < \frac{r-d}{2}$)

If $y \in B_\delta(x)$ we will be done by showing that $d(u, y) < r$ but

$$d(u, y) \leq d(u, x) + d(x, y) < d + \delta < r$$

☺

Corollary 8.0.1

By the result of the proof, we can then show...

Lemma 8.0.1

Suppose $\vec{v}_1, \dots, \vec{v}_n \in \mathbb{R}^n$ is subspace of \mathbb{R}^n .

Proposition 8.0.1

$1 + 1 = 2$.

Random

Definition 8.0.2: Normed Linear Space and Norm $\|\cdot\|$

Let V be a vector space over \mathbb{R} (or \mathbb{C}). A norm on V is function $\|\cdot\| : V \rightarrow \mathbb{R}_{\geq 0}$ satisfying

- ① $\|x\| = 0 \iff x = 0 \forall x \in V$
- ② $\|\lambda x\| = |\lambda| \|x\| \forall \lambda \in \mathbb{R}(\text{or } \mathbb{C}), x \in V$
- ③ $\|x + y\| \leq \|x\| + \|y\| \forall x, y \in V$ (Triangle Inequality/Subadditivity)

And V is called a normed linear space.

• Same definition works with V a vector space over \mathbb{C} (again $\|\cdot\| \rightarrow \mathbb{R}_{\geq 0}$) where ② becomes $\|\lambda x\| = |\lambda| \|x\| \forall \lambda \in \mathbb{C}, x \in V$, where for $\lambda = a + ib, |\lambda| = \sqrt{a^2 + b^2}$

Example 8.0.2 (p -Norm)

$V = \mathbb{R}^m, p \in \mathbb{R}_{\geq 0}$. Define for $x = (x_1, x_2, \dots, x_m) \in \mathbb{R}^m$

$$\|x\|_p = \left(|x_1|^p + |x_2|^p + \dots + |x_m|^p \right)^{\frac{1}{p}}$$

(In school $p = 2$)

Special Case $p = 1$: $\|x\|_1 = |x_1| + |x_2| + \dots + |x_m|$ is clearly a norm by usual triangle inequality.

Special Case $p \rightarrow \infty$ (\mathbb{R}^m with $\|\cdot\|_\infty$): $\|x\|_\infty = \max\{|x_1|, |x_2|, \dots, |x_m|\}$

For $m = 1$ these p -norms are nothing but $|x|$. Now exercise

Question 2

Prove that triangle inequality is true if $p \geq 1$ for p -norms. (What goes wrong for $p < 1$?)

Solution: For Property ③ for norm-2

When field is \mathbb{R} :

We have to show

$$\begin{aligned}\sum_i (x_i + y_i)^2 &\leq \left(\sqrt{\sum_i x_i^2} + \sqrt{\sum_i y_i^2} \right)^2 \\ \Rightarrow \sum_i (x_i^2 + 2x_i y_i + y_i^2) &\leq \sum_i x_i^2 + 2\sqrt{\left[\sum_i x_i^2 \right] \left[\sum_i y_i^2 \right]} + \sum_i y_i^2 \\ \Rightarrow \left[\sum_i x_i y_i \right]^2 &\leq \left[\sum_i x_i^2 \right] \left[\sum_i y_i^2 \right]\end{aligned}$$

So in other words prove $\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle$ where

$$\langle x, y \rangle = \sum_i x_i y_i$$

Note:-

- $\|x\|^2 = \langle x, x \rangle$
- $\langle x, y \rangle = \langle y, x \rangle$
- $\langle \cdot, \cdot \rangle$ is \mathbb{R} -linear in each slot i.e.

$$\langle rx + x', y \rangle = r\langle x, y \rangle + \langle x', y \rangle \text{ and similarly for second slot}$$

Here in $\langle x, y \rangle$ x is in first slot and y is in second slot.

Now the statement is just the Cauchy-Schwartz Inequality. For proof

$$\langle x, y \rangle^2 \leq \langle x, x \rangle \langle y, y \rangle$$

expand everything of $\langle x - \lambda y, x - \lambda y \rangle$ which is going to give a quadratic equation in variable λ

$$\begin{aligned}\langle x - \lambda y, x - \lambda y \rangle &= \langle x, x - \lambda y \rangle - \lambda \langle y, x - \lambda y \rangle \\ &= \langle x, x \rangle - \lambda \langle x, y \rangle - \lambda \langle y, x \rangle + \lambda^2 \langle y, y \rangle \\ &= \langle x, x \rangle - 2\lambda \langle x, y \rangle + \lambda^2 \langle y, y \rangle\end{aligned}$$

Now unless $x = \lambda y$ we have $\langle x - \lambda y, x - \lambda y \rangle > 0$ Hence the quadratic equation has no root therefore the discriminant is greater than zero.

When field is \mathbb{C} :

Modify the definition by

$$\langle x, y \rangle = \sum_i \bar{x}_i y_i$$

Then we still have $\langle x, x \rangle \geq 0$

Algorithms

Algorithm 1: what

Input: This is some input

Output: This is some output

/ This is a comment */*

```
1 some code here;
2  $x \leftarrow 0$ ;
3  $y \leftarrow 0$ ;
4 if  $x > 5$  then
5   |  $x$  is greater than 5 ;                                // This is also a comment
6 else
7   |  $x$  is less than or equal to 5;
8 end
9 foreach  $y$  in 0..5 do
10  |  $y \leftarrow y + 1$ ;
11 end
12 for  $y$  in 0..5 do
13  |  $y \leftarrow y - 1$ ;
14 end
15 while  $x > 5$  do
16  |  $x \leftarrow x - 1$ ;
17 end
18 return Return something here;
```
