17/8/2023

Deep Learning

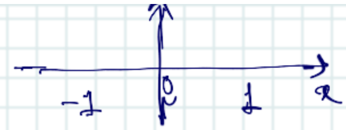
- Recap
- Math preliminaries
  - Metric, metric spaces
  - Linear spaces
  - Norms, normed linear space
  - Convergence
  - Cauchy sequences
  - Completeness

- Math prelims:
  - $D_s = \{ (x_1, y_1), (x_2, y_2) \ldots (x_N, y_N) \}$; $x \in \mathcal{X}$, $y \in \mathcal{Y}$ (supervised learning)

    input 1; label 1; # data points; input space; label space

  - $D_u = \{ x_1, x_2 \ldots x_N \}$; $x \in \mathcal{X}$ (unsupervised learning).
  - Assume that the datapoints are drawn from a fixed but unknown distribution $p(x,y)$
  - In this setting, we want to estimate $p(y|x) = \frac{p(x,y)}{p(x)}$
  - The machine learning model is a parameterized function that maps the input space to the label space, $f_\theta : \mathcal{X} \rightarrow \mathcal{Y}$; $f(x;\theta)$
  - In the probabilistic setting, the sup. learning problem is one of estimating $p(y|x)$ using $f(x;\theta)$
  - Notion of distance is important since we want to measure how close our model output is to the ground truth. $d(\cdot,\cdot) : \mathcal{Y} \times \mathcal{Y} \rightarrow \mathbb{R}$
  - Distance: Let X be an arbitrary set. A metric or a distance function is defined as $d : X \times X \rightarrow \mathbb{R}$ such that:
    - $d(x,y) \geq 0$ for all $x, y \in X$ (non-negativity)
    - $d(x,y) = d(y,x)$ for all $x, y \in X$ (symmetry)
    - $d(x,y) \leq d(x,z) + d(z,y)$ for all $x, y, z \in X$. (triangle inequality)
    - $d(x,y) = 0 \iff x = y$.

Ex1: $\mathcal{X} = \mathbb{R}$, $d(x, y) = \sqrt{(x-y)^2}$: let $z = 0$, $x = 1$, $y = -1$



$d(1, -1) = 4$,

$d(1, 0) = 1$, $d(0, -1) = 1$

Ex2: $\mathcal{X} = \mathbb{N}$; $d(x,y) = \begin{cases} 0 & \text{if } x = y \\ 1 & \text{if } x \neq y. \end{cases}$

Think of more examples.

<u>Metric space</u>: A metric space is a set $\mathcal{X}$ endowed with a distance d, i.e. $(\mathcal{X}, d)$.

- In the DL setting, the function $\mathcal{L}(\theta) = \sum_{i=1}^{n} d(y_i, \hat{y}_i) = \sum_{i=1}^{n} d(y_i, f(x_i; \theta))$

- <u>Linear space</u>: A non-empty set $\underline{\underline{\mathcal{X}}}$ over a scalar field $\mathbb{R}$ (or $\mathbb{C}$) that is closed under addition and scalar multiplication.

  - $x + y = y + x$
  - $x + (y + z) = (x + y) + z$
  - $x + (-x) = 0$
  - $(\lambda x) y = \lambda (x y)$?
  - $\lambda (x + y) = \lambda x + \lambda y$
  - $(\lambda + \gamma) x = \lambda x + \gamma x$