Homework 0

Spring 24, CS 446/ECE 449: Machine Learning

Instructor: Han Zhao and Shenlong Wang

1 Marginal Independence vs Joint Independence

Let X, Y, Z be three random variables.

1.1

Show that if X, Y, Z are jointly independent, then they are pairwise independent as well, i.e., $X \perp Y, Y \perp Z, Z \perp X$.

1.2

Show that the other direction is not true by constructing a counterexample.

2 ℓ_p Norms

Let $p \in \mathbb{N}$, for $x \in \mathbb{R}^d$, we define $\|x\|_p := \left(\sum_{i=1}^d |x_i|^p\right)^{1/p}$. For $d = 2$, plot the corresponding feasible regions of $\|x\|_p \leq 1$, for $p = 1, 2, \infty$, respectively.

2.1

For a fixed d, and any $x \in \mathbb{R}^d$, what is the relationship between $\|x\|_1$, $\|x\|_2$ and $\|x\|_\infty$? Could you provide a chain of inequalities to relate them together? More specifically, find the largest $C > 0$ and the smallest $c > 0$ such that the following inequality holds:

$$\forall x \in \mathbb{R}^d,\ C \cdot \|x\|_1 \leq \|x\|_2 \leq c \cdot \|x\|_\infty.$$

3 Shannon Entropy

For $p \in \mathbb{N}$, let Δ_p be the $p-1$ dimensional probability simplex, i.e., $\Delta_p := \{x \in \mathbb{R}^p : \sum_{i\in[p]} x_i = 1, x_i \geq 0\}$. For any $x \in \Delta_p$, we define the Shannon entropy of x as $H(x) := -\sum_{i\in[p]} x_i \log x_i$ (Note: by default, we treat $0 \log 0$ as 0.)

3.1

Show that $\forall x \in \Delta_p, 0 \leq H(x) \leq \log p$.