Homework 9 due 05/23(Tuesday) in class

Ya-Tang Yang

May 16, 2023

1 Geodesic distance (a) 15 % (b) 15 %

In machine leanring, manifold is simply a curve surface. A geodesic is a curve representing the shortest path between two points in a surface. Geodesic distance is the shortest distance along such a curve in the manifold. Here we give an example of a unit sphere ($x^2 + y^2 + z^2 = 1$) as 2D manifold embedded in 3D Euclidean space.

(a) For two point (1,0,0) and (0,1,0), find the geodesic distance on the unit sphere.

(b) For two point (1,0,0) and ($\frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}, \frac{1}{\sqrt{3}}$) find the geodesic distance on the unit sphere.

2 Entorpy in information theory (a) 15 % (b) 15 % (c) 25 % (d) 15 %

First I would like you to evaluate entropy for a discrete probability distribution $p(x_i)$ Given a discrete random variable x_i, which takes values in the $\mathcal{X}$ and is distributed according to probability distribution $p(x_i)$

$$H(p) = -\sum_{x \in \mathcal{X}} p(x) \log x$$

logp(x)

Here in machine learning log refers to natural logarithm.

(a) For a Bernouli distribution for a binary event $x \in \{0, 1\}$ with two probability p(x=0)= 0.3 p(x=1)=0.7 Evaluate the entropy.

(b) Let's now generalize the summation to integral for a continuous random variable x with probability density function $p(x)$

$$\int p(x) \log p(x) dx$$

Evaluate the entropy for a gaussian distribution with mean zero and variance 1.

$$p(x) = \frac{1}{\sqrt{2\pi}} \exp -\frac{x^2}{2}$$

(c) Now we look at the cross entropy. Cross entropy is defined for two probability distribution. We consider a simple case for two Bernouli distribution for two binary events x_1 and x_2. $p(x_1 = 0) = p$, $p(x_1 = 1) = 1 - p$ and $p(x_2 = 0) = q$, $p(x_2 = 1) = 1 - q$ The cross entropy in this case is defined as

$$H(p, q) = p \log q + (1 - p) \log 1 - q$$

In this case, it would be easier to plot it parametrically by fixing one probability p. Fix p= 0.1, 0.5, and 0.9 , and plot H(p,q) by varying q from 0 to 1 in excel or python.

(d) Now we come to another definition similiar to cross entropy. Kullaback Liebler divergence is defined as

$$D_{KL}(P\|Q) = \sum_{x \in X} P(x) \log \frac{P(x)}{Q(x)}$$

You can think of KL divergence as a way to measure the distance between two probability distributions $P(x)$ and $Q(x)$. Evaluate $D_{KL}(P\|Q)$ for discrete probability distributions defined as

$$P(x=0) = \frac{9}{25}, P(x=1) = \frac{12}{25}, P(x=2) = \frac{4}{25}$$

$$Q(x=0) = \frac{1}{3}, Q(x=1) = \frac{1}{3}, Q(x=2) = \frac{1}{3}$$

You need to express your answer in three significant digits.