HW9 solution

1.In machine learning, 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^1=1) as 2D manifold embedded in 3D Euclid space.

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

(b) For two point (1,0,0) and $(1/\sqrt{3},\ 1/\sqrt{3}, (1/\sqrt{3})$

answer:

(a) the angle between (1,0,0) and (0,1,0) is pi/4

the geodesic distance is there 2*pi/4 = pi/2

(b) you can take inner product of (1,0,0) and $(1/\sqrt{3},\ 1/\sqrt{3}, (1/\sqrt{3})$

cos theta = 1/ $\sqrt{3}$

theta = 0.9553

radius =1 , geodesic distance = 0.9553

cross entropy

1. First I would like you to evaluate entropy for a discrete probability distribution p(xi)

Given a discrete random variable X, which takes values in the X and is distributed according to probability distribution p(x)

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

Here in machine learning log refers to natural logarithm.

(a) For a Bernouli distribution with two probability

p(x=0)= 0.3

p(x=1)=0.7

Evaluate the entropy.

Answer: simply take the log and multiply by p

H(p)= 0.610

(b) Now we want to generalize to a continuous random variable and hence p(x) is the probability density function and summation is replaced with integral.

$$H(p) = \int p(x) \log p(x)\, dx$$

Evaluate the entropy for a Gaussian distribution (for sigma =1, and mu=0)

$$p(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp\left[-\frac{(x-\mu)^2}{2\sigma^2}\right]$$

Answer:

$$H(p) = \int \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{(x)^2}{2}\right] \left[\left(-\frac{(x)^2}{2}\right) + \log\frac{1}{\sqrt{2\pi}}\right] dx = \left[-\frac{1}{2} - \log\left(\sqrt{2\pi}\right)\right]$$

(the correct definition is $H(p) = -\int p(x) \log p(x)\, dx$. From this definition,)

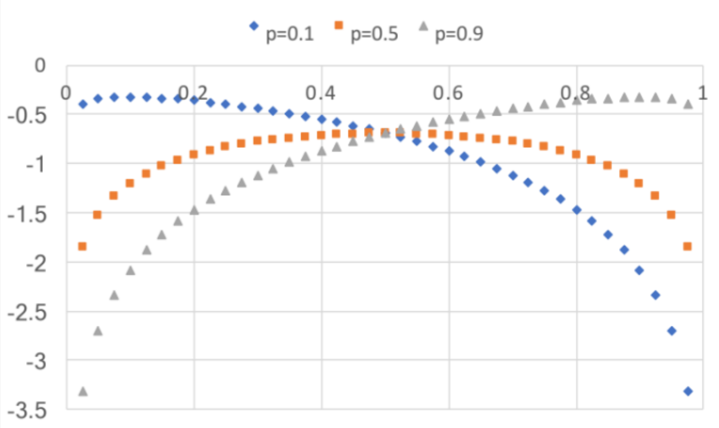
H(p)= $\frac{1}{2} + \log\left(\sqrt{2\pi}\right)$

(c) cross entropy plot

Answer: cross entropy is defined as

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

Fix p= 0.1, 0.5, and 0.9



You can see that p=0.1,0.5, and0.9 the cross entropy diverges at q=0 and q=1

(d) Kullaback Liebler divergence

$$\begin{aligned} D_{\mathrm{KL}}(P \parallel Q) &= \sum_{x \in \mathcal{X}} P(x) \ln\left(\frac{P(x)}{Q(x)}\right) \\ &= \frac{9}{25} \ln\left(\frac{9/25}{1/3}\right) + \frac{12}{25} \ln\left(\frac{12/25}{1/3}\right) + \frac{4}{25} \ln\left(\frac{4/25}{1/3}\right) \\ &= \frac{1}{25}\left(32\ln(2) + 55\ln(3) - 50\ln(5)\right) \approx 0.0852996 \end{aligned}$$