HW 10 (due 06/01/2023 in class)

1. (30%) We can define Kullaback Liebler divergence between two probability distribution P(x) and Q(x)

$$D_{KL} = -\int \mathrm{P(x)log}\frac{P(x)}{Q(x)}dx$$

Find formula for the Kullback Liebler divergence between two gaussian distribution P(x) and Q(x). Q(x) is a Gaussian distribution zero mean and standard deviation 1(for σ =1, and μ=0) and P(x) is a Gaussian distribution for mean μ and standard deviation σ

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

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

Problem 2 composition of sigmodal and cross entropy (20% each)

In the previous homework we define the cross entropy between two Bernouli distribution as

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

Now we would like you to plot the composition of such cross entropy and a sigmodal function

$$\sigma(z) = \frac{1}{\exp(-z)+1}$$

Define a new function g(z)= H(p=1, σ(z)) by setting p=1

(a)Plot g(z) for -5<= z <=0 with at least 30 data point

(b) Find the approximation formula when z approach -∞ for this function?

3. (10% each) (30%) Softmax calculation

In exploration and exploitation, it is possible to use the probability of taking action based on the following formula

$$p(a_i) = \exp(Q(a_i)/T) \,/\sum_i \exp\left(\frac{Q(a_i)}{T}\right)$$

The summation is sum over all possible actions, i.e, a_i

To gain some idea on this formula, we ask to compute the probability. The variable T is effective temperature. For simplicity, we consider the case of four possible actions a1, a2, a3, and a4. The corresponding Q values are given as

Q(a1) = 10 Q(a2)= 11 Q(a3)= 12 Q(a4)=13

For the probability distribution, p(a1), p(a2), p(a3), and p(a4) for various temperature

(a) low temperature, T=1

(b) intermediate temperature, T=10

(c) High temperature, T= 100

(Note: As an example, you may think of these actions as moving up, down, left and right in 2D grid word. Here, Q differs roughly by ~ 10 %.)