Deep Learning - HW2

Afek Adler & Or Wolkomir

January 19, 2020

1 Question 1 - distance measures between distributions

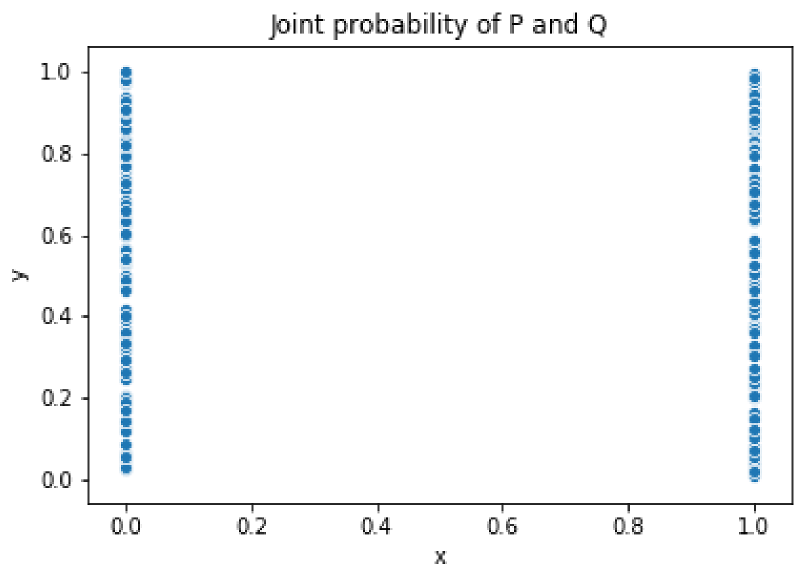

Figure 1: P on the left and Q on the right, without any loss of generality $\theta = 1$

(a) For $\theta \neq 0$

- The KL divergence is infinite since all X's in P are not in the support of Q.
- D_{js} is $log2$
- Wasserstein distance $= \theta$.

(b) For $\theta = 0$

- The KL divergence is 0.
- D_{js} is 0.

- Wasserstein distance is 0.

(c) The Wassertein distance is helpful where the previous methods fail (KL divergence and Jensen-Shannon Divergence). By producing a real number which is easy to work with and comparable to other solutions. Those more robust and 'better' for some optimization methods.

2 Question 2 - GAN 2d distribution visualization

Figure 2: P on the left and Q on the right, without any loss of generality $\theta = 1$

3 Question 3 - VAE

After doing grid search for best hyper-parameters (SVM) we discovered that the best kernel': 'rbf'. Convergence graphs and final results can be seen in the .ipynb notebook

4 Question 4 - DCCAN+WGAN

Convergence graphs and photo generation can be seen in the .ipynb notebooks