Deep Generative models

HW 1

Yosef Shimon Sella 318731205

* 1. Given vector with covariance matrix , and vector such that

.

The variance of is defined by:

The passes are made from linear algebra calculations and definitions of random and scalar vectors.

* 1. By definition a matrix is positive semi definite if for any vector the matrix

. In part a) we've shown that for vector that satisfies the condtion stated, the value is greater or equals 0, because the variance by definition is always greater or equals 0. Therefore the matrix is positive semi definite.

* 1. Lets see first :

Jensen's Inequality states:

For a convex function .

Since the function is concave, the function is convex for which Jensen's inequality holds.

Therefore we can write the following:

since we get

and overall:

Now lets focus on the private case, iff :

We can observe that the only way for the expression to be equal to 0 is iff , since

* 1. We can write the KL divergence as:

Therefore:

Which is the same as minimizing the negative log likelihood:

* 1. As defined previously the KL divergence is

In our case we will calculate the KL divergence when

The KL divergence in this case is:

Using logarithm rules we get:

We can see that for each we fix and therefore only contributes to that expression for every iteration and we get:

* 1. The KL divergence in the continuous case is:

And the probability density function of a gaussian is:

We can plug the probability density function into the KL divergence for two Gaussians

And obtain:

Lets simplify the term in the log:

Now lets revisit the KL divergence using the simplified log term:

Notice that the integral is the same as the expression for the expected value over

By definition:

Using mathematical manipulation we get:

Overall, we get: