Pattern Recognition and Machine Learning: Homework 13, Zhengzuo Liu

Problem 1

Answer:

The purpose of minimizing the second term D_{KL} is to regularise the latent space, that is, to enforce the distributions returned by the encoder to be close to a standard normal distribution.

Without this regularisation term, the model can learn, in order to minimize its reconstruction error (the first term), to behave like classic autoencoders (leading to overfitting). To do so, the encoder can either return distributions with tiny variances (that would tend to be punctual distributions) or return distributions with very different means (that would then be really far apart from each other in the latent space). In both ways, without explicit regularisation, some points of the latent space are "meaningless" once decoded.

Problem 2

Answer:

The code is completed in "GAN.py".