Pattern Recognition: Homework 13

Due date: 2023.5.30

Problem 1 (40 pt)

In our class, we have learned that the training objective of VAE is

$$\begin{aligned}\log p_\theta(x) \geq L_{\mathrm{ELBO}}(x, \theta, \phi) &= \mathbb{E}_{z \sim q_\phi(z|x)}\left[\log \frac{p_\theta(x, z)}{q_\phi(z \mid x)}\right] \\ &= \mathbb{E}_{z \sim q_\phi(z|x)}\left[\log p_\theta(x \mid z) + \log p(z) - \log q_\phi(z \mid x)\right] \\ &= \mathbb{E}_{z \sim q_\phi(z|x)}\left[\log p_\theta(x \mid z)\right] - D_{KL}\left(q_\phi(z \mid x) \| p(z)\right)\end{aligned}$$

The loss consists of two terms. The first term, "Reconstruction Loss", is easy to understand. Because we want to accurately reconstruct the input. The second term $D_{KL}\left(q_\phi(z \mid x) \| p(z)\right)$ is a little tricky. Please explain what will our encoder $q_\phi(\cdot)$ and decoder $p_\theta(\cdot)$ become if the training objective does not contain the second regularization term. What problems will it bring?

Problem 2 (60 pt)

We provide you with a very simple code base for GAN on MNIST dataset. Fill in the blank in `GAN.py` to let it run. We have provided you detailed instruction hints to finish, so please follow them carefully. You do not need to change any other code. After you finish the code, run `python main.py` and you should see generated images in `./results/mnist/GAN/`. Train you model and get satisfying digit images.