STAT3007/7007 Deep Learning, Tutorial 12

2022 Semester 2

1. (GAN) In GAN, the generator G constructs a fake example by first drawing z from a given noise distribution $p_Z(z)$, and then mapping z to the fake example $G(z)$. The discriminator D assigns to each (real or fake) example x a probability $D(x)$ that the example is real.

In the original formulation of GAN, the objective is to solve

$$\min_G \max_D L(D, G) = \frac{1}{2}\,\mathbb{E}_{x\sim p_{data}} \ln D(x) + \frac{1}{2}\,\mathbb{E}_{z\sim p_Z} \ln(1 - D(G(z))),$$

where p_{data} is the true data distribution.

In this question, we consider the following slightly modified objective function

$$\min_G \max_D L_\alpha(D, G) = \alpha\,\mathbb{E}_{x\sim p_{data}} \ln D(x) + (1-\alpha)\,\mathbb{E}_{z\sim p_Z} \ln(1 - D(G(z))),$$

where $\alpha \in (0, 1)$ is a constant. What is the optimal discriminator for a fixed G? What is the optimal G?

2. (Reinforcement learning)

(a) What is the difference between planning and reinforcement learning?

(b) What is the difference between on-policy and off-policy reinforcement learning algorithms?