GAN

Qingyun Li

July 20, 2018

1. Introduction

2. Related work

3. Adversarial nets

4. Theoretical results

The generator G implicitly defines a probability distribution p_g as the distribution of the samples G(z) obtained when z∼p(z). Therefore, we should like Algorithm 1 to converge a good estimator of p_{data}, if given enough capacity and training time. The result of this section are done in a nonparametric setting, e.g. we represent a model with infinite capacity by studying convergence in the space of probability density functions.

4.1. Global optimality of $p_g = p_{data}$

At first the authors consider the optimal discriminator D for any given generator G.

Proposition1. For G fixed, the optimal discriminator D is

$$D^*_G(x) = \frac{p_{data}(x)}{p_{data}(x) + p_g(x)} \tag{1}$$

The training criterion for the discriminator D, given any generator D, given any generator G, is to maximize the quantity V(G,D)

$$\begin{aligned} V(G,D) &= \int_x p_{data}(x) log(D(x)) dx \\ &+ \int_z p_z(z) log(1 - D(g(z))) dz \\ &= \int_x p_{data}(x) log(D(x)) \\ &+ p_g(x) log(1 - D(x)) dx \end{aligned} \tag{2}$$