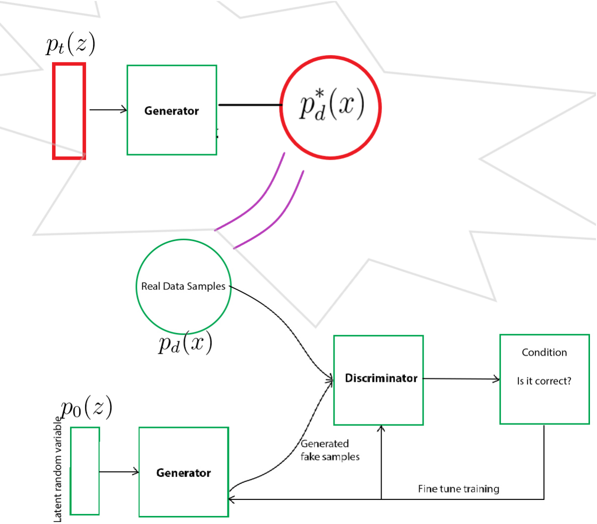
Your GAN is Secretly an Energy-based Model and You Should use Discriminator Driven Latent Sampling

Роберт Сафиуллин

Москва, 2022

$$V(G,D) = E_{p_d(x)}[logD(x)] + E_{p_0(z)}[log(1 - D(G(z))]$$

$$min_G max_D V(G,D)$$



Как? How?

Оптимальный дискриминатор

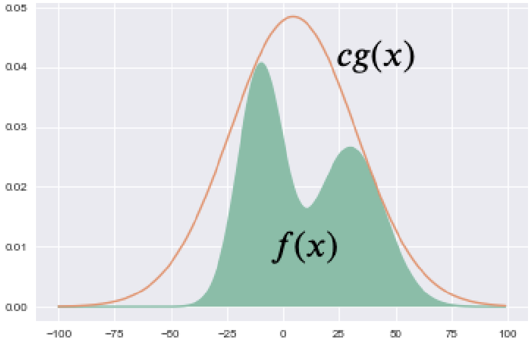
$$D^* = \frac{p_d(x)}{p_d(x)+p_g(x)} = \frac{1}{1+\frac{p_g(x)}{p_d(x)}}$$

$$D^* = \frac{1}{1+e^{-d(x)}} = \sigma(d(x))$$

$$d(x) = log\frac{p_d(x)}{p_g(x)}$$

$$p_d^*(x) = p_g(x)\frac{e^{d(x)}}{Z_0}$$

Rejection Sampling



$$x \sim g(x)$$
$$\text{С вероятностью } \frac{f(x)}{cg(x)} : x \sim f(x)$$

Теорема

- Взяв c $= max \frac{p_d^*(x)}{p_g(x)}$:
- $x \sim p_g(x), \frac{p_d^*(x)}{p_g(x)c} = \frac{e^{d(x)}}{Z_0 c} \to x \sim p_d^*(x)$
- $z \sim p_0, \frac{e^{d(G(z))}}{Z_0 c} \to G(z) \sim p_d^*$
- $z \sim p_0(z)e^{d(G(z))}/Z = p_t(z)$

Theorem 1. *Assume p_d is the data generating distribution, and p_g is the generator distribution induced by the generator $G : \mathcal{Z} \to \mathcal{X}$, where $\mathcal{Z}$ is the latent space with prior distribution $p_0(z)$. Define Boltzmann distribution $p_d^* = e^{\log p_g(x)+d(x)} / Z_0$, where Z_0 is the normalization constant.*

Assume p_g and p_d have the same support. We address the case when this assumption does not hold in Corollary 2. Further, let $D(x)$ be the discriminator, and $d(x)$ be the logit of D, namely $D(x) = \sigma(d(x))$. We define the energy function $E(z) = -\log p_0(z) - d(G(z))$, and its Boltzmann distribution $p_t(z) = e^{-E(z)} / Z$. Then we have:

1. *$p_d^* = p_d$ when D is the optimal discriminator.*
2. *If we sample $z \sim p_t$, and $x = G(z)$, then we have $x \sim p_d^*$. Namely, the induced probability measure $G \circ p_t = p_d^*$.*

Результаты

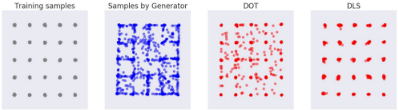


Figure 1: DDLS, generator alone, and generator + DOT, on 2d mixture of Gaussians distribution

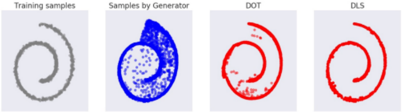


Figure 2: DDLS, the generator alone, and generator + DOT, on the swiss roll dataset.

Table 1: DDLS suffers less from mode dropping when modeling the 2d synthetic distribution in Figure 1. Table shows number of recovered modes, and fraction of “high quality” (less than four standard deviations from mode center) recovered modes.

	# recovered modes	% “high quality”	std “high quality”
Generator only	24.8 ± 0.2	70 ± 9	0.11 ± 0.01
DRS	24.8 ± 0.2	90 ± 2	$\mathbf{0.10 \pm 0.01}$
GAN w. DDLS	24.8 ± 0.2	$\mathbf{98 \pm 2}$	$\mathbf{0.10 \pm 0.01}$

Динамика

Figure 3: CIFAR-10 Langevin dynamics visualization, initialized at a sample from the generator (left column). The latent space Markov chain appears to mix quickly, as evidenced by the diverse samples generated by a short chain. Additionally, the visual quality of the samples improves over the course of sampling, providing evidence that DDLS improves sample quality.

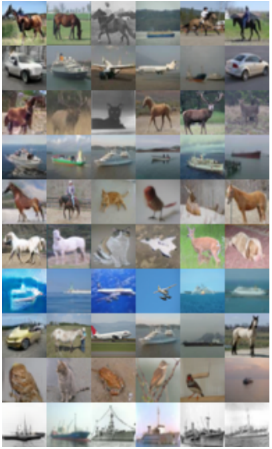


Figure 4: Top-5 nearest neighbor images (right columns) of generated CIFAR-10 samples (left column).

Использованная литература

Статья