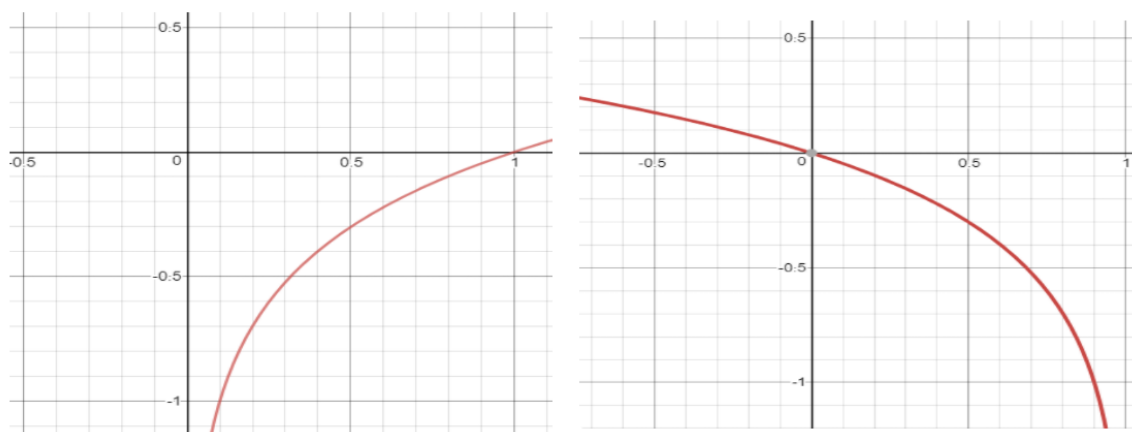


Non-Saturating GAN Loss

🕒 작성일시	@January 25, 2022 2:04 PM
☰ Keywords	DL GANs 스터디 그룹
👤 생성자	👤 김하연
☰ Note	
🕒 최종 편집	@February 25, 2022 8:24 PM
🔗 속성	
👤 최종편집자	👤 김하연

▼ $\log(1-x)$ 를 미분하면 $x=0$ 일 때, gradient가 1인데 왜 saturation?(상엽님 질문)

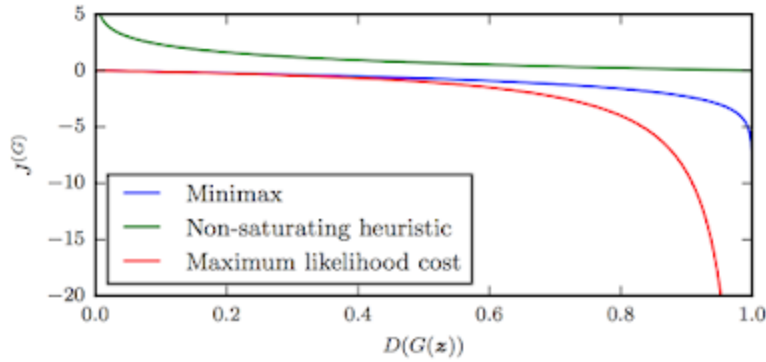


왼 : $\log(x)$, 오 : $\log(1-x)$

왼쪽은 $x=0$ 부근에서 무한대에 가까운 급격한 기울기를 가지고, 오른쪽은 0부근에서 1의 기울기를 가진다. 따라서 $D(G(z))$ 가 0 근처의 값에서 움직일 때, $\log(1-D(G(z)))$ 보다 $\log(D(G(z)))$ 의 기울기가 더 클 것입니다. 따라서, $\log(1-D(G(z)))$ 를 최소화 하기보다는, $\log(D(G(z)))$ 를 최대화 하는 것이 효율적입니다.

GAN Loss Function :

$$\min_G \max_D V(D, G) = E_{x \sim p_{data}(x)} [\log D(x)] + E_{z \sim p_z(z)} [\log(1 - D(G(z)))]$$



x축 : $D(G(z))$, y축 : $J(G) = \log(1-D(G(z)))$

위 그래프에서 $D(G(z))$ 가 0에 가까울 때, **파란색**의 그래프의 기울기는 거의 0이다 → **saturation**

따라서, (θ_G = G의 parameter)

$$\frac{\partial J(G)}{\partial \theta_G} = \frac{\partial J(G)}{\partial D(G(z))} \frac{\partial D(G(z))}{\partial (G(z))} \frac{\partial G(z)}{\partial \theta_G} \approx 0$$

즉, 여기서 알 수 있는 문제는 G가 학습 초기에는 학습을 거의 하지 못한다는 것 입니다. 이에 이 문제를 non-saturating 방식으로 해결할 수 있습니다.

Non-Saturating

$$\begin{aligned} \min_G V(D, G) &= E_z \sim p_z(z) [\log(1 - D(G(z)))] \\ \rightarrow \max_G V(D, G) &= E_z \sim p_z(z) [\log(D(G(z)))] \end{aligned}$$

다음과 같이 $\log(1-D(G(z)))$ 를 최소화 하는 것은 **$\log(D(G(z)))$ 를 최대화** 하는 것과 같은 말이므로, 이렇게 변경함으로써 saturation문제를 해결할 수 있습니다. 화살표 오른쪽 식의 그래프는 위 사진의 **초록색** 그래프와 동일한데, 학습 초기에 그래프의 기울기가 0이 아님을 확인 할 수 있습니다.

⇒ 따라서, non-saturating은 GAN의 generator의 loss를 min이 아닌 max관점으로 바꾸어 saturation을 해결할 수 있습니다.(초기에 학습을 잘 진행하고 어느 정도 진행된 후, 학습을 조금씩 진행하게 됩니다)

▼ 직관적으로 생각해보면 ?

G가 초기에는 이상한 이미지(노이즈가 많은 이미지)만 생성하기 때문에, D는 당연히 너무 쉽게 이를 구분하게 되면서 $\log(1-D(G(z)))$ 는 매우 saturate합니다 !

Reference

- [1] <http://sanghyeonna.blogspot.com/2018/08/generative-adversarial-networkgan-part3.html>
- [2] <https://dongsarchive.tistory.com/31>
- [3] 교수님 추천 : <https://danieltakeshi.github.io/2017/03/05/understanding-generative-adversarial-networks/>
- [4] **NIPS 2016 Tutorial: Generative Adversarial Networks** (Fig.16 부근)