

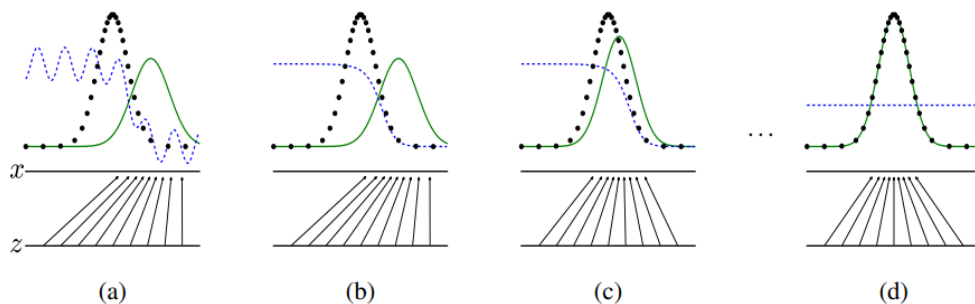
Paper Review : Generative Adversarial Nets (2014) by Goodfellow, et al.

- 현재까지의 딥러닝 기법과는 달리, 이 페이퍼에서 제안하는 adversarial nets framework는 생성모델의 적과의 경쟁관계를 이용하여 학습한다. 결과적으로 가장 진짜 같은 가짜를 생성해내는 것이 목표인 알고리즘이다.
- 비유: 생성모델은 더욱더 완벽한 위조지폐를 만드는 방향으로 학습하고 식별모델을 위조지폐를 더욱더 잘 구별하는 방향으로 학습한다.
- Loss function

$$\min_G \max_D V(D, G) = \mathbb{E}_{\mathbf{x} \sim p_{\text{data}}(\mathbf{x})} [\log D(\mathbf{x})] + \mathbb{E}_{\mathbf{z} \sim p_{\mathbf{z}}(\mathbf{z})} [\log(1 - D(G(\mathbf{z})))]$$

Discriminator (식별 모델)은 loss 함수가 커지도록 학습하며 Generator (생성 모델)은 loss 함수가 작아지도록 학습한다.

- Training process



파란색 선 : 식별 확률 분포

검정색 선 : 진짜 데이터의 분포

초록색 선 : 생성모델의 분포

(a)에서 (d)로 갈수록 (학습의 과정을 거치면서) 생성모델의 분포가 진짜 데이터의 분포에 가까워지며 식별 확률 분포가 안정적으로 1/2 확률에 수렴하는 것을 볼 수 있다.

- Algorithm 1

목표 : 좋은 estimator에 수렴 ($p_g = p_{\text{data}}$)

Possible Evaluation Methods for GANs

Universal gold-standard discriminator does not exist for GAN models, thus it is hard to evaluate them.

1. Inception Score (IS)

- 기준 : 생성된 이미지의 품질과 다양성
- 엔트로피 계산을 통해 식별하기 쉬운 이미지 및 식별된 레이블의 편차가 클수록 IS가 높게 출력 됨
- 엔트로피 : 예측하기 어려울수록 높음 (무작위성)

2. Fréchet Inception Distance (FID)

- 기준 : 생성된 이미지의 품질
- 집합 사이의 거리 (생성된 이미지와 정답 이미지)

$$FID(x, g) = \left\| \mu_x - \mu_g \right\|_2^2 + \text{Tr} \left(\Sigma_x + \Sigma_g - 2 \left(\Sigma_x \Sigma_g \right)^{\frac{1}{2}} \right)$$

- 활성화 분포 사이의 거리를 측정했을 때 유사한 이미지일수록 FID가 낮게 출력 됨

3. Multi-scale Structural Similarity (MS-SSIM)

- 기준 : 이미지 유사성
- 인간의 지각적 유사성 판단을 예측 -> 유사성을 정량적으로 평가
- 인간 인식에 중요하지 않은 이미지의 feature는 무시하는 기각 유사성 메트릭스의 다중 척도 변형
- 작은 loss는 발견하지 못함