

AI504 9강 정리

- Autoencoder
 - 입력 -encoding(압축)-> Code -decoding(압축 해제)-> 출력
- 새로운 샘플 만들기 (Generating New Samples)
 - 존재하는 z 를 교란
 - ◆ Code인 z 의 값에 약간의 변화를 주어 새로운 샘플 z' 를 만들
 - 구역에서 샘플을 추출
 - ◆ 하나의 종류로 분류된 것들이 모여있는 구역에서 샘플을 추출함
 - ◆ 하지만 구역이 명확한 구분으로 나뉜 것이 아니기에 추출하기 난감함
 - 두 방법 모두 정상적으로 작동할 것이라는 보장이 없음
- Variational Inference
 - 사후분포 (Posterior Distribution)
 - ◆ 데이터 x 와 관측되지 않은 변수 z
 - ◆ z 가 x 를 결정한다고 가정할 때 $P(z | x)$ 를 구하고 싶음
 - ◆ x 가 주어졌을 때 z 를 추론하는 것
 - ◆ Autoencoder에서 MNIST 이미지 x 와 그 Latent representation z
 - ◆ 일반적으로 사후확률은 다루기 힘들
 - 베이지 정리
 - $$P(z|x) = \frac{P(x|z)P(z)}{P(x)} = \frac{P(x|z)P(z)}{\int_z P(x,z) dz}$$
 - 대신 더 간단한 함수 Q 로 근사함 (주로 가우시안 사용)
 - $Q(z)$ 가 최대한 $P(z | x)$ 와 유사하길 원함
 - Kullback-Leibler Divergence(KL-Divergence) 최소화
 - $$D_{KL}(Q||P) \triangleq \sum_z Q(z) \log \frac{Q(z)}{P(z|x)}$$
 - ◆ 정리하면 $D_{KL}(Q||P) = E_z[\log Q(z) - \log P(z, x)] + \log P(x)$ (E는 기대값) -> ①
 - D_{KL} 은 대칭적이지 않음

- $D_{KL}(Q||P)$ -> Reverse KL
- ①에서 $\log P(X) = D_{KL}(Q||P) - E_Z[\log Q(Z) - \log P(Z, X)]$
 - ◆ 좌변은 고정되어 있음
 - ◆ $D_{KL}(Q||P)$ 을 최소화하려면 $-E_Z[\log Q(Z) - \log P(Z, X)]$ 를 최대화하면 됨
 - ◆ $-E_Z[\log Q(Z) - \log P(Z, X)]$ -> Evidence Lower Bound(ELBO) = H(Q)

- Variational Autoencoder

- 목표
 - ◆ $P(Z | X)$ 를 따르도록 x 를 z 로 압축
 - ◆ 다시 x 가 되도록 z 를 압축해제
- $P(X)$ 에서 $P(Z)$ 로 매핑 -> $P(Z)$ 에서 z 샘플 추출 가능, 추출한 z 를 x 로 변환
- VAE Loss
 - ◆ Reconstruction Loss: z 에서 x 를 reconstruction
 - ◆ Regularization Term: z 가 가우시안 분포를 따르도록 함
- VAE Posterior Distribution
 - ◆ 진짜 $P(Z | X)$ 를 구할 수는 없으므로 Variational Inference 사용
 - ◆ $Q(Z)$ 대신 $Q_\lambda(Z | X)$ 사용
 - x 에 의해 z 가 결정되기 때문
 - $Q_\lambda(Z | X)$ -> Encoder
 - λ 는 가우시안 Q 에 대한 (μ, σ) -> 인코더가 (μ, σ) 생성
 - ◆ $ELBO(\lambda) = E_q[\log p(x, z)] - E_q[\log q_\lambda(z|x)]$
 - 각 샘플 x_i 에 대해,

$$ELBO_i(\lambda) = E_{q_\lambda(z|x_i)}[\log p(x_i|z)] - KL(q_\lambda(z|x_i)||p(z))$$
- 학습과정
 - ◆ 랜덤 샘플 x_i 에 대한 예시
 - x_i 를 인코더에 넣고 (μ_i, σ_i) 를 구함

- $N(E_z[f(z)])$ 의 몬테카를로 예측)에서 z_i 샘플 추출
- z_i 를 디코더에 넣고 x_i' 구함
- squared error $\|x_i' - x_i\|^2$ 와 $D_{KL}(N(\mu_i, \sigma_i) || N(0, 1))$ 계산
- 역전파, 인코더와 디코더 업데이트

◆ $\|x_i' - x_i\|^2$ 의 계산

- z_i 와 연관된 식
- z_i 를 디코더의 입력으로 생각하면 디코더에서 역전파 가능
- 인코더 역전파에서 문제 발생
 - 랜덤하게 추출된 샘플인 z_i 가 문제를 발생시킴
 - 원래 z 는 $N(\mu, \sigma)$ 에서 랜덤으로 추출된 샘플
 - 대신 z 를 $\mu + \sigma * \epsilon$ 로 나타냄 (ϵ 는 $N(0, 1)$ 에서 랜덤으로 추출)

◆ 이렇게 나타내면 역전파에서 랜덤성이 없어짐

- β -VAE: 잠재 변수를 사용해 이미지를 특정 요소를 기준으로 나눔

◆ 이미지의 색상, 방향 등만 바꾸기 가능

- VAE를 사용하여 이미지 뿐 아니라 sequence, graph 등도 생성할 수 있음