



1. VAE의 개념



목표: VAE 개념에 대해서 논문과 다른 정보들을 통한 이해

VAE란?

- VAE는 **Generative Model**이다.
 - **Generative Model**:
 - Training data가 주어졌을 때 이 training data가 가지는 real 분포와 같은 분포에서 sampling 된 값으로 new data를 생성하는 model
- VAE는 input image X 를 잘 설명하는 feature를 추출하여 **Latent vector z** 에 담고, 이 Latent vector z 를 통해 X 와 유사하지만 완전히 새로운 데이터를 생성하는 것을 목표
- 이때 각 feature가 가우시안 분포를 따른다고 가정하고 latent z 는 각 feature의 평균과 분산값을 나타낸다.
 - **Q. 가우시안 분포를 따른다는 것은?**
 - 가우시안 분포는 널리 사용되는 확률 분포
 - 종 모양의 형태를 가지며 평균과 분산을 특징으로 함

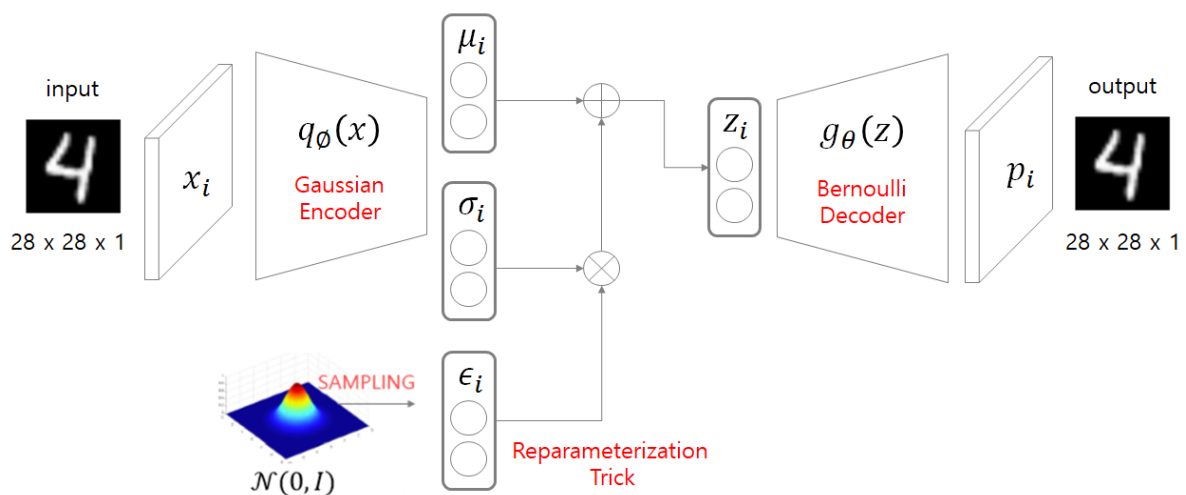
- 가우시안 분포는 평균을 중심으로 데이터가 모여있고, 평균에서 멀어질수록 데이터가 드물게 나타나는 형태를 가지고 있다.

$$z \rightarrow x$$

$$p_{\theta^*}(z) \quad p_{\theta^*}(x|z^{(i)})$$

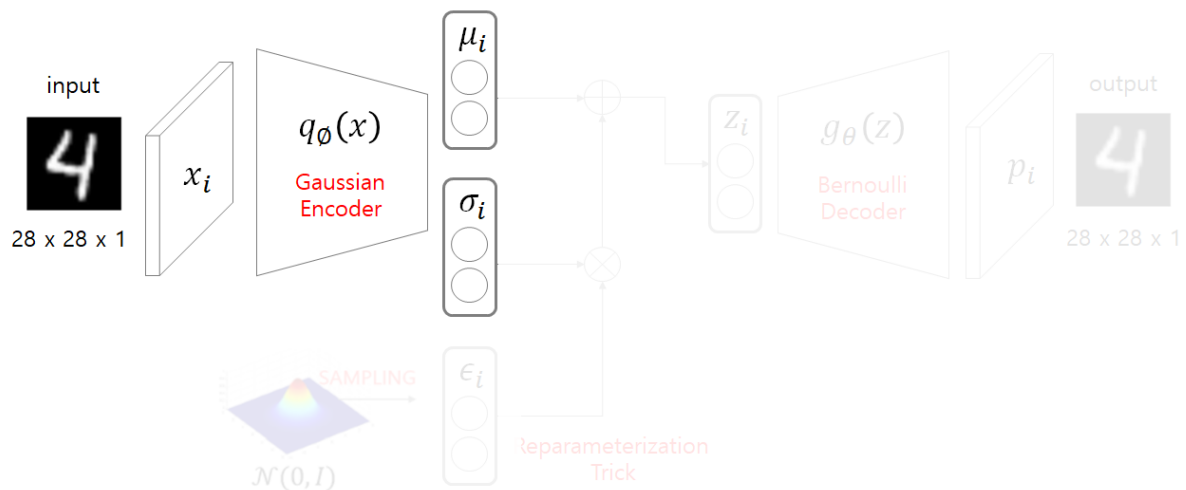
- $p(z)$: latent vector z 의 확률밀도함수. 가우시안 분포를 따른다고 가정
- $p(x|z)$: 주어진 z 에서 특정 x 가 나올 조건부 확률에 대한 확률밀도함수
- θ : 모델의 파라미터

VAE의 구조



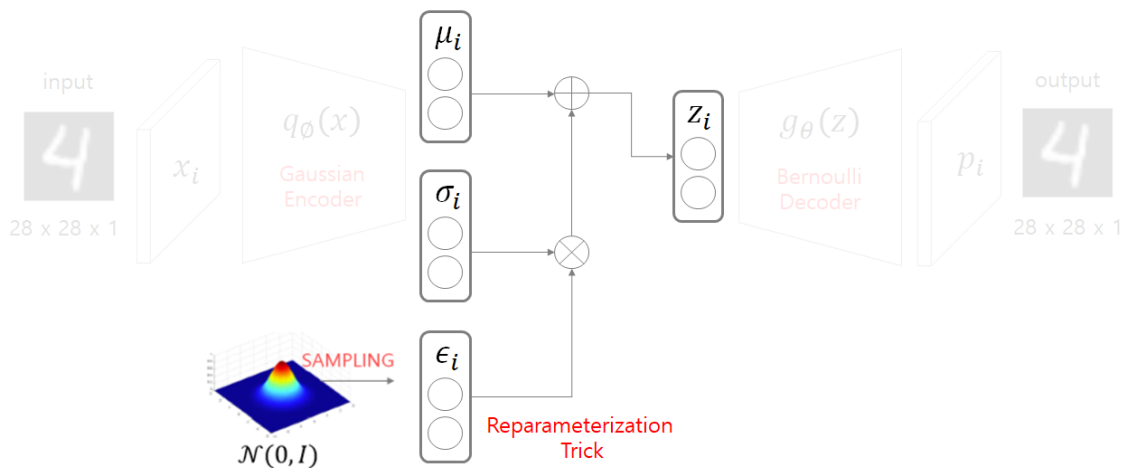
- Input data X 를 Encoder에 통과시켜 Latent vector z 를 구하고, Latent vector z 를 다시 Decoder에 통과시켜 기존 Input data X 와 비슷하지만 새로운 data X 를 찾아내는 구조
1. input: $x \rightarrow q_{\phi}(x) \rightarrow \mu_i, \sigma_i$
 2. $\mu_i, \sigma_i, \epsilon_i \rightarrow z_i$
 3. $z_i \rightarrow g_{\theta}(z_i) \rightarrow p_i$: output

1. Encoder : input: $x \rightarrow q_{\phi}(x) \rightarrow \mu_i, \sigma_i$



- X 라는 데이터의 Encoder Output 은 μ_i (평균), σ_i (표준편차) 이며 어떤 데이터의 특징을(latent variable) X를 통해 추측한다.
- 기본적으로 여기서 나온 특징들의 분포는 정규분포를 따른다고 가정.

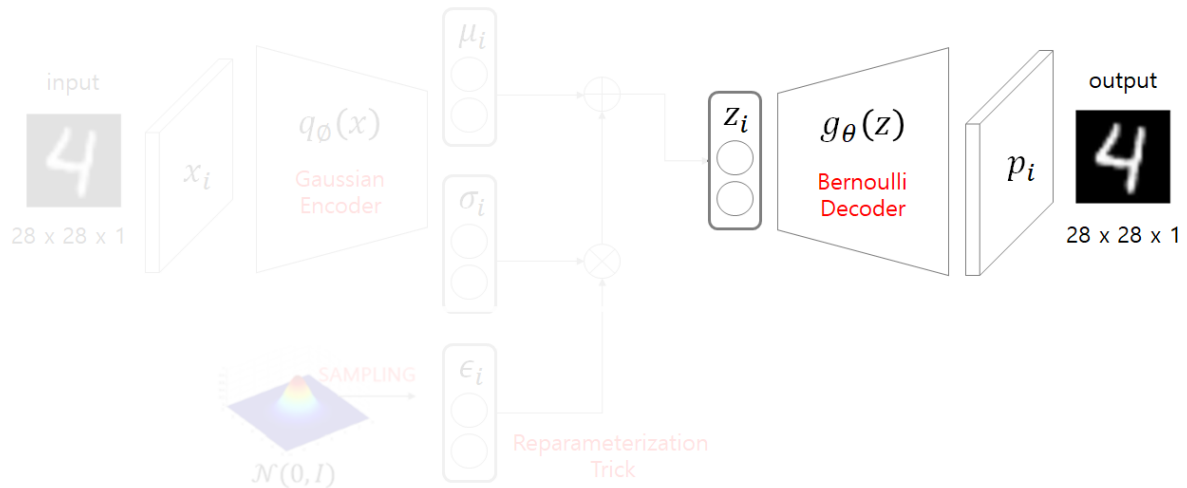
cf) Reparameterization Trick (Sampling)



- VAE에서 Latent vector z 를 **sampling** 하는 과정을 안정적으로 수행하기 위한 기법
 - 일반적으로 Encoder로부터 얻은 μ_i, σ_i 를 사용하여 Latent vector z 를 생성하는데 이 방법을 그대로 사용하면 Latent vector z 의 sampling 과정이 미분 불가능하여 gradient를 정확하게 전파할 수 없다는 문제가 발생!

- Latent vector z 를 sampling하기 위해 노이즈(ϵ)를 사용하고, 이 노이즈에 학습 가능한 파라미터인 μ_i 과 σ_i 를 곱하고 더함으로써 sampling을 수행

2. Decoder: $z_i \rightarrow g_\theta(z_i) \rightarrow p_i$: output



- Latent vector z 값을 임의로 샘플링하여 Decoder로 전달 새로운 Output을 생성한다.

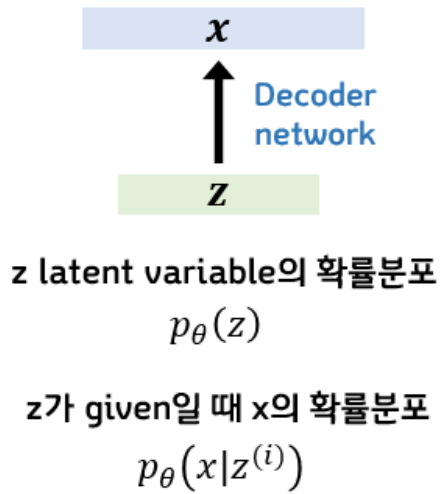
VAE의 수식적 증명

- 모델의 파라미터 θ 가 주어졌을 때 우리가 원하는 정답인 x 가 나올 확률이 $p_\theta(x)$ 높을 수록 좋은 모델이라고 할 수 있다.
 - \Rightarrow 즉 $p_\theta(x)$ 를 **최대화** 하는 방향으로 VAE의 파라미터 θ 를 학습



즉, 우리가 VAE를 통해 얻고자 하는 건 True parameter θ 이다!

Decoder



어떻게 학습?

네트워크의 출력값이 있을 때
우리가 원하는 정답 x가 나올 확률이 높길바람
= x의 likelihood를 최대화하는 확률분포 찾자



Maximize

$$p_{\theta}(x) = \int p_{\theta}(z)p_{\theta}(x|z)dz$$

- 여기서 알아야 하는 건 training data의 **Likelihood**를 최대화하는 모델의 **parameter**를 구해야하는 것이다. 이 과정에서 VAE는 **Intractable Density**를 구할 수 있으며, 최적화를 직접적으로 할 수 없기 때문에 **Lower Bound**로 유도하여 Likelihood를 구하게 된다.

▪ Intractability of VAE

Intractable to compute $p_\theta(x|z)$ for every z

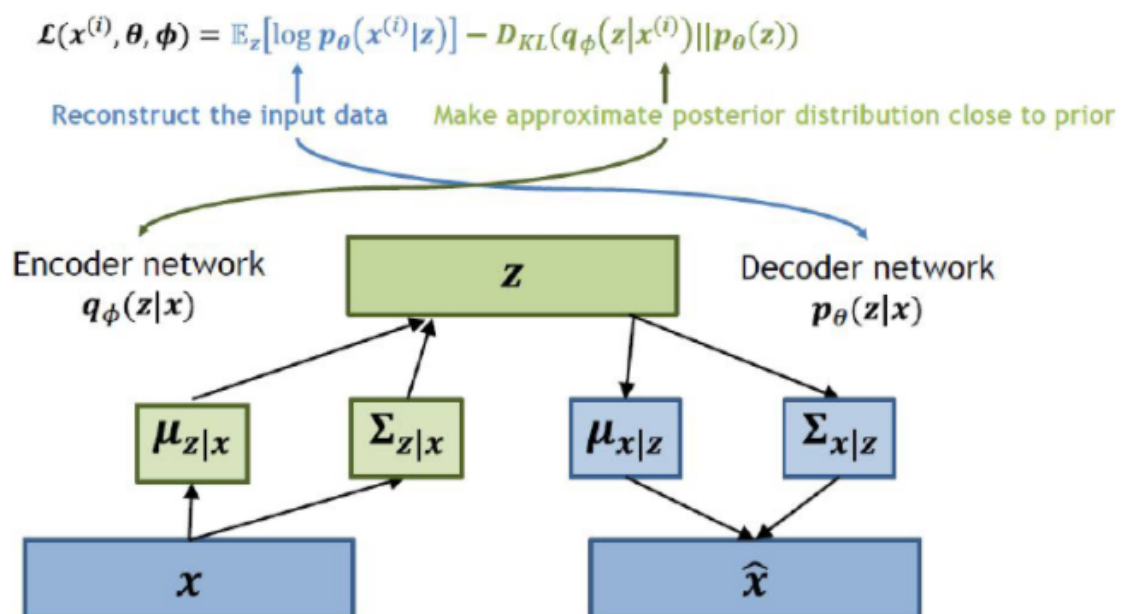
$$p_\theta(x) = \int p_\theta(z) p_\theta(x|z) dz$$

Simple Gaussian prior Decoder network

$$p_\theta(x) = \frac{p_\theta(x|z)p_\theta(z)}{p_\theta(z|x)}$$

Define additional encoder network $q_\phi(z|x)$ that approximates $p_\theta(z|x)$

This allows us to derive a lower bound on the data likelihood that is tractable, which we can optimize



References

- [Generative Models](#)
- [\[위키독스\] 한땀한땀 딥러닝 컴퓨터 비전 백과사전](#)
- [VARIATIONAL-AUTOENCODER와 ELBO\(EVIDENCE LOWER BOUND\)](#)
- [VAE : Variational Auto-Encoder를 직관적으로 이해하기!](#)

- KAIST 강남우 교수(전 숙명여대)의 딥러닝과 설계 강의