Auto-Encoding Variational Bayes

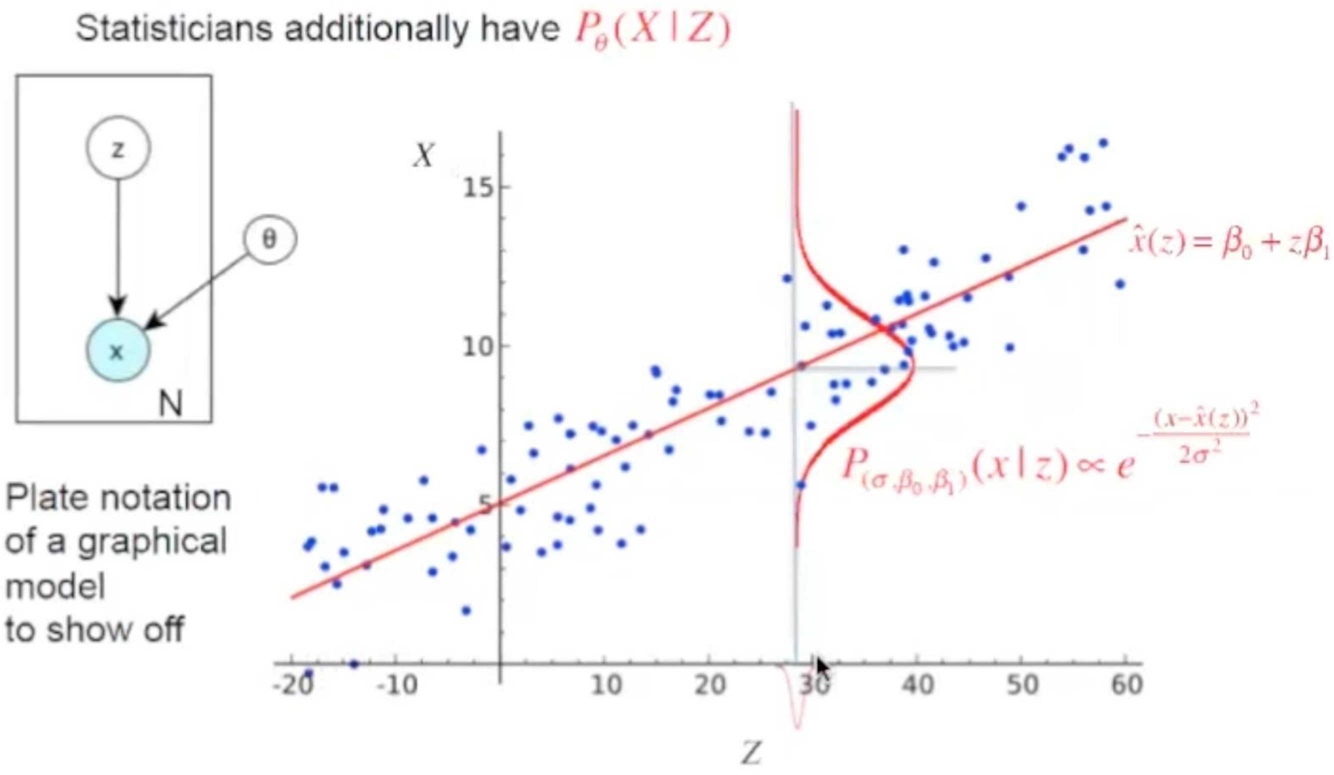
서상우, 유재창

Generative Model

- 이 논문에서 소개할 Variational Auto-Encoder는 일종의 Generative Model

- Generative Model은 입력 변수(*latent variable*) z로부터 결과물 x (가장 흔하게는 image)을 만들어내는 모델

Autoencoder?



Maximum likelihood

$$\arg\max_{\theta} \left[p_\theta(x) = \int_z p_\theta(x, z) = \int_z p_\theta(x|z) p_\theta(z) \right]$$

$$\int_z q(z|x) = 1$$

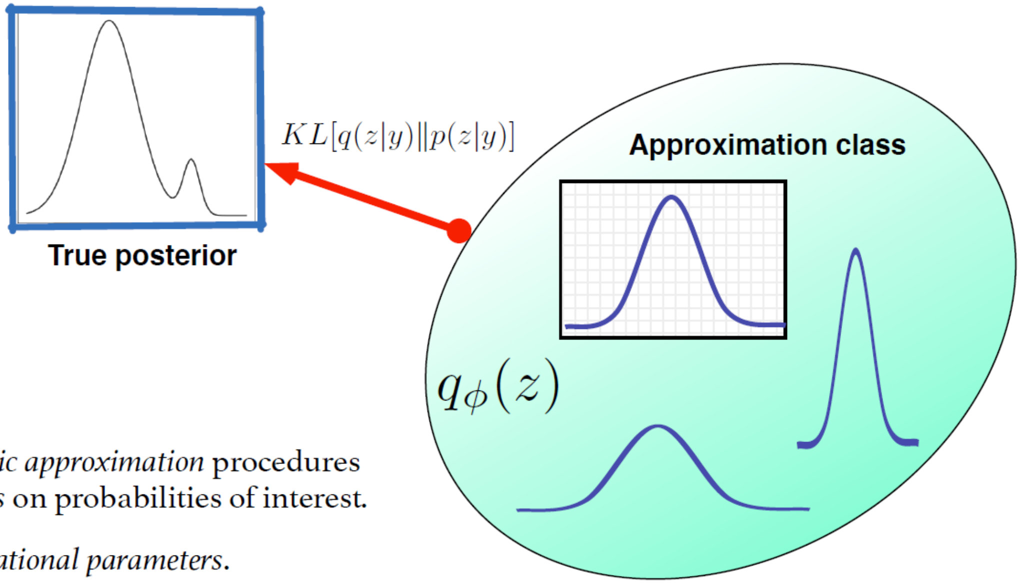
$$p(x) = \frac{p(z, x)}{p(z|x)}$$

$$\mathbb{E}[X] = \int x f(x)$$

$$D_{KL}(P||Q) = \int p(x) \log \frac{p(x)}{q(x)}$$

Variational Inference

- Variational inference의 기본 아이디어는 우리가 posterior inference를 어떻게 할 지 알고 있는 모델 $q_{ø}$를 가지고 inference를 하되 parameter ø를 잘 조정해서 P에 최대한 가깝게 만들자.



Deterministic *approximation* procedures with *bounds* on probabilities of interest.

Fit the *variational parameters*.

$$q_\phi(z|x) \approx p_\theta(z|x)$$

Kullback Leibler divergence

- 두 분포가 "가깝다"는 것을 어떻게 표현??
- 두 분포의 차이를 계산해주는 KL divergence!
- Z가 연속하면 적분으로 표현가능

$$KL(Q_\phi(Z|X)||P(Z|X)) = \sum_{z \in Z} q_\phi(z|x) \log \frac{q_\phi(z|x)}{p(z|x)}.$$

Variational Inference

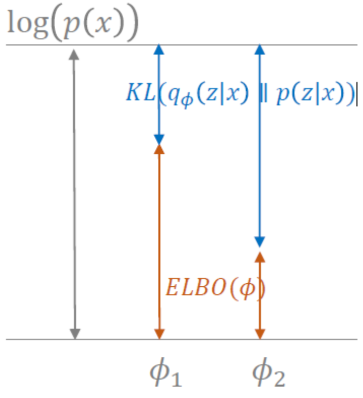
$$\log(p(x)) = \int \log(p(x)) q_\phi(z|x)dz \qquad \leftarrow \int q_\phi(z|x)dz = 1$$

$$= \int \log\left(\frac{p(x,z)}{p(z|x)}\right) q_\phi(z|x)dz \quad \leftarrow p(x) = \frac{p(x,z)}{p(z|x)}$$

$$= \int \log\left(\frac{p(x,z)}{q_\phi(z|x)} \cdot \frac{q_\phi(z|x)}{p(z|x)}\right) q_\phi(z|x)dz$$

$$= \underbrace{\int \log\left(\frac{p(x,z)}{q_\phi(z|x)}\right) q_\phi(z|x)dz}_{ELBO(\phi)} + \underbrace{\int \log\left(\frac{q_\phi(z|x)}{p(z|x)}\right) q_\phi(z|x)dz}_{KL\left(q_\phi(z|x) \parallel p(z|x)\right)}$$

두 확률분포 간의 거리≥0



KL을 최소화하는 $q_\phi(z|x)$의 ϕ값을 찾으면 되는데 $p(z|x)$를 모르기 때문에,
KL최소화 대신에 ELBO를 최대화하는ϕ값을 찾는다.

Variational Inference - ELBO

$$\log\big(p(x)\big) = ELBO(\phi) + KL\big(q_\phi(z|x)|p(z|x)\big)$$

$$q_{\phi^*}(z|x) = \underset{\phi}{\operatorname{argmax}}\, ELBO(\phi)$$

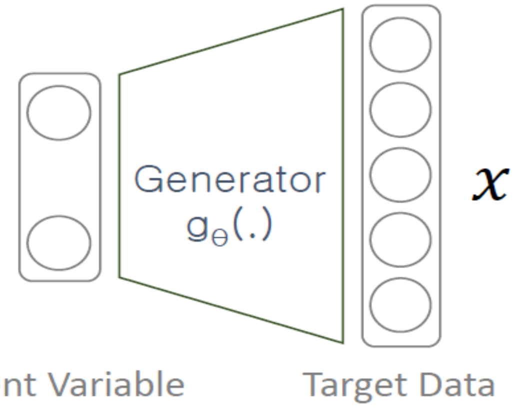
$$\begin{aligned} ELBO(\phi) &= \int \log\left(\frac{p(x,z)}{q_\phi(z|x)}\right) q_\phi(z|x)dz \\ &= \int \log\left(\frac{p(x|z)p(z)}{q_\phi(z|x)}\right) q_\phi(z|x)dz \\ &= \int \log\big(p(x|z)\big)q_\phi(z|x)dz - \int \log\left(\frac{q_\phi(z|x)}{p(z)}\right) q_\phi(z|x)dz \\ &= \mathbb{E}_{q_\phi(z|x)}\big[\log\big(p(x|z)\big)\big] - KL\big(q_\phi(z|x)||p(z)\big) \end{aligned}$$

앞 슬라이드에서의 KL과 인자가 다른 것에 유의

식을 위와 같이 변형하면
KL부분을 쉽게 구할 수 있다.

Variational Inference - loss

$$p(z|x) \approx q_\phi(z|x) \sim z$$



Optimization Problem 1 on ϕ: Variational Inference

$$\log\big(p(x)\big) \geq \mathbb{E}_{q_\phi(z|x)}\big[\log\big(p(x|z)\big)\big] - KL\big(q_\phi(z|x)\big|\big|p(z)\big) = ELBO(\phi)$$

Optimization Problem 2 on θ: Maximum likelihood

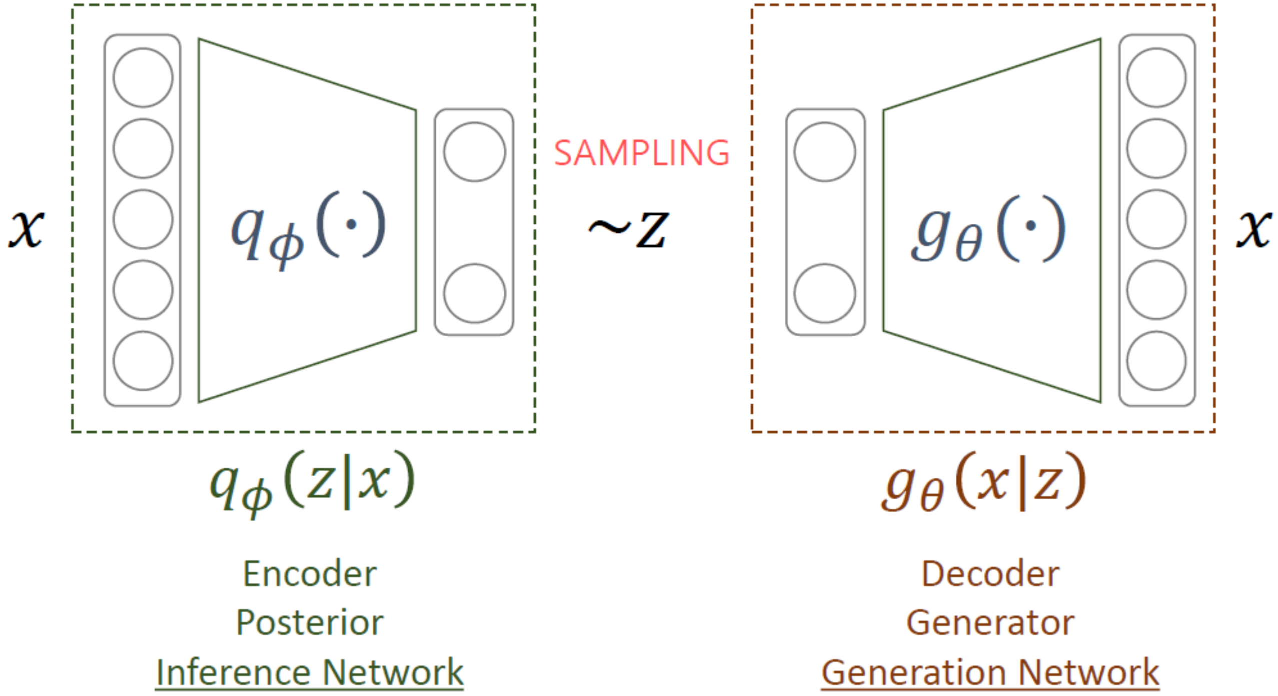
$$-\sum_i \log\big(p(x_i)\big) \leq -\sum_i \Big\{\mathbb{E}_{q_\phi(z|x_i)}\big[\log\big(p(x_i|g_\theta(z))\big)\big] - KL\big(q_\phi(z|x_i)\big|\big|p(z)\big)\Big\}$$

Final Optimization Problem

$$\arg\min_{\phi,\theta} \sum_i -\mathbb{E}_{q_\phi(z|x_i)}\big[\log\big(p(x_i|g_\theta(z))\big)\big] + KL\big(q_\phi(z|x_i)\big|\big|p(z)\big)$$

Variational Inference

$$\arg\min_{\phi,\theta} \sum_i \underbrace{-\mathbb{E}_{q_\phi(z|x_i)}\left[\log\left(p(x_i|g_\theta(z))\right)\right] + KL\left(q_\phi(z|x_i)||p(z)\right)}_{L_i(\phi,\theta,x_i)}$$



The mathematical basis of VAEs actually has relatively little to do with classical autoencoders

Variational Inference

$$\arg\min_{\phi,\theta} \sum_i \underbrace{-\mathbb{E}_{q_\phi(z|x_i)}\left[\log\left(p(x_i|g_\theta(z))\right)\right] + KL\left(q_\phi(z|x_i)||p(z)\right)}_{L_i(\phi,\theta,x_i)}$$

원 데이터에 대한 likelihood

Variational inference를 위한 approximation class 중 선택

다루기 쉬운 확률 분포 중 선택

$$L_i(\phi,\theta,x_i) = \underbrace{-\mathbb{E}_{q_\phi(z|x_i)}\left[\log\left(p(x_i|g_\theta(z))\right)\right]}_{\text{Reconstruction Error}} + \underbrace{KL\left(q_\phi(z|x_i)||p(z)\right)}_{\text{Regularization}}$$

Reconstruction Error

- 현재 샘플링 용 함수에 대한 negative log likelihood
- x_i에 대한 복원 오차 (AutoEncoder 관점)

Regularization

- 현재 샘플링 용 함수에 대한 추가 조건
- 샘플링의 용의성/생성 데이터에 대한 통제성을 위한 조건을 prior에 부여 하고 이와 유사해야 한다는 조건을 부여

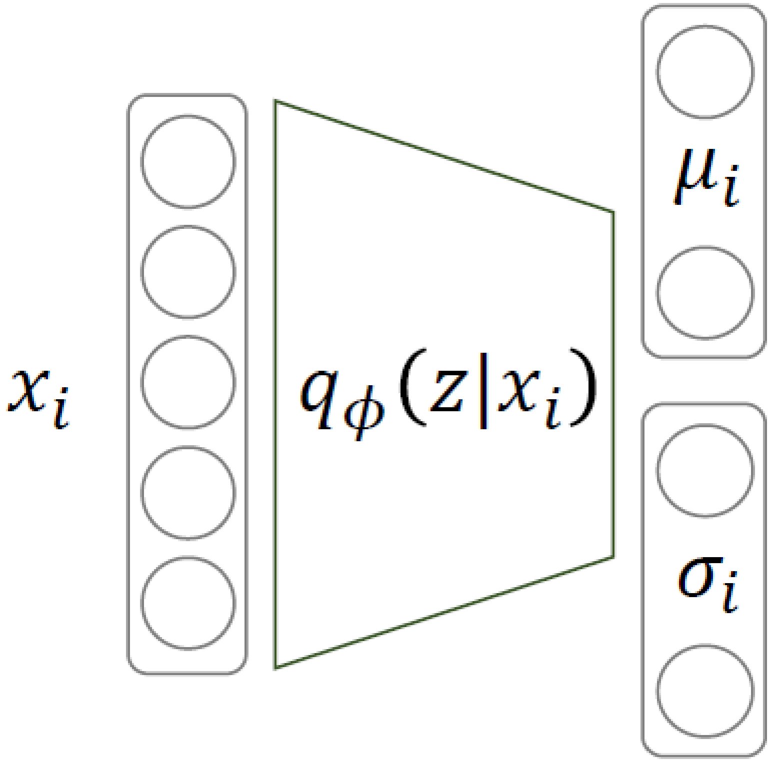
Regularization 부분계산

- Assumption 1
  - q_{ϕ}는 Gaussian distrinution

$$q_{\phi}(z|x_i) \sim N(\mu_i, \sigma_i^2 I)$$

- Assumption 2
  - P(z)는 normal distribution

$$p(z) \sim N(0, I)$$



Regularization 부분계산

$$KL\left(q_\phi(z|x_i)||p(z)\right) = \frac{1}{2}\left\{tr\left(\sigma_i^2 I\right) + \mu_i^T\mu_i - J + \ln\frac{1}{\prod_{j=1}^{J}\sigma_{i,j}^2}\right\}$$

$$= \frac{1}{2}\left\{\sum_{j=1}^{J}\sigma_{i,j}^2 + \sum_{j=1}^{J}\mu_{i,j}^2 - J - \sum_{j=1}^{J}\ln\left(\sigma_{i,j}^2\right)\right\}$$

$$= \frac{1}{2}\sum_{j=1}^{J}\left(\mu_{i,j}^2 + \sigma_{i,j}^2 - \ln\left(\sigma_{i,j}^2\right) - 1\right)$$ Easy to compute!!

Kullback–Leibler divergence [edit]

The Kullback–Leibler divergence from $\mathcal{N}_0(\boldsymbol{\mu}_0, \boldsymbol{\Sigma}_0)$ to $\mathcal{N}_1(\boldsymbol{\mu}_1, \boldsymbol{\Sigma}_1)$, for non-singular matrices Σ_0 and Σ_1, is:[8]

posterior prior

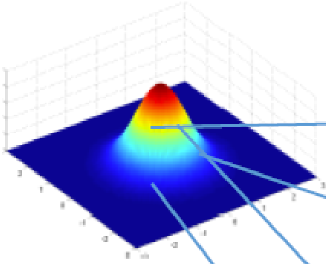
$$D_{\text{KL}}(\mathcal{N}_0\|\mathcal{N}_1) = \frac{1}{2}\left\{\text{tr}\left(\boldsymbol{\Sigma}_1^{-1}\boldsymbol{\Sigma}_0\right) + (\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0)^{\text{T}}\boldsymbol{\Sigma}_1^{-1}(\boldsymbol{\mu}_1 - \boldsymbol{\mu}_0) - k + \ln\frac{|\boldsymbol{\Sigma}_1|}{|\boldsymbol{\Sigma}_0|}\right\},$$

where k is the dimension of the vector space.

Reconstruction 부분계산

$$\mathbb{E}_{q_\phi(z|x_i)}[\log(p_\theta(x_i|z))] = \int \log(p_\theta(x_i|z))q_\phi(z|x_i)dz$$

Monte-carlo technique → $$\approx \frac{1}{L}\sum_{z^{i,l}} \log\left(p_\theta(x_i|z^{i,l})\right)$$



SAMPLING

$z^{i,1}$ $\log p_\theta(x_i|z^{i,1})$

$z^{i,2}$ $\log p_\theta(x_i|z^{i,2})$

$\vdots$

$z^{i,l}$ $\log p_\theta(x_i|z^{i,l})$

$\vdots$

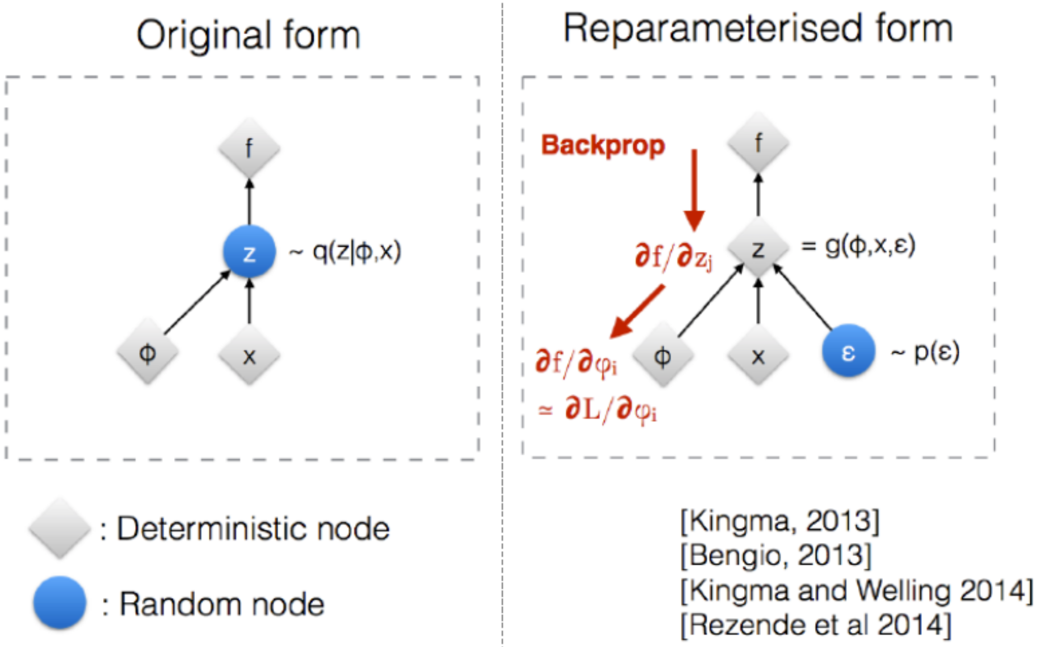
$z^{i,L}$ $\log p_\theta(x_i|z^{i,L})$

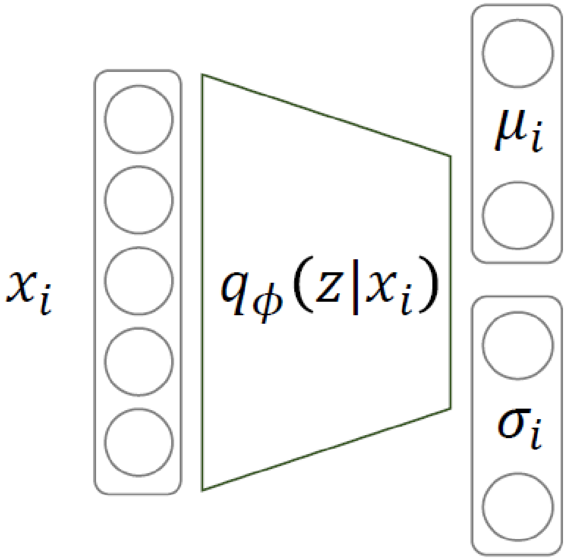
mean → **Reconstruction Error**

- L is the number of samples for latent vector
- Usually L is set to 1 for convenience

Reconstruction 부분계산

- 근데 Backpropagation이 불가능...





$$z^{i,l} \sim N(\mu_i, \sigma_i^2 I)$$

$$z^{i,l} = \mu_i + \sigma_i^2 \odot \epsilon$$
$$\epsilon \sim N(0, I)$$

Reconstruction 부분계산

$$\int \log(p_\theta(x_i|z))q_\phi(z|x_i)dz \approx \frac{1}{L}\sum_{z^{i,l}} \log\left(p_\theta\left(x_i|z^{i,l}\right)\right) \approx \log\left(p_\theta\left(x_i|z^i\right)\right)$$

Monte-carlo technique

L=1

Assumption 3-2

[Decoder, likelihood]

multivariate bernoulli or gaussain distribution

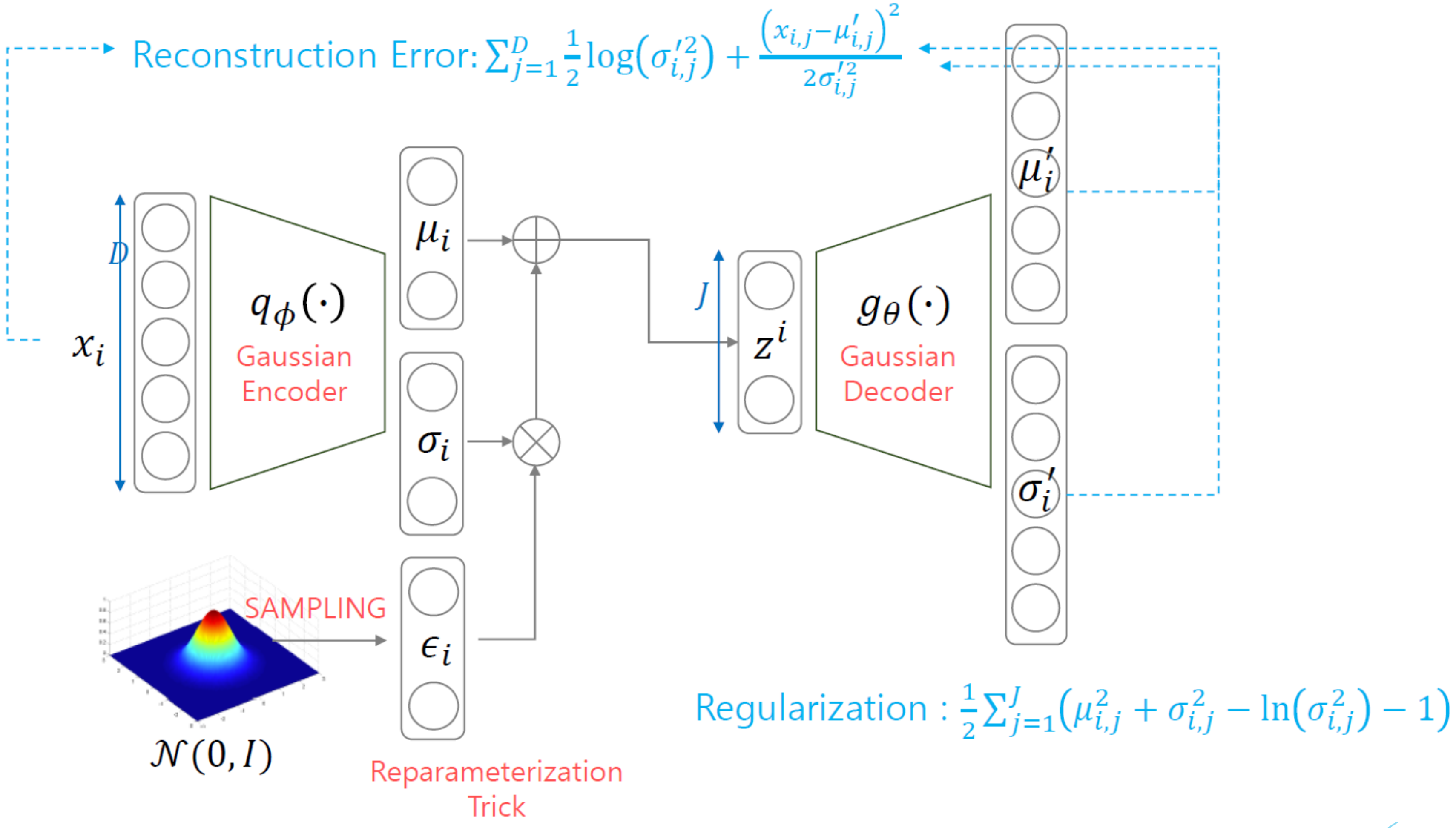
$$\log\left(p_\theta\left(x_i|z^i\right)\right) = \log(N(x_i;\mu_i,\sigma_i^2 I))$$

$$= -\sum_{j=1}^{D} \frac{1}{2}\log(\sigma_{i,j}^2) + \frac{(x_{i,j}-\mu_{i,j})^2}{2\sigma_{i,j}^2}$$

For gaussain distribution with identity covariance

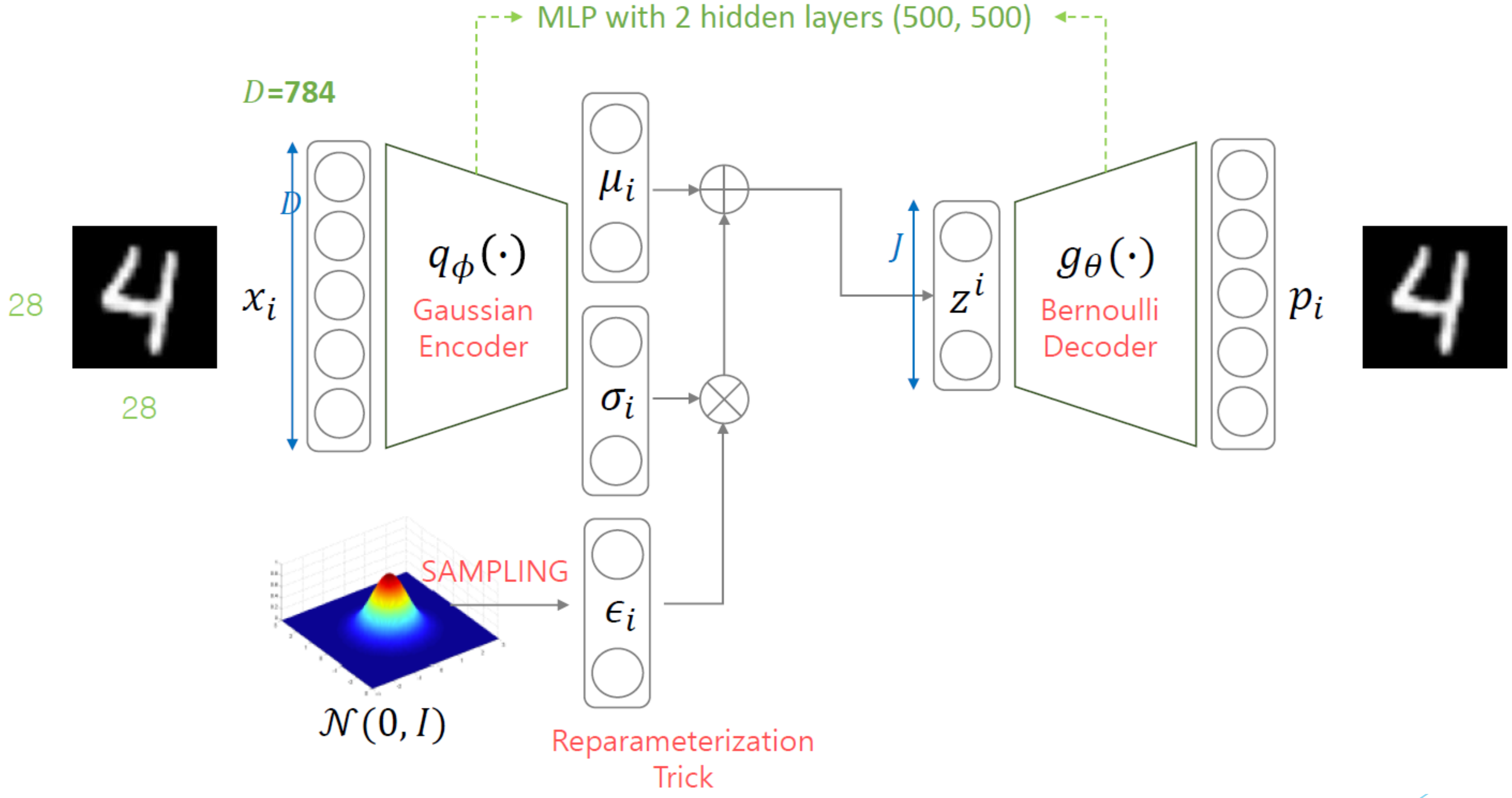
$$\log\left(p_\theta\left(x_i|z^i\right)\right) \propto -\sum_{j=1}^{D}(x_{i,j}-\mu_{i,j})^2$$ ← Squared Error

전체 구조



Reconstruction Error: $\sum_{j=1}^{D}\frac{1}{2}\log(\sigma'^2_{i,j}) + \frac{(x_{i,j}-\mu'_{i,j})^2}{2\sigma'^2_{i,j}}$
D
x_i
$q_\phi(\cdot)$
Gaussian
Encoder
μ_i
σ_i
J
z^i
$g_\theta(\cdot)$
Gaussian
Decoder
μ'_i
σ'_i
SAMPLING
ϵ_i
$\mathcal{N}(0, I)$
Reparameterization
Trick
Regularization : $\frac{1}{2}\sum_{j=1}^{J}(\mu^2_{i,j} + \sigma^2_{i,j} - \ln(\sigma^2_{i,j}) - 1)$

전체 구조



MLP with 2 hidden layers (500, 500)
D=784
D
28
28
x_i
$q_\phi(\cdot)$
Gaussian Encoder
μ_i
σ_i
ϵ_i
SAMPLING
$\mathcal{N}(0, I)$
Reparameterization Trick
J
z^i
$g_\theta(\cdot)$
Bernoulli Decoder
p_i

실험 결과