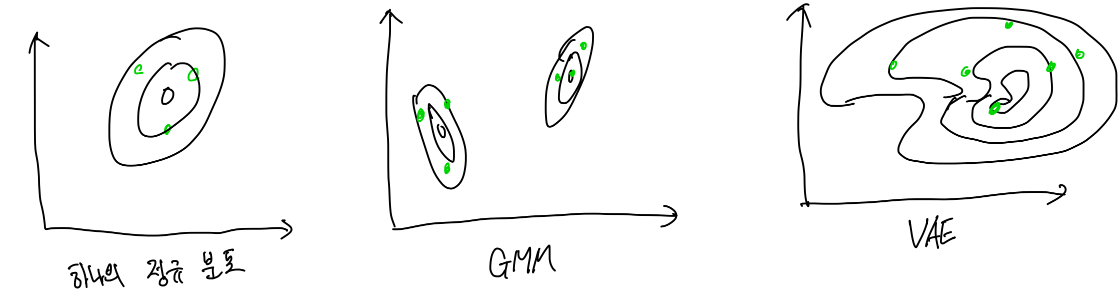
Ch 07. 변이형 오토인코더

1. VAE와 디코더
2. VAE의 인코더
3. ELBO 최적화
4. VAE 구현

7.1 VAE와 디코더



하나의 정규분포는?

매개변수 $= \theta = \{\mu, \Sigma\}$, 데이터 $= x$ ⇒ 확률 분포 $= p_\theta(x) = N(x; \theta)$ 이고?

N개 관측데이터 $D = \{x^{(1)}, x^{(2)}, \dots, x^{(N)}\} \xrightarrow{\text{로그 가능도}} \log p_\theta(D) = \log\left(p_\theta(x^{(1)})\, p_\theta(x^{(2)}) \dots p_\theta(x^{(N)})\right)$

GMM은?

1. 범주형 분포에 따라 K개의 정규분포 중 하나를 선택
2. 선택한 정규분포에서 데이터를 생성

⇒ ELBO를 목적함수로? ⟹ EM 알고리즘 수행?

이때? ELBO는

$$\log p_\theta(D) \ge \sum_{n=1}^{N} \sum_{z^{(n)}} q^{(n)}(z^{(n)}) \log \frac{p_\theta(x^{(n)}, z^{(n)})}{q^{(n)}(z^{(n)})}$$

VAE는?

VAE의 데이터 생성 과정

1. 잠재 변수 z를 고정된 정규 분포에서 생성?
2. 신경망을 통해 z를 x로 변환?

즉? → $\boxed{\text{고정된 정규분포}} \xrightarrow{\text{샘플링}} z \to \text{신경망} \to \hat{x}$

디코더(decoder)

↓ 따라서?

목표 : 관측 변수 x의 확률분포 $p(x)$를 얻기?

이를 위해 신경망 $p(x|z)$를 모델링?

$(z \to x)$

따라서!

출력 / 입력 / 신경망의 매개변수

$$\hat{x} = \text{NeuralNet}(z; \theta)$$

$$p_\theta(x|z) = N(x; \hat{x}, I)$$

⟹ x는 $\hat{x}$을 평균 벡터로 삼는 정규분포를 따른다?

VAE의 디코더의 수행 순서

$$\boxed{\mathcal{N}(z;0,I)} \xrightarrow{\text{샘플링}} z \xrightarrow{\text{입력}} \boxed{\text{디코더}} \xrightarrow{\text{출력}} \hat{x} \longrightarrow \boxed{\mathcal{N}(x;\hat{x},I)} \cdots\rightarrow \hat{x}$$

그럼 EM으로 할 수 있나? → NO!

VAE에서 E-스텝 → $q^{(n)}(z) = p_\theta(z|x^{(n)}) = \dfrac{p_\theta(x^{(n)}, z)}{p_\theta(x^{(n)})}$

$$= \frac{p_\theta(x^{(n)}, z)}{\int p_\theta(x^{(n)}, z)\, dz}$$

GMM : 이산 → Σ → 계산량 ↓

VAE : 연속 → $\int$ → 계산량 ↑

7.2 VAE와 인코더

EM 알고리즘은!

$$\log p_\theta(x) = \int q(z)\log\frac{p_\theta(x,z)}{q(z)}dz + \underbrace{D_{KL}(q(z)\,\|\,p_\theta(z|x))}_{\text{KL발산}}$$

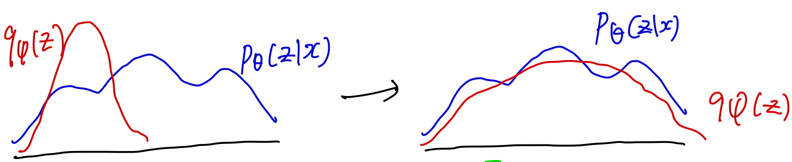
$$\ge \underbrace{\int q(z)\log\frac{p_\theta(x,z)}{q(z)}}_{\text{ELBO}}$$

EM 알고리즘
- GMM
  - → $q(z)$ 갱신 → $q(z) = p_\theta(z|x)$ (사후 분포 계산)
  - → θ 갱신 → ELBO 최대화
- VAE
  - → 사후 분포 계산 어려움 → $q(z)$를 간단한 확률분포로 제한! (→ 정규분포)
  - → 한정된 확률 분포 내에서 ELBO 최대화

즉! $q_\psi(z) = N(z; \mu, \Sigma)$ → 한정!

따라서 → $$\log p_\theta(x) = \underbrace{\int q_\psi(z)\log\frac{p_\theta(x,z)}{q_\psi(z)}dz}_{\text{ELBO!}} + D_{KL}(q_\psi(z)\,\|\,p_\theta(z|x))$$

ELBO 최대화 → $$\mathrm{ELBO}(x;\theta,\psi) = \int q_\psi(z)\log\frac{p_\theta(x,z)}{q_\psi(z)}dz$$



중요 $p_\theta(z|x)$와 $q_\psi(z)$가 멀리 않아!

즉! 근사!를 해야해!

이걸! 변분 근사 (variational approximation)
or
변분 베이즈 (variational bayes)
라고 해!

2절 전체 데이터에서는 ?!?!

$$\sum_{n=1}^{N} ELBO(x^{(n)}; \phi, \psi^{(n)}) = \sum_{n=1}^{N} \int q_{\psi}^{(n)}(z) \log \frac{p_\theta(x^{(n)}, z)}{q_{\psi^{(n)}}(z)} dz$$

$x^{(n)}$은 .. N개 ...

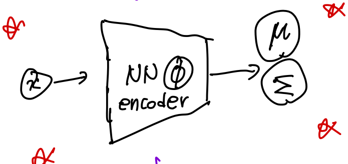
→ 근데 이걸 $q_{\psi}^{(n)}(z)$가 N개 필요..

너무 많아!

↓

그래서! 신경망!으로 대체!

↓



↓

즉! 근사 사후 분포의 매개변수(ψ)를 신경망으로 계산하는 방법은!!

분할 상환 추론!

(amortized inference)

즉!

$$x = \begin{pmatrix} x_1 \\ x_2 \\ \vdots \\ x_H \end{pmatrix}, \quad \mu = \begin{pmatrix} \mu_1 \\ \mu_2 \\ \vdots \\ \mu_H \end{pmatrix}, \quad \sigma = \begin{pmatrix} \sigma_1 \\ \sigma_2 \\ \vdots \\ \sigma_H \end{pmatrix}, \quad z = \begin{pmatrix} z_1 \\ z_2 \\ \vdots \\ z_H \end{pmatrix}$$

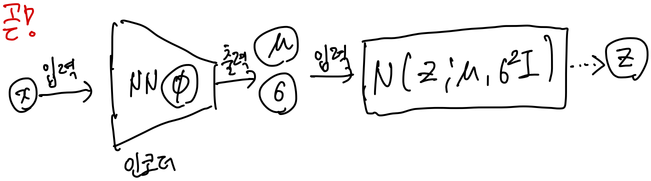
따라서! VAE 인코더가 하는 일은!

$$\mu, \sigma = \mathrm{NeuralNet}(x; \phi)$$

$$q_\phi(z|x) = N(z; \mu, \sigma^2 I)$$

Σ → 공분산 행렬

곧!



7.3 ELBO 최적화

VAE 디코더

$p(z) = N(z; 0, I)$

$\hat{x} = \text{NeuralNet}(z; \theta)$

$p_\theta(x|z) = N(x; \hat{x}, I)$

VAE 인코더

$\mu, \sigma = \text{NeuralNet}(x; \phi)$

$q_\phi(z|x) = N(z; x, \mu, \sigma^2 I)$

하나의 데이터에 대한 ELBO

$$\text{ELBO}(x; \theta, \phi) = \int q_\phi(z|x) \log \frac{p_\theta(x, z)}{q_\phi(z|x)} dz$$

여러 개의 데이터에 대한 ELBO

$$\sum_{n=1}^{N} \text{ELBO}(x^{(n)}; \theta, \phi) = \sum_{n=1}^{N} \int q_\phi(z|x^{(n)}) \log \frac{p_\theta(x^{(n)}, z)}{q_\phi(z|x^{(n)})} dz$$

그럼 ELBO 최대화 공식은 ?!?!

$$\begin{aligned}
\text{ELBO}(x; \theta, \phi) &= \int q_\phi(z|x) \log \frac{p_\theta(x, z)}{q_\phi(z|x)} dz \\
&= \int q_\phi(z|x) \log \frac{p_\theta(x|z)\, p(z)}{q_\phi(z|x)} dz \\
&= \int q_\phi(z|x) \log p_\theta(x|z)\, dz - \int q_\phi(z|x) \log \frac{q_\phi(z|x)}{p(z)} dz \\
&= \underbrace{E_{q_\phi(z|x)}\left[\log p_\theta(x|z)\right]}_{J_1} - \underbrace{D_{KL}\left(q_\phi(z|x) \,\|\, p(z)\right)}_{J_2}
\end{aligned}$$

J_1 = 기댓값
= 재구성 오차
(reconstruction error)
or
재구성 오류

J_2 = KL 발산
= 정규화
(regularization)
or
일관성 항

J_1 계산 → 몬테카를로!

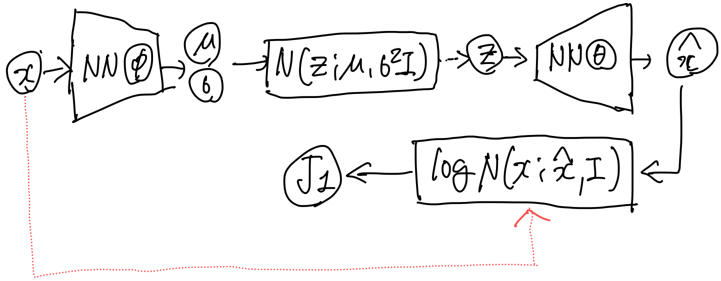
즉! $\mu, \sigma = \text{NeuralNet}(x; \phi)$

$z \sim N(z; \mu, \sigma^2 I)$

$\hat{x} = \text{NeuralNet}(z; \theta)$

$J_1 \approx \log N(x; \hat{x}, I)$

⇒



$$J_1 \approx \log N(x; \hat{x}, I)$$

$$= \log\left(\frac{1}{\sqrt{(2\pi)^D |I|}} \exp\left(-\frac{1}{2}(x-\hat{x})^T I^{-1} (x-\hat{x})\right)\right)$$

$$= -\frac{1}{2}(x-\hat{x})^T(x-\hat{x}) + \underbrace{\log\frac{1}{\sqrt{(2\pi)^D}}}_{\text{const: 상수}}$$

$$= -\frac{1}{2}\sum_{d=1}^{D}(x_d - \hat{x}_d)^2 + \text{const}$$

J_2 계산

$q_\phi(z|x) = N(z; \mu, \sigma^2 I)$

$p(z) = N(z; 0, I)$

⇒

$$J_2 = D_{KL}(q_\phi(z|x) \,\|\, p(z))$$

$$= -\frac{1}{2}\sum_{h=1}^{H}\left(1 + \log\sigma_h^2 - \mu_h^2 - \sigma_h^2\right)$$

즉! ELBO는!

$$\text{ELBO}(x; \theta, \phi) \approx -\frac{1}{2}\sum_{d=1}^{D}(x_d - \hat{x}_d)^2 + \frac{1}{2}\sum_{h=1}^{H}\left(1 + \log\sigma_h^2 - \mu_h^2 - \sigma_h^2\right) + \text{const}$$

드디어 ㅠㅠ

이때! $z \sim N(x; \mu, \sigma^2 I)$에서 z에서 미분이 곤란!

따라서, 재매개변수화 트릭 사용! ⇒ $\epsilon \sim N(\epsilon; 0, I)$ ⇒ $z = \mu + \sigma \odot \epsilon$

(reparameterization trick)

아다마르 곱

(Hadamar product)

7.4 VAE 구현

VAE 목표 = ELBO 최대화

$$ELBO(x;\theta,\phi) \approx -\frac{1}{2}\sum_{d=1}^{D}(x_d-\hat{x}_d)^2+\frac{1}{2}\sum_{h=1}^{H}(1+\log\sigma_h^2-\mu_h^2-\sigma_h^2)+const$$

이는 즉 손실 함수 최소화 따라서

$$Loss(x;\theta,\phi)\approx\sum_{d=1}^{D}(x_d-\hat{x}_d)^2-\sum_{h=1}^{H}(1+\log\sigma_h^2-\mu_h^2-\sigma_h^2)$$

즉 VAE의 학습 계산 그래프는