어떻게 하면 연속적인 latent variables와 다루기 힘든 posterior distributions, 그리고 거대한 데이터셋에서, directed probabilistic model을 효과적으로 추론하고 학습시킬 수 있을까?

우리는 stochastic variational inference와 학습 알고리즘을 소개할 것이다. 큰 데이터셋에 맞고, 약간 다른 조건 하에 있으며, 심지어 다루기 힘든 경우에서도 작동한다.

우리의 기여는 두가지로 구성되어 있다.

첫번째, 우리는 variational lower bound의 reparameterization을 소개한다. 이것이 standard stochastic gradient method를 사용한 직접적으로 optimized될 수 있는 lower bound estimator를 산출한다.

두번째, datapoint마다 연속적인 잠재 변수를 갖는 독립 항등 분포 (independent identically distributed)를 따르는 dataset을 소개할 것이다. 우리가 제안할 1) **lower bound estimator**를 사용해, 2) **approximate inference model**을 intractable한 posterior에 fit하도록 해 효과적으로 posterior 추론을 할 수 있습니다.

\* i.i.d -> 확률변수가 여러 개 있을 때 (X1, X2, …, Xn) 이들이 상호독립적이며, 모두 동일한 확률부포 f(x)를 가짐.

이론적 장점은 실험에 증명(반영)되었습니다.

**Introduction**

**<목적>**

어떻게 하면? Continuous latent variable과 parameter가 intractable posterior distribution을 갖는 directed probabilistic models로 효과적인 approximate inference와 learning을 수행할 수 있을까?

**<해답>**

**Variational Bayesian(VB) approach**는 intractable posterior에 대한 approximation의 optimization도 포함한다. 불운히, mean-field approach는 approximate posterior에 대한 **expectation의 분석적 해답**을 요구한다 (일반적으로 intractable하다).

우리는 “**variational lower bound**의 **reparameterization”**이 lower bound의 차별화되는 간단한 unbiased estimator를 내놓는지 보일 것이다.

이 **SGVB(Stochastic Gradient Variational Bayes)** estimator는 효과적인 approximate posterior inference에 사용될 수 있습니다. Continuous latent variable과 parameter가 있는 대부분의 모델에. 그리고 standard stochastic gradient ascent 기술을 사용해 opmtimize

I.i.d. dataset과 데이터포인트 마다의 continuous latent variables의 경우에 대해, 우리는 AEVB 알고리즘을 제안합니다.

**AEVB 알고리즘**에서, 우리는 SGVB estimator를 사용해 인식 모델을 최적화할 때 효과적인 inference와 learning을 배웁니다. **(SGVB를 사용함)**

우리가 ancestral sampling을 사용해 효과적인 approximate posterior inference를 수행할 수 있도록 합니다.

\*ancestral sampling: **미리 정의한 조건부 확률 분포로 구성된 ∏Ni=1pθ(xi−1|xi) ∏ i = 1 N p θ ( x i − 1 | x i ) 으로부터 이미지를 샘플링하는 방법**

이것은 우리가 데이터포인트마다 비용이 많이 드는 반복적인 inference scheme을 사용하지 않고도, 효과적으로 model parameter를 학습시키도록 합니다.

**<사용 목적>**

학습된 approximate posterior inference model은 recognition, denoising, representation 그리고 visualization 목적과 같은 task에 사용될 수 있습니다.

<variational auto-encoder>

신경망이 recognition model에 사용될 때, 우리는 variational auto-encoder를 사용하는 겁니다.

**METHOD**

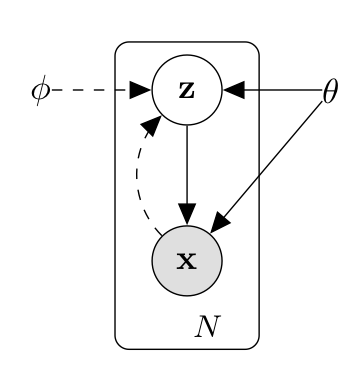
이 AEVB는 continuous latent variable을 갖는 다양한 directed graphical model을 위한 lower bound estimator에 사용될 수 있습니다. datapoint마다의 latent variable을 갖는 iid dataset을 갖고, global parameter와 latent variable에 대한 variational inference에 ML, MAP inference 사용할 수 있는 경우로 제한합니다.

ML -> prior가 균일하다고 가정. Likelihood가 최고인 점을 최적해 θ라 함

MAP -> prior가 균일하지 않아, 최적해 θ에 영향을 미침

<https://gaussian37.github.io/ml-concept-mle-and-map/>

예를 들어,



Solid line -> generative model **pθ(z)pθ(x|z)**

(**pθ(z)**는 **prior, pθ(x|z)**는 **likelihood(관측확률))**

Dashed line -> intractable posterior인 **pθ(z|x)**를 근사하는 variational approximation **qφ(z|x)**

(즉, **pθ(z|x)**를 알기 힘드므로 근사치인 **qφ(z|x)**을 구하겠다.**)**

**(왜? pθ(z)pθ(x|z)와 pθ(x)를 다 알지 못하기 때문에)**

**(Variational parameter φ는 generative model parameter인 θ와 같이 학습됩니다.)**

이 시나리오를 global parameter에 variational inference를 수행하는 개별 경우로 확장하기 위해, 이 알고리즘은 appendix ~~~ 논문의 목차에 대한 자세한 설명

**Problem Scenario**