

< VAE >

[ Abstract ]

본 논문에서는 **대규모의 데이터셋에서도 확장 가능하고, 미분 가능성 조건이 충족되는 경우에 사후 분포가 계산 불가능한 경우에도 작동하는 Stochastic Variational Inference** 및 학습 알고리즘을 제안한다.

첫 번째로, 변분 하한(ELBO)를 다시 파라미터화하여 표준 확률 경사 하강법으로 쉽게 최적화가 가능한 새로운 하한 추정기를 제안한다.

두 번째로, 각 데이터 포인트마다 연속 잠재 변수를 갖는 데이터셋의 경우 인식 모델이라고 불리는 근사 추론 모델을 학습하여 사후 분포를 빠르게 추정할 수 있게 한다.

[ Introduction ]

전통적인 방법인 Variational Bayesian은 사후 분포를 근사하여 최적화하는 방식이다.

그러나 일반적인 Mean-field 방식은 근사 분포에 대한 기댓값을 해석적으로 계산하는 과정이 필요한데, 대부분 계산이 불가능하다.

이에 본 논문은 **ELBO를 다시 파라미터화하여 단순하고, 미분 가능한 편향 없는 하한 추정기(SGVB)를 만들 수 있다**는 것을 보인다.

SGVB는 거의 대부분의 연속 잠재변수 모델에 적용 가능하고, 표준 확률 경사 하강법으로 쉽게 최적화할 수 있다.

특히, 각각의 데이터 포인트들마다 연속 잠재변수를 가지는 데이터셋의 경우 **Auto-Encoding VB(AEVB)**를 제안한다.

AEVB의 경우 **SGVB를 사용하여 인식 모델을 최적화**하고 이를 통해 단순한 샘플링 방식만으로 매우 효율적으로 근사 추론이 가능하며, MCMC와 같은 고비용의 반복 추론 과정 없이도 파라미터 학습이 가능하다.

학습된 근사 사후 추론 모델은 인식, 노이즈 제고, 표현 학습, 데이터 시각화 등의 작업에 활용 가능하다.

이때, 인식 모델로 신경망을 사용하는 경우를 VAE 구조라고 칭한다.