

23.07.20 _ 6주차

딥러닝 논문 요약 및 구현 스터디

발표자
CV

Variational Auto Encoder

Experiments

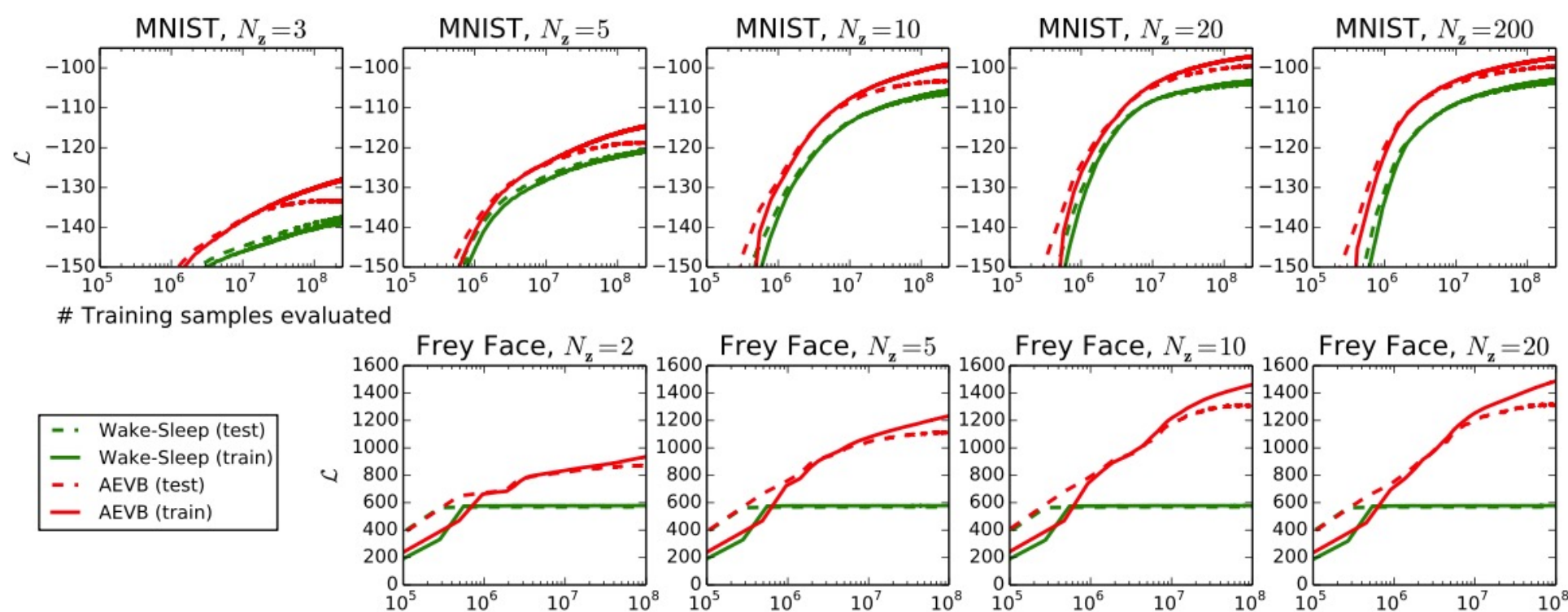


Figure 2: Comparison of our AEVB method to the wake-sleep algorithm, in terms of optimizing the lower bound, for **different dimensionality of latent space** (N_z). Our method converged considerably faster and reached a better solution in all experiments. Interestingly enough, more latent variables does not result in more overfitting, which is explained by the regularizing effect of the lower bound. Vertical axis: the estimated average variational lower bound per datapoint. The estimator variance was small (< 1) and omitted. Horizontal axis: amount of training points evaluated. Computation took around 20-40 minutes per million training samples with a Intel Xeon CPU running at an effective 40 GFLOPS.

ELBO

- AEVB가 더 빠르고 더 높은 값을 기록하는 것을 볼 수 있음.
- latent space의 크기를 달리 함에도 불구하고 일정하게 높은 성능을 관찰할 수 있음.
- overfitting도 없다고 하는데, lower bound의 정규화로 인한 것(왜인지는 모르겠습니다.)

Experiments

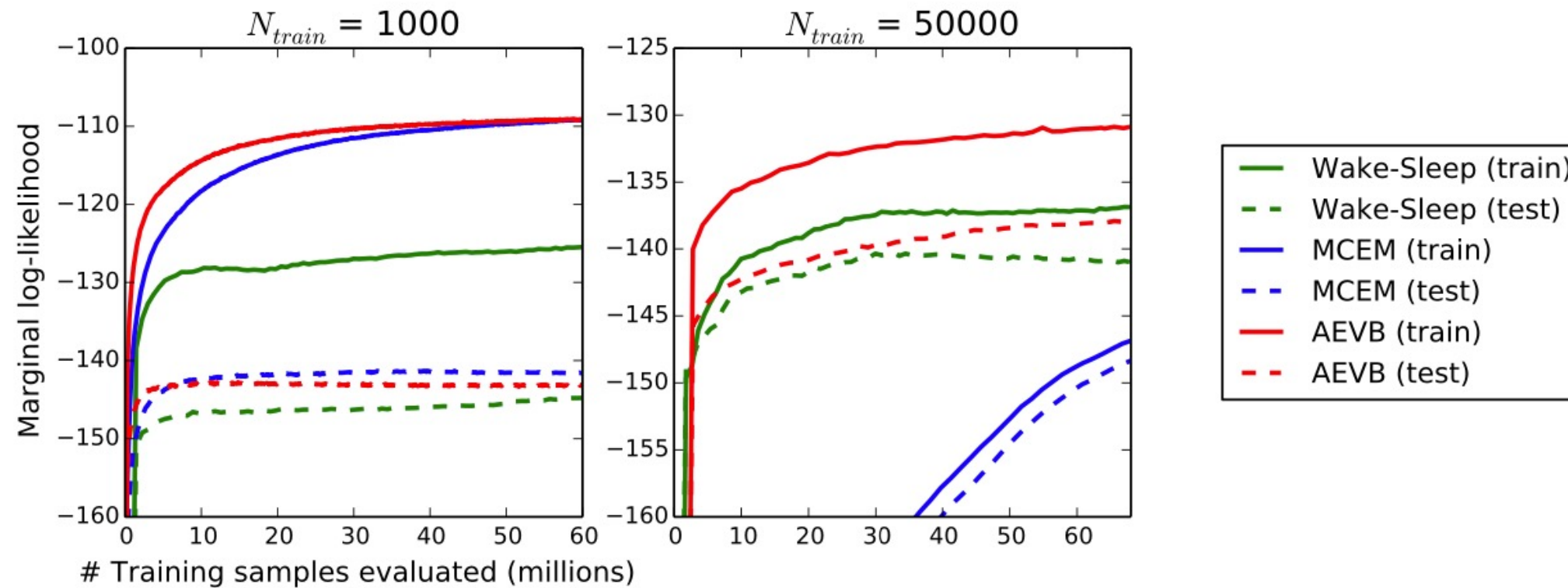


Figure 3: Comparison of AEVB to the wake-sleep algorithm and Monte Carlo EM, in terms of the estimated marginal likelihood, for a different number of training points. Monte Carlo EM is not an on-line algorithm, and (unlike AEVB and the wake-sleep method) can't be applied efficiently for the full MNIST dataset.

Marginal Log likelihood

- MCEM은 online algorithm이 아니기 때문에 전체 dataset이 아니기 때문에 효과가 떨어짐을 보임.

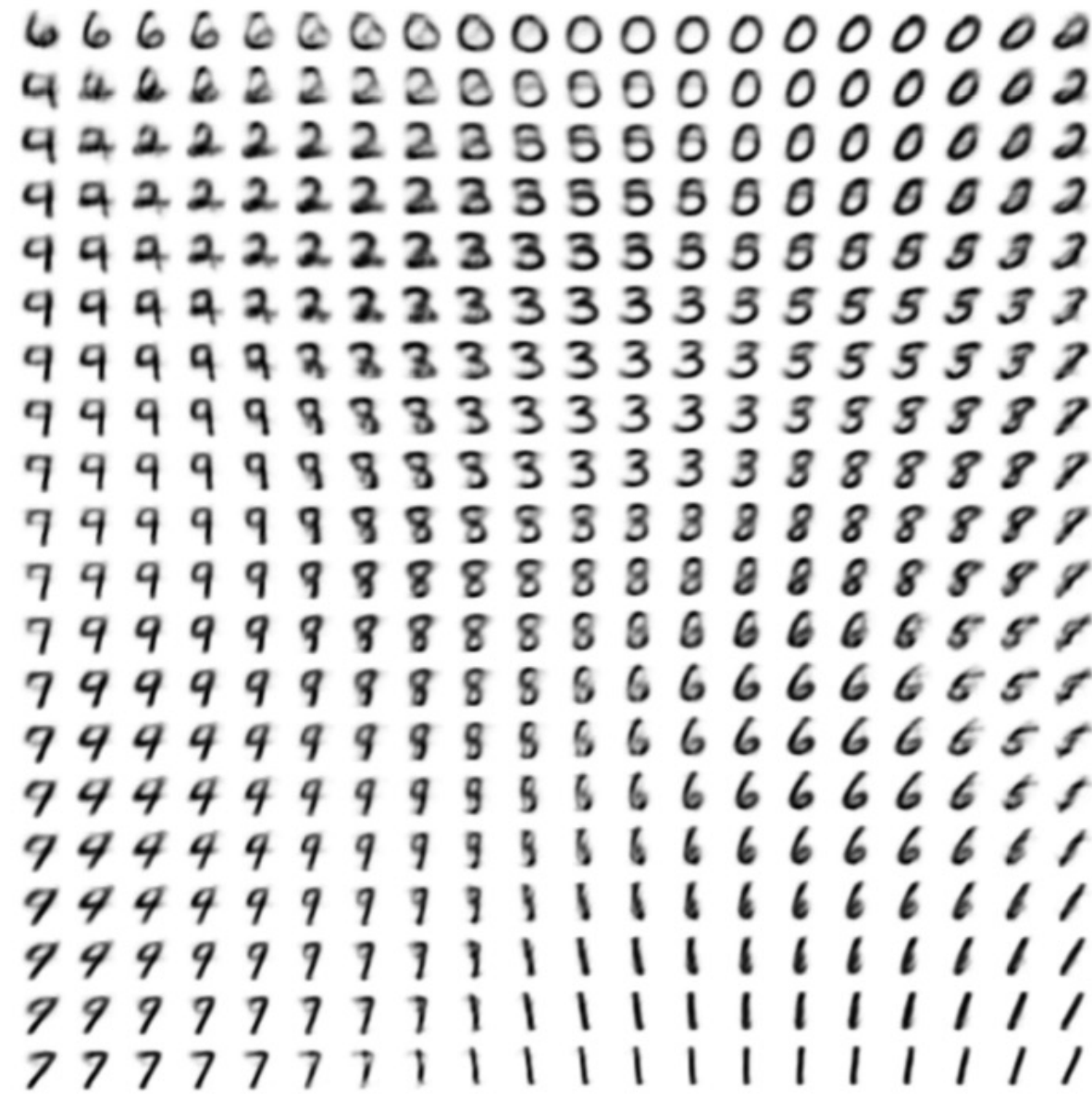
cf) 온라인 알고리즘: 시작할 때 모든 입력 정보를 가지고 있지 않고, 입력을 차례로 받아들이면서 처리하는 알고리즘.

오프라인 알고리즘: 풀고자 하는 문제의 모든 데이터를 가지고 시작해야만 문제를 해결할 수 있다.

Experiments



(a) Learned Frey Face manifold



(b) Learned MNIST manifold

Visualisation of high-dimensional data

- Frey Face와 MNIST dataset에 대해서, 시각화를 시도해 본 결과

Conclusion & Future work

Conclusion

- SGVB(Stochastic Gradient VB: standard stochastic gradient method를 사용해 간단하게 차별화된 최적화를 할 수 있다.
- AEVB(auto encoder VB): 독립 분포 데이터셋 및 데이터 포인트 당 연속된 잠재 변수의 경우 효율적인 추론 및 학습을 위한 알고리즘

Future work

- SGVB, AEVB는 연속 잠재 변수가 있는 거의 모든 추론과 학습 문제에 적용할 수 있으므로 향후 다양하게 활용될 수 있다.