

몬테 카를로 목표를 위한 분산 감소 방법

김건형, 장영수, 이종민, 김기웅

한국과학기술원

{ghkim, ysjang, jmlee}@ai.kaist.ac.kr, keeung.kim@kaist.edu

Variance Reduction Approach for Monte Carlo Objectives

Geon-Hyeong Kim, Youngsoo Jang, Jongmin Lee, Kee-Eung Kim
KAIST

요약

본 논문은 이산 잠재 변수를 사용하는 딥 생성 모델을 효율적으로 학습하기 위한 방법에 대한 연구이다. 본 논문에서는 다양한 학습 목표를 통해 딥 생성 모델을 학습할 때, 강화학습과의 대응을 제시한다. 이를 통해 기존 강화학습에서의 다양한 분산 감소 방법들을 적용할 수 있게 되어, 이를 통해 성능 향상을 꾀한다.

I. 서론

최근에는 잠재 변수 (latent variable)를 사용하는 딥 생성 모델 (deep generative model)을 학습하는 것을 통해 복잡한 데이터들을 생성하는 것에 대한 연구가 많다. 이는 주로 인공신경망 (neural network; NN)을 활용하여 구현되며, 고차원의 이미지 [1, 2]뿐만 아니라 음악이나 대화와 같은 소리 데이터 [3, 4], 그리고 영상 [5] 등을 생성하는 등 다양한 방향으로 연구가 진행 중에 있다. 이러한 생성 모델의 학습에 있어 가장 중요한 것은 어떤 목표 (objective)를 어떤 경사도 추정 (gradient estimation)을 사용해서 학습하느냐에 따라 성능이 크게 달라진다.

본 논문에서는 다양한 몬테 카를로 목표 (Monte Carlo objective; MCO) [4, 6]들을 사용해 시계열 구조를 갖는 딥 생성 모델과 강화학습 (reinforcement learning; RL)과의 대응을 제시한다. 그리고 이러한 대응을 통해 잠재 변수가 이산 (discrete) 값일 때에도 적용할 수 있는 방법론을 제시할 것이다. 이를 통해 기존 RL 에서 연구되었던 다양한 분산 감소 방법들을 적용할 수 있게 된다.

II. 본론

최근 변분 추론 (variational inference)을 사용하여 딥 생성 모델을 학습할 때에 주로 쓰이는 MCO 로는 데이터에 대한 한계 가능성도 (marginal likelihood)의 하한 (lower bound)인 ELBO (evidence lower bound) [1]나 IWAE (importance weighted autoencoder) [7]를 주로 사용하며, 이외에도 시계열 생성 모델의 경우 FIVO (filtering variational objectives) [4]등의 다양한 MCO 들이 사용된다.

이러한 MCO 들은 경사도 추정을 이용하는 확률 경사도 하강 (stochastic gradient descent; SGD)를 주로

사용하며, 이 때 재매개변수화 (reparameterization) [1, 8, 9]를 통한 방법이나 스코어 함수 추정 (score function estimator; SFE) [10, 11]이 일반적으로 널리 쓰인다. 재매개변수화는 매개변수화된 확률 분포 대신에 같은 매개변수로 매개변수화된 함수와 이와는 독립된 고정된 확률 분포로 나타내는 방법이다. SFE 를 사용하는 방법은 log derivative trick 이라고 알려진 방법을 통해 미분하는 방법이다. 재매개변수화는 일반적으로 경사도 추정 값의 분산이 작은 대신 연속된 잠재 변수가 필요한 등의 제약조건이 있고, SFE 의 경우 광범위하게 적용할 수 있지만 경사도 추정 값의 분산이 큰 단점이 있다.

기존에 [12]와 같이 ELBO 를 사용하는 생성 모델 학습과 RL 사이의 관계를 연결 지은 연구는 존재하지만, 다양한 MCO 에 대해 강화학습을 활용하는 연구는 부족하다. 이에, 본 논문에서는 IWAE 를 비롯한 다양한 MCO 와 RL 사이의 연결을 제시하고자 한다. 우선, 일반적인 MCO 는 관측된 데이터 x , 이산 잠재 변수 b , 매개변수 ϕ , b 에 대한 매개변수화된 분포 q_ϕ , 그리고 한계 가능성도에 대한 불편추정량 (unbiased estimator) f 를 통해 다음과 같이 정의할 수 있다:

$$\mathcal{L}_{\text{MCO}}(x) = \mathbb{E}_{b \sim q_\phi(b|x)} [\log f(b, x)]$$

또한 이 식의 ϕ 대한 경사도는 다음과 같이 얻어진다:

$$\mathbb{E}_{b \sim q_\phi(b|x)} [\nabla_\phi \log f(b, x) + \log f(b, x) \nabla_\phi \log q_\phi(b|x)]$$

이러한 식에서, 기대값 내부의 첫 번째 항은 두 번째 항과 비교해 분산이 작다고 알려져 있다. 따라서 두 번째 항에 대한 분산 감소 기법이 필요하다. 이 때, 우리는 시계열 모델을 다루기 때문에 이를 반영하기 위해 잠재 변수 b 를 총 시간 단계 (time step) T 에 대해 $b_{1:T}$ 와 같이 T 개의 변수로 분해할 수 있고, q_ϕ 와 f 또한 시간 단계에 대해 [4]과 유사하게

$$q_\phi(b|x) = \prod_{t=1}^T q_\phi(b_t | b_{1:t-1}, x),$$

$$f(\mathbf{b}, \mathbf{x}) = \prod_{t=1}^T f(\mathbf{b}_{1:t}, \mathbf{x})$$

와 같이 인수분해 될 수 있다는 가정을 할 수 있다. 이러한 가정과 함께 $\log q$ 의 경사도의 q 에 대한 기댓값이 0 이라는 사실을 활용하면 두 번째 항을 다음과 같이 다시 쓸 수 있다:

$$\sum_{t=1}^T \nabla_{\phi} \log q_{\phi}(\mathbf{b}_t | \mathbf{b}_{1:t-1}, \mathbf{x}) \sum_{t'=t}^T \log f(\mathbf{b}_{1:t'}, \mathbf{x})$$

이러한 식은 강화학습에서의 경사도 추정 값과 매우 유사하고, 이전 시간 단계에서의 모든 잠재 변수값을 상태, 현재 시간 단계에서의 잠재 변수값을 행동, 매개변수화된 분포를 행동정책, 현재 시간 단계에서의 불편추정량의 일부 요소가 보상으로 대응하면 강화학습의 각 요소와 대응이 된다. 이러한 대응 하에서, 두 번째 항의 경사도 추정 값에 다양한 강화학습에서의 방법들을 적용할 수 있고, 이를 첫 번째 항의 경사도 추정 값과 더해 원래 MCO 에 대한 경사도 추정 값을 구할 수 있다. 이를 통해 보다 낮은 분산을 가지면서도, 효율적인 딥 생성 모델 학습이 가능하다.

III. 결론

본 논문에서는 기존의 ELBO 를 활용하여 딥 생성 모델을 학습하는 문제와 RL 사이의 연결을 확장하여, 다양한 MCO 를 활용하여 딥 생성 모델을 학습하는 문제와 RL 사이의 연결을 제시하였다. 이를 통해 강화학습에서의 다양한 접근들을 활용할 수 있다. 하지만 이는 단순히 RL 에서의 기법을 적용한 것으로 딥 생성 모델만의 특성을 살리지는 못하였다. 이를 극복하기 위해 RL 과 딥 생성 모델과의 차이점에 유의하여 적절한 변형이 필요하다. 추후 연구로, 현재 강화학습과의 대응을 통한 분산 감소 기법과 함께 [6, 13]등에서 제시된 딥 생성 모델에 적용 가능한 다른 분산 감소 기법을 결합하는 등 단순한 적용을 확장하는 방안을 생각해 볼 수 있다. 이를 통해 보다 뛰어난 딥 생성 모델을 학습할 수 있을 것으로 기대된다.

ACKNOWLEDGMENT

본 연구는 과학기술정보통신부 및 정보통신기획평가원의 대학 ICT 연구센터지원사업의 연구결과로 수행되었음 (IITP-2019-2016-0-00464)

참 고 문 헌

[1] Diederik P Kingma and Max Welling. "Auto-encoding variational bayes". In *Second International Conference on Learning Representations*, 2014.

[2] Rezende, Danilo, and Shakir Mohamed. "Variational Inference with Normalizing Flows." *International Conference on Machine Learning*. 2015.

[3] Chung, Junyoung, et al. "A recurrent latent variable model for sequential data." *Advances in neural information processing systems*. 2015.

[4] Maddison, Chris J., et al. "Filtering variational objectives." *Advances in Neural Information Processing Systems*. 2017.

[5] Denton, Emily, and Rob Fergus. "Stochastic Video Generation with a Learned Prior." *International Conference on Machine Learning*. 2018.

[6] Mnih, Andriy, and Danilo Rezende. "Variational Inference for Monte Carlo Objectives." *International Conference on Machine Learning*. 2016.

[7] Yuri Burda, Roger Grosse, and Ruslan Salakhutdinov. Importance weighted autoencoders. *ICLR*, 2016.

[8] Titsias, Michalis, and Miguel Lázaro-Gredilla. "Doubly stochastic variational Bayes for non-conjugate inference." *International conference on machine learning*. 2014.

[9] Rezende, Danilo Jimenez, Shakir Mohamed, and Daan Wierstra. "Stochastic Backpropagation and Approximate Inference in Deep Generative Models." *International Conference on Machine Learning*. 2014.

[10] Williams, Ronald J. "Simple statistical gradient-following algorithms for connectionist reinforcement learning." *Machine learning* 8.3-4 (1992): 229-256.

[11] Fu, Michael C. "Gradient estimation." *Handbooks in operations research and management science* 13 (2006): 575-616.

[12] Weber, Theophane, et al. "Reinforced variational inference." *Advances in Neural Information Processing Systems (NIPS) Workshops*. 2015.

[13] Jang, Eric, Shixiang Gu, and Ben Poole. "Categorical Reparameterization with Gumbel-Softmax." *ICLR*. 2017.