Gradient Estimation Notes

June 30, 2020

1 Introduction

Gradient estimation

[3] Mohamed et al. [3]

2 Comparing Estimators

Owen [4]

3 Learning

3.1 Bias and Variance

4 Estimators

4.1 Relaxations

4.2 Score-function Estimator

5 Variance Reduction

[4]

5.1 Control Variates

5.2 Antithetic Sampling

5.3 Conditioning

Exponential Races, Poisson Processes, and Gumbel Perturbations

1 Gumbel Basics

Maddison [2]

Papers

1 RELAX

In this section we review Grathwohl et al. [1].

Motivation The aim of this paper is to find argmin $\theta \mathbb{E}_{p(b|\theta)}[f(b)]$ via gradient descent using a Monte Carlo approximation of the gradient. The focus is finding a low-variance unbiased estimator by combining the score function estimator with the reparameterization trick.

The two applications considered in this paper are gradient estimation for latent variable models and RL.

Contributions Grathwohl et al. [1] introduce a differentiable surrogate for f, and for discrete b a reparameterizable relaxation $p_\theta(z)$ such that $H(z) = b$. They use the differentiable surrogate (and $p_\theta(z)$ when applicable) as a control variate and directly minimize the variance of the score function estimator. The result is an unbiased and seemingly low-variance estimator.

Additionally, for discrete b, they resample the relaxed $\tilde{z} \mid b$, which allows them to further reduce variance by increasing the correlation between their control variate and the score function estimator.

More concretely, let $b \sim \mathrm{Cat}(\theta), \theta \in \Delta^{n-1}$. The score function estimator is

$$\nabla_\theta \mathbb{E}_b[f(b)] = \mathbb{E}_b[f(b)\nabla_\theta \log p(b \mid \theta)] \tag{1}$$

The DLAX estimator is given by introducing differentiable surrogate $c_\phi(z)$ with Gumbel-perturbed logits $z = T(\theta, \epsilon), \epsilon \sim \mathrm{Gumbel}(0, 1)$.

$$\nabla_\theta \mathbb{E}_b[f(b)] = \mathbb{E}_\epsilon \left[f(b)\nabla_\theta \log p(b \mid \theta) - c_\phi(z)\nabla_\theta \log p(z \mid \theta) + \underbrace{\nabla_\theta c_\phi(z)}_{\text{Reparam}} \right] \tag{2}$$

where $b = H(z), z = T(\theta, \epsilon), \epsilon \sim \mathrm{Gumbel}(0, 1)$ and H is a deterministic transformation. For the case of $z \sim \mathrm{Gumbel}(\theta)$, $H(z) = \mathrm{argmax}_i(z_i)$.

The RELAX estimator is obtained by further resampling $\tilde{z} \mid b$. With some rewriting, we obtain

$$\nabla_\theta \mathbb{E}_b[f(b)] = \mathbb{E}_\epsilon \left[(f(b) - \mathbb{E}_{\tilde{z}|b}[c_\phi(\tilde{z})])\nabla_\theta \log p(b \mid \theta) - \nabla_\theta \mathbb{E}_{\tilde{z}|b}[c_\phi(\tilde{z})] + \nabla_\theta c_\phi(z) \right] \tag{3}$$

See Appendix A for the derivation. The second to last term is of interest, as it entails resampling $\tilde{z} \mid b$. This term is also assumed to be reparameterizable. In total, the RELAX estimator requires one evaluation of f at b and two evaluations of c_ϕ at $z, \tilde{z}$.

In summary:

1. Introduce differentiable surrogate c_ϕ as reparameterizable control variate
2. Condition on b to increase correlation of control variate with estimand

Results Show better training and validation performance on a mixture density model for MNIST Better training but overfits on omniglot Demonstrate faster convergence and lower variance gradient estimators than A2C for three RL tasks: cart-pole, lunar lander, and inverted pendulum

Limitations

- No empirical validation of variance reduction in DLAX to RELAX
- Likely requires combination with other variance reduction techniques when dealing with large, unstructured distributions (unavoidable).

Questions / Comments

- I would have liked to see experiments analyzing the effect of parameter sharing between f and c_θ when possible.

2 Stochastic Softmax Tricks

[5]

3 Variance reduction properties of the reparameterization trick

[6]

References

[1] Will Grathwohl, Dami Choi, Yuhuai Wu, Geoffrey Roeder, and David Duvenaud. Backpropagation through the void: Optimizing control variates for black-box gradient estimation. *CoRR*, abs/1711.00123, 2017. URL `http://arxiv.org/abs/1711.00123`.

[2] Chris J. Maddison. A poisson process model for monte carlo. *ArXiv*, abs/1602.05986, 2016.

[3] Shakir Mohamed, Mihaela Rosca, Michael Figurnov, and Andriy Mnih. Monte carlo gradient estimation in machine learning. *ArXiv*, abs/1906.10652, 2019.

[4] Art B. Owen. *Monte Carlo theory, methods and examples.* 2013.

[5] Max B. Paulus, Dami Choi, Daniel Tarlow, Andreas Krause, and Chris J. Maddison. Gradient estimation with stochastic softmax tricks. *ArXiv*, abs/2006.08063, 2020.

[6] Ming Xu, Matias Quiroz, Robert Kohn, and Scott Anthony Sisson. Variance reduction properties of the reparameterization trick. In *AISTATS*, 2018.