DisARM: An Antithetic Gradient Estimator for Binary Latent Variables

Zhe Dong (zhedong@google.com)[1] Andriy Mnih(amnih@)[2] George Tucker(gjt@)[1]
[1]Google Research, Brain Team [2]DeepMind | https://arxiv.org/abs/2006.10680

Contributions

DisARM – a gradient estimator for binary latent variables

- Improves ARM (Yin and Zhou 2018) using *analytical integration*.
- *Outperforms* ARM and a strong independent sample baseline on variance and log-likelihood.

A multi-sample version of DisARM for the *multi-sample variational bound* (IWAE). It outperforms VIMCO, the current state-of-the-art method.

Problem

Fit a model with discrete latent variables, by optimizing

$$\mathbb{E}_{q_\theta(\mathbf{b})}\left[f_\theta(\mathbf{b})\right]$$

where the gradient is

$$\nabla_\theta E_{q_\theta(\mathbf{b})}\left[f_\theta(\mathbf{b})\right] = E_{q_\theta(\mathbf{b})}\left[f_\theta(\mathbf{b})\nabla_\theta \log q_\theta(\mathbf{b}) + \nabla_\theta f_\theta(\mathbf{b})\right]$$

This can be estimated by Monte-Carlo with high variance.

Q: *How to efficiently estimate the gradients?*

Idea

Bernoulli → (Increases variance) → Logistic Reparamization → (Reduces variance) → Antithetic coupling → (Reduces variance) → Bernoulli

ARM

DisARM

ARM uses continuous antithetic Logistic random variables $z = \epsilon + \alpha_\theta$ and $\tilde{z} = -\epsilon + \alpha_\theta$, where $\epsilon \sim \mathrm{Logistic}(0,1)$ but our objective only depends on the discrete values

$$(b, \tilde{b}) = (\mathbb{1}_{z>0}, \mathbb{1}_{\tilde{z}>0})$$

$$g_{\mathrm{DisARM}}(b, \tilde{b}) := \mathbb{E}_{q(z|b,\tilde{b})}\left[g_{\mathrm{ARM}}\right]$$

DisARM: An Antithetic Gradient Estimator for Binary Latent Variables

DisARM

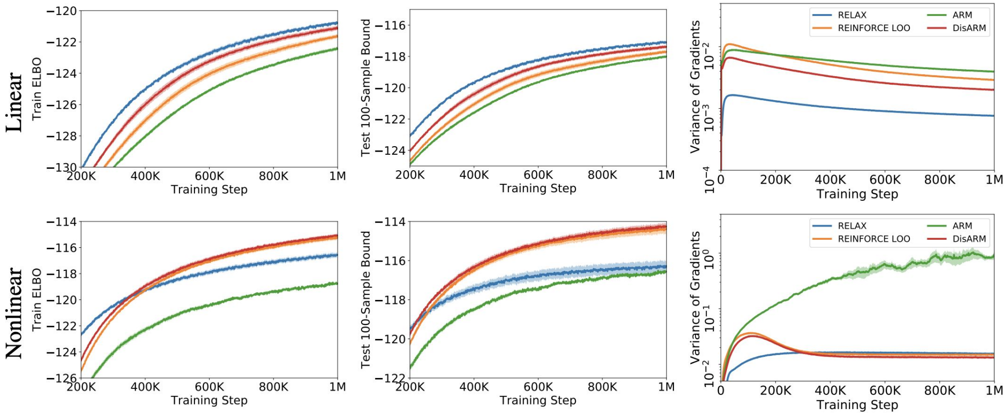
For scalar $(b, \tilde{b}) = (\mathbb{1}_{z>0}, \mathbb{1}_{\tilde{z}>0})$, DisARM integrates out z conditioned on $(b, \tilde{b})$.

$$g_{\text{DisARM}}(b, \tilde{b}) := \mathbb{E}_{q(z|b,\tilde{b})}\left[g_{\text{ARM}}\right] = \frac{1}{2}\mathbb{E}_{q(z|b,\tilde{b})}\left[(f(\mathbb{1}_{z>0}) - f(\mathbb{1}_{\tilde{z}>0}))\nabla_\theta \log q_\theta(z)\right]$$
$$= \frac{1}{2}(f(b) - f(\tilde{b}))\mathbb{E}_{q(z|b,\tilde{b})}\left[\nabla_\theta \log q_\theta(z)\right]$$
$$= \frac{1}{2}(f(b) - f(\tilde{b}))\left((-1)^{\tilde{b}}\mathbb{1}_{b\neq\tilde{b}}\sigma(|\alpha_\theta|)\right)\nabla_\theta\alpha_\theta.$$

For the multi-dimensional case,

$$g_{\text{DisARM}}(\mathbf{b}, \tilde{\mathbf{b}}) = \sum_i \left(\frac{1}{2}(f(\mathbf{b}) - f(\tilde{\mathbf{b}}))\left((-1)^{\tilde{\mathbf{b}}_i}\mathbb{1}_{\mathbf{b}_i\neq\tilde{\mathbf{b}}_i}\sigma(|(\alpha_\theta)_i|)\right)\nabla_\theta(\alpha_\theta)_i\right)$$

Omniglot



We evaluate the gradient estimators on three benchmark generative modeling datasets: MNIST, FashionMNIST and Omniglot. We use dynamic binarization to avoid confounding due to overfitting.

	REINFORCE LOO	ARM	DisARM	RELAX
Train ELBO				
Dynamic MNIST				
Linear	-116.57 ± 0.15	-117.66 ± 0.04	$\mathbf{-116.30 \pm 0.08}$	-115.93 ± 0.15
Nonlinear	$\mathbf{-102.45 \pm 0.12}$	-107.32 ± 0.28	$\mathbf{-102.56 \pm 0.19}$	-102.53 ± 0.15
Fashion MNIST				
Linear	-256.33 ± 0.14	-256.80 ± 0.16	$\mathbf{-255.97 \pm 0.07}$	-255.83 ± 0.03
Nonlinear	$\mathbf{-237.66 \pm 0.11}$	-241.30 ± 0.10	$\mathbf{-237.77 \pm 0.08}$	-238.23 ± 0.17
Omniglot				
Linear	-121.66 ± 0.10	-122.45 ± 0.10	$\mathbf{-121.15 \pm 0.12}$	-120.79 ± 0.09
Nonlinear	$\mathbf{-115.26 \pm 0.15}$	-118.76 ± 0.05	$\mathbf{-115.08 \pm 0.11}$	-116.56 ± 0.15
Test 100-sample bound				
Dynamic MNIST				
Linear	$\mathbf{-109.25 \pm 0.09}$	-109.70 ± 0.05	$\mathbf{-109.13 \pm 0.04}$	-108.76 ± 0.06
Nonlinear	$\mathbf{-97.41 \pm 0.09}$	-101.15 ± 0.39	$\mathbf{-97.52 \pm 0.11}$	-97.76 ± 0.11
Fashion MNIST				
Linear	-252.55 ± 0.12	-252.66 ± 0.07	$\mathbf{-252.30 \pm 0.05}$	-252.13 ± 0.06
Nonlinear	$\mathbf{-236.94 \pm 0.09}$	-239.37 ± 0.15	$\mathbf{-237.02 \pm 0.07}$	-237.95 ± 0.16
Omniglot				
Linear	-117.70 ± 0.10	-118.01 ± 0.06	$\mathbf{-117.39 \pm 0.09}$	-117.10 ± 0.08
Nonlinear	$\mathbf{-114.39 \pm 0.21}$	-116.56 ± 0.07	$\mathbf{-114.26 \pm 0.14}$	-116.28 ± 0.26

DisARM: An Antithetic Gradient Estimator for Binary Latent Variables

Multi-sample Variational Bound

The K sample training objective is

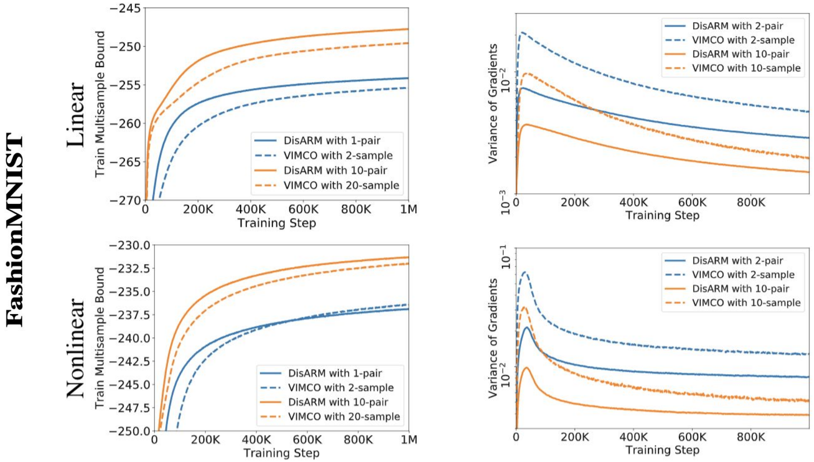
$$\mathcal{L} := \mathbb{E}_{\prod_k q_\theta(\mathbf{b}^k)}\left[\log \frac{1}{K}\sum_k w(\mathbf{b}^k)\right]$$

The gradient of *i*-th dimension and *k*-th sample is

$$\frac{1}{4}\left(f_{\mathbf{b}^{-k}}(\mathbf{b}^k) - f_{\mathbf{b}^{-k}}(\tilde{\mathbf{b}}^k) + f_{\tilde{\mathbf{b}}^{-k}}(\mathbf{b}^k) - f_{\tilde{\mathbf{b}}^{-k}}(\tilde{\mathbf{b}}^k)\right)\left(\mathbb{1}_{\mathbf{b}_i^k \neq \tilde{\mathbf{b}}_i^k}(-1)^{\tilde{\mathbf{b}}_i^k}\sigma(|(\alpha_\theta)_i|)\right)$$

$$\text{let } f_{\mathbf{b}^{-k}}(\mathbf{d}) = \log\frac{1}{K}\left(\sum\nolimits_{\mathbf{c}\in\mathbf{b}^{-k}} w(\mathbf{c}) + w(\mathbf{d})\right)$$

$$\text{with } \mathbf{b}^{-k} := (\mathbf{b}^1, \ldots, \mathbf{b}^{k-1}, \mathbf{b}^{k+1}, \ldots, \mathbf{b}^K)$$



To provide a computationally fair comparison between VIMCO 2K-samples and DisARM K-pairs, we report the 2K-sample bound for both, even though DisARM optimizes the K-sample bound.

	REINFORCE LOO	ARM	DisARM	RELAX
Train ELBO				
Dynamic MNIST				
Linear	-116.57 ± 0.15	-117.66 ± 0.04	$\mathbf{-116.30 \pm 0.08}$	-115.93 ± 0.15
Nonlinear	$\mathbf{-102.45 \pm 0.12}$	-107.32 ± 0.28	$\mathbf{-102.56 \pm 0.19}$	-102.53 ± 0.15
Fashion MNIST				
Linear	-256.33 ± 0.14	-256.80 ± 0.16	$\mathbf{-255.97 \pm 0.07}$	-255.83 ± 0.03
Nonlinear	$\mathbf{-237.66 \pm 0.11}$	-241.30 ± 0.10	$\mathbf{-237.77 \pm 0.08}$	-238.23 ± 0.17
Omniglot				
Linear	-121.66 ± 0.10	-122.45 ± 0.10	$\mathbf{-121.15 \pm 0.12}$	-120.79 ± 0.09
Nonlinear	$\mathbf{-115.26 \pm 0.15}$	-118.76 ± 0.05	$\mathbf{-115.08 \pm 0.11}$	-116.56 ± 0.15
Test 100-sample bound				
Dynamic MNIST				
Linear	$\mathbf{-109.25 \pm 0.09}$	-109.70 ± 0.05	$\mathbf{-109.13 \pm 0.04}$	-108.76 ± 0.06
Nonlinear	$\mathbf{-97.41 \pm 0.09}$	-101.15 ± 0.39	$\mathbf{-97.52 \pm 0.11}$	-97.76 ± 0.11
Fashion MNIST				
Linear	-252.55 ± 0.12	-252.66 ± 0.07	$\mathbf{-252.30 \pm 0.05}$	-252.13 ± 0.06
Nonlinear	$\mathbf{-236.94 \pm 0.09}$	-239.37 ± 0.15	$\mathbf{-237.02 \pm 0.07}$	-237.95 ± 0.16
Omniglot				
Linear	-117.70 ± 0.10	-118.01 ± 0.06	$\mathbf{-117.39 \pm 0.09}$	-117.10 ± 0.08
Nonlinear	$\mathbf{-114.39 \pm 0.21}$	-116.56 ± 0.07	$\mathbf{-114.26 \pm 0.14}$	-116.28 ± 0.26