Adversarial Variational Optimization of Non-Differentiable Simulators

Gilles Louppe[1] and **Kyle Cranmer**[1]

[1] *New York University*

In this note, ... [GL: todo.]

I. INTRODUCTION

[GL: Prescribed vs. implicit. See case of non-diff models in Balaji et al.]

II. PROBLEM STATEMENT

Consider a family of conditional densities $p(\mathbf{x}|\theta)$ defined implicitly through the simulation of a stochastic generative process, where $\mathbf{x} \in \mathbb{R}^d$ is the observed data and θ are the parameters of interest. The simulation may involve some complicated latent process, such that

$$p(\mathbf{x}|\theta) = \int p(\mathbf{x}|\mathbf{z}, \theta) p(\mathbf{z}) d\mathbf{z} \tag{1}$$

where $\mathbf{z} \in \mathbb{R}^m$ is a latent variable providing an external source of randomness.

We assume that we already have an accurate simulation of the stochastic process underlying $p(\mathbf{x}|\mathbf{z}, \theta)$, specified as a function $g(\cdot;\theta) : \mathbb{R}^m \to \mathbb{R}^d$. That is,

$$p(\mathbf{x}|\theta) = \frac{\partial}{\partial x_1} \cdots \frac{\partial}{\partial x_d} \int_{\{\mathbf{z}:g(\mathbf{z};\theta)\leq \mathbf{x}\}} p(\mathbf{z}) d\mathbf{z}. \tag{2}$$

In particular, the simulator g is assumed to be a non-invertible function, that can only be used to generate data in forward mode. For this reason, evaluating the integral in Eqn. 2 is intractable. Importantly, and as increasingly found in science, we consider the additional constraint that g is a non-differentiable model, e.g. as specified by computer program.

Given some observed data $\{\mathbf{x}_i | i = 1, \ldots, N\}$, our goal is the inference of the parameters of interest θ^* that maximize the (marginal) log-likelihood of the observations:

$$\theta^* = \arg\max_\theta \sum_i \log p(\mathbf{x}_i|\theta). \tag{3}$$

[GL: Or rather, define the goal as the minimization of a divergence?]

III. METHOD

A. Variational optimization

B. Generative adversarial networks

C. Adversarial variational optimization

IV. EXPERIMENTS

A. Toy problem

B. Physics example

V. RELATED WORKS

[GL: Implicit generative models.] [GL: ABC.] [GL: carl [1].] [GL: Wood's papers.] [GL: CMA-ES.]

VI. SUMMARY

ACKNOWLEDGMENTS

KC and GL are both supported through NSF ACI-1450310, additionally KC is supported through PHY-1505463 and PHY-1205376.

[1] Cranmer, K., Pavez, J., and Louppe, G. Approximating likelihood ratios with calibrated discriminative classifiers. *arXiv preprint arXiv:1506.02169* (2015).