Simulation-based Inference

A Summary Written by Stefan Wezel
stefan.wezel@student.uni-tuebingen.de

Machine Learning for and with Dynamical Systems
Summer Term 2021

Abstract

Simulators are used across different scientific domains where serve as valuable tools to encode empirical knowledge about a system of interest. Inference in this setting often boils down to finding parameters of the simulator for a given, real-world observation, which typically cannot be computed analytically.Papamakarios and Murray [2016] propose to learn a posterior over a simulators parameters by using neural density estimators. Here, we give an overview of their work and SBI in general.

Problem Setting and Related Work

Scientific fields, such as population genetics, particle physics, epidemiology, astrophysics among others make use of sophisticated simulators to model observed systems [Brehmer and Cranmer, 2020, de Witt et al., 2020, Delaunoy et al., 2020, Cranmer et al., 2020, Pritchard et al., 1999]. These simulators are able to encode prior knowledge like causal relations, hierarchies of variables, ... Performing inference in this setting, however, would require a finding a posterior over parameters θ for a given observation x_0. This can be stated in closed form as

$$p(\theta|x = x_0) = \frac{p(x|\theta)\,p(\theta)}{p(x)} = \frac{\overbrace{\int p(x,z|\theta)\,\mathrm{d}z}^{intractable}\,p(\theta)}{\underbrace{\int p(x|\theta)\,p(\theta)\,\mathrm{d}\theta}_{intractable}}, \tag{1}$$

where z is a nuisance parameter. If likelihood and evidence are intractable, which is typically the case, this cannot be computed analytically. Intuitively, it can be thought of as inverting the simulator.

A collective term for methods to solve this inverse problem is Approximate Bayesian Computation (ABC). The most simple approach is rejection ABC [Pritchard et al., 1999] where samples produced by a simulator are discarded if notis within a ϵ-ball around the observation of interest. For small values of ϵ this method may take many steps to even produce a single matching parameter set. For large ϵ it is not precise. Moreover, this method does not produce a full posterior over the parameter but merely yields point estimates with confident intervals. With Sampling ABC [Marjoram et al., 2003] and Sequential ABC [Beaumont et al., 2009, Bonassi et al., 2015] improvements over rejection ABC have been proposed. These methods produce samples more efficiently. While this can lead to faster convergence, it does not produce a full posterior over parameters that is conditioned on the actual observation, but rather a sample that is close to it.

Neural Approach

Papamakarios and Murray [2016] build on recent advances in deep learning to learn a posterior over parameters directly from sample produced by a simulator. They use Deep Neural Networks (DNN) to parameterize a Gaussian Mixture Model (GMM). They use the simulator to perform n simulations and storing pairs of samples x and corresponding parameters θ. This gives the set $\{(x_i, \theta_i)\}_{i \in n}$ which they use as training data. DNNs are fitted to parameterize a GMM

$$\sum_k \alpha_k \mathcal{N}(\theta | m_k, S_k) = \sum_k f_1(x)_k \mathcal{N}(\theta | f_2(x)_k, diag(f_3(x))_k) \tag{2}$$

, where k is the number of components, m the mean and S the covariance matrix of a Gaussian, and $f_{\{1,2,3\}}$ are MLPs.

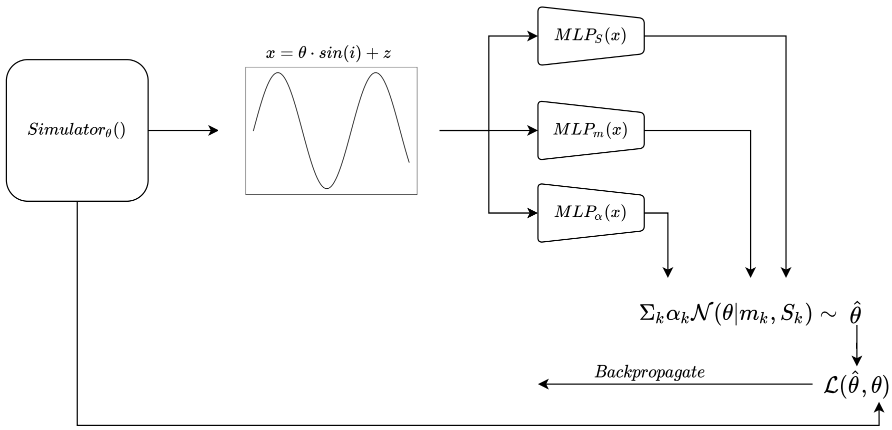

Figure 1: Given samples by the simulator, Multilayer perceptrons (MLP) are trained to parameterize a distribution over parameters generating the sample. A training signal is created by sampling from this distribution and comparing the sample to the actual parameters which are known.

Possible Enhancements/Open Ends

Conclusion

References

M. A. Beaumont, J.-M. Cornuet, J.-M. Marin, and C. P. Robert. Adaptive approximate bayesian computation. *Biometrika*, 96(4):983–990, 2009.

F. V. Bonassi, M. West, et al. Sequential monte carlo with adaptive weights for approximate bayesian computation. *Bayesian Analysis*, 10(1):171–187, 2015.

J. Brehmer and K. Cranmer. Simulation-based inference methods for particle physics. *arXiv preprint arXiv:2010.06439*, 2020.

K. Cranmer, J. Brehmer, and G. Louppe. The frontier of simulation-based inference. *Proceedings of the National Academy of Sciences*, 117(48):30055–30062, 2020.

C. S. de Witt, B. Gram-Hansen, N. Nardelli, A. Gambardella, R. Zinkov, P. Dokania, N. Siddharth, A. B. Espinosa-Gonzalez, A. Darzi, P. Torr, et al. Simulation-based inference for global health decisions. *arXiv preprint arXiv:2005.07062*, 2020.

A. Delaunoy, A. Wehenkel, T. Hinderer, S. Nissanke, C. Weniger, A. R. Williamson, and

G. Louppe. Lightning-fast gravitational wave parameter inference through neural amortization. *arXiv preprint arXiv:2010.12931*, 2020.

P. Marjoram, J. Molitor, V. Plagnol, and S. Tavaré. Markov chain monte carlo without likelihoods. *Proceedings of the National Academy of Sciences*, 100(26):15324–15328, 2003.

G. Papamakarios and I. Murray. Fast epsilon-free inference of simulation models with bayesian conditional density estimation. *arXiv preprint arXiv:1605.06376*, 2016.

J. K. Pritchard, M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman. Population growth of human y chromosomes: a study of y chromosome microsatellites. *Molecular biology and evolution*, 16(12):1791–1798, 1999.