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 they serve as valuable tools to encode empirical knowledge about a system of interest. Inference in this setting often means finding parameters of the simulator for a given, real-world observation, which typically cannot be computed analytically.Papamakarios and Murray [2016] and Lueckmann et al. [2017] propose to learn a posterior over a simulators parameters by using Bayesian 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 serve as forward models with strong inductive biases, encoding prior knowledge. 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

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

where z is a nuisance parameter and x is observation data from experiments or simulations. This term typically cannot be computed analytically.
A collective term for methods that aim to solve this inverse problem is Approximate Bayesian Computation (ABC). The most simple approach is rejection ABC [Pritchard et al.,

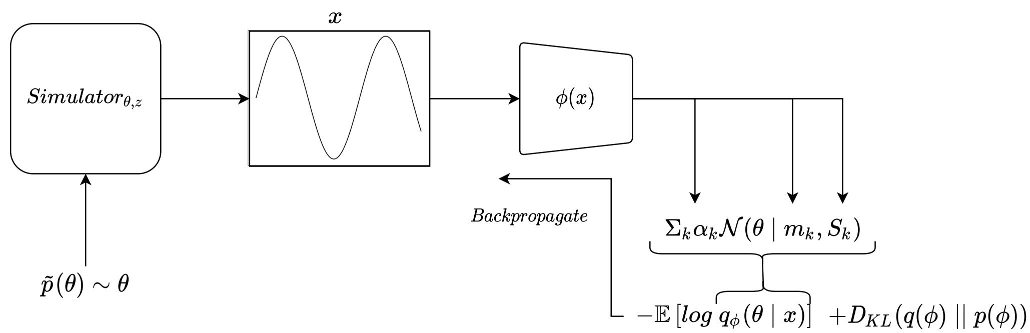

Figure 1: Given samples by the simulator, Multilayer perceptrons (MLP) are trained to parameterize a distribution over parameters generating the sample. A training signal is created by sampling from this distribution and comparing the sample to the actual parameters which are known.

1999] where samples produced by a simulator are discarded if not within an ϵ-ball around an observation x_0. For small values of ϵ this method is inefficient. For large ϵ it is not precise [Papamakarios and Murray, 2016]. Moreover, this method does not produce a full posterior over the parameter but merely yields point estimates that can be used to approximate a posterior. Note that this posterior is then conditioned on an ϵ-ball $\| x - x_0 \| < \epsilon$ around the observation rather than the actual observation x_0.
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. While these methods produce samples more efficiently, they suffer from the latter problem as well, as they only produce point estimates for parameters that produce something in ϵ-range of x_0.

Neural Approach

Papamakarios and Murray [2016] build on recent advances in deep learning and propose to learn a posterior over parameters directly from samples produced by a simulator. They parameterize a Gaussian Mixture Model (GMM) with a Bayesian Neural Network (BNN) with parameters $\phi = \mathcal{N}(\phi_m, exp^{\frac{1}{2}\phi_S})$ with mean m and Covariance S. They perform n simulations and store pairs of observations x_i and corresponding parameters θ_i, sampled from a proposal prior $\tilde{p}(\theta)$ to obtain the training set $\{(x_i, \theta_i)\}_{i \in n}$. To find a suiting ϕ, $x_{0,\ldots,n}$ are passed to it. Each respective output is used to parameterize a GMM $q_\phi(\theta \mid x) = \sum_k \alpha_k \mathcal{N}(\theta \mid m_k, S_k)$ with k components, mixing coefficient α_k, mean m_k, and covariance S_k.

An expectation $\frac{1}{N}\sum_n \mathbb{E}\left[log\ q_\phi(\theta \mid x)\right]$ is estimated empirically through sampling. An regularization term $\frac{1}{N}D_{KL}(q^t(\phi) \mid\mid p(\phi))$, where $p(\phi)$ is a isotropic, zero-centered prior over weights is computed analytically. These terms form a loss $\mathscr{L}$ which returns a scalar value that can be backpropagated so that the first term is maximized and the latter minimized. This completes an 'inner' training loop. Papamakarios and Murray [2016] further propose an 'outer' training loop, where once ϕ is fitted to n observations, $\frac{p(\theta)}{\tilde{p}(\theta)} q(\theta \mid x = x_0)$ replaces the proposal prior for a specific observation of interest. This, however, requires $\tilde{p}(\theta)$ to be of a form so that the importance weights $\frac{p(\theta)}{\tilde{p}(\theta)}$ can be computed analytically.

Application in Neuroscience

Lueckmann et al. [2017] propose further enhancements to Papamakarios and Murray [2016]'s method. They apply SBI to forward models used in neuro-science, such as the Hodgkin-Huxley model Hodgkin and Huxley [1952]. Simulators in this field frequently produce nonsensical observations [Lueckmann et al., 2017]. To alleviate this, Lueckmann et al. [2017] additionally train a classifier that learns to detect parameter sets that will produce such 'bad' simulations, which are then discarded. They state that finding useful summary statistics of the often high-dimensional data is challenging. By using a recurrent neural network to extract features from observations, they find an informative, learned summary statistic. Moreover, Lueckmann et al. [2017] formulate a loss such that it includes the importance weighting. This allows the use of more complex proposal priors, as they don't need to be computed analytically in their method. In the regularization term of their loss, they replace the isotropic, zero-centered Gaussian as prior in the regularization term through the posterior over weights from previous round to implicitly store information.
Lueckmann et al. [2017] evaluate their method for different models and find that it reliably finds matching parameter settings for synthetic data. On in-vitro experiments they find parameter settings that match empirical data accurately.

Conclusion

Simulators are an important tool of reasarch scientiest across domains. Papamakarios and Murray [2016] and Lueckmann et al. [2017] leverage recent advances in deep learning and variational inference to efficiently perform simulation-based inference in challenging settings. An afterthought of our summary is that SBI methods can be applied in different settings but may require adaptions to suit the task.

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.

A. L. Hodgkin and A. F. Huxley. A quantitative description of membrane current and its application to conduction and excitation in nerve. *The Journal of physiology*, 117(4): 500–544, 1952.

J.-M. Lueckmann, P. J. Goncalves, G. Bassetto, K. Öcal, M. Nonnenmacher, and J. H. Macke. Flexible statistical inference for mechanistic models of neural dynamics. *arXiv preprint arXiv:1711.01861*, 2017.

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.