Variational autoencoder with weighted samples for high-dimensional non-parametric adaptive importance sampling

Julien Demange-Chryst[†,1,2], François Bachoc[§,2], Jérôme Morio[§,1], Timothé Krauth[1,3]

[†] PhD student (presenting author). [§] PhD supervisor

PhD expected duration: Oct. 2021 – Sep. 2024

[1] ONERA/DTIS, Université de Toulouse, F-31055 Toulouse, France
{julien.demange-chryst,jerome.morio}@onera.fr

[2] Institut de Mathématiques de Toulouse, UMR5219 CNRS, 31062 Toulouse, France
francois.bachoc@math.univ-toulouse.fr

[3] Zurich University of Applied Sciences, Centre for Aviation, Winterthur, Switzerland
timothe.krauth@zhaw.ch

Abstract

Importance sampling (IS) is a well-known uncertainty quantification method, classically used as a variance-reduction technique for Monte Carlo integration including rare event estimation [6], or for generating points from a target probability distribution known up to a constant [5]. The common denominator of every importance sampling procedure is that they all require to estimate a target probability distribution with weighted samples, and obviously, the accuracy of the algorithm depends on the quality of the estimation of the distribution. Moreover, we also need to be able to not only sample from the built auxiliary distribution, but also to have access to its PDF values.

A first way to perform this density estimation is to use non-parametric models, such as kernel smoothing. These models are flexible, but despite some improvements they strongly suffer from the curse of dimensionality since the size of the required sample to have a good approximation of the target distribution exponentially grows with the dimension. Another solution is to use parametric families of distributions, such as the Gaussian or Gaussian mixture ones, which are more robust in medium-high dimension. However, they sometimes require some prior knowledge on the target distribution, and their lack of flexibility and the huge number of parameters to estimate can negatively impact the quality of the estimation when the dimension is high.

In order to combine both flexibility and robustness faced to the dimension, we propose to use as the auxiliary sampling distribution a distribution parameterised by a variational autoencoder [4], whose main principle has been introduced in the last decade. Variational autoencoders are deep generative models for approximating high-dimensional complex distributions of observed data and generating new samples. The specific feature of a variational autoencoder compared to other density estimation methods is that it performs a dimensionality reduction into a lower dimensional latent space in order to facilitate the estimation. Moreover, in opposition to other dimensionality reduction techniques such as principal component analysis or autoencoders, variational autoencoders have good generation properties and give explicitly the approximating distribution, allowing to perform Monte Carlo simulations. This tool is now popular in the machine learning community but not so much in uncertainty quantification.

In the present communication [3], we extend the existing framework of variational autoencoders to the case of weighted samples by introducing a new objective function. The resulting IS auxiliary

distribution is close to an infinite mixture of Gaussian distributions. Then, its flexibility makes it as expressive as a non-parametric model, and despite the very high number of parameters to estimate, it is much more efficient in high dimension than the classical Gaussian or Gaussian mixture families. Moreover, in order to add even more flexibility to the model and to be able to learn multimodal distributions, we consider a learnable prior distribution for the variational autoencoder latent variables. We also introduce a new pre-training procedure for the variational autoencoder to find good starting weights of the neural networks to prevent as much as possible the posterior collapse phenomenon to happen.

At last, we explicit how the resulting distribution can be combined with importance sampling. Indeed, the existing procedure [7] to compute the PDF values of the resulting distribution of a variational autoencoder leads to a biased and non-convergent importance sampling estimator. In order to keep an unbiased and consistent estimator, we introduce a new way to compute the PDF values. Then, we show how to integrate the whole suggested procedure into existing reliability algorithms, such as the cross-entropy algorithm, for rare event estimation. Finally, we illustrate the practical interests of the previous efforts on two multimodal rare-event-estimation problems. The code to reproduce the numerical experiments is publicly available at: `https://github.com/Julien6431/Importance-Sampling-VAE.git`.

Short biography

I graduated in 2021 from the engineering school ISAE-SUPAERO in Toulouse and I also obtained a MSc degree in applied mathematics from Toulouse III - Paul Sabatier University. Then, I performed my final-year internship at *ONERA* in Toulouse which led to my current PhD thesis co-funded by *ONERA* and *Toulouse III - Paul Sabatier University*. My first two topics of research of my PhD on reliability-oriented sensitivity analysis with dependant inputs and the common estimation of multiple expectations lead to two journal publications [2, 1].

References

[1] Julien Demange-Chryst, François Bachoc, and Jérôme Morio. Efficient estimation of multiple expectations with the same sample by adaptive importance sampling and control variates. *Statistics and Computing*, 33(5):103, 2023.

[2] Julien Demange-Chryst, François Bachoc, and Jérôme Morio. Shapley effect estimation in reliability-oriented sensitivity analysis with correlated inputs by importance sampling. *International Journal for Uncertainty Quantification*, 13(3), 2023.

[3] Julien Demange-Chryst, François Bachoc, Jérôme Morio, and Timothé Krauth. Variational autoencoder with weighted samples for high-dimensional non-parametric adaptive importance sampling. *arXiv preprint arXiv:2310.09194*, 2023.

[4] Diederik P. Kingma and Max Welling. Auto-Encoding Variational Bayes. *Article accepted in the 2nd International Conference on Learning Representations 2014*, 2014.

[5] Jean-Michel Marin, Pierre Pudlo, and Mohammed Sedki. Consistency of adaptive importance sampling and recycling schemes. *Bernoulli*, 25(3):1977 – 1998, 2019.

[6] Reuven Y Rubinstein and Dirk P Kroese. *The cross-entropy method: a unified approach to combinatorial optimization, Monte Carlo simulation, and machine learning*, volume 133. Springer, 2004.

[7] Hechuan Wang, Mónica F Bugallo, and Petar M Djurić. Adaptive importance sampling supported by a variational auto-encoder. In *2019 IEEE 8th International Workshop on Computational Advances in Multi-Sensor Adaptive Processing (CAMSAP)*, pages 619–623. IEEE, 2019.