Improvement of the cross-entropy method in high dimension through a one-dimensional projection without gradient estimation

Maxime El Masri
ONERA and ISAE-SUPAERO (Université de Toulouse)

Supervisor(s): Florian Simatos (ISAE), Jérôme Morio (ONERA)

Ph.D. expected duration: Oct. 2018 - Sep. 2021

Adress: 2 avenue Edouard Belin, FR-31055 TOULOUSE CEDEX 4

Email: melmasri@onera.fr

Abstract:

An important topic in reliability analysis is to determine the failure probability of complex systems, such as $P = \mathbb{P}_f(\varphi(X) > 0)$, where X is a random vector, possibly high dimensional, drawn from the probability density f, and $\varphi : \mathbb{R}^n \mapsto \mathbb{R}$ is a deterministic black-box function. This probability can be difficult to evaluate when $\{\varphi(x) > 0\}$ is a rare event and calls to the function φ are computationally expensive. For these reasons, deterministic methods and the classical stochastic method of Monte Carlo are inefficient to evaluate such probabilities. Some stochastic techniques have been developed to efficiently estimate rare event probabilities ([3, 1]). In the thesis, we focus on the Cross-entropy (CE) method [6], a particular case of Adaptive Importance Sampling (AIS) (see [4]), consisting in sampling from an auxiliary density for which the failure event $\{\varphi(x) > 0\}$ is more frequent. CE is based on the minimisation of the Kullback-Leibler divergence between the optimal IS density and a chosen parametric family of auxiliary densities, in order to find an efficient auxiliary density and estimate the probability. A convenient choice of parametric family is the Gaussian family $\mathcal{N}(m, \Sigma)$, with mean vector m, and covariance matrix Σ. Thus the basic CE algorithm estimates the optimal parameters m^* and Σ^*, with $\hat{m}$ and $\hat{\Sigma}$, then draws a sample from $\mathcal{N}(\hat{m}, \hat{\Sigma})$ and finally estimates the probability.

However, CE and more generally IS techniques, often fail to estimate probabilities, when the dimension of the parameters is large (see for instance [5, 8, 7]). The main contribution of the thesis [2] is the development of a simple algorithm improving the CE method in high dimension and for unimodal problems, by estimating the variance only in a one-dimensional subspace. Indeed, if n is the dimension of the input X, the space of Gaussian parameters (m, Σ) is of dimension $n + n(n + 1)/2$, which means that we estimate more than $n^2/2$ parameters in the CE algorithm. By estimating the covariance matrix in a one-dimensional subspace, we can reduce the number of estimated parameters to only $n + 1$. This subspace is chosen as the subspace spanned by the CE-optimal mean m^*, $\text{span}(m^*)$, which gives an influential direction in the covariance estimation (i.e. a direction where the covariance varies significantly during the CE algorithm). The new algorithm, called CE-m^*, gives accurate probability estimation in high dimension, without any additional computation, and without gradient information, provided that we are in a unimodal case. Figure 1 shows the efficiency of CE-m^* and the degradation of CE when the dimension grows.

Although the projection on $\text{span}(m^*)$ gives efficient results, this direction is not always optimal. Thus, to pursue this work, we aim to find other directions to project the covariance matrix Σ. The minimisation of the Kullback-Leibler divergence suggests that some eigenvectors of Σ can be optimal projections. This work is still in progress but already gives hopeful results. Moreover, this can be applied to any expectation estimation in general, not only to rare event probabilities.

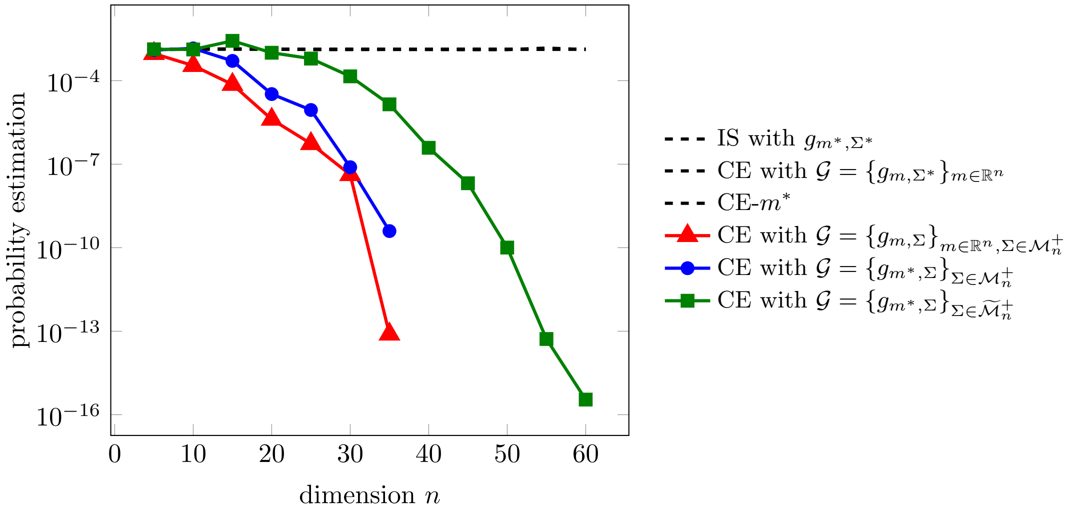

Figure 1: Evolution of the estimation of the probability $P = \mathbb{P}_f(\varphi(X) > 0)$ as the dimension increases for φ a linear function and for different algorithms. In dashed are three algorithms which provide accurate and indistinguishable results: IS with the optimal density g_{m^*,Σ^*}, CE with auxiliary family $\mathcal{G} = \{g_{m,\Sigma^*}\}_{m\in\mathbb{R}^n}$ where one uses the optimal variance and estimates the mean, and CE-m^*. In red with triangle is CE with $\{g_{m,\Sigma}\}_{m\in\mathbb{R}^n,\Sigma\in\mathcal{M}_n^+}$ (both the mean and the full covariance matrix are estimated). In blue with circles is CE with $\{g_{m^*,\Sigma}\}_{\Sigma\in\mathcal{M}_n^+}$ (one uses the optimal mean m^* and estimates the full covariance matrix). In green with squares is CE with $\{g_{m^*,\Sigma}\}_{\Sigma\in\widetilde{\mathcal{M}}_n^+}$ (one uses the optimal mean m^* and estimates 3/4-th of the covariance matrix).

References

[1] Jean-Marc Bourinet. Reliability analysis and optimal design under uncertainty-focus on adaptive surrogate-based approaches. 2018. Université Clermont Auvergne. Habilitation à diriger des recherches.

[2] Maxime El Masri, Jérôme Morio, and Florian Simatos. Improvement of the cross-entropy method in high dimension through a one-dimensional projection without gradient estimation. *arXiv preprint arXiv:2012.11241*, 2020.

[3] Jérôme Morio, Mathieu Balesdent, Damien Jacquemart, and Christelle Vergé. A survey of rare event simulation methods for static input–output models. *Simulation Modelling Practice and Theory*, 49:287–304, December 2014.

[4] Art Owen and Yi Zhou. Safe and Effective Importance Sampling. *Journal of the American Statistical Association*, 95(449):135–143, March 2000.

[5] Reuven Y. Rubinstein and Peter W. Glynn. How to Deal with the Curse of Dimensionality of Likelihood Ratios in Monte Carlo Simulation. *Stochastic Models*, 25(4):547–568, November 2009.

[6] Reuven Y Rubinstein and Dirk P Kroese. *The cross-entropy method: a unified approach to combinatorial optimization, Monte-Carlo simulation and machine learning*. Springer, New York; London, 2011. OCLC: 1064531163.

[7] Felipe Uribe, Iason Papaioannou, Youssef M. Marzouk, and Daniel Straub. Cross-entropy-based importance sampling with failure-informed dimension reduction for rare event simulation. *arXiv:2006.05496 [stat]*, June 2020. arXiv: 2006.05496.

[8] Ziqi Wang and Junho Song. Cross-entropy-based adaptive importance sampling using von Mises-Fisher mixture for high dimensional reliability analysis. *Structural Safety*, 59:42–52, March 2016.

Short biography – I graduated from University of Bordeaux in 2018 with a Master's degree in Statistics and Stochastic Simulation. I began my PhD thesis in October 2018 on sampling methods in reliability analysis, specifically rare event probability estimation in high dimension. The PhD is funded by ONERA and ISAE-SUPAERO.