Sensitivity analysis using multilevel Monte Carlo and surrogate-based control variates

Matthias, De Lozzo
Olivier, Sapin
IRT Saint Exupéry, Toulouse, France

Paul, Mycek
IRT Saint Exupéry, Toulouse, France
CECI, CNRS - CERFACS, Toulouse, France

Sensitivity analysis (SA) has become an essential part of the engineer's toolbox to analyze the variability of the output of a model and explain it from the different sources of uncertainty. However, the cost of simulators can be an obstacle to their use as SA techniques often implies to estimate statistics, *e.g.*, expectation, variance or sensitivity indices, using Monte Carlo (MC) sampling. An alternative consists of replacing the simulator by a surrogate model but when the input dimension increases, the curse of dimensionality degrades its quality and challenges its use. The same concern arises when using a lower fidelity model. Given these limitations, we propose to combine the best of both worlds: MC techniques to guarantee unbiasedness and multifidelity models and/or surrogate models to reduce the variance of the estimators.

Given a collection of numerical simulators with increasing accuracy and computational cost, [1] proposed the multilevel Monte Carlo (MLMC) technique to estimate the expectation unbiasedly. Then, MLMC methods have been extended to other statistics with algorithms designed to achieve a given precision. In [2], we proposed a unified MLMC framework where the unbiased MC estimator of the quantity of interest based on the finest level can be written as the telescoping sum of unbiased MC estimators. We applied this framework to the estimation of the covariance term of the pick-and-freeze estimator of a Sobol' index [3] and proposed an algorithm to allocate the sampling cost to the different fidelity models. The allocation rule is driven by the target computational cost, which may be more appropriate for engineering studies where one looks to reach the best accuracy under the constraint that the total simulation time is lower than a given requirement.

In [4], we proposed to combine MLMC techniques with control variates (CV) based on surrogate models to reduce the variance of the estimator. The CV method corrects the MC estimator with a term derived from auxiliary random variables that are highly correlated with the original random variable and we proved that using several control variates could not increase the variance. Based on this, we proposed to use the outputs of surrogate models as control variates, *e.g.*, Taylor polynomials (TP), Gaussian process (GP) regressors or polynomial chaos expansions (PCE) and illustrated this approach on an academic use case for which the simple use of a first-order TP can already improve the quality of the MC estimator of the expectation. We also proposed three extensions of this surrogate-based CV strategy to the multilevel framework. MLCV is presented as an extension of CV where the correction terms devised from surrogate models for simulators of different levels add up. MLMC-CV improves the MLMC estimator by using a CV based on a surrogate of the correction term at each level. Further variance reduction is achieved by using the surrogate-based CVs of all the levels in the MLMC-MLCV strategy. Although these techniques can be applied to an arbitrary statistic, we provided specific expressions for the expectation and the variance. In the case of the expectation, we also compared them in terms of accuracy and computational cost, depending on whether the construction of the surrogates, and the associated computational cost, precede the evaluation of the estimator.

Building on [4], we extend such estimators to sensitivity indices, *e.g.*, input-output correlation coefficients, Sobol' indices, derivative global sensitivity measures (DGSM) [5] and Hilbert-Schmidt Independence Criterion (HSIC) [6], as well as a generic framework for a wide family of estimators of sensitivity indices and a technique for the joint estimation of several sensitivity indices [7]. Finally, a Nastran-based mechanical use case with 50 uncertain parameters is used to assess these surrogate-based CV estimators of sensitivity indices and demonstrate the value of these variance reduction methods in engineering.

References:

[1] M.B. Giles. Multilevel Monte Carlo path simulation, Operations Research, **3**, pp. 607–617, 2008.

[2] P. Mycek and M. De Lozzo, "Multilevel Monte Carlo Covariance Estimation for the Computation of Sobol' Indices", *SIAM/ASA JUQ*, **7**(4):1323–1348, 2019.

[3] A. Janon, T. Klein, A. Lagnoux, M. Nodet, and C. Prieur, "Asymptotic normality and efficiency of two Sobol index estimators", ESAIM: PS, **18**, pp. 342–364, 2014.

[4] E.M. El Amri, P. Mycek, S. Ricci and M. De Lozzo, "Multilevel Surrogate-based Control Variates", *https://doi.org/10.48550/arXiv.2306.10800*, 2024.

[5] S. Kucherenko and B. Iooss. "Derivative-Based Global Sensitivity Measures". In: Ghanem, R., Higdon, D., Owhadi, H. (eds) Handbook of Uncertainty Quantification. Springer, Cham., 2015.

[6] S. Da Veiga. "Global sensitivity analysis with dependence measures", Journal of Statistical Computation and Simulation, **85**(7), 1283–1305, 2015.

[7] R. Y. Rubinstein and R. Marcus, "Efficiency of Multivariate Control Variates in Monte Carlo Simulation", Operations Research, **33**(3), 661–677, 1985.

[Matthias De Lozzo; IRT Saint Exupéry, Toulouse, France]
[`matthias.delozzo@irt-saintexupery.com` – `www.irt-saintexupery.com`]