Sampling strategy for stochastic simulators with heterogeneous noise.

Loïc Le Gratiet [†] [‡], Josselin Garnier [†]
[‡] CEA, DAM, DIF, F-91297 Arpajon, France
[†] Université Paris Diderot-Paris 7 75205 Paris Cedex 13

1 Presentation

I am a PhD student at the university of Denis-Diderot Paris VII and the CEA (the french atomic energy authority). I have a master degree at the university of Jean-Monnet of Saint-Etienne and I am an engineer graduate of the "Ecole Nationale Supérieure des Mines de Saint-Etienne". My PhD thesis - started on october 2010 - is in applied mathematics, statistics and probability area. My PhD advisor is Josselin Garnier, professor at the university of Denis-Diderot Paris VII and my CEA advisor is Claire Cannaméla, a doctor in applied mathematics from the university of Denis-Diderot Paris VII.

My PhD thesis - entitled "Multifidelity metamodelling and experimental design" - deals with the approximation of the output of large computer codes or costly real experiments, in order to design complex physical systems. Indeed, at CEA, some complex computer simulations last weeks or months and, in this case, building a classical surrogate model requires too many computer experiments to be reasonable.

2 Abstract

Kriging-based approximation is a useful tool to approximate the output of a Monte-Carlo simulator. Such simulator has the particularity to have a known relationship between its accuracy and its computational cost. Our objective is to find a sampling strategy optimizing the tradeoff between the fidelity and the number of simulations given a limited computational budget when the fidelity of the M-C simulator depends on the value of the input space parameter. This poster joins within the framework of the multi-fidelity metamodelling. Actually, at CEA, computer codes can usually be run at different levels of complexity and a hierarchy of s levels of code can hence be obtained. The aim of multi-fidelity metamodelling is to study the use of several levels of a code to predict the output of the most expensive one. The reader is refered to [Le Gratiet (2011)] for further detail about the multi-fidelity metamodelling.

The output provided by a Monte-Carlo simulator has the following form:

$$f_{N_j}(x) = \frac{1}{N_j}\sum_{i=1}^{N_j} Y_i(x)$$

where $Y_i(x)$ are independant random variables of variance $\tau^2(x)$. We therefore have $\mathrm{var}\left(f_N(x)\right) = \frac{\tau^2(x)}{N}$. The kriging predictive variance is given by :

$$s^2(x) = k(x,x) - k^T(x)(K+\Delta)^{-1}k(x)$$

with $k(x,y)$ a continuous symmetric positive definite kernel, $k^T(x) = (k(x,x_1),\dots,k(x,x_n))$, $D = (x_1,\dots,x_n)$ the experimental design and $(K+\Delta)$ is the matrix given by:

$$K+\Delta = [k(x_i,x_j)]_{1\le i,j\le n} + \left[\frac{\tau^2(x_i)}{N_i}\delta_{i=j}\right]_{1\le i,j\le n}$$

A widely used sampling design strategy consists in minimizing the Integrated Mean Squared Error:

$$\mathrm{IMSE} = \int_D s^2(x) d\mu(x)$$

where $\mu(x)$ is a measure on the input random space. If we consider a constant budget $T = \sum_{i=1}^n N_i$ and a given experimental design D, our first objective is to find the best allocation $(N_1, \ldots, N_n)$. In fact, we proved that under certain restricted conditions (i.e., when K is diagonal) the optimal allocation is given by:

$$N_i = \frac{1}{k(x,x)} \left(\frac{\sqrt{\tau^2(x_i)}}{\sum_{i=1}^n \sqrt{\tau^2(x_i)}} \left(k(x,x)T + \sum_{i=1}^n \tau^2(x_i) \right) - \tau^2(x_i) \right)$$

We numerically observe that this allocation remains efficient in more general cases although it is not anymore optimal.

Our second objective is to determine the model improvement if we increase the budget T. In fact if we consider the Karhunen-Loève decomposition of $k(x,y)$:

$$k(x,y) = \sum_{p \geq 0} \lambda_p \phi_p(x) \phi_p(y)$$

and if the eigenvalues λ_p satisfy the asymptotic behavior $\lambda_p = \mathcal{O}\left(\frac{1}{p^\alpha}\right)$, when $p >> 1$ with $\alpha > 1$. Then the IMSE decreases as $\frac{1}{T^{1-\frac{1}{\alpha}}}$. Therefore, the improvement depends on the kernel function $k(x,y)$. The presented results are a generalization in the non-degenerate case of the results presented in the thesis of [V. Picheny (2009)] and in [Rasmussen & Williams (2006)].

Examples

- For a fractional Brownian motion with Hurst coefficient H, we have $\alpha = 2H + 1$.
- For a 1-D Matèrn kernel of regularity parameter ν, we have $\alpha = 2\nu$.
- For a 1-D Gaussian kernel we have the asymptotic $\lambda_p = \mathcal{O}(\exp(-p))$ and the IMSE decreases as $\frac{\log(T)}{T}$.

We see that the IMSE decay rate depends on the regularity of the kernel function. The asymptotic behaviour of the eigenvalues for a fractional Brownian motion is presented in [Bronski (2003)], for the Matèrn class covariance function in [Pusev (2011)] and for the Gaussian kernel function in [Schwab & Todor (2006)].

References

[Le Gratiet (2011)] Le Gratiet, L. 2011 Bayesian analysis of hierarchical multi-fidelity codes. *Arxiv preprint arXiv:1112.5389, 2011*

[Schwab & Todor (2006)] Schwab, C. and Todor, R.A. 2006 Karhunen-Loève Approximation of Random Fields by Generalized Fast Multipole Methods. *Research Report No. 2006-01*

[Bronski (2003)] Bronski, J.C 2003 Asymptotics of Karhunen-Loeve Eigenvalues and Tight Constants for Probability Distributions of Passive Scalar Transport. *Commun. Math. Phys.* **238**, 563-582.

[Pusev (2011)] Pusev R.S. 2011 Small Deviation Asymptotics for Matèrn Processes and Fields under Weighted Quadratic Norm. *Theory Probab. Appl.* **55**, No. 1, 164-172.

[V. Picheny (2009)] V. Picheny, *Improving Accuracy and Compensating for uncertainty in Surrogate Modeling*, Ecole Nationale Supérieure des Mines de Saint Etienne, 2009.

[Rasmussen & Williams (2006)] Rasmussen, C. E. & Williams, C. K. I. 2006 *Gaussian Processes for Machine Learning*. the MIT Press.