Bayesian analysis of hierarchical codes with different levels of accuracy.
Analyse bayésienne de codes hiérarchiques avec différents niveaux de précision.

Loïc Le Gratiet
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

We focus here on the Gaussian process regression metamodel and on its extension to multiple response models. 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 this presentation is to study the use of several levels of a code to predict the output of the most expensive one. The method presented here can also be used to predict the result of an experiment by considering it as a complex code and by using the simulation as a lower complexity code.

A first metamodel for multi-level computer codes was built by [Kennedy & O'Hagan (2000)] using a spatial autocorrelation structure. This multi-stage model is a particular case of co-kriging which is a well known geostatistical method. Then, [Qian et al. (2006)] built an extension to this model in a case of non spatial stationarity and [Forrester, Sobester & Keane (2007)] went into more detail about the estimation of the model parameters. Furthermore, Forrester *et al.* presented the use of co-kriging for multi-fidelity optimization based on the EGO (Efficient Global Optimization) algorithm. A Bayesian approach was also proposed by [Qian & Wu (2008)].

We will present here a new approach to estimate the parameters of the model which is effective in the case of non-spatial stationarity and for large value of s. Furthermore, this approach allows us to consider prior information in the estimation of the parameters. Then, when the value of s is large, it could be an issue to invert the covariance matrix of te co-kriging. We will provide a solution to this problem which shows that the inverse can be easily calculated. Finally, we know that with a non-Bayesian approach, the variance of the predictive distribution could be underestimated

[Kennedy & O'Hagan (2000)]. We will suggest a Bayesian modelling different from the one presented in [Qian & Wu (2008)] and which avoids prohibitive implementation.

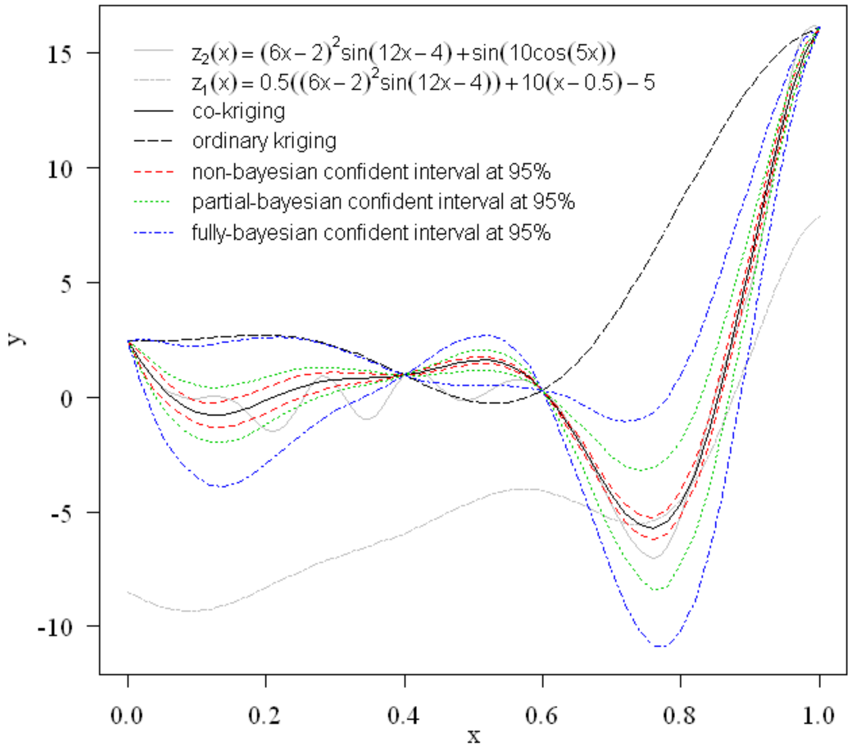

Figure 1: A co-kriging example. $z_2(.)$ represents the expensive code and $z_1(.)$ represents the cheap code. The co-kriging significantly improves the ordinary kriging metamodel.

3 Important result

For a non-Bayesian s-levels co-kriging, we proved that building an s-levels co-kriging is equivalent to build s independent krigings. This result is very important since it solves two of the most important key issues of the co-kriging which are the inversion of the covariance matrix and the estimation of the parameters.

References

[Kennedy & O'Hagan (2000)] M. C. Kennedy and A. O'Hagan, *Predicting the output from a complex computer code when fast approximations are available*, Biometrika, 87, 1, pp. 1-13, 2000.

[Forrester, Sobester & Keane (2007)] A. I. J. Forrester, A. Sobester and A. J. Keane, *Multi-fidelity optimization via surrogate modelling*, Proc. R. Soc. A 463, pp. 3251-3269, 2007.

[Qian et al. (2006)] Z. Qian, C. C. Seepersad, V. R. Joseph, J. K. Allen and C. F. Jeff Wu, *Building Surrogate Models Based on Detailed and Approximate Simulations*, ASME, Vol. 128, July, 2006.

[Qian & Wu (2008)] P.Z.G. Qian and C. F. Jeff Wu, *Bayesian Hierarchical Modeling for Integrating Low-accuracy and High-accuracy Experiments*, Technometrics, Vol.50, NO. 2, 2008.