

PhD thesis offer: Metamodelling under inequality constraints

Chair in Applied Mathematics OQUAIDO

Context

This PhD thesis offer is associated to the Chair in Applied Mathematics OQUAIDO of Mines Saint-Étienne (2016 – 2019). The Chair, which follows the ReDICE consortium (www.redice-project.org), gathers partners from academia, technological research and industry around innovative mathematical methods for the Design and Analysis of Computer Experiments (DACE). The partners of the Chair are:

- Mines Saint-Etienne, Ecole Centrale de Lyon, Univ. of Nice, Univ. of Toulouse (academia)
- BRGM, CEA, IFPEN, IRSN, Safran, Storengy (technological research and industry)

The DACE research domain is devoted to exploit industrial computer codes that are costly to evaluate. Such computer codes typically model complex physical phenomena such as encountered in the development of new technologies or risk studies. The partners of the Chair use them in various sectors such as energy, environment and transport.

The Chair proposes an original and dynamical framework to develop innovative research, and fosters interaction between all its members: researchers, engineers, post-doctoral and PhD students. Expected developments are methodological, but guided by case studies from real-life problems. The Chair is funding 1 post-doc and 2 PhD thesis starting in 2016. This offer is one of them.

Research lines and aim

Among the research lines with strong methodological interest and promising application potential, the Chair scientific committee identified the case of models with inequality constraints, such as boundary or monotonicity. Indeed in practice, underlying physical phenomena are often known to be monotonic with respect to one or several variables, or to belong to a range of possible values. Taking into account that information may lead to more accurate results. In addition to this main line, another axis regarding constraints can be brought by the PhD candidate, depending on its master work or past experience in DACE.

The aim of this 3-year research project is twofold:

- Methodology: To improve existing works for real-life problems, involving tens of input variables.
- Application: To apply the proposed new methodology to at least 1 case study of the Chair.

Among starting points on inequality constraints, the PhD works should investigate the extension of a promising Gaussian process-based model (Maatouk, Bay, 2014), which satisfies inequality constraints (boundary, inequality, convexity) everywhere in the space, contrarily to other approaches based on spatial discretization. At this stage, the model has two limitations: Its application is limited to a small number of constraints (1, 2 or 3) because of the tensorization technique used, and parameter estimation is based on a time-consuming cross-validation. More generally, the PhD methodological developments may develop Gaussian process-based models in 3 main directions:

1. Estimation under inequality constraints.
2. Extension to higher dimensions.
3. Case of a high number of observations (consequence of the previous point).

Supervising team

The supervising team gathers researchers from 2 sites :

- Olivier Roustant and Nicolas Durrande, Mines Saint-Etienne
- François Bachoc, Institut de Mathématiques de Toulouse

Applicant profile

Candidates should have completed a Master in applied mathematics, statistics, machine learning, or related disciplines. The applicant should demonstrate both theoretical and computational skills. Implementations in R are expected.

CV and motivation letter in English or French should be sent to the coordinators of the Chair (Olivier Roustant and Nicolas Durrande) using the e-mail address: oquaido@emse.fr

Conditions

- Date/duration: The position is a 3-year contract, expected to start in September 2016.
- Location: The main location is Mines Saint-Étienne, France. Visits at the Institut de Mathématiques de Toulouse are expected.
- Participation to teaching: 40 hours / year (in French or English)
- Net salary: 1630 €/month

References

- S. Da Veiga and A. Marrel (2012), “Gaussian process modeling with inequality constraints”, *Annales de la Faculté des Sciences de Toulouse*, **21**, p. 529–555.
- S. Golchi, D. R. Bingham, H. Chipman, and D. A. Campbell (2015), “Monotone Emulation of Computer Experiments”, *SIAM/ASA Journal on Uncertainty Quantification*, **3** (1), p. 370-392
- H. Maatouk and X. Bay (2014), “Gaussian Process emulators for computer experiments with inequality constraints”, in revision for *SIAM/ASA Journal on Uncertainty Quantification*.
- Robert B. Gramacy & Daniel W. Apley (2015), “Local Gaussian process approximation for large computer experiments”, *Journal of Computational and Graphical Statistics*, **24** (2), p. 561-578.
- Deisenroth, Marc Peter and Ng, Jun Wei (2015), “Distributed Gaussian Processes”, *Proceedings of the 32nd International Conference on Machine Learning*, Lille, France. *JMLR: W&CP* volume 37.
- François Bachoc (2014), “Asymptotic analysis of the role of spatial sampling for covariance parameter estimation of Gaussian processes”, *Journal of Multivariate Analysis*, **125**, p. 1-35.



PhD offer:

High-dimensional optimization of the internal and external aerodynamics of a vehicle

Period: October 2016 - October 2019

Location: Saint-Etienne, frequent stays in the Paris and Toulouse areas.

Funding:

- CIFRE Groupe PSA (Industrial Agreement for Training through Research) - The CIFRE fellow will sign a 3 years full time work contract with Groupe PSA.
- Year gross salary: of the order of 30Keuros.

Advisors: Rodolphe Le Riche (CNRS and Ecole des Mines de St-Etienne), Victor Picheny (INRA Toulouse), Benoît Eaux and Clément Dumand (Groupe Peugeot Société Anonyme)

Contacts: Rodolphe Le Riche (leriche@emse.fr), Victor Picheny (Victor.Picheny@toulouse.inra.fr), Benoît Eaux (benoit.eaux@mpsa.com).

Applicant Profile: Candidates should have completed a Master in applied mathematics, statistics, machine learning, or engineering with a mathematical background. The applicant should demonstrate both theoretical and computational skills. Implementations in R are expected.

Context: Committed to reducing the fuel consumption and emissions of its vehicles, Groupe PSA innovates continuously to develop ever cleaner and energy-efficient models. As an integral part of its advanced research process for developing new engines and external aerodynamics solutions, Groupe PSA has been using 3D numerical simulations for many years. To handle both complexity of the phenomena that occurs in these simulations and optimization of antagonist objectives, multi-criteria optimization methods are needed.

Objectives of the PhD: The objective of this PhD is to devise and validate a multi-objective optimization method of 3D aerodynamic simulations. In this context, optimization problems are particularly challenging to tackle, due to the non-linearity of the simulation outputs, the computational cost of a single simulation (from 12 hours to three days), the need to handle complex constraints, or the large number of parameters to optimize over (20 to 40 unknowns).

Recent approaches based on Gaussian processes have shown promising results, but address this problem only in part: see [1,2,3,4] for examples. Projections in subspaces have been proposed to handle large number of unknowns [5]. For this PhD, the investigated methodology will follow this line of work, and will focus in particular on:

- efficient sampling strategy in the presence of several objectives and constraints,
- an automated reduction in dimension in order to deal with many unknowns with a restricted number of simulations,
- the possibility to distribute calculations in an asynchronous fashion on about 10 computing nodes, and
- dealing with the event of simulation failures.

The approach developed will be validated on two industrial cases dealing with shape optimization:

- optimization of the chamber design of an internal combustion engine,
- optimization of the external shape of a vehicle [6,7,8].

In these cases, the unknowns are design parameters (via CAD tools), and drag coefficient or combustion efficiency are examples of objectives and constraints. The method will be programmed in the R language.

Bibliography

- [1] M. Binois and V. Picheny, "GPareto: An R Package for Gaussian-Process Based Multi-Objective Optimization and Analysis", 2016, https://cran.r-project.org/web/packages/GPareto/vignettes/GPareto_vignette.pdf
- [2] V. Picheny, "Multiobjective optimization using Gaussian process emulators via stepwise uncertainty reduction", *Statistics and Computing*, Volume 25, Issue 6, pp. 1265-1280, 2015
- [3] Paul Feliot, Julien Bect, Emmanuel Vazquez, "A Bayesian approach to constrained single- and multi-objective optimization", *Learning and Intelligent Optimization: 9th International Conference, LION 9, Lille, France, 2015*
- [4] Wagner, Tobias, et al. "On expected-improvement criteria for model-based multi-objective optimization." *Parallel Problem Solving from Nature, PPSN XI*. Springer Berlin Heidelberg, 2010. 718-727.
- [5] Z. Wang, M. Zoghi, F. Hutter, D. Matheson and N. De Freitas, "Bayesian Optimization in High Dimensions via Random Embeddings", *IJCAI conf.*, 2013.
- [6] L. Cui, T. Wang, Z. Lu, M. Jia and Y. Sun, "Full-Parameter Approach for the Intake Port Design of a Four-Valve Direct-Injection Gasoline Engine". *ASME. J. Eng. Gas Turbines Power*. 2015;137(9).
- [7] M. Yagoubi, L. Thobois, M. Schoenauer, "An asynchronous steady-state NSGA-II algorithm for multi-objective optimization of Diesel combustion", *Int. Conference of Engineering Optimization*, 2010.
- [8] John B. Heywood. "Internal Combustion Engine Fundamentals". *McGraw-Hill*, 1988.

PhD Offer

Metamodeling and robust optimization – application to ideotype design under climatic uncertainty

General information

- Period : October 2016 – October 2019
- Location: MIA-T, INRA (Toulouse)
- Advisors: Victor Picheny and Robert Faivre (INRA Toulouse) and Aurélien Garivier (IMT, Université Paul Sabatier, Toulouse)
- Contacts : victor.picheny@toulouse.inra.fr, robert.faivre@toulouse.inra.fr, aurelien.garivier@math.univ-toulouse.fr
- Funding : of the order of 24K€ / year gross salary

Applicant Profile

Candidates should have completed a Master in applied mathematics, statistics, machine learning, or engineering with a mathematical background. The applicant should demonstrate both theoretical and computational skills. Implementations in R are expected.

Key-words

Surrogate model – Gaussian process- Bayesian optimization – EGO – computer experiments – noise

Context and motivation

Using complex numerical models for prediction, risk assessment or design has become an essential tool in numerous fields, including engineering, economy or in natural sciences. Often times, the modeled systems depend on uncertain environments that greatly influence the quantities of interest : typically in agronomy, an annual yield or environmental impact depend on the climatic conditions at hand [1]. It is then necessary to design methods that account for uncertain environments, in particular by integrating risk-aversion concepts.

This PhD focuses on metamodel-based optimization (also referred to as Bayesian optimization), and in particular on Gaussian process regression, which efficiency has been empirically demonstrated numerous times over the last 15 years [2]. The research domain is hence multi-disciplinary, at the intersection between applied mathematics, statistics and *machine learning*.

PhD objectives

Metamodel-based optimization under uncertainty has retained a growing attention over the past few years, following the emergence of complex simulators that depend on stochastic processes. However, if several solutions have been proposed when average performances are optimized [3], solutions are scarce as soon as risk-aversion strategies are sought.

The objective of this PhD is first to propose new metamodeling approaches allowing to consider either several risk levels at once, or the entire distribution of the quantity of interest. Among possible starting points, the student may focus on vectorial prediction for several risk indices (*co-kriging*, [4]), or functional predictors [5]. Secondly, sampling strategies based on these metamodels may be developed to achieve optimization. Those approaches could either be heuristic (in the spirit of [2,3]) or based on recent convergence results [6,7].

The approaches developed may be validated on an agronomical problem : the ideotype design of sunflower crops, based on the SUNFLO model [8].

References

- [1] Debaeke P. et Quilot-Turion B. (eds) (2014). Conception d'idéotypes de plantes pour une agriculture durable. Collection École-chercheurs INRA, FormaSciences, FPN, INRA-CIRAD (ISBN 2-7380-1347-3), 254p.
- [2] Jones, D. R. (2001). A taxonomy of global optimization methods based on response surfaces. *Journal of global optimization* 21, 345–383
- [3] Picheny, V., Wagner, T., Ginsbourger, D., (2013). A benchmark of kriging-based infill criteria for noisy optimization. *Structural and Multidisciplinary Optimization*, 48(3), 607-626
- [4] Rougier, J., (2008). Efficient emulators for multivariate deterministic functions. *Journal of Computational and Graphical Statistics*, 17(4), 827-843
- [5] Plumlee, Matthew, and Rui Tuo. "Building accurate emulators for stochastic simulations via quantile Kriging." *Technometrics* 56.4 (2014): 466-473.
- [6] Srinivas N., Krause A., Kakade S. and Seeger M. (2012), Information-Theoretic Regret Bounds for Gaussian Process Optimization in the Bandit Setting, *IEEE Transactions on Information Theory*, 58(5),3250-3265
- [7] Bubeck S., Munos R., Stoltz G., Szepesvari C. (2011), X-armed bandits, *Journal of Machine Learning Research*, 12, 1655-1695
- [8] Casadebaig P., Guilioni L., Lecoeur J., Christophe A , Champolivier L., Debaeke P., (2010). SUNFLO, a model to simulate genotype-specific performance of sunflower crop in contrasting environments. *Agricultural Forest Meteorology*, 151(2), 163-178

PhD thesis offer (CIFRE)

Optimization under uncertainty – Application to robust design

Entity

Safran is an international high-technology group, a leading manufacturer in the fields of Aerospace (propulsion and equipment), Defense and Security. Operating worldwide, the Safran Group has 69,000 employees and logged sales of 15.4 billion euros in 2014. Comprising a number of companies, the Group holds, alone or in partnership, world or European leadership positions in its markets. To keep pace with its fast-evolving markets, Safran undertakes extensive research and development programs which, in 2014, totaled investments of €2 billion. Safran is listed on Euronext Paris and on the CAC 40 index.

Context

In the general field of optimization, optimization under uncertainty aims at accounting for the potential variability of some design parameters in the original optimization problem. Such variability may come from limited knowledge on physical parameters (preliminary design stage) or uncertainty on external phenomena (environment, precise properties of materials). The worst-case approach was originally introduced in the 50's, where all the uncertain parameters are fixed at their worst-case value. However the conservatism of the solution will discard designs that are less optimal in the worst-case sense but are superior in the most likely scenarios. In this PhD thesis, the goal is to adopt an innovative point of view for optimization under uncertainty.

Research focus

The goal of the PhD thesis is to address several practical limitations arising when an industrial robust design problem must be modeled as a mathematical optimization problem. More precisely, two main lines of research have been identified:

1) *Formulation of a robust design problem*

The biggest challenge in optimization under uncertainty for industrial systems consists in expressing the operational needs into a standard problem with objectives and constraints. For example, does robust optimization imply optimizing the mean of the performance criterion when the uncertain parameters vary? Its standard deviation? Or maybe a compromise between the two? Depending on this choice, the optimal design will certainly be different.

We propose here to investigate multi-objective formulations, which would account for several robust criteria. Ideally the obtained Pareto front would include solutions to various formulations, thus permitting to identify which formulations lead to particular areas of the Pareto front and what kind of optimal designs they imply. The objective will be to build such a multi-objective problem and study from a theoretical point of view the link between the Pareto front and the robust formulations.

2) *Sensitivity analysis*

In industrial cases, the number of design variables and uncertain parameters is usually large, ranging from a few tens to several hundreds. In order to facilitate both the exploration and the optimization of such systems, it is generally mandatory to reduce the number of variables to keep the most influential ones only.

Recent advances make it possible to perform this selection efficiently when the goal is to identify which variables impact the mean level of a criterion. However they perform poorly for optimization problems, where it would be more relevant to detect the variables that influence the areas where the criterion tends to be optimal. Similarly, in robust optimization the question is to define the impact of an uncertain parameter on the probability constraints. This research line will focus on the definition of new sensitivity indices dedicated to optimization and extend them to the optimization under uncertainty framework.

Supervising team & practical details

The PhD thesis is a CIFRE contract proposed by Safran Tech, the new corporate research center of Safran located in Magny-Les-Hameaux, "Plateau de Saclay" area near Paris, France.

The supervising team will include

- Sébastien Da Veiga, researcher at Safran
- Rodolphe Le Riche & Olivier Roustant, Mines Saint-Etienne

Applicant profile

Candidates should have completed a Master in Applied Mathematics, Operational Research or related disciplines. The applicant should demonstrate both theoretical and computational skills. CV and cover letter in English or French should be sent to sebastien.daveiga@safran.fr