UQOP 2019
Uncertainty Quantification & Optimization Conference
18-20 March, Paris, France

Optimization under Uncertainty of High Dimensional Problems using Quantile Bayesian Regression

Christian Sabater[1], Olivier Le Maître [2] and Pietro Congedo [3]

1 DLR, German Aerospace Center, christian.sabatercampomanes@dlr.de
2 LIMSI, Laboratoire d'Informatique pour la Mécanique et les Sciences de l'Ingénieur
3 DeFI Team (INRIA Saclay IDF, Ecole Polytechnique), CMAP Laboratory

Keywords: Bayesian Optimization, Quantile Regression, Optimization under Uncertainty.

ABSTRACT

The use of robust optimization techniques is increasing in popularity in order to come up with configurations less sensitive to aleatory uncertainties. Instead of optimizing the Quantity of Interest, QoI, an statistic of the QoI is sought. The quantile is a flexible statistic to be chosen as objective function in engineering problems. For example a high quantile can be minimized to reduce extreme events. If the median is chosen, day to day events are minimized.

Traditionally, an uncoupled approach can be used, in which at each iteration of the optimization process a complete uncertainty quantification is performed to obtain the statistic of the QoI. To reduce the computational cost, the stochastic space can be characterized through surrogate models. Then, a large number of Quasi Monte Carlo samples can be cheaply evaluated in the surrogate.

However, when dealing with a large number of uncertainties, the required number of training samples to construct an accurate surrogate increases exponentially. Also, surrogate models are not suitable to model non-parametric uncertainties. Another disadvantage is that the uncoupled approach does not fully explore the intrinsic relationship between uncertainties and design parameters. Samples used to construct the surrogate at each optimization are not reused. The combination of uncertainties with design parameters is expected to improve the efficiency of the optimization. The objective of this paper is to develop coupled framework for optimization under uncertainty that is insensitive to the number of uncertainties by means of a Bayesian approach.

A global Design of Experiments, DoE, is selected, covering both the design and uncertainty space. This initial sampling is chosen by Quasi Monte Carlo because of its good capability of filling the design space. All the realizations are then projected into a single plane, in order to obtain the response of all the realizations only as a function of the design parameters, X.

Then, quantile regression is applied to the initial DoE sampling [1]. The objective is to obtain an explicit relation between the quantile and the design parameters. This relation is parametrized through a Thin Plane Spline, TPS, [2] that is characterized through the X and Y coordinates of its control points. These control points are uniformly spaced in X, the design space. The Y coordinates

are selected according to the minimization of a loss function (maximization of the likelihood) following quantile regression. As a result the TPS provides an initial predictor of the quantile as a function of the design parameters.

Due to the trade-off between model complexity and model accuracy, there is an optimum number of control points to define the TPS. By increasing the number of control points, the loss function would be minimized more and more. However, risk of overfitting could occur and additional sampling points would not be properly predicted. The initial number of control is then selected by k-fold cross validation on the loss function of quantile regression.

In addition, it is desired to have an estimation of the error of the quantile predicted by the TPS at any given point. The formulation of quantile regression can be considered from a Bayesian point of view through the Asymmetric Laplace Distribution and the assumption of an improper uniform prior [3]. The posterior distribution of the Y coordinates of the control points is obtained by sampling the likelihood function through Markov Chain Monte Carlo. In particular, an adaptive Metropolis-Hastings algorithm is used. The covariance of the step function is adapted according to the covariance of the chain. From the MCMC sampling it is possible to obtain the posterior distribution of the Y coordinate of each control point of the TPS.

The addition of infill samples should balance exploration in areas with large uncertainty in the estimation of the quantile with exploitation in locations close to the minimum. Once the model is built, several realizations from the MCMC chain are randomly chosen. New samples are placed in the minimum location of each realization (in the design space), at a random location in the stochastic space. After the infill, the predictor is recomputed again until convergence occurs in either the optimum location of the design parameter or in the optimum quantile value.

In addition, it is possible to increase the model complexity by adding additional control points to build the TPS in locations closer to the local minima. This should be done as far as the increase in model accuracy does not penalize the model complexity.

The framework has been validated in a one dimensional test function with two local minima for the optimization of the 80% quantile. The use of cross validation results in an optimum number of 21 points to define the TPS. Bayesian quantile regression is a powerful tool to globally estimate the quantile and its error as a function of the design parameters. The predicted quantile converges globally to the real one as the number of samples increases. The infill criteria can be used improve the accuracy in the region of the local minima and accelerate the optimization process.

In conclusion, an optimization under uncertainty framework insensitive to the number of uncertainties has been presented. The application of Bayesian Quantile Regression makes possible the estimation of the quantile and its prediction error at any given design point. This can be used for sequential optimization under uncertainty. In the future, the framework will be applied to an engineering problem with a large number of manufacturing tolerances

References

[1] Brian S. Cade and Barry R. Noon. A gentle introduction to quantile regression for ecologists. *Frontiers in Ecology and the Environment*, 1(8):412–420, oct 2003.

[2] Jean Duchon. Splines minimizing rotation-invariant semi-norms in sobolev spaces. In *Constructive Theory of Functions of Several Variables*, pages 85–100. Springer Berlin Heidelberg, 1977.

[3] Keming Yu and Rana A. Moyeed. Bayesian quantile regression. *Statistics & Probability Letters*, 54(4):437–447, oct 2001.