Deep Gaussian processes for analysis and optimization of aerospace systems.

A. Hebbal [1,2], L. Brevault [1], M.Balesdent[1]
E-G. Talbi[2], N. Melab[2]

[1]ONERA - The French Aerospace Lab
[2]Université de Lille, CNRS/CRIStAL, Inria Lille

Journée des doctorants ONERA 2020

ONERA
THE FRENCH AEROSPACE LAB

Inria
INVENTEURS DU MONDE NUMÉRIQUE

Université de Lille
1 SCIENCES ET TECHNOLOGIES

Table of Contents

Introduction

Review on Bayesian Optimization

Bayesian optimization for non stationary functions

Multi-objective Bayesian Optimization

Multi fidelity with varying input spaces

Conclusions

Context

Analysis and optimization of design problems characterized by :

- Black box and computationally expensive functions,
- Non-stationary functions,
- Multi-objective problems with correlated objectives,
- Multi-fidelity problems with varying input space dimensions.

Context

Analysis and optimization of design problems characterized by :

- Black box and computationally expensive functions,
- Non-stationary functions,
- Multi-objective problems with correlated objectives,
- Multi-fidelity problems with varying input space dimensions.

✓ Analysis and optimization using surrogate models

Context

Analysis and optimization of design problems characterized by :

- Black box and computationally expensive functions,
- Non-stationary functions,
- Multi-objective problems with correlated objectives,
- Multi-fidelity problems with varying input space dimensions.

Abrupt change in the function

Non stationary modelisation of a constraint

Variation of the smoothness along the input space

Non stationary objective function

Context

Analysis and optimization of design problems characterized by :

- Black box and computationally expensive functions,
- Non-stationary functions,
- Multi-objective problems with correlated objectives,
- Multi-fidelity problems with varying input space dimensions.

Maximize f_1 : The payload value.

Minimize f_2 : The gross lift off-weight value.

Context

Analysis and optimization of design problems characterized by :

- Black box and computationally expensive functions,
- Non-stationary functions,
- Multi-objective problems with correlated objectives,
- Multi-fidelity problems with varying input space dimensions.

Model of launch vehicle thrust frame.

Wing geometry and its representation in LF and HF codes [Tao et al., 2019]

Table of Contents

Introduction

Review on Bayesian Optimization
- Gaussian Processes
- BO framework

Bayesian optimization for non stationary functions

Multi-objective Bayesian Optimization

Multi fidelity with varying input spaces

Conclusions

Gaussian Process Regression

Gaussian process [Rasmussen, 2004]

A Gaussian Process is used to describe a distribution over function. It is a collection of infinite random variables, **any finite number of which have a joint Gaussian distribution**.
It is defined by its mean function and covariance function (Kernel) :
$f(.) \sim \mathcal{GP}(\mu(.), k^{\Theta}(.))$

BO framework

Bayesian Optimization (BO)

Limits of GPs

Handling non stationarity

- Abrupt change in the response
- Different behavior according to the input

2-D Function with an abrupt change

GP Approximation

Limits of GPs

Handling non stationarity

- Abrupt change in the response
- Different behavior according to the input

1-D Function with a different behaviour according to the input space

Table of Contents

Introduction

Review on Bayesian Optimization

Bayesian optimization for non stationary functions
- Review of methods for handling non stationarity in BO
- Deep Gaussian Processes
- Coupling of BO and DGPs
- Experimentations

Multi-objective Bayesian Optimization

Multi fidelity with varying input spaces

Conclusions

Review of methods of handling non stationarity in BO

- Direct formulation of non-stationary covariance function [Higdon et al., 1999] [Paciorek and Schervish, 2006]
- Local stationary covariance function [Haas, 1990] [Rasmussen and Ghahramani, 2002]
- Non-linear mapping [Xiong et al., 2007]
- Curse of dimensionality,
- Scarce data.

Review of methods of handling non stationarity in BO

- Direct formulation of non-stationary covariance function [Higdon et al., 1999] [Paciorek and Schervish, 2006]
- Local stationary covariance function [Haas, 1990] [Rasmussen and Ghahramani, 2002]
- Non-linear mapping [Xiong et al., 2007]
- Curse of dimensionality,
- Scarce data.

Deep Gaussian Processes

Deep Gaussian Processes [Damianou and Lawrence, 2013]

DGPs are a class of surrogate models based on the structure of neural networks, where each layer is a GP. It considers that the statistical relationship between the inputs and the response is expressed by a functional composition of GPs :

$$y = f_L(\mathbf{f}_{L-1}(...(\mathbf{f}_1(\mathbf{f}_0(\mathbf{x}))) + \epsilon_L$$

Deep Gaussian Processes

$\mathbf{f}_0 \sim \mathcal{GP}(0, K_{\mathcal{X}\mathcal{X}} + \sigma_0^2 \mathbf{I})$ $\mathbf{f}_1 \sim \mathcal{GP}(0, K_{\mathbf{h}_1\mathbf{h}_1} + \sigma_1^2 \mathbf{I})$ $\mathbf{f}_L \sim \mathcal{GP}(0, K_{\mathbf{h}_L\mathbf{h}_L} + \sigma_L^2 \mathbf{I})$

A deterministic observed variable

A distribution with **Non**-observed instantiations

A distribution with observed instantiations

Deep Gaussian Processes

GP approximation of a non-stationary 1-D function. The GP model can not capture the stability of the region $[0.4, 1]$ and continues to oscillate

DGP approximation of a non-stationary 1-D function. The DGP model appropriately capture the two regions with different smoothness

Coupling of BO and DGPs

BO and DGPs

DGPs allow a flexible way of Bayesian kernel construction through input warping and dimensionality expansion to better fit data. Hence, its use for BO of non-stationary problems can be interesting. However, the integration of DGPs into the framework of BO is not direct.

Coupling of BO and DGPs

BO and DGPs

DGPs allow a flexible way of Bayesian kernel construction through input warping and dimensionality expansion to better fit data. Hence, its use for BO of non-stationary problems can be interesting. However, the integration of DGPs into the framework of BO is not direct.

The limitation of DGPs : **uncertainty quantification**

Standard deviation on the prediction given by a 2 hidden layers DGP model trained using Adam Optimizer.

Coupling of BO and DGPs

$\mathbf{f}_0 \sim \mathcal{GP}(0, K_{\mathcal{X}\mathcal{X}} + \sigma_0^2 \mathbf{I})$ $\mathbf{f}_1 \sim \mathcal{GP}(0, K_{\mathbf{h}_1\mathbf{h}_1} + \sigma_1^2 \mathbf{I})$ $\mathbf{f}_L \sim \mathcal{GP}(0, K_{\mathbf{h}_L\mathbf{h}_L} + \sigma_L^2 \mathbf{I})$

Combining natural and stochastic gradient for training

✗ Issue : Parameter space not euclidian ($q(\mathbf{u})$) $\rightarrow$ stochastic ordinary gradient may not be adapted

✓ Proposed optimization process adapted to the parameter space :

Loop between :

- step using stochastic ordinary gradient (Adam Optimizer) with respect to $\left(\{\Theta\}_{l=1}^{s}, \{Z_l\}_{l=1}^{s}\right)$
- step using natural gradient with respect to all the variational distributions $\left(q\{(\mathbf{u}_l)\}_{l=1}^{s}\right)$.

Coupling of BO and DGPs

Combining natural and stochastic gradient for training

✗ Issue : Parameter space not euclidian ($q(\mathbf{u})$) $\rightarrow$ stochastic ordinary gradient may not be adapted

✓ Proposed optimization process adapted to the parameter space :

Loop between :

- step using stochastic ordinary gradient (Adam Optimizer) with respect to $\left(\{\Theta\}_{l=1}^{s}, \{Z_l\}_{l=1}^{s}\right)$
- step using natural gradient with respect to all the variational distributions $\left(q\{(\mathbf{u}_l)\}_{l=1}^{s}\right)$.

.

Hebbal, A., Brevault, L., Balesdent, M., Talbi, E. G., and Melab, N. (2019). Bayesian Optimization using Deep Gaussian Processes. arXiv preprint arXiv :1905.03350.

Hebbal, A., Brevault, L., Balesdent, M., Talbi, E., and Melab, N., Multi-fidelity modeling using DGPs : Improvements and a generalization to varying input space dimensions à 4th workshop on Bayesian Deep Learning (NeurIPS 2019), Vancouver, Canada.

Coupling of BO and DGPs

Standard deviation on the prediction given by a 2 hidden layers DGP model trained using Adam Optimizer.

Standard deviation on the prediction given by a 2 hidden layers DGP model trained using the proposed optimization approach.

Coupling of BO and DGPs

BO with DGPs

Hebbal, A., Brevault, L., Balesdent, M., Talbi, E. G., and Melab, N. (2019). Bayesian Optimization using Deep Gaussian Processes. arXiv preprint arXiv :1905.03350.

Hebbal, A., Brevault, L., Balesdent, M., Talbi, E., and Melab, N., Efficient Global Optimization with Deep Gaussian Processes," Evolutionary Computation (CEC), 2018 IEEE Congress.

Application to industrial test case : design of aerospace vehicle

FIGURE – Optimization problem of a solid-propellant booster engine.

Application to industrial test case : design of aerospace vehicle

A sectional view of the change in velocity according to the diameters of the nozzle.

A sectional view of a constraint according to the diameters of the nozzle.

FIGURE – Sectional view of the non-stationary behaviors of some functions involved in the booster problem

Application to industrial test case : design of aerospace vehicle

FIGURE – Convergence curve of $-\Delta V$ using DEGO and BO & GP.

Table of Contents

Introduction

Review on Bayesian Optimization

Bayesian optimization for non stationary functions

Multi-objective Bayesian Optimization
- MO-BO with independent objectives
- MO-DGP model
- Experimentations

Multi fidelity with varying input spaces

Conclusions

MO-BO with DGPs

Maximize f_1 : The payload value.

Minimize f_2 : The gross lift off-weight value.

MO-BO with DGPs

Maximize f_1 : The payload value.

Minimize f_2 : The gross lift off-weight value.

MO-BO with DGPs

Maximize f_1 : The payload value.

Minimize f_2 : The gross lift off-weight value.

Multi-Objective Bayesian Optimization (MO-BO) [Emmerich et al., 2006]

Initial DoE | 30 added points | 50 added points

Hebbal, A., Brevault, L., Balesdent, M., Talbi, E., and Melab, N., Multi-objective optimization using Deep Gaussian Processes : Application to Aerospace Vehicle Design, AIAA Scitech 2019.

MO-BO with DGPs

Problem : Optimization of a set of objectives for a booster engine.
Objectives : Minimization of the Gross Lift-off Weight ($GLOW$) and maximization of the change in velocity (ΔV).

MO-BO with DGPs

Convergence plot of BO with different architecture of DGPs and a regular GP.

MO-DGP model

Actual approach : A model for each objective is trained independently.

Proposed approach : A model for each objective is trained independently.

MO-DGP model

- Derivation of an evidence lower bound,
- A covariance function to capture the correlations in the objective space as well as in the original input space.

Proposed approach : A model for each objective is trained independently.

Hebbal, A., Brevault, L., Balesdent, M., Talbi, E., and Melab, N., A Deep Gaussian Process based model for Multi-Objective optimization. The 13th International Conference on Multiple Objective Programming and Goal Programming (MOPGP) 2019.

Experimentations

Kursawe 3D

Exact Pareto Front

Solutions of the different MO-BO approaches.

. *Work in progress*

Experimentations

Kursawe 10D

Exact Pareto Front

Solutions of the different MO-BO approaches.

. *Work in progress*

Table of Contents

Introduction

Review on Bayesian Optimization

Bayesian optimization for non stationary functions

Multi-objective Bayesian Optimization

Multi fidelity with varying input spaces
- Review of GP based MF models
- Multi-fidelity with varying input spaces
- Experimentations

Conclusions

Multi-fidelity modeling

- High-fidelity (HF) models $\rightarrow$ **accurate data** but on a **limited dataset**.
- Low-fidelity (LF) models $\rightarrow$ **large amount of data** of **approximated data**.
- Use of multi-fidelity model to exhibit correlations between datasets of s increasing levels of fidelity $(X^t, \mathbf{y}^t), \forall 1 \leq t \leq s$, to improve prediction accuracy.

[Choi et al., 2004]

Review of GP based MF models

Multi-fidelity models based on Gaussian Processes (GPs) :

- **Auto-Regressive model (AR1)** [Kennedy and O'Hagan, 2001] :

$$f_t(\mathbf{x}) = \rho f_{t-1}(\mathbf{x}) + \gamma_t(\mathbf{x})$$

| Scaling factor $\rho \rightarrow$ capture only linear correlations between fidelities.

- **Non-linear Autoregressive model (NARGP)** [Perdikaris et al., 2017] :

$$f_t(\mathbf{x}) = g_t\left(f^*_{t-1}(\mathbf{x})\right) + \gamma_t(\mathbf{x})$$

| The posterior mean GP f^*_{t-1} is used $\rightarrow$ Sequential optimization of each fidelity $\rightarrow$ overfitting.

Review of GP based MF models

Multi-fidelity models based on Gaussian Processes (GPs) :

- **Multi-fidelity Deep Gaussian Process model (MF-DGP)** [Cutajar et al., 2019] :

$$f_t(\mathbf{x}) = g_t\left(f_{t-1}(\mathbf{x})\right) + \gamma_t(\mathbf{x})$$

| Keeping the prior f_{t-1} comes back to a Deep Gaussian Process.

Multi-fidelity with varying input spaces

Multi-fidelity with varying input spaces

In some industrial problems each fidelity can be characterized by its own input space due to :

- Different modeling approaches in each fidelity,
- Omission of some variables in the lower fidelities.

Model of launch vehicle thrust frame.

Wing geometry and its representation in LF and HF codes [Tao et al., 2019]

Multi-fidelity with varying input spaces

Classic approach for multi-fidelity with varying input spaces [Tao et al., 2019].

- Disjoint optimization of the mapping parameters and the multi-fidelity model,
- Training only on the lower fidelity input space

Proposed model : A two-level DGP

MF-DGP-IP graphical representation

- ✓ Uncertainty quantification on the input space mapping,
- ✓ Joint optimization of the mapping and the multi-fidelity model,
- ✓ Correlation computed in the original input space of each fidelity.

Proposed model : A two-level DGP

MF-DGP-IP graphical representation

- ✓ Uncertainty quantification on the input space mapping,
- ✓ Joint optimization of the mapping and the multi-fidelity model,
- ✓ Correlation computed in the original input space of each fidelity.

Hebbal, A., Brevault, L., Balesdent, M., Talbi, E., and Melab, N., Multi-fidelity modeling using DGPs : Improvements and a generalization to varying input space dimensions. 4th workshop on Bayesian Deep Learning (NeurIPS 2019), Vancouver, Canada.

Experimentations

Problem : HF 3D and LF 2D with different parametrization.

	Problem		
Algorithms	MNLL	RMSE	std RMSE
HF model	86.96	0.829	0.382
MF-DGP-IP	**1.02**	**0.663**	**0.11**
MF-DGP IMC	5.562	0.7632	0.167
MF-DGP nominal	2.701	0.8996	0.196
AR1	98.31	0.736	0.148
NARGP	2221.	0.81	0.272

Performance using 20 repetitions with different DoE (30 inputs data on LF and 5 inputs data on HF). RMSE refers to the root mean squared error and MNLL to the mean negative test log likelihood.

. *Work in progress*

Table of Contents

Introduction

Review on Bayesian Optimization

Bayesian optimization for non stationary functions

Multi-objective Bayesian Optimization

Multi fidelity with varying input spaces

Conclusions

Conclusions

Contributions :

- Proposition of a training approach for DGPs to improve uncertainty quantification,
- Coupling of DGP and BO and MO-BO and application to aerospace vehicle design problems,
- Proposition of MO-DGP model taking into account the correlations in the objective space,
- Proposition of a MF-DGP model for varying input space MF problems.

Conclusions

Future works :

- Generalization of the Expected Hypervolume Improvement criterion to dependent objectives,
- Application of the proposed MF-DGP and MO-DGP models to aerospace vehicle design problems.
- Multi-fidelity bayesian optimization using MF-DGP.

Publications

Publications :

- Accepted : Brevault, L., Balesdent, M., Hebbal, A. *Surrogate model-based multi-objective MDO approach for partially Reusable Launch Vehicle design.* **Journal of Spacecraft and Rockets**.
- Book Chapter : Hebbal, A., Pelamatti, J., Brevault, L., Balesdent, M., Talbi, E., and Melab, N. *MDO related issues multi-objective and mixed continuous/discrete optimization* in **Aerospace System Analysis and Optimization in Uncertainty**. *Edited by Springer.*
- Submitted : Hebbal, A., Brevault, L., Balesdent, M., Talbi, E., and Melab, N. *Bayesian Optimization using Deep Gaussian Processes.*, **Journal of Optimization and Engineering**.

Publications

Communications : Présentées :

- Hebbal, A., Brevault, L., Balesdent, M., Talbi, E., and Melab, N., Efficient Global Optimization with Deep Gaussian Processes," **Evolutionary Computation (CEC), 2018 IEEE Congress.**
- Hebbal, A., Brevault, L., Balesdent, M., Talbi, E., and Melab, N., Bayesian Optimization using Deep Gaussian Processes for non-stationary problems, **PGMO days 2018**.
- Hebbal, A., Brevault, L., Balesdent, M., Talbi, E., and Melab, N., *Multi-objective optimization using Deep Gaussian Processes : Application to Aerospace Vehicle Design*, **AIAA SciTech 2019.**
- Brevault, L., Balesdent, M., Hebbal, A., and Patureau De Mirand, A., *Surrogate model-based multi-objective MDO approach for partially Reusable Launch Vehicle design*, **AIAA SciTech 2019.**
- Hebbal, A., Brevault, L., Balesdent, M., Talbi, E., and Melab, N., *A Deep Gaussian Process based model for Multi-Objective optimization* à **The 13th International Conference on Multiple Objective Programming and Goal Programming (MOPGP) 2019**.
- Hebbal, A., Brevault, L., Balesdent, M., Talbi, E., and Melab, N., *Multi-fidelity modeling using DGPs : Improvements and a generalization to varying input space dimensions* à **4th workshop on Bayesian Deep Learning (NeurIPS 2019), Vancouver, Canada.**

Acceptées :

- Multi-Fidelity modeling with varying input space dimensions using Deep Gaussian Processes : Application to a blended wing body. **SIAM Conference on Uncertainty Quantification (UQ20)**

Prévue :

- A Deep Gaussian Process model for bayesian multi-objective optimization with dependent objectives. **NeurIPS 2020**

Merci pour votre attention !

Bibliography I

Choi, S., Alonso, J., Kroo, I., and Wintzer, M. (2004).
Multi-fidelity design optimization of low-boom supersonic business jets.
In *10th AIAA/ISSMO Multidisciplinary Analysis and Optimization Conference*, page 4371.

Cutajar, K., Pullin, M., Damianou, A., Lawrence, N., and González, J. (2019).
Deep gaussian processes for multi-fidelity modeling.
arXiv preprint arXiv :1903.07320.

Damianou, A. and Lawrence, N. (2013).
Deep gaussian processes.
In *Artificial Intelligence and Statistics*, pages 207–215.

Emmerich, M. T., Giannakoglou, K. C., and Naujoks, B. (2006).
Single-and multiobjective evolutionary optimization assisted by gaussian random field metamodels.
IEEE Transactions on Evolutionary Computation, 10(4) :421–439.

Haas, T. C. (1990).
Kriging and automated variogram modeling within a moving window.
Atmospheric Environment. Part A. General Topics, 24(7) :1759–1769.

Higdon, D., Swall, J., and Kern, J. (1999).
Non-stationary spatial modeling.
Bayesian statistics, 6(1) :761–768.

Jones, D. R., Schonlau, M., and Welch, W. J. (1998).
Efficient global optimization of expensive black-box functions.
Journal of Global optimization, 13(4) :455–492.

Bibliography II

Kennedy, M. C. and O'Hagan, A. (2001).
Bayesian calibration of computer models.
Journal of the Royal Statistical Society : Series B (Statistical Methodology), 63(3) :425–464.

Paciorek, C. J. and Schervish, M. J. (2006).
Spatial modelling using a new class of nonstationary covariance functions.
Environmetrics, 17(5) :483–506.

Perdikaris, P., Raissi, M., Damianou, A., Lawrence, N., and Karniadakis, G. E. (2017).
Nonlinear information fusion algorithms for data-efficient multi-fidelity modelling.
Proceedings of the Royal Society A : Mathematical, Physical and Engineering Sciences, 473(2198) :20160751.

Rasmussen, C. E. (2004).
Gaussian processes in machine learning.
In *Advanced lectures on machine learning*, pages 63–71. Springer.

Rasmussen, C. E. and Ghahramani, Z. (2002).
Infinite mixtures of gaussian process experts.
In *Advances in neural information processing systems*, pages 881–888.

Tao, S., Apley, D. W., Chen, W., Garbo, A., Pate, D. J., and German, B. J. (2019).
Input mapping for model calibration with application to wing aerodynamics.
AIAA Journal, pages 2734–2745.

Xiong, Y., Chen, W., Apley, D., and Ding, X. (2007).
A non-stationary covariance-based kriging method for metamodelling in engineering design.
International Journal for Numerical Methods in Engineering, 71(6) :733–756.