



HAL
open science

On Gaussian Processes for Spatial Modeling and Optimization

Didier Rullière, Rodolphe Le Riche

► **To cite this version:**

Didier Rullière, Rodolphe Le Riche. On Gaussian Processes for Spatial Modeling and Optimization. Joint HUST-LIMOS online Workshop on Operations Research, AI and Networks (ORAIN 2023), May 2023, Clermont - Ferrand, France. hal-04208916

HAL Id: hal-04208916

<https://hal.science/hal-04208916v1>

Submitted on 15 Sep 2023

HAL is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

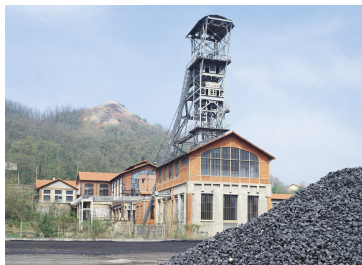


Distributed under a Creative Commons Attribution 4.0 International License

On Gaussian Processes for Spatial Modeling and Optimization

Didier Rullière and Rodolphe Le Riche
may 16-17, 2023

LIMOS-HUST online workshop on Operations Research, Artificial Intelligence and Networks (ORAIN 2023)



picture: mining headframe (chevalement) at Saint-Etienne

.....

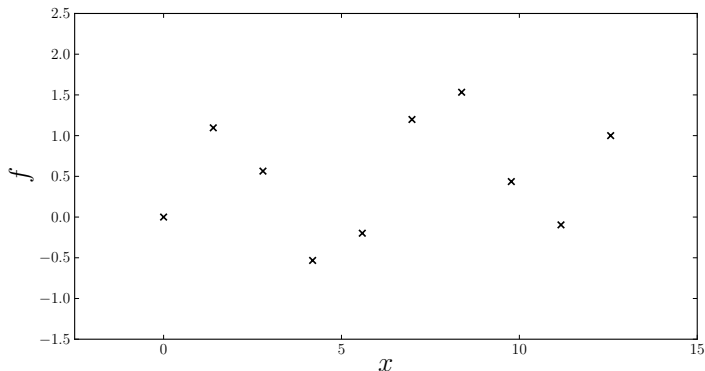
Outline

- 1 Gaussian Process Interpolation
- 2 GPs in decision theory: case studies

Gaussian Process Interpolation

Gaussian process regression (1/4)

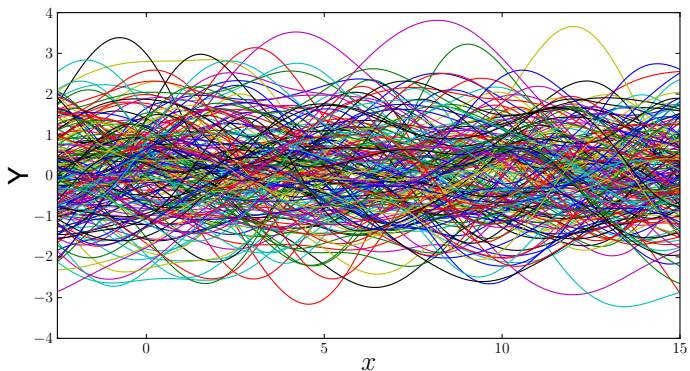
Assume we have observed a function $f(\cdot)$ over a set of points $\mathbb{X} = (x_1, \dots, x_n)$:



Here in 1D. The vector of observations is $\mathbf{y} = f(\mathbb{X})$, i.e. $y_i = f(x_i)$.

Gaussian process regression (2/4)

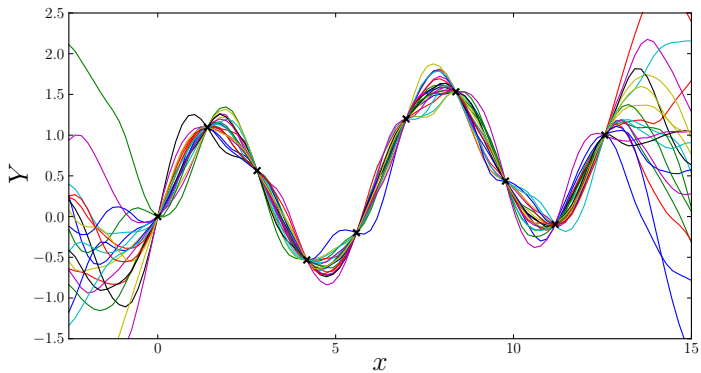
Since $f(\cdot)$ is unknown, we make the general assumption that it is close to the sample path of a Gaussian process $Y \sim \mathcal{N}(\mu(\cdot), k(\cdot, \cdot))$:



(here $\mu(x) = 0$)

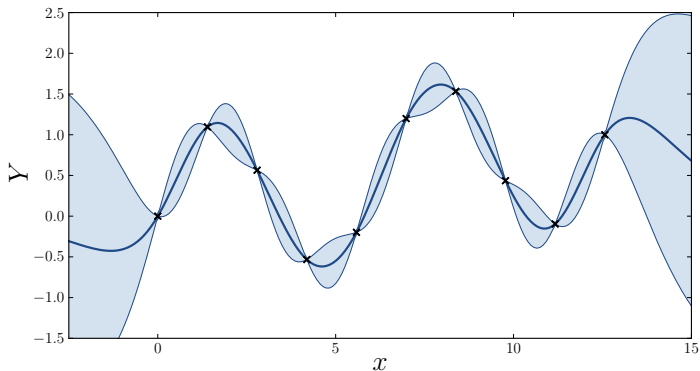
Gaussian process regression (3/4)

If we remove all the samples that do not interpolate the observations we obtain:



Gaussian process regression (4/4)

It can be summarized by a mean function and 95% confidence intervals.



... it is easy to extend to higher dimensions. Formally the mean predictor at one location x is the Best Linear Unbiased Prediction (BLUP) using observations $y_i = f(x_i)$.

You can play here: https://durrande.shinyapps.io/gp_playground/

Kriging equations

The conditional distribution can be obtained analytically:

By definition, (Y_x, \mathbf{Y}) is multivariate normal. Formulas on the conditioning of Gaussian vectors give the distribution of $Y_x | \mathbf{Y} = \mathbf{y}$. It is $\mathcal{N}(m(\cdot), c(\cdot, \cdot))$ with :

$$\begin{aligned}m(x) &= \mu(x) + k(x, \mathbb{X})k(\mathbb{X}, \mathbb{X})^{-1}(\mathbf{y} - \mu(\mathbb{X})) \\c(x, x') &= k(x, x') - k(x, \mathbb{X})k(\mathbb{X}, \mathbb{X})^{-1}k(\mathbb{X}, x')\end{aligned}$$

where $k(\cdot, \cdot)$ is the covariance function of the prior Gaussian Process.

Simple Kriging, Gaussian case

For a centered process, when $\mu(x) = 0$ for all x , the simple Kriging predictor mean and variance are

$$\begin{cases} \mathbb{E}[Y_x | \mathbf{Y} = \mathbf{y}] &= m(x) \\ \text{Var}[Y_x | \mathbf{Y} = \mathbf{y}] &= c(x, x) \end{cases}$$

Problems

Here are few selected problems:

- **Model selection:**
how to choose prior process, prior covariance function family, prior covariance function parameters ?
- **Computation:**
how to compute the predictor when the matrices are huge?
- **Adaptation:**
how to adapt to specific settings (monotony, uncertainty, high dimension...)

Because kriging (GP regression) predicts a full, spatially varying, distribution, it is of great use in decision making.

⇒ a stepping stone to address decision-theoretic problems.

GPs in decision theory: case studies

Input uncertainty

Case of Energy Performance Certificate: measurement are done at an unknown location on a specific area (not averaged on the area)

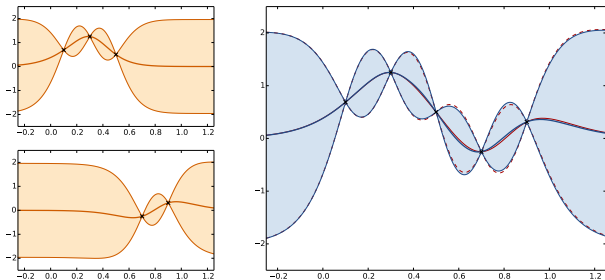


Figure: Left: Prescribed vignette appearing on the French energy performance certificate up to 2021. Label A refers to energy efficient dwellings, G refers to energy intensive dwellings. Right: Map of French inventoried EPCs over a neighborhood of Lyon city.

Adaptation of Kriging to mixtures of Gaussian rv. *Grossouvre et al., 2023*

High Dimension

Large observation number: with (sub)models combinations *Rullière et al., 2018*



Kriging sub-models

aggregation in blue, full model in red.

Figure: Aggregation of two Gaussian process regression models.

Large dimension: with (sub)models combinations *Appriou et al., 2022*

is it better and faster to combine few submodels than to find the best model?

Constrained & robust Bayesian Optimization I

Problem addressed : optimization, with expensive constraints and objectives, in the presence of uncertainties. Cost is twice an issue!

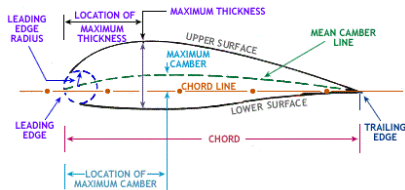
Example : compressor rotor design

Maximize rotor polytropic efficiency

by changing its (parameterized) shape, x

with geometrical, inflow and rotation speed uncertainties, u ,

under flow constraints (angles, no separation, Mach number) indexed by p .



Constrained & robust Bayesian Optimization II

Principle of the method:

- Build GPs of the objective in (x, u) space and of the constraints in the (x, u, p) space.
- Use the GPs to define a figure of merit : the expected feasible improvement (EFI)
- Sequentially (carefully) choose the next variable, uncertainties and constraints to be computed

$$x^{n+1} = \arg \max_x \text{EFI}(x)$$

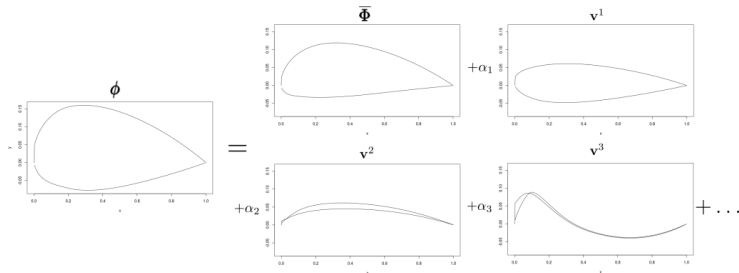
$$u^{n+1}, p^{n+1} = \arg \min_{u,p} \text{Variance of one-step-ahead EFI}(x^{n+1})$$

⇒ faster convergence. More explanations in [Pelamatti et al., 2023](#).

Costly shape optimization I

Example: airfoil minimizing drag. “Infinite” dimensional, and CFD is costly.

Eigendecomposition of the shapes, then work in α -space:



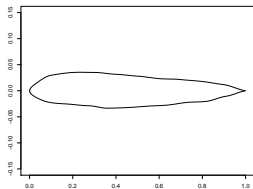
Costly shape optimization II

Principle of the method:

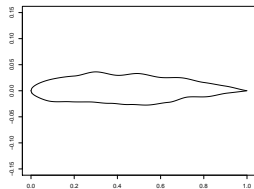
- Learn active components of α from known (inputs, outputs) by maximizing a likelihood plus a term that favors sparsity.
- Build a GP on active components (dimension reduction) + a simple GP on less active components. Perform Bayesian optimization on active space + a random 1D embedding for less active components.
- Project α^{n+1} on manifold of feasible shapes.

⇒ faster convergence, smoother airfoils

(here 100 calls to CFD)



BO on active reduced space



BO on traditional CAD space

More explanations in [Gaudrie et al., 2020](#).

Related recent works by the same team

Quantization of rare events : *Sire et al., 2023*

More Bayesian optimization :

with mixed discrete-continuous variables : *Cuesta Ramirez et al., 2022*

with HSIC-based variable selection : *Spagnol et al., 2019*

with trust regions for higher dimensions : *Diouane et al., 2022*

with constraints (monotonicity, convexity...) : *Maatouk et al., 2023*

testing Bayesian optimization : *Le Riche and Picheny, 2021*

References I

- [1] Appriou, T., Rullière, D., and Gaudrie, D. (2022). Combination of High-Dimensional Kriging Sub-models. working paper or preprint.
- [2] Cuesta Ramirez, J., Le Riche, R., Roustant, O., Perrin, G., Durantin, C., and Glière, A. (2022). A comparison of mixed-variables bayesian optimization approaches. *Advanced Modeling and Simulation in Engineering Sciences*, 9(1):6.
- [3] Diouane, Y., Picheny, V., Le Riche, R., and Scotto Di Perrotolo, A. (2022). Trego: a trust-region framework for efficient global optimization. *Journal of Global Optimization*, pages 1–23.
- [4] Gaudrie, D., Le Riche, R., Picheny, V., Enaux, B., and Herbert, V. (2020). Modeling and optimization with gaussian processes in reduced eigenbases. *Structural and Multidisciplinary Optimization*, 61:2343–2361.
- [5] Grossouvre, M., Rullière, D., and Villot, J. (2023). Spatial interpolation using mixture distributions: A Best Linear Unbiased Predictor. working paper or preprint.
- [6] Le Riche, R. and Picheny, V. (2021). Revisiting bayesian optimization in the light of the coco benchmark. *Structural and Multidisciplinary Optimization*, 64(5):3063–3087.
- [7] Maatouk, H., Rullière, D., and Bay, X. (2023). Bayesian analysis of constrained Gaussian processes. working paper or preprint.

References II

- [8] Pelamatti, J., Le Riche, R., Helbert, C., and Blanchet-Scalliet, C. (2023). Coupling and selecting constraints in Bayesian optimization under uncertainties. *Optimization and Engineering*. accepted for publication in April 2023, also available as arXiv 2204.00527.
- [9] Rullière, D., Durrande, N., Bachoc, F., and Chevalier, C. (2018). Nested kriging predictions for datasets with a large number of observations. *Statistics and Computing*, 28(4):849–867.
- [10] Sire, C., Le Riche, R., Rullière, D., Rohmer, J., Pheulpin, L., and Richet, Y. (2023). Quantizing rare random maps: application to flooding visualization. *Journal of Computational and Graphical Statistics*, 0(ja):1–31.
- [11] Spagnol, A., Le Riche, R., and Veiga, S. D. (2019). Global sensitivity analysis for optimization with variable selection. *SIAM/ASA Journal on uncertainty quantification*, 7(2):417–443.