Gaussian processes for computer experiments with monotonicity information

Hassan MAATOUK*

Thesis supervised by X. BAY*, L. GRAMMONT(1), Y. RICHET (2), O. ROUSTANT*

(*) Ecole Nationale Supérieure des Mines de Saint-Etienne, France

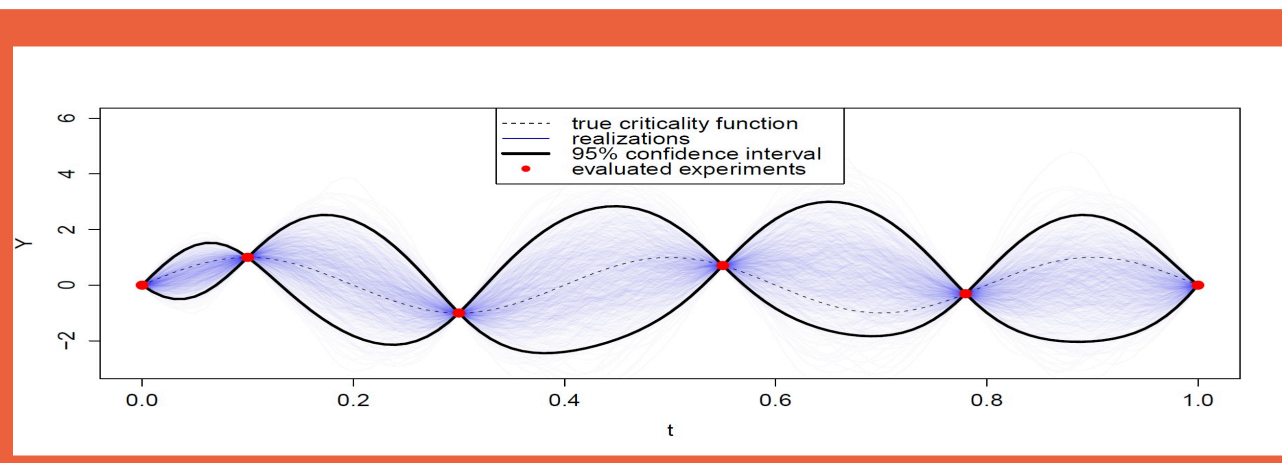
(1) Université Jean Monnet, Saint-Etienne, Institut Camille Jordan (Lyon 1)

(2) Institut de Radioprotection et de Sûreté Nucléaire (IRSN, Paris)

hassan.maatouk@mines-stetienne.fr

Kriging and monotone Kriging

- Complex phenomena are often studied by time-consuming computer codes.
- Kriging is an interpolation method by conditioning Gaussian process at observation data.
- Kriging with monotonicity information is needed when the true function in the physical system is known to be increasing or decreasing with respect to some or all input variables.



1 – Methodology

Let $y = f(x)$ be an increasing function defined for $x \in [0,1]$. We consider a set of computer experiments $\{(x_i, y_i) \mid i = 1, \cdots, n\}$ with

$$y_i = f(x_i),\ 1 \le i \le n. \tag{1}$$

Denote $\mathcal{M}$ the space of increasing functions and $(Y_x)_{x\in[0,1]}$ a centered Gaussian process (GP) with a covariance kernel $K(x, x')$, defined as

$$K(x, x') = cov(Y_x, Y_{x'}) = E[Y_x Y_{x'}].$$

- **Formulation of the general problem** : to find the conditional distribution of the GP Y, given data and monotonicity information

$$\begin{aligned} &Y_{x_i} = y_i,\ 1 \le i \le n, \\ &Y \in \mathcal{M}. \end{aligned} \tag{2}$$

- **Solution** : to approximate Y by a finite-dimensional GP Y^N

$$Y_x^N = \sum_{j=0}^{N} \xi_j \phi_j(x), \tag{3}$$

in which $\xi = (\xi_0 \cdots \xi_N)^\top$ is $\mathcal{N}(0, \Gamma_N)$ and the basis functions ϕ_j are in $\mathcal{M}$. The covariance matrix Γ_N is chosen to ensure the convergence of the GP Y^N to the GP Y when N tends to infinity.

Property of the basis functions

We choose the basis functions ϕ_j such that

$$Y_x^N \text{ is increasing} \iff \xi_j \ge 0;\ 0 \le j \le N.$$

A simple case is given by the primitive functions of B-splines of order 2.

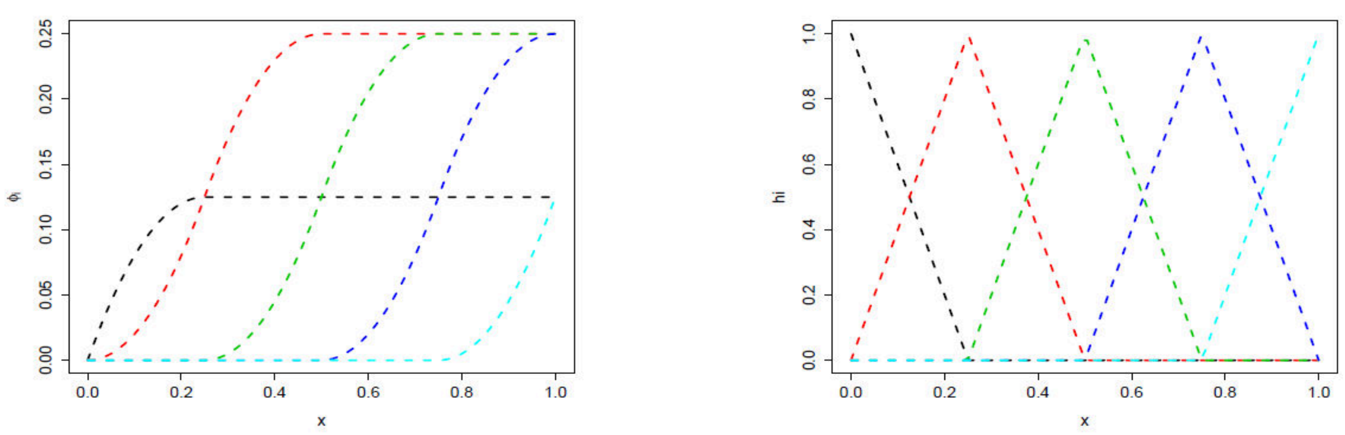


FIGURE : Basis functions and B-spline functions for $N = 4$.

Problem reformulation

To simulate the conditional Gaussian vector ξ such that

$$\begin{cases} \sum_{j=0}^{N} \xi_j \phi_j(x_i) = y_i,\ 1 \le i \le n & \text{(n interpolation linear equations)} \\ \xi_j \ge 0,\ 0 \le j \le N & \text{(N+1 inequality conditions)} \end{cases} \tag{$\mathcal{C}$}$$

- Hence, we need to simulate a truncated multi-normal distribution.

2 – A new rejection technique to simulate a truncated multivariate normal distribution

Let $\mathcal{C}$ be a convex set in R^d, and f and g are the probability density functions of the two multivariate normal distributions centered at 0 and μ^*, respectively, where μ^* is the mode of f restricted to $\mathcal{C}$. We define

$$\tilde{f}(x) = f(x \mid 0, \Sigma)1_{x\in\mathcal{C}} \quad and \quad \tilde{g}(x) = g(x \mid \mu^*, \Sigma)1_{x\in\mathcal{C}},$$

By rejection sampling, we use $\tilde{g}$ to simulate f in $\mathcal{C}$.

Property

The optimal constant such that $\tilde{f}(x) \le k\tilde{g}(x)$ for all x in $\mathcal{C}$, is

$$k = e^{-\frac{1}{2}(\mu^*)^T\Sigma^{-1}\mu^*}.$$

The proposed algorithm

1. Generate X with pseudo density function $\tilde{g}$.
2. Generate U uniformly on $[0,1]$. If $U \le e^{(\mu^*)^T\Sigma^{-1}\mu^* - X^T\Sigma^{-1}\mu^*}$ accept X ; otherwise, go back to step 1.

Illustration

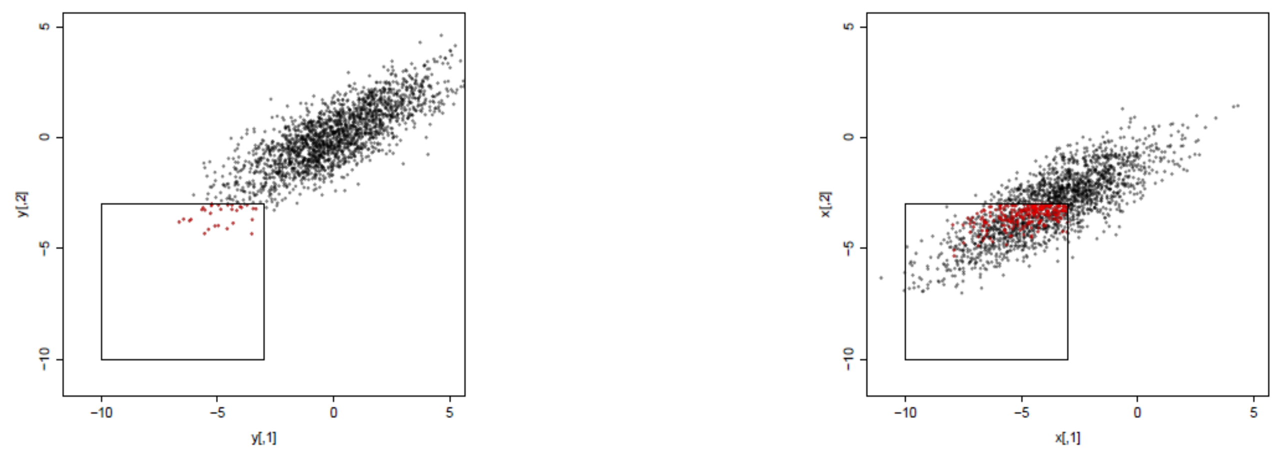


FIGURE : Crude rejection sampling with 2% acceptance rate (left figure) and the so-called rejection sampling from the mode with 20% acceptance rate (right figure).

3 – Monotone simulations

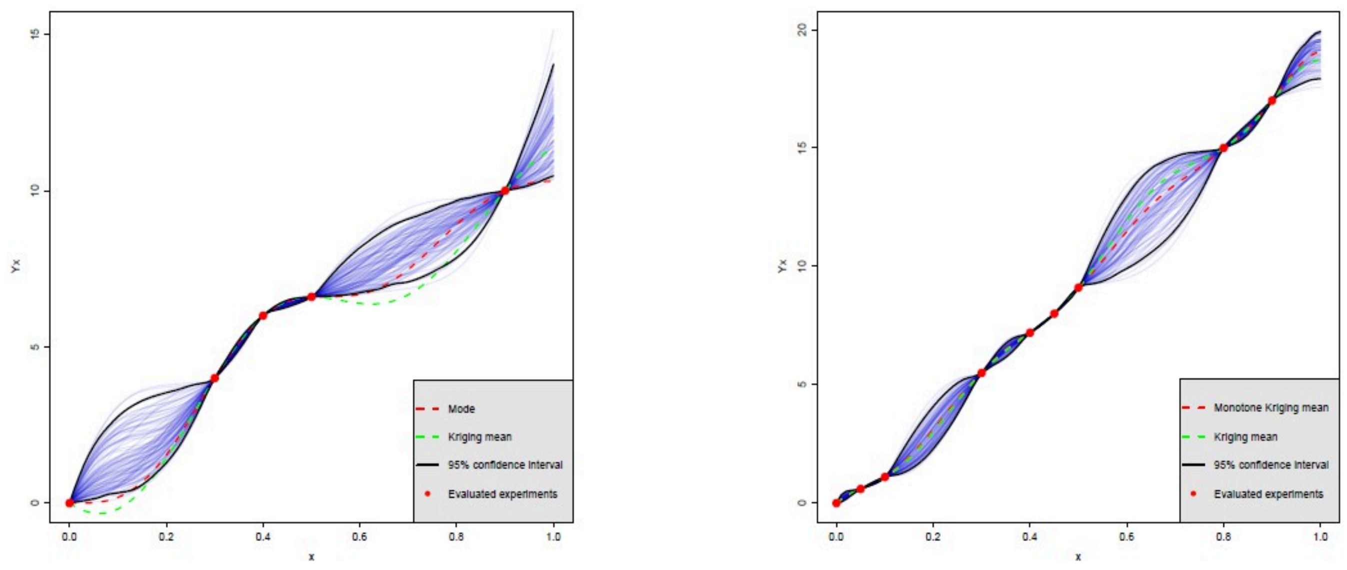


FIGURE : 100 sample paths taken from the conditional GP with respect to monotonicity information and output values at 5 and 9 design points (left and right respectively). Notice that these paths are monotone over the whole domain.

4 – Conclusion and future work

1 - We propose a new methodology based on GP metamodeling to sample from posterior distribution including monotonicity information. Notice that all the paths are monotone over the whole domain.

2 - We develope a new rejection sampling technique to simulate a truncated multivariate Gaussian random variable in a convex set.

Question 1 : Convergence of the finite-dimensional GP to Y when N tends to infinity ?

Question 2 : Generalization to higher dimensions ?

[1] Da Veiga, S. and Marrel, A. (2012). Gaussian process modeling with inequality constraints. Annales de la faculté des sciences de Toulouse Mathématiques, 21(3):529–555.

[2] Golchi, S., Bingham, D. R., Chipman, H., and Campbell, D. A. (2013). Monotone Function Estimation for Computer experiments. http://adsabs.harvard.edu/abs/2013arXiv1309.3802G.

[3] Robert, C. (1995). Simulation of truncated normal variables. Statistics and Computing, 5(2)