Kriging models with Gaussian processes - covariance function estimation and impact of spatial sampling

François Bachoc

former PhD advisor: Josselin Garnier
former PhD co-advisor: Jean-Marc Martinez

Department of Statistics and Operations Research, University of Vienna

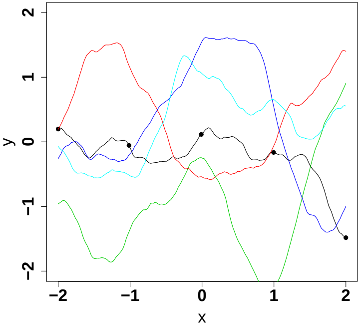
(Presented work performed while PhD student at French Atomic Energy and Alternative Energies Commission and Paris Diderot University)

NRC 2014 - Harvard University - August 2014

Kriging model with Gaussian processes

Kriging model

Study of a **single realization** of a Gaussian process $Y(x)$ on a domain $\mathcal{X} \in \mathbb{R}^d$



Goal

Predicting the continuous realization function, from a finite number of **observation points**

Covariance function

Covariance function

The function $K : \mathcal{X}^2 \to \mathbb{R}$, defined by $K(x_1, x_2) = cov(Y(x_1), Y(x_2))$

- Most important characteristic of the Gaussian process

Parameterization

Covariance function model $\{K_\theta, \theta \in \Theta\}$ for the Gaussian Process Y

- θ is a multidimensional covariance parameter. K_θ is a covariance function

Observations

Y is observed at $x_1, ..., x_n \in \mathcal{X}$, yielding the Gaussian vector $y = (Y(x_1), ..., Y(x_n))$

Estimation

Objective : build estimator $\hat{\theta}(y)$

Cross Validation and Maximum Likelihood for estimation

Maximum Likelihood estimation

Numerical optimization of the explicit Gaussian likelihood function for the observation vector y

General principle of Cross Validation

Leave-One-Out prediction errors

$$\hat{y}_{\theta,i,-i} = \mathbb{E}_\theta(Y(x_i)|y_1, ..., y_{i-1}, y_{i+1}, ..., y_n)$$

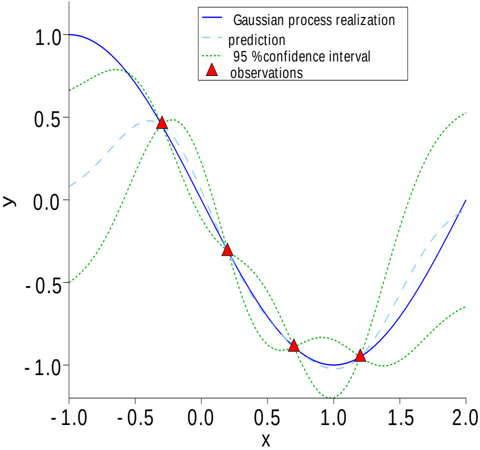
Leave-One-Out criterion we study

$$\hat{\theta}_{CV} \in \underset{\theta \in \Theta}{\operatorname{argmin}} \sum_{i=1}^{n} (y_i - \hat{y}_{\theta,i,-i})^2$$

$\Longrightarrow$ Numerical optimization with same computational cost as Maximum Likelihood

Two-step approach

1. Estimation of the covariance function
2. Assume the covariance function is known and equal to its estimate. Use Gaussian conditioning theorem for prediction of the Gaussian process realization given the observed values



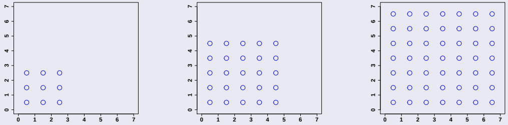
$\Longrightarrow$ Widely applied to computer experiments, where a computationally expensive numerical model is represented by a realization of a Gaussian process. E.g. in nuclear engineering, aeronautic...

In the Poster : expansion-domain asymptotics for covariance function estimation

Asymptotics (number of observations $n \to +\infty$) for covariance function estimation is an active area of research

Increasing-domain asymptotics

A minimum spacing exists between the observation points $\longrightarrow$ infinite observation domain

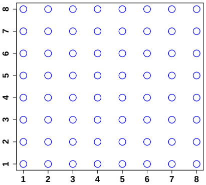


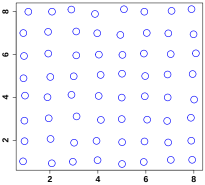
We study a randomly perturbed regular grid of observation points :

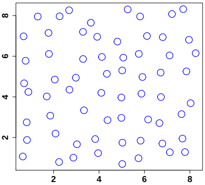
- We show that ML and CV are consistent and have a $\sqrt{n}$ rate of convergence
- We show that ML has a smaller asymptotic variance

- Spatial sampling : initial design of experiments for Kriging
- It has been shown that irregular spatial sampling is often an advantage for covariance parameter estimation

$\Longrightarrow$ We confirm this finding in our asymptotic framework with the randomly perturbed regular grid







Thank you for your attention !

Reference paper :

F. Bachoc, Asymptotic analysis of the role of spatial sampling for covariance parameter estimation of Gaussian processes, *Journal of Multivariate Analysis 125 (2014) 1-35*.