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

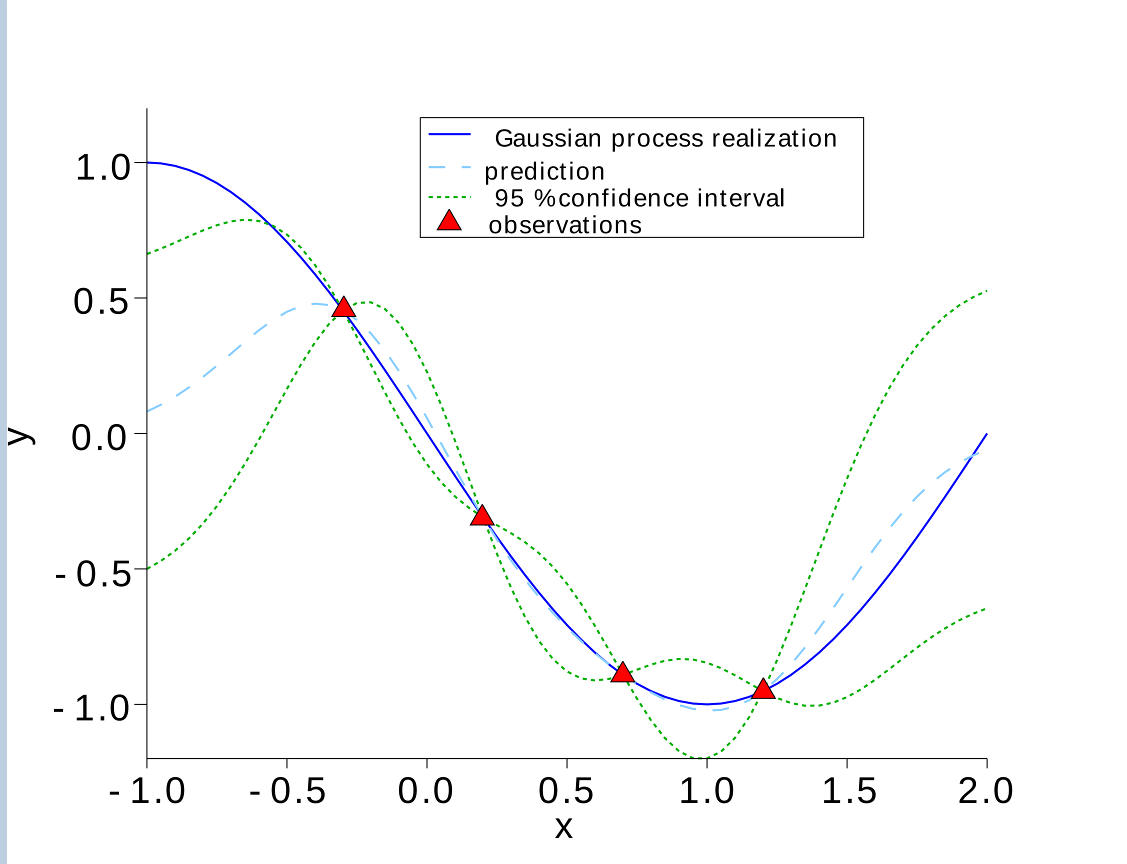
François Bachoc
francois.bachoc@univie.ac.at

Kriging models

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

Two-step approach: covariance function estimation and prediction

Widely applied to computer experiments. E.g. in nuclear engineering, aeronautic...



Covariance function estimation

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

Estimator $\hat{\theta}(y)$ for a vector of observations $y = (Y(x_1), ..., y(x_n))$

Maximum likelihood: optimization of the explicit Gaussian likelihood function for the observation vector y

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

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

Numerical optimization for both methods with same computational cost

Main objectives

Study the consistency and asymptotic distribution of the Cross Validation estimator

Confirm that, asymptotically, Maximum Likelihood is more efficient

Study the influence of the irregularity of the spatial sampling on the estimation

References

[1] Bachoc F, Cross Validation and Maximum Likelihood estimations of hyper-parameters of Gaussian processes with model misspecification, *Computational Statistics and Data Analysis 66 (2013) 55-69*

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

[3] Mardia K, Marshall R, Maximum likelihood estimation of models for residual covariance in spatial regression, *Biometrika 71 (1984) 135-146*

[4] Zhu Z, Zhang H, Spatial sampling design under the infill asymptotics framework, *Environmetrics 17 (2006) 323-337*

Acknowledgements

This work was performed while the author was a PhD student, supervised by **Jean-Marc Martinez** (French Atomic Energy commission) and **Josselin Garnier** (Paris Diderot University).

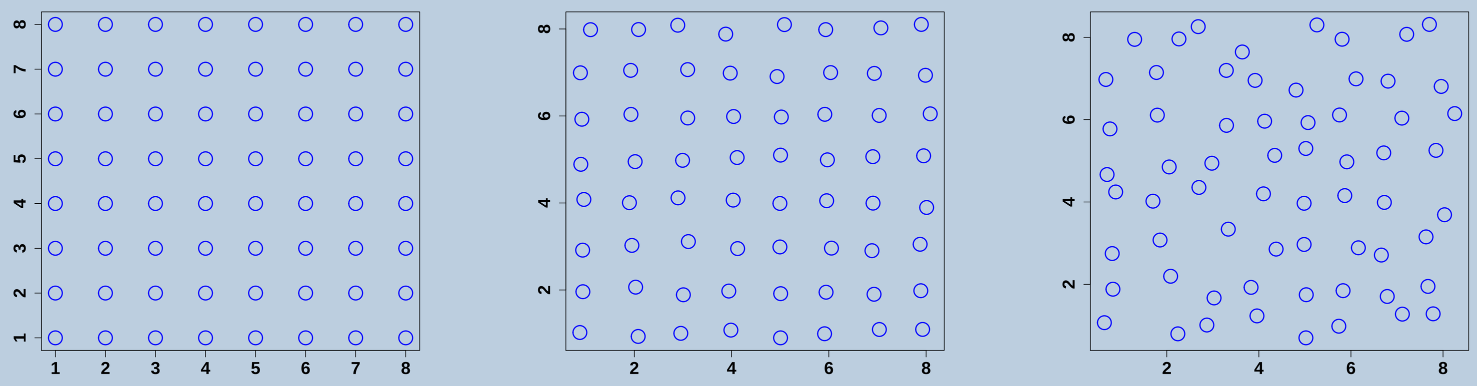
Randomly perturbed regular grid

Observation point i ($i = 1, ..., n$):

$$\boldsymbol{v}_i + \epsilon X_i$$

- $(\boldsymbol{v}_i)_{i \in \mathbb{N}^*}$: infinite regular square grid of step one in dimension d
- $(X_i)_{i \in \mathbb{N}^*}$: *iid* with uniform distribution on $[-1, 1]^d$
- $\epsilon \in (-\frac{1}{2}, \frac{1}{2})$ is the regularity parameter of the spatial sampling

Illustration with $d = 2$, $n = 64$ and $\epsilon = 0, \frac{1}{8}, \frac{3}{8}$:



$\Longrightarrow$ This is an increasing-domain asymptotic framework

Consistency and asymptotic normality

Almost-sure convergence of the random Fisher information matrix to a $p \times p$ deterministic matrix $\mathbf{I}_{ML}$ as $n \to +\infty$.

For Maximum likelihood: with $\Sigma_{ML} = \mathbf{I}_{ML}^{-1}$,

$$\sqrt{n}\left(\hat{\boldsymbol{\theta}}_{ML} - \boldsymbol{\theta}_0\right) \to \mathcal{N}(0, \Sigma_{ML})$$

For Cross Validation: Same result with more complex expressions for asymptotic covariance matrix Σ_{CV}

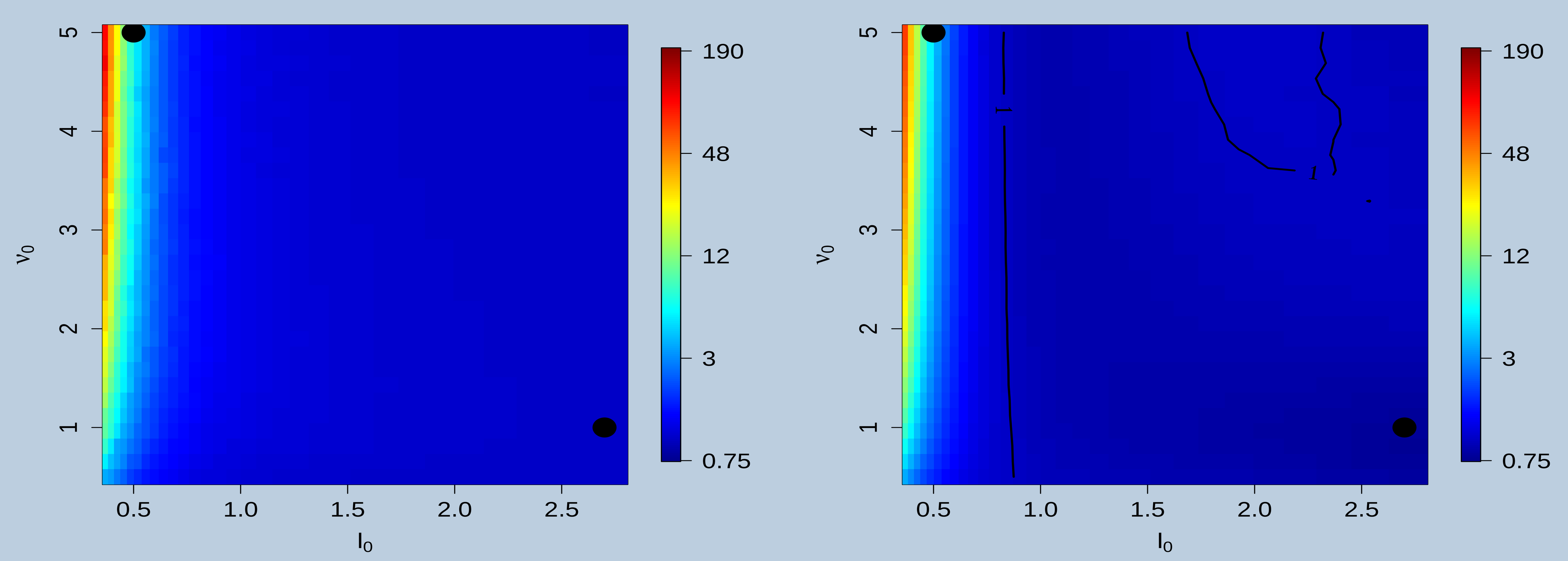
Irregular spatial sampling is beneficial to estimation

The asymptotic covariance matrices $\Sigma_{ML,CV}$ depend only on the regularity parameter ϵ.

We study the estimation of either the correlation length ℓ or the smoothness parameter ν in the Matérn model in dimension 1

Level plot of $[\Sigma_{ML,CV}(\epsilon = 0)] \,/\, [\Sigma_{ML,CV}(\epsilon = 0.45)]$ in $\ell_0 \times \nu_0$

Estimation of ℓ when ν_0 is known for ML (left) and CV (right)



Estimation of ν when ℓ_0 is known for ML (left) and CV (right)