Gaussian Process (Kriging)

Hannah Bartmann, Ferry Heinzelmann, Nils Hichert

Case Studies II

Prof. Dr. Katja Ickstadt
Dr. Henrike Weinert
Yassine Talleb
Carmen van Meegen

Department of Statistics, TU Dortmund

02.12.2025

Model Idea

A zero mean Gaussian Process is a collection of random variables where any finite subset has a joint Gaussian distribution:

$$Z(X) \sim \mathcal{N}(\mathbf{0}, \mathsf{K}(X, X)).$$

The output of the 'black box' function

$$f(X) = (f(\boldsymbol{x}_1), \ldots, f(\boldsymbol{x}_n))^\top =: \boldsymbol{f}$$

can be modelled by a linear regression

$$\boldsymbol{f} = h(X)^\top \beta + Z(X), \quad \beta \in \mathbb{R}^{p+1}$$

where $h(X)$ is some known basis function, eg. $h(X) = \mathbf{1}_{p+1}$ or $h(\boldsymbol{x}_i) = (1, x_{i1}, \ldots, x_{ip})$, $i = 1 \ldots, n$

Types of Kriging

The basis function h is the choice of the statistician!

Different types of Kriging:

- Simple Kriging: $h(X)^\top \beta \equiv \text{const}$, β known
- Ordinary Kriging: $h(X)^\top \beta \equiv \text{const}$, β unknown
- Universal Kriging: $h(X)^\top \beta$ linear in β, β unknown

For simplicity we continue with simple Kriging with $h(X)^\top \beta = 0$

- model reduces to $\boldsymbol{f} = Z(X)$
- ordinary or universal Kriging requires estimating β which would necessitate a lot more notation

Covariance structure

To predict the function values $\boldsymbol{f}^*$ at some new inputs $X^* = (\boldsymbol{x}_1^* \dots \boldsymbol{x}_{n^*}^*)$ we need

- the vector of some known function values $\boldsymbol{f}$
- the covariance function $\mathrm{K}(\cdot, \cdot) \rightarrow$ statistician's choice

An example for K is given by the radial base function

$$K(\boldsymbol{x}, \boldsymbol{x}^*) = \exp\left\{-\left(\varepsilon \cdot ||\boldsymbol{x} - \boldsymbol{x}^*||\right)^2\right\}$$

for some $\varepsilon > 0$

- K must be chosen in a way that it produces non singular covariance matrices

Prediction

The conditional distribution of the function value at $\boldsymbol{x}^*$ is given by

$$f^*|\boldsymbol{f}, X, \boldsymbol{x}^* \sim \mathcal{N}\big(\, \mathrm{K}(\boldsymbol{x}^*, X)\, \mathrm{K}(X, X)^{-1}\boldsymbol{f},\\ \mathrm{K}(\boldsymbol{x}^*, X) - \mathrm{K}(\boldsymbol{x}^*, X)\, \mathrm{K}(X, X)^{-1}\, \mathrm{K}(X, \boldsymbol{x}^*)\big)$$

The best linear unbiased predictor (BLUP) for $f(\boldsymbol{x}^*)$ is given by

$$\hat{f}(\boldsymbol{x}^*) = \mathrm{E}(f(\boldsymbol{x}^*)) = \mathrm{K}(\boldsymbol{x}^*, X)\, \mathrm{K}(X, X)^{-1}\boldsymbol{f}$$

Properties of BLUP

- linear in $\boldsymbol{f}$ and smallest variance of all linear estimators
- the BLUP interpolates the training data X
  - $\mathrm{E}(\hat{f}(\boldsymbol{x}^*) \,|\, f(\boldsymbol{x}),\, X = \boldsymbol{x},\, \boldsymbol{x}^* = \boldsymbol{x}) = f(\boldsymbol{x})$
  - $\mathrm{Cov}(f(\boldsymbol{x}^*) \,|\, \boldsymbol{f},\, X = \boldsymbol{x},\, \boldsymbol{x}^* = \boldsymbol{x}) = 0$

Software packages

`R`:

- `rkriging` [1]:
  - `Fit.Kriging(X =` X`, y =` $\boldsymbol{y}$`, kernel =` K_{XX}`, ...)`
  - Ordinary and Universal Kriging

`Python`:

- `PyKrige` [2]
  - First result when googling 'kriging python'

Bibliography I

[1] Chaofan Huang and V. Roshan Joseph. *rkriging: Kriging Modeling*. R package version 1.0.2. 2025. DOI: 10.32614/CRAN.package.rkriging.

[2] Benjamin Murphy, Roman Yurchak, and Sebastian Müller. *GeoStat-Framework/PyKrige: v1.7.3*. Version v1.7.3. Oct. 2025. DOI: 10.5281/zenodo.17372225.

[3] Carl Edward Rasmussen and Christopher K. I. Williams. *Gaussian processes for machine learning*. en. 1st ed. Cambridge, Mass.: MIT Press, 2006. ISBN: 026218253X.

[4] Thomas J Santner, Brian J Williams, and William I Notz. *The design and analysis of computer experiments*. 2nd ed. Springer Series in Statistics. New York, NY: Springer, 2019. DOI: 10.1007/978-1-4939-8847-1.