Bayesian Methods in Geostatistics: Using prior knowledge about trend parameters for Kriging and Design of Experiments

Albrecht Gebhardt[*]
agebhard@uni-klu.ac.at

Claudia Gebhardt[†]
cgebhard@uni-klu.ac.at

Submitted to useR! 2004

Abstract

This paper covers an approach to incorporate prior knowledge about trend parameters into Kriging estimation. The implementation in R covers prediction as well as monitoring network design.

1 Introduction

Kriging, the central method for spatial prediction of a regionalised variable $Z(\underline{x}), \underline{x} \in \mathbb{R}^d$, is based on some prerequisites: translation invariant covariance structure (second order stationarity)

$$\mathrm{Cov}\left(Z(\underline{x}), Z(\underline{x}+\underline{h})\right) = C(\underline{h}), \quad \underline{h} \in \mathbb{R}^d \tag{1}$$

and knowledge about an underlying trend. Depending on the existence or absence of a trend function one has to choose between Ordinary or Universal Kriging. Additional prior knowledge about the trend parameter can be incorporated into the prediction using a Bayesian approach. This approach has been implemented in `R` for prediction and monitoring network design purposes.

2 Bayesian Kriging

Given a data set $\underline{Z} = (Z(\underline{x}_i))_{i=1,\dots,n}$ and using a linear trend model

$$Z(\underline{x}) \quad = \quad \underline{f}(\underline{x})^\top \underline{\theta} + \varepsilon(\underline{x}), \qquad \underline{\theta} \in \Theta \subseteq \mathbb{R}^r \tag{2}$$

Universal Kriging takes a weighted average $\hat{Z}(\underline{x}_0) = \underline{w}^\top \underline{Z}$ of the data as estimator at location $\underline{x}_0$. It assumes unbiasedness $\mathrm{E}\,\hat{Z}(\underline{x}_0) = \mathrm{E}\,Z(\underline{x}_0)$ which leads to the so called universality constraint $\mathbf{F}^\top \underline{w} = \underline{f}_0$. Choosing the weights $\underline{w}$ which minimise the variance of the prediction yields finally the Universal Kriging estimator

$$\hat{Z}(\underline{x}_0) \quad = \quad \underline{c}_0^\top \mathbf{C}^{-1}(\underline{Z} - \mathbf{F}\underline{\hat{\mu}}) + \underline{f}_0^\top \underline{\hat{\theta}} \tag{3}$$

using the notations $\underline{f}_0 = \underline{f}(\underline{x}_0)$, $\mathbf{F} = (\underline{f}(\underline{x}_1), \underline{f}(\underline{x}_2), \dots, \underline{f}(\underline{x}_n))^\top$, $(\mathbf{C})_{ij} = C(\underline{x}_i - \underline{x}_j)$ and $(\underline{c}_0)_i = C(\underline{x}_i - \underline{x}_0)$ $i, j = 1, \dots, n$ where $\underline{\hat{\theta}} = (\mathbf{F}^\top \mathbf{C}^{-1}\mathbf{F})^{-1}\mathbf{F}^\top \mathbf{C}^{-1}\underline{Z}$ corresponds to the generalised least squares estimator of $\underline{\theta}$.

[*]University of Klagenfurt, Institute of Mathematics, Austria
[†]University of Klagenfurt, University Library, Austria

Bayesian Kriging requires knowledge about the prior distribution of the trend parameter θ. This approach (see Omre (1987)) makes only use of the first two moments $\mathrm{E}\,\underline{\theta} = \underline{\mu}$ and $\mathrm{Cov}\,\underline{\theta} = \mathbf{\Phi}$. In contrast to Universal Kriging the condition of unbiasedness gets now thrown away in favour of a modified Bayesian Kriging estimator involving a bias component w_0

$$\hat{Z}(\underline{x}_0) \quad = \quad \underline{w}^\top \, \underline{Z} + w_0 \; . \tag{4}$$

Again the weights are chosen to minimise the prediction variance. The solution of the corresponding Lagrange system is

$$\hat{Z}(\underline{x}_0) \quad = \quad \underline{\tilde{c}}_0^\top \, \tilde{\mathbf{C}}^{-1} \, (\underline{Z} - \mathbf{F}\underline{\mu}) + \underline{f}_0^\top \underline{\mu} \tag{5}$$

which involves the following modified covariance terms

$$\tilde{C}_0 = C_0 + \underline{f}_0^\top \, \mathbf{\Phi} \, \underline{f}_0 \quad \underline{\tilde{c}}_0 = \underline{c}_0 + \mathbf{F}\,\mathbf{\Phi}\,\underline{f}_0 \quad \tilde{\mathbf{C}} = \mathbf{C} + \mathbf{F}\,\mathbf{\Phi}\,\mathbf{F}^\top \; . \tag{6}$$

3 Monitoring network design

Monitoring networks evolve over time. After starting with an initial set of measurement sites it is a common task to expand (or shrink) this network in an optimal way. Kriging does not only return estimated values but also the associated prediction variance. As this prediction variance only depends on the (planned) measurement locations it is possible to use G- and I-optimality criteria minimising maximum or mean prediction variance to choose an optimal subset fo a given set of candidate measurement points.

4 Implementation Details and Example

The implementation (see Gebhardt (2003)) is primarily based on the `sgeostat` library. In a first step a library `rgeostat` implementing universal kriging estimation in Fortran code based on LAPACK subroutines using the same data structures as in `sgeostat` has been written. Another library `baykrig` implements the above shown Bayesian approach based on empirical prior data (see Pilz (1991)). Finally library `kdesign` puts all things together and provides functions for optimal network design based on complete enumeration among a finite set of candidate measurement points. The usage of these libraries will be demonstrated in a short example.

Preliminary versions of these libraries are available at `ftp://ftp.uni-klu.ac.at/pub/R`.

References

C. Gebhardt. *Bayessche Methoden in der geostatistischen Versuchsplanung.* PhD thesis, University of Klagenfurt, 2003.

H. Omre. Bayesian Kriging - merging observations and qualified guesses in kriging. *Mathematical Geology*, 19:25–39, 1987.

J. Pilz. Ausnutzung von a-priori-Kenntnissen in geostatistischen Modellen. In G. Peschel, editor, *Beiträge zur Mathematischen Geologie und Geoinformatik*, volume 3, pages 74–79, Köln, 1991. Verlag Sven von Loga.