Numerical study on various estimators of the parameters of a Matérn covariance function for kriging

S. Petit
L2S, CentraleSupélec, University of Paris-Saclay, Safran Aircraft Engines

Supervisor(s): J. Bect (CentraleSupélec), S. Da-Veiga (Safran Tech), P. Feliot (Safran Aircraft Engines), E. Vazquez (CentraleSupélec)

Ph.D. expected duration: Apr. 2019 - Mar. 2022

Address: 8-10 Rue Joliot Curie, 91190 Gif-sur-Yvette

Email: sebastien.petit@centralesupelec.fr

Abstract:

This work studies empirically a variety of estimators of the parameters of the Matérn class of covariances functions.

The context of our work is that of industrial simulators, which are more time consuming and harder to interpret as the complexity increases. As an example, Safran Aircraft Engines faces this problem for designing turbomachines with complex physical simulations that can take several hours to run for the physics of a fan blade. As a consequence, there is a growing need in *surrogate* models that are able to capture information from a few runs of a simulator. Kriging methods [1, 2] have emerged as powerful tools for solving such tasks. The advantages of those methods are their flexibility and their nice probabilistic interpretation for Gaussian processes that allows to report analytically full posterior predictive distributions of unknown points from observed ones.

Kriging is often used to make predictions for random fields with unknown covariance function. In such cases, the covariance function is often supposed to belong to some parametric family. Kriging accuracy then relies criticaly on the estimation of those parameters. As pointed out in [3], the fact that we are only capable in general to observe one trajectory limits our ability to estimate processes parameters and encourages us to suppose that some ergodicity-related conditions are met as for example weak stationarity. This suggests that the input training data distribution has an influence on the accuracy of the estimators [4, 5].

The Matérn class of covariance functions has received considerable attention in the literature [4]. One reason for that is its parametrization that allows to represent efficiently approximation properties. Especially, its regularity and input scaling parameters account for the asymptotical and small sample predictor approximation errors.

A large variety of methods exists to estimate the parameters of a Gaussian process model of an unknown function. Most popular methods are *maximum likelihood estimation* and related techniques [6]. One alternative that has been studied is *cross-validation* which can be computed analytically thanks to the analytical Gaussian posterior formulas available [7, 8]. Most of the literature on this topic focus on *leave-one-out mean square predictor error*. As the prediction is probabilistic, a broad variety of other *cross-validation scores* can be used to estimate the parameters among which *log predictive density* and *continous rank probability scores* [9]. Another procedure called *generalized cross validation* was also proposed in the case of a *mean square predictor error* loss [8]. To our knowledge, this method has not been tested yet for the estimation of input scaling parameters such as the one of Matérn covariance functions. Other metrics such as *kernel alignment* [10] or RKHS norms (see *native spaces* norm in [11]) have been proposed in the literature of *function approximation* and have not been tested in the context of Matérn covariance functions parameters estimation to our knowledge.

This work aims at comparing the accuracies of those estimators. We run numerical tests in different settings such as isotropic or anisotropic fields, short or large input scale fields, different regularities and varying training sets sizes and inputs distributions. We also compare them in the case of a mispecified covariance function. The contribution of this work is to enrich numerical experiments that are presented in the literature [12, 13]. For instance we want to investigate the role of the regularity parameter in the asymptotical behaviour of the errors on the estimated parameters. We hope for this study to provide useful guidelines in various settings.

References

[1] G. Matheron. Principles of geostatistics. *Economic Geology*, 58(8):1246–1266, 12 1963. ISSN 0361-0128.

[2] T. J. Santner, Williams B., and Notz W. *The Design and Analysis of Computer Experiments, Second Edition.* Springer-Verlag, 2018.

[3] G. Matheron. *Estimer et choisir : essai sur la pratique des probabilités.* 1978.

[4] M. L. Stein. *Interpolation of spatial data.* Springer Series in Statistics. Springer-Verlag, New York, 1999. ISBN 0-387-98629-4. Some theory for Kriging.

[5] F. Bachoc. Asymptotic analysis of the role of spatial sampling for covariance parameter estimation of Gaussian processes. January 2013.

[6] P. K. Kitanidis. Statistical estimation of polynomial generalized covariance functions and hydrologic applications. *Water Resources Research*, 19(4):909–921, 1983.

[7] O. Dubrule. Cross validation of kriging in a unique neighborhood. *Journal of the International Association for Mathematical Geology*, 15:687–699, 1983.

[8] P. G. Craven and G. Wahba. Smoothing noisy data with spline functions: estimating the correct degree of smoothing by the method of generalized cross-validation. 1978.

[9] T. Gneiting and A. Raftery. Strictly proper scoring rules, prediction, and estimation. *Journal of the American Statistical Association*, 102:359–378, 03 2007.

[10] N. Cristianini, J. Shawe-Taylor, A. Elisseeff, and J. Kandola. On kernel-target alignment. In *Proceedings of the 14th International Conference on Neural Information Processing Systems: Natural and Synthetic*, NIPS'01, pages 367–373, Cambridge, MA, USA, 2001. MIT Press.

[11] H.Wendland. Scattered data approximation. 2004.

[12] E. Vazquez. *Kernel-based non-linear systems black-box modeling and applications.* Theses, Université Paris Sud - Paris XI, May 2005.

[13] F. Bachoc. Cross validation and maximum likelihood estimation of hyper-parameters of gaussian processes with model misspecification. *Computational Statistics and Data Analysis*, 66, 01 2013.

Short biography – With a background in Applied Mathematics, Sébastien began his PhD in April 2019 with CentraleSupélec and Safran Aircraft Engines. This PhD thesis is funded by Safran Aircraft Engines, under the scope of a CIFRE Agreement, and aims at further developing Safran Aircraft Engines *Bayesian optimization* tools.