Regularized estimation for multivariate Gaussian random fields

Gabriel Riffo
Former Student: Magíster en Matemática, Departamento de Matemática, Universidad Técnica Federico Santa María
Advisor: Francisco Cuevas
Departamento de Matemática, Universidad Técnica Federico Santa María

Abstract

The formulation of covariance models for multivariate Gaussian random fields was of utmost importance, as they formed the basis for modeling and prediction. This motivated the extension of univariate models such as the linear model of coregionalization, the Kummer model, the Matérn model, among others. The extension of these models generated constraints in the parameter space and had to be handled with care. However, current estimation methods presented problems by providing solutions that did not necessarily respect the parameter space. Additionally, these extensions were achieved through an increase in the number of parameters involved, which led to an increase in computation time and memory usage due to the computational costs that estimation and prediction processes entailed. For these reasons, at the time of writing this text, parametric models had only been implemented for a low number of variables (at most 5 variables). Therefore, it was necessary to explore appropriate estimation techniques to work with data with a large number of variables.

In this work, LASSO-type penalization approaches were addressed for the estimation of multivariate random fields. This strategy reduced the computational burden while allowing the identification of significant attributes of the random field, such as determining when two variables were correlated, thus reducing the problem's dimension. For the estimation, we implemented a projected gradient algorithm that ensured the estimates respected the parameter space constraints. This methodology could be applied to various parametric covariance models. In this work, we focused on the estimation of the multivariate Matérn model, applying a penalization on the parameter that modeled the correlation between the different variables. For the estimation, we used the likelihood functions and the composite likelihood function for multivariate random fields, which helped us reduce the computational complexity of the problem.