

Covariance functions for massive space-time data set.¹

Viktor P. Zastavnyi[†] and Emilio Porcu^{‡ 2}

[†] Donetsk National University

[‡] Universitat George August, Goettingen

Abstract: We discuss alternative covariance-based procedures allowing for considerable computational gains when dealing with space-time data sets. On the one hand tapering, on the other hand the use of compactly supported covariance functions can both represent powerful tools for this purpose.

Keywords: Covariance, Positive definiteness, Tapering

1. Introduction and Motivation

Recent literature persistently emphasizes the use of approximation methods and new methodologies for dealing with massive spatial data set. When dealing with spatial data, calculation of the inverse of the covariance matrix becomes a crucial problem. For instance, the inverse is needed for best linear unbiased prediction (alias kriging), and is repeatedly calculated in the maximum likelihood estimation or the Bayesian inferences. Thus, large spatial sample sizes traduce into big challenges from the computational point of view.

A natural idea that made proselytes in the last year is to make the covariances exactly zero after certain distance so that the resulting matrix has a high proportion of zero entries and is therefore a sparse matrix. Operations on sparse matrices take up less computer memories and run faster. However, this should be done in a way to preserve positive definiteness of the resulting covariance matrix. The idea goes under the name of covariance *tapering*, by meaning that the true covariance is multiplied pointwise with a compactly supported and radial correlation function. This operation is technically justified by the fact that the Schur product preserves positive definiteness.

The effects of tapering in terms of estimation and interpolation have been recently inspected by Du, Zhang and Mandrekar (3), where general conditions are given in order to ensure that tapering does not affect the efficiency of the maximum likelihood estimator. For spatial interpolation, (4) show that under some regularity conditions, tapering procedures yield asymptotically optimal prediction. In order to assess these properties, the asymptotic framework adopted by the authors is of the infill type, and the tool allowing to evaluate the performances of tapering is the equivalence of Gaussian measures, for which a comprehensive theory can be found in the seminal work by Yadrenko (8).

These points fix very briefly the state of the art and we refer the reader to (2) for an excellent survey on the topic.

Although tapering has been well understood in the spatial framework, there is nothing done, to the knowledge of the authors, for the spatio-temporal case. In particular, the use

¹Emilio Porcu acknowledges DFG-SNF Research Group FOR916, subproject A2

²Address of correspondence: Universitat Goettingen, Institut fur Mathematische Stochastik, Goldschmidtstrasse 7, 37077 Goettingen, Germany. email: porcu@mat.uni.gwdg.de

of tapering is at least questionable in space-time, since the same type of asymptotics does not apply and thus it is not easy to evaluate its performances.

But a deeper look at this problem also highlights the non existence, in the literature, of space-time covariance functions that are compactly supported over space, time or both. These facts motivate the research documented in this manuscript.

2. Review of space-time covariance functions

Geostatistical approaches have been developed to fit Random Function (RF) models in continuous space and time, based on a limited number of spatially and/or temporally dispersed observations. These approaches model the observations as a partial realization of a spatio-temporal (typically Gaussian) Random Function.

Let $Z(\mathbf{x}, t)$, $\mathbf{x} \in \mathbb{R}^d$ and $t \in \mathbb{R}$, be a stationary real valued Gaussian spatio-temporal RF with constant mean, which does not depend on the space-time coordinates, and with constant and finite variance. Then, the function $C(\mathbf{h}, u) = \text{cov}(Z(\mathbf{x}_i, t_i), Z(\mathbf{x}_j, t_j))$ is defined for $(\mathbf{h}, u) = (\mathbf{x}_i - \mathbf{x}_j, t_i - t_j) \in \mathbb{R}^d \times \mathbb{R}$ and is called the *stationary* spatio-temporal covariance function of the process, as it exclusively depends on the spatial and temporal separation vectors, respectively $\mathbf{h}$ and u . As well known, a real valued function C defined on the product space $\mathbb{R}^d \times \mathbb{R}$ is the covariance function associated to a stationary spatio-temporal RF if and only if

$$\sum_{i=1}^n \sum_{j=1}^n a_i a_j C(\mathbf{x}_i - \mathbf{x}_j, t_i - t_j) \geq 0,$$

for all finite sets of real coefficients a_i and points $(\mathbf{x}_i, t_i) \in \mathbb{R}^d \times \mathbb{R}$, $i = 1, \dots, n$.

This property is called semi-positive definiteness. Bochner's theorem sets up the equivalence between covariance and characteristic functions of non-negative finite measures. Namely, a continuous function C as defined above is positive definite if and only if

$$C(\mathbf{h}, u) = \int_{\mathbb{R}^d} \int_{\mathbb{R}} e^{i(\omega \mathbf{h} + \tau u)} dF(\omega, \tau), \quad (1)$$

with F a non-negative measure on $\mathbb{R}^d \times \mathbb{R}$. A stationary covariance function is called isotropic if it is rotation and translation-invariant, *i.e.*

$$C(\mathbf{h}, u) = \tilde{C}(\|\mathbf{h}\|, |u|),$$

where with $\|\cdot\|$ we denote the usual Euclidean norm. In the merely spatial or temporal isotropic case, the Fourier transform in (1) simplifies to the so called Hankel transform, *i.e.* an integral with respect to mixtures of Bessel functions. In this case, covariance functions (or equivalently positive definite *radial functions*) are the characteristic functions of spherically symmetric distributed random vectors. For similar problems we refer the reader to Zastavnyi (9) (10), (11) and (12).

The Gneiting class of covariance functions (5) has been very popular in the last years: for $(\mathbf{h}, u) \in \mathbb{R}^{d+1}$, the function

$$(\mathbf{h}, u) \mapsto K(\mathbf{h}, u) := h(|u|^2)^{-d/2} \varphi\left(\frac{\|\mathbf{h}\|^2}{h(|u|^2)}\right) \quad (2)$$

is positive definite, for φ completely monotone on the positive real line and h a Bernstein function. A tempting choice for a compactly supported function, at least over space, would be to replace the function φ with another that is compactly supported, that is it vanishes outside the unit ball. Zastavnyi and Porcu ((13)) find some necessary conditions for this class to be permissible over space and time. The sufficient part still is an open problem.

We believe that this kind of functions represents a more natural choice than the tapering proposed by the previously mentioned authors.

Another aspect related to tapering regards the possibility of tapering the variogram instead of the covariance function. As well known, the product of two variograms is not necessarily a variogram. The results in Porcu and Schilling (7) can open new avenues in this direction.

References

- [1] Chiles, J.P., and Delfiner, P. (1999). *Geostatistics: Modeling Spatial Uncertainty*. Wiley, New York.
- [2] Du, J., Zhang, H. (2009). Covariance tapering in spatial statistics. *Positive Definite Functions: from Schoenberg to Space-Time Challenges*. J. Mateu and E. Porcu (Eds). Editorial Universitat Jaume I. Department of Mathematics.
- [3] Du, J., Zhang, H., and Mandrekar, V. (2007). Infill asymptotic properties of tapered maximum likelihood estimators. *Annals of Statistics*, to appear.
- [4] Furrer, R., Genton, M. G., and Nychika, D. (2006). Covariance tapering for interpolation of large spatial datasets. *Journal of Computational and Graphical Statistics*, **15**(3), 502–523.
- [5] Gneiting, T. (2002). Nonseparable, stationary covariance functions for space-time data. *Journal of the American Statistical Association*, **97**, 590-600.
- [6] Porcu, E., Gregori, P. and Mateu, J. (2006). Nonseparable stationary anisotropic space-time covariance functions. *Stochastic Environmental Research and Risk Assessment*, **21**, 2, 113-122.
- [7] Porcu, E. and Schilling, R. (2009). From Schoenberg to Pick-Nevanlinna: : towards a complete picture of the variogram class
- [8] Yadrenko, M. (1983). *Spectral Theory of Random Fields*. Optimization Software, New York.
- [9] Zastavnyi, V. P. (1991). Positive definite functions depending on a norm. Solution of a problem of Schoenberg. Preprint N1-35, *Inst. Appl. Math. Mech. Acad. Sci. Ukraine, Donetsk* (in Russian).
- [10] Zastavnyi, V. P. (1992). Positive definite functions depending on a norm. *Dokl. Russian Acad. Nauk* **325** 901-903 (English translation in *Russian Acad. Scie. Dokl. Math.* (1993) **46**(1) 112-114).
- [11] Zastavnyi, V. P. (1993). Positive definite functions depending on the norm. *Russian J. Math. Physics* **1**(4) 511-522.
- [12] Zastavnyi V.P. (2000). On positive definiteness of some functions. *Journal of Multivariate Analysis*, **73**, p. 55-81.
- [13] Zastavnyi, V.P., and Porcu, E. (2009). Characterization Theorems for Space-Time covariances and their Relations to Schoenberg Problem. Submitted.