Liet. matem. rink., T. 43, spec. nr., 515–519

Estimation of covariance functions for spatio-temporal data

Edita LESAUSKIENĖ, Kęstutis DUČINSKAS (KU)
e-mail: duce@gmf.ku.lt, l.edita@one.lt

1. Introduction

A large number of environmental phenomena may be regarded as realizations of spatio-temporal random process [1, 2] Geostatistics offers a variety of methods to model spatial data: however, applying such space approaches to spatio-temporal random processes, may lead to the loss of valuable information in the time dimension.

One obvious solution to this problem is to consider the spatio-temporal phenomenon as a realization of a random process defined in $\Re^{d+1}$ (i.e., d is the space dimension plus one time dimension). This approach demands the extension of the existing spatial techniques into the space-time domain. Despite the straightforward appearence of this extension, there are a number of theoretical and practical problems that should be addressed prior to any successfful application of geostatistical methods to spatio-temporal data.

Let $\{Z(s;t)\colon s \in D \subset \Re^d;\ t \in [0, \infty]\}$ denote a spatio-temporal random process. Optimal prediction (in space and time) of the unobserved parts of the process, based on the observated part of process is often the ultimate goal, but to achieve this goal, a model is needed for how various parts of the process co-vary in space and time. In what follows, we assume that the spatio-temporal process $Z(s;t)$ satisfies the regularity condition $var(Z(s;t)) < \infty$, for all $s \in D$, $t \geqslant 0$. Then we denote the mean function as

$$\mu(s;t) \equiv E(Z(s;t)),$$

and covariance function as

$$K(s, r; t, q) \equiv cov(Z(s;t), Z(r;q)); \quad s, r \in D,\ t > 0,\ q > 0.$$

Let $\{Z_{it}\}$ denote observations of $Z(s,t)$ at spatial locations from the set $\{s_i\colon i = 1, \ldots, m_t\}$ and time moments $t = 1, \ldots, T$. Suppose that we have data consisting of $N = \sum_{t=1}^{T} m_t$ observations of $Z(s,t)$. Here m_t denote the number of spatial locations observed at the time t. Set $Z = (z_{11}, \ldots, z_{m_1 1}, \ldots; z_{1T}, \ldots, z_{m_T T})'$.

Furthemore, the optimal [i.e., minimum MSPE (see, e.g., [1, 3])] linear predictor of $Z(s_0;t_0)$ is

$$Z^*(s_0;t_0) = \mu(s_0;t_0) + C(s_0;t_0)'\Sigma^{-1}(Z - \mu), \tag{1}$$

where $\Sigma \equiv cov(Z)$, $C(s_0;t_0)' \equiv cov(Z(s_0;t_0),Z)$, and $\mu \equiv E(Z)$. The MSPE of $Z^*(s_0;t_0)$ is $C(s_0;t_0)'\Sigma^{-1}C(s_0;t_0)$. In the rest of this article, we assume that the covariance function is stationary in space and time, namely

$$K(s,r;t,q) = C(s-r;t-q), \tag{2}$$

for a certain functions C. This assumption helped us to estimate the covariance function from real data. For any $(r_1;q_1),\dots,(r_m;q_m)$, any real $a_1,\dots,a_m$, and any positive integer m, C must satisfy positive definitness condition

$$\sum_{i=1}^{m}\sum_{j=1}^{m} a_i a_j C(r_i - r_j; q_i - q_j) \geqslant 0. \tag{3}$$

To ensure (3) one often specifies the covariance function C to belong to a parametric family whose members are known to be positive definite. That is, one assumes that

$$cov(Z(s;t), Z(s+h_s;t+h_t)) = C^0(h_s;h_t|\theta), \tag{4}$$

where C^0 satisfies (3) for all $\theta \in \Theta \in R^p$.

While there are no difficulties in extending the various kriging estimators and the kriging equations to the space-time setting, there has been a lack of known valid space-time covariances and variograms.

Ussualy defined two parametric families C^0 for (4), i.e., separable and nonseparable covariance function families.

In the paper [4], we have described the separable covariance functions family, where we have produced some exemples for this case. There we have presented the results in case for separate product model.

Let C_s be a covariance function on R^m and C_t be a covariance function on T, then the separate product model is

$$C_{st}(h_s;h_t) = C_s(h_s)C_t(h_t). \tag{5}$$

Other family is nonseparable function, when we can't separate the covariance functions for space and for time. In general case, nonseparable stationary covariance functions that model space-time interactions are in great demand. Using simple stochastic partial differential equations over space and time, Jones and Zhang [5] have developed a four-parameter family of spectral densities that implicitly yield such stationary covariance functions, although not in closed form.

Cressie and Huang [6] have presented a new methodology for developing whole class of nonseparable spatio-temporal stationary covariance functions in closed form. But this class of covariance functions cover ones that satisfied (4) and is described in terms of complicated integrals of spectral density functions.

One of the objectives of this paper is to perform how to avoid the usage of complicated covariance functions.

In the case of temporal independence statio-temporal covariance function is of the form

$$C_{st}(h_s; h_t) = C^*(h_s). \tag{6}$$

Usually it is necessary to carry out the prediction for each season separately. Then one had constructed covariance functions $\{C_i^*(h_s|\theta),\ i = 1, \dots, 4\}$, corresponding four seasons.

After averaging of spatio-temporal covariance models for all seasons we have seasonal average model:

$$C^0(h_s; h_t|\theta) = \sum_{i=1}^{4} \gamma_i C_i^*(h_s|\theta) = C(h_s). \tag{7}$$

2. Main results

After the covariance function estimation, the interpoliation between the measurement points was carried out. For this purpose, different geostatistical methods were used.

The kriging equations for space-time kriging are the same as for purely spatial problems, the difference is in the usage of a space-time covariance instead of a purely spatial covariance. In the case of regression model of mean function

$$\mu(s; t) = E(Z(s; t)) = X_{s;t}^T \beta \tag{8}$$

the optimal prediction is called universal kriging (see, e.g., [1]), where $X_{s;t}$ – regressors matrix, and β – regression parameters.

Assume that, $m_t = m,\ t = 1, \dots, T$. Then covariance matrix of Z has the form $C_t \otimes C_s$, where C_t is $T \times T$ temporal covariance function, and C_s is $m \times m$ spatial covariance function.

Lemma. *Optimal linear prediction equation for product covariance function, defined in* (5), *is*

$$\widehat{Z}_{UK}(s_0, t_0) = x_{s_0,t_0}^T \hat{\beta} + \delta^T [Z - X_{s,t} \hat{\beta}], \tag{9}$$

where

$$\hat{\beta} = \left(X_{s,t}^T (C_t^{-1} \otimes C_s^{-1}) X_{s,t}\right)^{-1} X_{s,t}^T (C_t^{-1} \otimes C_s^{-1}) Z, \tag{10}$$

$$\delta = (C_t^{-1} \otimes C_s^{-1})(C_{t0} \otimes C_{s0}), \tag{11}$$

where $\otimes$ is the Kronecker product and C_{s0} and C_{t0} are vectors of spatial and temporal covariances between predicted point with observed points.

Proof. Expressions (10) and (11) were obtained by using (5) in universal kriging equation.

Then mean squared prediction error for the predictor, given in (9), is of the form

$$MSPE_{UK} = \delta(0) - 2b^T(C_{t0} \otimes C_{s0}) + b^T(C_t^{-1} \otimes C_s^{-1})b, \tag{12}$$

where

$$\begin{aligned} b^T &= x_{s0,t0}^T(X_{s,t}^T(C_t^{-1} \otimes C_s^{-1})X_{s,t})^{-1}X_{s,t}^T(C_t^{-1} \otimes C_s^{-1}) \\ &\quad +(C_{t0}^T \otimes C_{s0}^T)(C_t^{-1} \otimes C_s^{-1})\left(I - X_{s,t}((X_{s,t}^T(C_t^{-1} \otimes C_s^{-1})X_{s,t}\right)^{-1} \\ &\quad \times X_{s,t}^T(C_t^{-1} \otimes C_s^{-1}) \end{aligned}$$

and $\delta(0) = C_s(0)C_t(0)$.

3. Example

In this section we apply the spatio-temporal stationary covariance functions to the problem of prediction at the unobserved locations. The spatio-temporal data, used in this article, was collected in the Baltic sea, where the number of observations is taken regularly in three monthly during the period (1994–1998) at 6 stations in the coastal zone. Solinity is the observed feature. Exponential spatio-temporal covariance model (5) was considered in [4]. For this data, fitted exponential covariance model (6) under the assumption of temporal independence, gives the expression

$$C^*(h_s) = 0.0973698 \exp\{-13.572|h_s|\}. \tag{13}$$

Using prediction equation(9) and mean squared prediction error MSPE equation (12) for the prediction at an unobserved location we have the following results

$$\widehat{Z}_{UK}(s_0) = 6.73264,$$

$$MSPE(\widehat{Z}_{UK}(s_0)) = 0.10826.$$

Taking into account seasonality one can choose four spatio-temporal covariance models, prediction of solinity and MSPE for spring, summer, autumn, and winter are given by

$$C^*(h_s) = 0.210409 \exp\{-13.572|h_s|\}, \tag{14}$$
$$\widehat{Z}_{UK}(s_0) = 6.8512, \quad MSPE(\widehat{Z}_{UK}(s_0)) = 0.10453;$$
$$C^*(h_s) = 0.200872 \exp\{-13.572|h_s|\}, \tag{15}$$
$$\widehat{Z}_{UK}(s_0) = 6.74233, \quad MSPE(\widehat{Z}_{UK}(s_0)) = 0.10813;$$
$$C^*(h_s) = 0.016463 \exp\{-13.572|h_s|\}, \tag{16}$$

$$\begin{aligned} &\widehat{Z}_{UK}(s_0) = 6.48502, \quad MSPE(\widehat{Z}_{UK}(s_0)) = 0.10994; \\ &C^*(h_s) = 0.035765\exp\{-13.572|h_s|\}, \\ &\widehat{Z}_{UK}(s_0) = 6.15312, \quad MSPE(\widehat{Z}_{UK}(s_0)) = 0.10322, \end{aligned} \tag{17}$$

respectively.

After averaging of exponential spatio-temporal covariance model for all seasons we have seasonal average model

$$C^*(h_s) = 0.115877\exp\{-13.572|h_s|\}. \tag{18}$$

Using this model for real data we get prediction of solinity at an unobserved location and MSPE:

$$\widehat{Z}_{UK}(s_0) = 6.727926, \quad MSPE(\widehat{Z}_{UK}(s_0)) = 0.1288901.$$

Thus, on the base of (13–18) we can claim that for this real data the covariance model (13), which yielded by assumption of temporal independence, is optimal.

References

[1] N. Cressie, *Statistics for Spatial Data*, Wiley, New York (1993).

[2] I. Rodriguez–Iturbe, J.M. Mejia, The design of rainfall networks in time and space, *Water Resources Reseach*, **10**, 713–729 (1974).

[3] K. Dučinskas, E. Lesauskienė, J. Šaltytė, Universal kriging, *Lith. Math. J.*, **40**, 277–280 (2000).

[4] K. Dučinskas, E. Lesauskienė, Separable spatio-temporal covariance functions, *Lith. Math. J.*, **42**, 497–500 (2002).

[5] R.H. Jones, Y. Zhang, Models for continuous stationary space-time processes, in: *Modelling Longitudinal and Spatially Correlated Data* (*Lecture Notes in Statistics*, **122**), Springer, New York (1997).

[6] N. Cressie, H. Huang, Classes of nonseparable, spatio-temporal stationary covariance functions, *Journal of the American Statistical Association*, **94**(448), *Theory and Methods*, 1330–1340 (1999).

Universalaus krigingo metodas duomenims, kintantiems erdvėje ir laike

K. Dučinskas, E. Lesauskienė

Straipsnyje gautos analitinės išraiškos UK (universalaus krigingo) ir MSPE (vidutinės kvadratinės prognozės klaidos), kai erdvės-laiko kovariacinė funkcija yra gaunama eliminuojant laiko įtaką bei imant vidutinę kovariacinę funkciją sezonų atžvilgiu. Paėmus realius duomenis (1994–1998 metais kas tris mėnesius Baltijos jūros tyrimų centro rinkti duomenys apie druskingumo kiekį šešiose pakrantės zonos stotyse), buvo įvertinti erdvės ir laiko kovariacijų parametrai ir atlikta optimali prognozė laisvai pasirinktame taške.