Liet. matem. rink., T. 40, spec. nr., 277–280

Universal kriging

Kęstutis DUČINSKAS, Edita LESAUSKIENĖ, Jūratė ŠALTYTĖ (KU)
e-mail: duce@gmf.ku.lt; jsaltyte@gmf.ku.lt

Suppose the spatial data $Z(s_1), \ldots, Z(s_n)$ observed at spatial locations $\{s_1, \ldots, s_n\}$ are modelled as a collection of random variables generated by the random field

$$Z(s) = x'(s)\beta + \delta(s), \tag{1}$$

where $x'(s) = (x_1(s), \ldots, x_q(s))$ is a $q \times 1$ vector of nonrandom regressors and $\beta = (\beta^1, \ldots, \beta^q)' \in B$ are parameter vectors, B being an open subset of R^q. Assume, that $\{\delta(s) : s \in D \subset \mathbf{R}^2\}$ is a zero-mean second-order stationary random Gaussian field with spatial covariance defined by a parametric model $\operatorname{cov}\{\delta(s), \delta(t)\} = C(s - t; \theta)$ for all $s, t \in D$, where $\theta \in \Theta$ is a $p \times 1$ parameter vector, Θ being an open subset of R^p.

Spatial prediction refers to predicting the unobserved value of $Z(s_0)$ at known spatial location $s_0 \in D$ from data $Z \equiv (Z(s_1), \ldots, Z(s_n))'$. Denote any predictor by $p(s_0)$ and let $\Sigma = \|\operatorname{cov}(Z(s_i), Z(s_j))\|_{i,j=1,\ldots,n}$, $c' = (C(s_0 - s_1; \theta), \ldots, C(s_0 - s_n; \theta))$. The prime always will denotes vector transpose in this paper.

DEFINITION. The mean-squared prediction error (MSPE) for any spatial predictor $p(s_0)$ is defined by

$$MSPE(p(s_0)) = E\big(Z(s_0) - p(s_0)\big)^2. \tag{2}$$

Using the model given by (1) we can write

$$Z = X\beta + \delta,$$

where X is an $n \times q$ matrix whose (i, j)th element is $x_j(s_i)$, $\delta \equiv (\delta(s_1), \ldots, \delta(s_n))'$.

Assume that the spatial dependence parameter θ is known. Then the best linear unbiased estimator of β is generalised-least-squares (GLS) estimator

$$\widehat{\beta}_{gls} = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Z,$$

that is, the value of β that minimizes

$$(Z - X\beta)'\Sigma^{-1}(Z - X\beta).$$

If $\Sigma = \sigma^2 I$, then the appropriate minimizer is

$$\widehat{\beta}_{ols} = (X'X)^{-1}X'Z,$$

which is the ordinary-least-squares (OLS) estimator of β. Because the OLS estimator does not require kowledge of Σ for many practical situations it was used even when $\Sigma \neq \sigma^2 I$.

Suppose it is desired to predict $Z(s_0)$ linearly from data Z using a uniformly unbiased predictor. That is, the predictor is of the form

$$p(s_0) = \lambda' Z, \quad \text{for} \quad \lambda' X = x', \tag{3}$$

where $\lambda' = (\lambda_1, \ldots, \lambda_n)$, $x' = (x_1(s_0), \ldots, x_q(s_0))$.

Universal kriging predictor $p_{uk}(s_0) = \lambda_u' Z$, where

$$\lambda_u' = \{c + X(X'\Sigma^{-1}X)^{-1}(x - X'\Sigma^{-1}c)\}' \Sigma^{-1}, \tag{4}$$

is the optimal predictor among ones defined in (3), which minimizes the MSPE (see, e.g., Cressie, p. 154).

Then minimum mean-squared predictor error or kriging variance is

$$MSPE(p_{uk}(s_0)) = C(0) - 2\lambda_u' c + \lambda_u' \Sigma \lambda_u. \tag{5}$$

Define two unbiased linear predictors for $Z(s_0)$ in the following way

$$p_{ols}(s_0) = x'\widehat{\beta}_{ols} = \eta' Z, \tag{6}$$

where

$$\eta' = x' \left(X'X\right)^{-1} X, \tag{7}$$

and

$$p_{gls}(s_0) = x'\widehat{\beta}_{gls} = \gamma' Z, \tag{8}$$

where

$$\gamma' = x' \left(X'\Sigma^{-1}X\right)^{-1} X'\Sigma^{-1}. \tag{9}$$

It is obvious that linear unbiased predictors $p_{ols}(\cdot)$ and $p_{gls}(\cdot)$ are less optimal than $p_{uk}(\cdot)$ in the sense of minimum of $MSPE$. But as we see from (4), (7) and (9) they are simpler and require less computation time in practical realizations. Comparision of $MSPE$ for predictors $p_{ols}(\cdot)$ and $p_{gls}(\cdot)$ with kriging variance given in (4) are presented in the following lemma.

Lemma. *Suppose that for the prediction of value of the random field $Z(s)$ defined in (1) at the location s_0 from observed data $Z' = (Z(s_1), \ldots, Z(s_n))$ we use three linear unbiased predictors $p_{ols}(s_0)$, $p_{gls}(s_0)$ and $p_{uk}(s_0)$. Then*

$$MSPE(p_{gls}(s_0)) = MSPE(p_{uk}(s_0)) + c'\Sigma^{-1}\left(I - X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\right)c, \quad (10)$$

$$\begin{aligned} MSPE(p_{ols}(s_0)) = {} & MSPE(p_{uk}(s_0)) + c'\Sigma^{-1}\left(I - X(X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}\right)c \\ & + 2x'\left((X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1} - (X'X)^{-1}X'\right)c \\ & + x'\left((X'X)^{-1}X'\Sigma^{-1}X(X'X)^{-1} - (X'\Sigma^{-1}X)^{-1}\right)x. \quad (11) \end{aligned}$$

Proof. Since considered predictors are unbiased then

$$MSPE(p_{ols}(s_0)) = D(Z(s_0) - p_{OLS}(s_0)) = C(0) - 2\eta'c + \eta'\Sigma\eta, \quad (12)$$

$$MSPE(p_{GLS}(s_0)) = D(Z(s_0) - p_{GLS}(s_0)) = C(0) - 2\gamma'c + \gamma'\Sigma\gamma, \quad (13)$$

Substituting (7), (9) respectively in (12), (13) and using (4), (5) we complete the proof of the stated theorem.

For comparision of these methods of spatial prediction we use the quantities

$$\kappa_1 = \frac{MSPE_{gls} - MSPE_{uk}}{MSPE_{uk}}, \qquad \kappa_2 = \frac{MSPE_{ols} - MSPE_{uk}}{MSPE_{uk}}.$$

These quantities can be easily calculated by (10) and (11) in the case of known spatial dependence parameter θ. But in practical situations usually true value of the parameter θ is unknown. Then often predicting proceeds in the following way, by assuming that estimated value $\widehat{\theta}$ is the true value. Coefficients κ_1 and κ_2 then are calculated by formulae (10), (11) with θ substituted by $\widehat{\theta}$.

EXAMPLE. Data from Klaipeda Marine Research Centre surveyed at 9 locations in the costal zone of Baltic sea are used. PH is considered as a response variable and depth, temperature and salinity are considered as regressors. Isotropic exponential covariance model (see, e.g., Ripley, 1981, p. 56) is used for fitting. Fitted by least squares exponential model of covariance function is

$$C(h; \widehat{\theta}) = 10.447e^{-0.008h},$$

where $\widehat{\theta}' = (10.447; 0.008)$.

For fitted covariance model $\kappa_1 = 17.909$ and $\kappa_2 = 8.536$. That shows the significant advantage of the universal kriging prediction over the two proposed unbiased predictors.

References

[1] N.A.C. Cressie, *Statistics for Spatial Data*, Wiley Sons, NY (1993).
[2] B.D. Ripley, *Spatial Statistics*, Wiley Sons, NY (1981).

Universalus krigingas

K. Dučinskas, E. Lesauskienė, J. Šaltytė

Straipsnyje palyginami tiesinės nepaslinktos prognozės metodai su optimaliu universalaus krigingo metodu stacionarių Gauso laukų atveju. Pateikiamos analitinės išraiškos šių metodų vidutinei kvadratinei prognozės klaidai skaičiuoti. Šių formulių pagalba palyginami prognozės metodai realiems duomenims, naudojant eksponentinį kovariacijų modelį.