Liet. matem. rink., T. 42, spec. nr., 577–583

Discriminant analysis of spatial-temporal data

Jūratė ŠALTYTĖ–BENTH, Kęstutis DUČINSKAS (KU)
e-mail: jsaltyte@gmf.ku.lt, duce@gmf.ku.lt

1. Introduction

In the last 40 years an increasing amount of data is collected in the form of images. The analysis of images is related with the reconstruction and interpretation of images influenced by noises or even by certain (non-linear) transformations [1], [2]. Images include data collected from agricultural field trials, using remote sensing, microscopy, etc. When analysing image data, pattern recognition techniques are of great importance. And one of the most important problems in pattern recognition is the classification problem [3]. For example, in the recognition of electrocardiograms the classes are disease categories plus the class of normal subjects. There are a lot of methods of classification, and discriminant analysis (DA) is one of them.

The notion that data close together, in time and space, are likely to be correlated is natural. Here could be mentioned problems related to the pollution of atmosphere with chemical wastes, changing of meteorological conditions, etc. And DA in such areas is very usefull.

When classes are completely specified, an optimal classification rule in the sense of minimum classification error is the Bayesian classification rule. In practice, however, the complete description of classes usually is not possible and for the estimation of probabilistic characteristics of each class the training samples are required. When estimators of unknown parameters are used, the expressions for the expected error rate are very cumbersome even for the simplest procedures of DA. Therefore, asymptotic expansions of the expected error rate are especially important.

2. Model and problem

Suppose the model of $Z(\mathrm{s};t)$ in population Ω_l is $Z(\mathrm{s};t) = \mu_l + \varepsilon_l(\mathrm{s};t)$, where μ_l is the mean and $\{\varepsilon_l(\mathrm{s};t)\colon \mathrm{s} \in D \subset R^2, t \in [0,\infty)\}$ is a zero-mean intrinsically stationary Gaussian random field with stationary (in space and time) spatial-temporal covariance function defined by a parametric model $cov\{\varepsilon_l(\mathrm{s};t), \varepsilon_l(\mathrm{u};v)\} = \sigma(\mathrm{s}-\mathbf{u}, t-v;\theta_l)$ for all $\mathrm{s},\mathbf{u} \in D$, $t > 0$, $v > 0$, where $\theta_l \in \Theta$ is a $m \times 1$ parameter vector, Θ being an open subset of R^m, $l = 1,2$. We restrict our attention to the homoscedastic models, i.e., $\sigma(\mathbf{0},0;\theta) = \sigma$, for each $\theta \in \Theta$. Then the spatial-temporal covariance function in Ω_l is $cov\{\varepsilon_l(\mathrm{s};t), \varepsilon_l(\mathrm{u};v)\} = c(\mathrm{s}-\mathbf{u}, t-v;\theta_l)\sigma^2$, where $c(\mathrm{s}-\mathbf{u}, t-v;\theta_l)$

is the spatial-temporal correlation function, $l = 1, 2$. It is assumed that the function $c(\mathbf{s} - \mathbf{u}, t - v; \theta_l)$ is positive definite [4]. Assume that, for all $\mathbf{s}, \mathbf{u} \in D$, $t > 0, v > 0$, $\mathbf{s} \neq \mathbf{u}, t \neq v$, $cov\left\{\varepsilon_1(\mathbf{s}; t), \varepsilon_2(\mathbf{u}; v)\right\} = 0$.

Consider the problem of classification of observation $Z(\mathbf{r}; w) = z(\mathbf{r}; w)$, with $\mathbf{r} \in D_0 \subset D$, $w > 0$, into one of two populations specified above. Under the assumption that the populations are completely specified and for known prior probabilities of populations $\pi_1(\mathbf{r}; w), \pi_2(\mathbf{r}; w)$ ($\pi_1(\mathbf{r}; w)+\pi_2(\mathbf{r}; w) = 1$), the Bayes classification rule (BCR) $d_B(\cdot)$ minimizing the probability of misclassification (PMC) is

$$d_B(z(\mathbf{r}; w)) = \arg \max_{\{l=1,2\}} \pi_l(\mathbf{r}; w) p_l(z(\mathbf{r}; w)), \tag{1}$$

where $\pi_l(\mathbf{r}; w)$ is a prior probability and $p_l(z(\mathbf{r}; w))$ is a probability density function of Ω_l, $l = 1, 2$.

Denote by P_B the PMC of BCR, usually called Bayes error rate.

In practical applications the parameters of density function are usually not known. Then the estimators of unknown parameters can be found from training samples T_1 and T_2 taken separately from Ω_1 and Ω_2, respectively. When estimators of unknown parameters are used, the plug-in version of BCR is obtained. The performance of the plug-in version of the BCR when parameters are estimated from training samples with independent observations or time series observations or observations, which are spatially correlated is widely investigated by many authors (see, e.g., [5], [6], [7], [8]). In this paper, we shall consider the performance of the plug-in linear DF when the parameters are estimated from training samples being realisations of spatial-temporal Gaussian random field. Here the maximum likelihood estimators of unknown means and common variance assuming spatial-temporal correlation to be known are considered.

Suppose the spatial-temporal random field is observed at N_l spatial-temporal coordinates in region $D_1 \subset D$, $D_1 \cap D_0 = \emptyset$, i.e., we observe the training sample $T = \{T_1, T_2\}$ with $T_l = \{Z_{l1}, \ldots, Z_{lN_l}\}$, where $Z_{l\alpha} = Z\left(\mathbf{s}^l_\alpha; t^l_\alpha\right)$ denotes the α'th observation from Ω_l, $\alpha = 1, \ldots, N_l$, $l = 1, 2$. Assume that D_1 is beyond the zone of influence of D_0. Then $Z(\mathbf{r}; w)$ is independent on T.

Let $\widehat{\mu}_l$ and $\widehat{\sigma}^2$ be the estimators of μ_l and σ^2, respectively, based on T.

The plug-in rule $d_B\left(z(\mathbf{r}; w); \widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}^2\right)$ is obtained by replacing the parameters in (1) with their estimators. Then the corresponding sample linear discriminant function $\widehat{L}$ is defined as

$$\widehat{L} = \left(z(\mathbf{r}; w) - \frac{1}{2}(\widehat{\mu}_1 + \widehat{\mu}_2)\right)(\widehat{\mu}_1 - \widehat{\mu}_2)\frac{1}{\widehat{\sigma}^2} + g(\mathbf{r}; w),$$

where $g(\mathbf{r}; w) = \ln \frac{\pi_1(\mathbf{r};w)}{\pi_2(\mathbf{r};w)}$.

DEFINITION 1. The actual error rate for $d_B\left(z(\mathbf{r}; w); \widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}^2\right)$ is defined as

$$P(\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}^2) = \sum_{l=1}^{2} \pi_l(\mathbf{r}; w) \int \left(1 - \delta\left(l, d_B\left(z(\mathbf{r}; w); \widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}^2\right)\right) p_l\left(z(\mathbf{r}; w); \mu_l, \sigma^2\right)\right) \mathrm{d}z(\mathbf{r}; w),$$

where $\delta(\cdot,\cdot)$ is the Kronecker's delta.

In the considered case the actual error rate for $d_B\left(z(\mathbf{r};w);\widehat{\mu}_1,\widehat{\mu}_2,\widehat{\sigma}^2\right)$ can be rewritten as

$$P\left(\widehat{\mu}_1,\widehat{\mu}_2,\widehat{\sigma}^2\right)=\sum_{l=1}^{2}\pi_l(\mathbf{r};w)\,\Phi\Big((-1)^l\,\frac{\left(\mu_l-\frac{1}{2}\left(\widehat{\mu}_1+\widehat{\mu}_2\right)\right)\left(\widehat{\mu}_1-\widehat{\mu}_2\right)+\widehat{\sigma}^2 g(\mathbf{r};w)}{\widehat{\sigma}\sigma\sqrt{\left(\widehat{\mu}_1-\widehat{\mu}_2\right)^2}}\Big),$$

where $\Phi(\cdot)$ is the standard normal distribution function.

DEFINITION 2. The expectation of the actual error rate with respect to the distribution of T, designated as $E_T\left\{P\left(\widehat{\mu}_1,\widehat{\mu}_2,\widehat{\sigma}^2\right)\right\}$, is called the expected error rate (EER) for the $d_B\left(z(\mathbf{r};w);\widehat{\mu}_1,\widehat{\mu}_2,\widehat{\sigma}^2\right)$.

The goal of this paper is to find an asymptotic expansion of EER associated with plug-in LDF. Here we present the asymptotic expansion up to the order $\left(N^{-2}\right)$, where $N=N_1+N_2$, for the EER of classifying spatially distributed Gaussian observation with different means and common spatially factorised covariance. Terms of higher order are omitted from the asymptotic expansion since their contribution is in general negligible [9]. The ML estimators of means and common variance are used in the plug-in version of the BCR. A set of calculations for a certain training sample structure and two separable spatial-temporal correlation models is performed in order to estimate the plug-in BCR and compare (in sense of EER) considered correlation functions. Separable models are often chosen for convenience rather than for their ability to fit the data well; at least they are guaranteed to satisfy positive definitness condition and hence are valid [4].

3. Asymptotic expansion

The expectation vector and covariance matrix of the vectorised training sample T_l defined by $T_l^V=(Z_{l1},\ldots,Z_{lN_l})^T$ are

$$\mu_l^V=\mathbf{1}_{N_l}\cdot\mu_l \text{ and } \Sigma_l^V=\sigma^2 C_l,$$

respectively, where $\mathbf{1}_{N_l}$ is the N_l-dimensional vector of ones and C_l is the spatial-temporal correlation matrix of order $N_l\times N_l$, whose (α,β)'th element is $c_{l;\alpha\beta}=c\left(\mathrm{s}_\alpha-\mathrm{s}_\beta,t_\alpha-t_\beta;\theta_l\right),\alpha,\beta=1,\ldots,N_l,l=1,2$. Suppose that C_l is known and $\widehat{\mu}_l$, $\widehat{\sigma}^2$ are the maximum likelihood estimator of μ_l and σ^2, respectively, based on T_l, $l=1,2$. Let $C_l^{-1}=\left(c_l^{\alpha\beta}\right)$.

In [10], [11] it is shown, that, for $l=1,2$, the maximum likelihood estimators of μ_l and σ^2 are

$$\widehat{\mu}_l=\frac{1}{c_l^{\cdot\cdot}}\sum_{\alpha=1}^{N_l}c_l^{\cdot\alpha}Z_{l\alpha}, \tag{2}$$

and

$$\widehat{\sigma}^2 = \frac{1}{N}\sum_{l=1}^{2}\sum_{\alpha,\beta=1}^{N_l} c_l^{\alpha\beta}\left(Z_{l\alpha} - \widehat{\mu}_l\right)\left(Z_{l\beta} - \widehat{\mu}_l\right), \tag{3}$$

respectively, where $c_l^{\cdot\alpha} = \sum_{\beta=1}^{N_l} c_l^{\alpha\beta}$ and $c_l^{\cdot\cdot} = \sum_{\alpha,\beta=1}^{N_l} c_l^{\alpha\beta}$.

MLE under spatial sampling of Gaussian random fields was studied by Mardia, Marshall [12]. They gave the regularity conditions which ensure consistency and asymptotic normality of parameter estimators. We assume that in our case these conditions also hold.

Put $\Delta\widehat{\mu}_l = \widehat{\mu}_l - \mu_l, l = 1, 2, \Delta\widehat{\sigma}^2 = \widehat{\sigma}^2 - \sigma^2$ and let $\Delta^2\left(\mathbf{r}; w\right) = \frac{1}{\sigma^2}\left(\mu_1 - \mu_2\right)^2$ be the square of Mahalanobis distance. Let $\varphi\left(\cdot\right)$ denote the standard normal distribution density function. Let $P_l^{(1)} = \frac{\partial P()}{\partial\widehat{\mu}_l}$, $P_{\widehat{\sigma}^2}^{(1)} = \frac{\partial P()}{\partial\widehat{\sigma}^2}$, $P_{k,l}^{(2)} = \frac{\partial^2 P()}{\partial\widehat{\mu}_k\partial\widehat{\mu}_l}$, $P_{(\widehat{\sigma}^2)^2}^{(2)} = \frac{\partial^2 P()}{\partial(\widehat{\sigma}^2)^2}$ and $P_{l,\widehat{\sigma}^2}^{(2)} = \frac{\partial^2 P()}{\partial\widehat{\mu}_l\partial\widehat{\sigma}^2}$ be the partial derivatives of $P\left(\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}^2\right)$ up to second order with respect to the corresponding parameters evaluated at $\widehat{\mu}_1 = \mu_1, \widehat{\mu}_2 = \mu_2$ and $\widehat{\sigma}^2 = \sigma^2$, $l, k = 1, 2$.

Theorem. *Suppose* $\frac{1}{c_l^{\cdot\cdot}} \to 0$, *as* $N_l \to \infty$, $l = 1, 2$. *Then the asymptotic expansion of EER for the* $d_B\left(z\left(\mathbf{r}; w\right); \widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}^2\right)$ *is*

$$\begin{aligned} E_T\left\{P\left(\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}^2\right)\right\} &= \sum_{l=1}^{2}\pi_l\left(\mathbf{r}; w\right)\Phi\Big(-\frac{\Delta\left(\mathbf{r}; w\right)}{2} + (-1)^l\frac{g\left(\mathbf{r}; w\right)}{\Delta\left(\mathbf{r}; w\right)}\Big) \\ &+ \frac{\pi_1\left(\mathbf{r}; w\right)}{\Delta\left(\mathbf{r}; w\right)}\varphi\Big(-\frac{\Delta\left(\mathbf{r}; w\right)}{2} - \frac{g\left(\mathbf{r}; w\right)}{\Delta\left(\mathbf{r}; w\right)}\Big)\Big(\sum_{l=1}^{2}\frac{1}{2c_l^{\cdot\cdot}}\Big(-\frac{\Delta\left(\mathbf{r}; w\right)}{2} + (-1)^l\frac{g\left(\mathbf{r}; w\right)}{\Delta\left(\mathbf{r}; w\right)}\Big)^2 \\ &+ \frac{g^2\left(\mathbf{r}; w\right)}{N-2}\Big) + \mathrm{O}\left(M^{-2}\right), \end{aligned} \tag{4}$$

where $M = \min\left(c_1^{\cdot\cdot}, c_2^{\cdot\cdot}, N - 2\right)$.

Proof. Without loss of generality we use the convenient canonical form of $\mu_1 = -\mu_2 = \frac{\Delta(\mathbf{r}; w)}{2}$ and $\sigma^2 = 1$ (see, e.g., [13]). Expanding $P\left(\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}^2\right)$ in Taylor series about the true values of parameters we have

$$\begin{aligned} P\left(\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}^2\right) &= P_B + \sum_{l=1}^{2}P_l^{(1)}\Delta\widehat{\mu}_l + P_{\widehat{\sigma}^2}^{(1)}\Delta\widehat{\sigma}^2 \\ &+ \frac{1}{2}\Big(\sum_{l,k=1}^{2}P_{k,l}^{(2)}\Delta\widehat{\mu}_k\Delta\widehat{\mu}_l + P_{(\widehat{\sigma^2})^2}^{(2)}\left(\Delta\widehat{\sigma}^2\right)^2 + \sum_{l=1}^{2}P_{l,\widehat{\sigma^2}}^{(2)}\Delta\widehat{\mu}_l\Delta\widehat{\sigma}^2\Big) + O_3, \end{aligned} \tag{5}$$

where

$$P_B = \sum_{l=1}^{2}\pi_l\left(\mathbf{r}; w\right)\Phi\Big(-\frac{\Delta\left(\mathbf{r}; w\right)}{2} + (-1)^l\frac{g\left(\mathbf{r}; w\right)}{\Delta\left(\mathbf{r}; w\right)}\Big),$$

and O_3 is the third and higher order terms of $\Delta\widehat{\mu}_l$ and $\Delta\widehat{\sigma}^2$ and their products.

Since $P\left(\widehat{\mu}_1, \widehat{\mu}_2, \widehat{\sigma}^2\right)$ is minimised at $\mu_l = (-1)^{l+1}\frac{\Delta(\mathbf{r};w)}{2}$, $l = 1, 2$, and $\sigma^2 = 1$, then

$$P_l^{(1)} = P_{\widehat{\sigma}^2}^{(1)} = 0. \tag{6}$$

It is easy to show, that $E\left\{\Delta\widehat{\mu}_l\right\} = 0,\ E\left\{\Delta\widehat{\sigma}^2\right\} = 0,$

$$E\{\,(\Delta\widehat{\mu}_l)^2\,\} = \frac{1}{c_l^{\cdot\cdot}}, \quad E\{\,(\Delta\widehat{\sigma}^2)^2\,\} = \frac{2}{N-2}, \tag{7}$$

$$E\left\{\Delta\widehat{\mu}_k\Delta\widehat{\mu}_l\right\} = E\left\{\Delta\widehat{\mu}_l\Delta\widehat{\sigma}^2\right\} = 0. \tag{8}$$

Note, that, for $l = 1, 2$,

$$P_{l,l}^{(2)} = \frac{\pi_1(\mathbf{r};w)}{\Delta(\mathbf{r};w)}\varphi\Big(-\frac{\Delta(\mathbf{r};w)}{2} - \frac{g(\mathbf{r};w)}{\Delta(\mathbf{r};w)}\Big)\Big(-\frac{\Delta(\mathbf{r};w)}{2} + (-1)^l\frac{g(\mathbf{r};w)}{\Delta(\mathbf{r};w)}\Big)^2, \tag{9}$$

and

$$P_{(\Delta\widehat{\sigma}^2)^2}^{(2)} = \frac{\pi_1(\mathbf{r};w)}{\Delta(\mathbf{r};w)}g^2(\mathbf{r};w)\varphi\Big(-\frac{\Delta(\mathbf{r};w)}{2} - \frac{g(\mathbf{r};w)}{\Delta(\mathbf{r};w)}\Big). \tag{10}$$

By substituting estimators (2), (3) in (5), taking the expectation of the right side of (5) and using (6)–(10) we complete the proof of the theorem.

As the contribution of higher order terms in the presented asymptotic expansion is in generally negligible [9], for the evaluation of the performance of LDF the asymptotic expected error regret ($AEER$)

$$AEER = \frac{\pi_1(\mathbf{r};w)}{\Delta(\mathbf{r};w)}\varphi\Big(-\frac{\Delta(\mathbf{r};w)}{2} - \frac{g(\mathbf{r};w)}{\Delta(\mathbf{r};w)}\Big)$$
$$\times\Big(\sum_{l=1}^{2}\frac{1}{2c_l^{\cdot\cdot}}\Big(-\frac{\Delta(\mathbf{r};w)}{2} + (-1)^l\frac{g(\mathbf{r};w)}{\Delta(\mathbf{r};w)}\Big)^2 + \frac{g^2(\mathbf{r};w)}{N-2}\Big)$$

is used. Minimum of $AEER$ could also be used as a criterion for optimal training sample design.

The numerical comparison of $AEER$ obtained using two different separable spatial-temporal correlation functions is given in the example below.

4. Example

As an example consider the integer regular 2-dimensional lattice of size 15 × 15. We use two different designs for training samples from populations Ω_1 and Ω_2. In the sample T_1 there are 8 spatial locations, as in the sample T_2 – 11. Suppose, that the observations at

Table 1

Values of $AEER$'s with $\pi_1 = \pi_2 = 0.5$

Δ	P_B	$AEER_E$	$AEER_{OU}$	$AEER_{IND}$
1.50	0.3229	0.0846	0.2126	0.9347
2.00	0.2233	0.0287	0.0721	0.0659
2.50	0.1468	0.0111	0.0278	0.0109
3.00	0.0918	0.0045	0.0114	0.0043
3.50	0.0546	0.0019	0.0048	0.0015
4.00	0.0308	0.0008	0.0021	0.0005
4.50	0.0165	0.0003	0.0008	0.0002

each spatial location were taken 3 times, let's say once every decade. Thus, the sample sizes are $N_1 = 24$ and $N_2 = 33$.

A lot of valid spatial and temporal correlation models are readily available (e.g., [2]) and hence they can be combined in product form to give valid spatial-temporal correlation models. These are so called separable models.

Here we will consider two separable correlation models:

1. $c_l^1(\mathbf{h}, v) = \frac{1}{\sigma^2} \exp\left(-\kappa_1^l |\mathbf{h}| - \eta_1^l |v|\right)$, $\kappa_1^l > 0$, $\eta_1^l > 0$, $l = 1, 2$.

2. $c_l^2(\mathbf{h}, v) = \frac{1}{\sigma^2} \exp\left(-\kappa_2^l |\mathbf{h}|^2 - \eta_2^l |v|^2\right)$, $\kappa_2^l > 0$, $\eta_2^l > 0$, $l = 1, 2$.

In the first case, the correlation model consists of two exponential correlation functions, as in the second one – of two Ornstein–Uhlenbeck correlation functions. Both of them are isotropic. Consider $\kappa_1^1 = 0.1$, $\eta_1^1 = 0.2$, $\kappa_1^2 = 0.3$, $\eta_1^2 = 0.4$, $\kappa_2^1 = 0.2$, $\eta_2^1 = 0.3$, $\kappa_2^2 = 0.1$ and $\eta_2^2 = 0.2$.

In the Table 1 values of $AEER$ for considered two correlation functions are presented. Denote by $AEER_E$ and $AEER_{OU}$ the values of $AEER$, when exponential and Ornstein–Uhlenbeck correlation functions, respectively, are used. The values of $AEER$ in the case of independent observations are denoted by $AEER_{IND}$. Also the values of P_B are given.

The values of $AEER$'s approach zero when the distance between classes increases. As it was expected, the asymptotic expected error regret in the case of independent observations is the smallest one. The comparision of $AEER$ also shows, that exponential correlation function is better (gives smaller $AEER$) than Ornstein–Uhlenbeck correlation function for the considered neighbourhood.

Since the asymptotic EER for the case of dependent observations is bigger than that in the case of independence, it is very important take into consideration the spatial dependence factor, when practical problems are solved.

References

[1] N.A.C. Cressie, Geostatistics, *Amer. Stat. Assoc., Amer. Stat.*, 43(4), 197–202 (1989).

[2] N.A.C. Cressie, *Statistics for Spatial Data*, Wiley & Sons, New York (1993).

[3] A.K. Jain, R.P.W. Duin, J. Mao, Statistical pattern recognition: a review, *IEEE Trans. Pattern Anal. Machine Intell.*, **22**(1), 4–38 (2000).

[4] N.A.C. Cressie, H.–C. Huang, Classes of nonseparable, spatio-temporal stationary covariance functions, *Amer. Stat. Assoc., Theory and Methods*, **94**(448), 1330–1340 (1999).

[5] M. Okamoto, An asymptotic expansion for the distribution of the linear discriminant function, *Ann. Math. Statist.*, **34**, 1286–1301 (1963).

[6] C.R.O. Lawoko, G.J. McLachlan, Discrimination with autocorrelated observations, *Pattern Recognition*, **18**(2), 145–149 (1985).

[7] K. Ducinskas, An asymptotic analysis of the regret risk in discriminant analysis under various training schemes, *Lith. Math. J.*, **37**(4), 337–351 (1997).

[8] K. Ducinskas, J. Saltyte, The effect of spatial autocorrelation on the error rates of the linear discriminant functions, *Lith. Math. J.*, **42**(2), 169–178 (2002).

[9] M.J. Scherwish, Asymptotic expansions for the means and variances of error rates, *Biometrika*, **65**(1), 286–299 (1981).

[10] J. Saltyte, K. Ducinskas, Statistical classification based on observations of random Gaussian fields, *Mathematical Modelling and Analysis*, **4**, 153–162 (1999).

[11] J. Saltyte, K. Ducinskas, Comparison of ML and OLS estimators in discriminant analysis of spatially correlated observations, *Informatica*, **13**(2), 227–238 (2001).

[12] K.V. Mardia, R.J. Marshall, Maximum likelihood estimation of models for residual covariance and spatial regression, *Biometrika*, **71**, 135–146 (1984).

[13] G.J. McLachlan, *Discriminant Analysis and Statistical Pattern Recognition*, Wiley & Sons, New York (1992).

Erdvės-laiko duomenų diskriminantinė analizė

J. Šaltytė–Benth, K. Dučinskas

Sprendžiamas vidiniai stacionaraus vienmačio atsitiktinio erdvės-laiko Gauso lauko realizacijos priskyrimo vienai iš dviejų populiacijų su skirtingais vidurkiais ir faktorizuotomis kovariacijų matricomis uždavinys. Nežinomi vidurkiai ir bendra dispersija įvertinami pagal laike ir erdvėje koreliuotas mokymo imtis, tariant, kad erdvės-laiko koreliacijų funkcijos yra žinomos. Pateikiamas tikėtinos klasifikavimo klaidos pirmos eilės asimptotinis skleidinys. Asimptotinis klaidos prieaugis įvertinamas ir skaitiškai, naudojant dvi erdvės-laiko separabilias koreliacijų funkcijas.