

Comparing tests appear in model-check for normal regression with spatially correlated observations

Wayan Somayasa and Gusti A. Wibawa

Department of Mathematics, Halu Oleo University, Kendari 93232, Indonesia

E-mail: wayan.somayasa@uho.ac.id and gustiadihi.wibawa@uho.ac.id

Abstract. The problem of investigating the appropriateness of an assumed model in regression analysis was traditionally handled by means of F test under independent observations. In this work we propose a more modern method based on the so-called set-indexed partial sums processes of the least squares residuals of the observations. We consider throughout this work univariate and multivariate regression models with spatially correlated observations, which are frequently encountered in the statistical modelling in geosciences as well as in mining. The decision is drawn by performing asymptotic test of statistical hypothesis based on the Kolmogorov-Smirnov and Cramér-von Mises functionals of the processes. We compare the two tests by investigating the power functions of the test. The finite sample size behavior of the tests are studied by simulating the empirical probability of rejections of H_0 . It is shown that for univariate model the KS test seems to be more powerful. Conversely the Cramér-von Mises test tends to be more powerful than the KS test in the multivariate case.

1. Introduction

Checking the appropriateness of an assumed model becomes an important step in a serious empirical statistical model building in order to obtain a suitable model. By this reason the problem of model check or goodness-of-fit in regression analysis has obtained much attention. Classical F test studied in standard text book is frequently applied in the practice, see Graybill [10] and Seber and Lee [15] for the regression model with single response and Christensen [7] and Johnson and Wichern [11] for that with multiple responses. Asymptotic procedures have been proposed by authors such as Arnold [2] established F_n test for univariate model, whereas Arnold [3] and Zellner [22] proposed asymptotic method for multivariate model. The study for obtaining a better but more simple procedure was extended by Stute [19, 20] leading to the pioneering study using the empirical processes marked by the residuals.

The application of the set-indexed partial sums processes of the residuals in asymptotic model check for spatial regression was investigated in the work of Somayasa [16], Somayasa and et al [17, 18]. In these references the limit process was obtained by generalizing the geometric method proposed in Bischoff [5] and Bischoff and Somayasa [6]. The main idea of the study was motivated by the attempt of McNeill and Jandhyalla [12] and Xie and McNeill [21] to detect the existence of change in the regression functions by embedding the spatial residuals to stochastic process using the partial sums operator.

The partial sums approaches established in the literatures mentioned above were derived under a common assumption in that the observations are stochastically independent. However prediction method in geostatistics such as ordinary and universal kriging are mainly derived



under spatially correlated observations, see e.g. Cressie [8]. Kriging for isotropic observations was studied in Schabenberger and Gotway [14]. By this reason in this paper we study the application of the set-indexed partial sums technique in investigating the correctness of the assumed model when the correlation between the observations are inherent. However the underlying distribution model of the sample is still assumed to be normal. This normality assumption is important in order to keep the independence among the observations holds true, see Section 2. The limit process for univariate and multivariate case are established in Section 3. Simulation study for comparing the finite sample power functions of the tests is presented in Section 4. The paper is closed with conclusion and some remark for future research, see Section 5.

2. Model definition and estimation

We consider univariate and multivariate models with correlated observations. In univariate case the response consists of one measurement only, whereas in multivariate case two or more variables are measured simultaneously on each individual or point. That why the analysis must be carried out by taking into account not only the correlation between the observations (points) but also among the response variables in each point. That why the mathematical derivation of the method becomes more complicated.

2.1. Univariate case

Throughout this paper we consider second-order stationary Gaussian spatial process

$$Y := \{Y(\mathbf{t}) : \mathbf{t} := (t_1, \dots, t_d)^\top \in \Pi_{i=1}^d [a_i, b_i] := [\mathbf{a}, \mathbf{b}]\},$$

with the property $\mathbf{E}(Y(\mathbf{t})) = g(\mathbf{t})$ for an unknown real valued function g on $[\mathbf{a}, \mathbf{b}]$ and $Cov(Y(\mathbf{t}_j), Y(\mathbf{t}_k)) = \sigma(\mathbf{t}_j - \mathbf{t}_k)$, for every $\mathbf{t}_j, \mathbf{t}_k \in [\mathbf{a}, \mathbf{b}]$. That is Y has the covariance function which depends only on the difference $(\mathbf{t}_j - \mathbf{t}_k)$. Then the process Y admits a decomposition

$$Y(\mathbf{t}) = g(\mathbf{t}) + \varepsilon(\mathbf{t}), \quad \mathbf{t} \in [\mathbf{a}, \mathbf{b}], \quad (1)$$

where $\varepsilon(\mathbf{t})$ is the random error which is normally distributed with $\mathbf{E}(\varepsilon(\mathbf{t})) = 0$ and $Cov(\varepsilon(\mathbf{t}_j), \varepsilon(\mathbf{t}_k)) = \sigma(\mathbf{t}_j - \mathbf{t}_k)$. In the isotropic spatial regression model, the covariance function of Y is assumed to depend only on the Euclidean distance between two points. As for example Schabenberger and Gotway [14] assume a process with the covariance function given by $Cov(Y(\mathbf{t}_j), Y(\mathbf{t}_k)) = \sigma^2 \exp\{-\|\mathbf{t}_j - \mathbf{t}_k\|\}$, for any $\mathbf{t}_j, \mathbf{t}_k \in [\mathbf{a}, \mathbf{b}]$, where σ^2 is unknown positive constant. In the references cited in Section 1 and also in the usual theory of regression analysis the covariance function of the process Y was assumed to be constant, i.e. $Cov(Y(\mathbf{t}_j), Y(\mathbf{t}_k)) = \sigma_{jk} \delta_{jk}$, where δ_{jk} is the Kronecker delta. In contrast, for our result we do not need such assumption.

Let $\mathbf{W} := [f_1, \dots, f_p]$ be a subspace in $L_2(\lambda^d, [\mathbf{a}, \mathbf{b}])$, where $f_1, \dots, f_p$ are known real valued regression functions on $[\mathbf{a}, \mathbf{b}]$ and $L_2(\lambda^d, [\mathbf{a}, \mathbf{b}])$ is the space of square integrable functions with respect to the Lebesgue measure λ^d on $[\mathbf{a}, \mathbf{b}]$. We aim to test the hypothesis of the form

$$H_0 : g \in \mathbf{W} \text{ versus } H_1 : g \notin \mathbf{W}. \quad (2)$$

This means that under H_0 the unknown function g can be represented as a linear combination $g(\mathbf{t}) = \sum_{i=1}^p \beta_i f_i(\mathbf{t})$ for some unknown parameters $\beta_1, \dots, \beta_p$, while under H_1 we consider a nonparametric model.

Suppose Model 1 is observed over an experimental design given by a regular lattice $\Xi_{n_1 \dots n_d}$ with $n_1 \times \dots \times n_d$ experimental condition (design points), where $n_1 \geq 1, \dots, n_d \geq 1$. That is

$$\Xi_{n_1 \dots n_d} := \{(t_{n_1 j_1}, \dots, t_{n_d j_d}) : 1 \leq j_k \leq n_k, \quad k = 1, \dots, d\} \subset [\mathbf{a}, \mathbf{b}],$$

where for $k = 1, \dots, d$, $t_{n_k j_k} := a_k + \frac{j_k}{n_k}(b_k - a_k)$. If we have an observation on every point of $\Xi_{n_1, \dots, n_d}$ then we get an array of observed responses $\{Y_{j_1 \dots j_d} : 1 \leq j_1 \leq n_1, \dots, 1 \leq j_d \leq n_d\}$ that satisfies the regression model

$$Y_{j_1 \dots j_d} = g_{j_1 \dots j_d} + \varepsilon_{j_1 \dots j_d}, \quad 1 \leq j_1 \leq n_1, \dots, 1 \leq j_d \leq n_d,$$

where $Y_{j_1 \dots j_d}$ denotes the observation in the point $(t_{n_1 j_1}, \dots, t_{n_d j_d})$. Furthermore, $Y_{j_1 \dots j_d}$ is normally distributed with $\mathbf{E}(Y_{j_1 \dots j_d}) = g_{j_1 \dots j_d}$ and $Cov(Y_{j_1 \dots j_d}, Y_{j'_1 \dots j'_d}) = \sigma(\mathbf{t}_{j_1 \dots j_d} - \mathbf{t}_{j'_1 \dots j'_d})$ for every point $\mathbf{t}_{j_1 \dots j_d} := (t_{n_1 j_1}, \dots, t_{n_d j_d})$ and $\mathbf{t}_{j'_1 \dots j'_d} := (t_{n_1 j'_1}, \dots, t_{n_d j'_d})$ in $\Xi_{n_1 \dots n_d}$.

In order to establish estimation procedure for the parameters of the model we need to describe the model in vectors and matrices forms. Let $\mathbf{Y}_{n_1 \dots n_d}$ be the $n_1 \dots n_d$ -dimensional vector of random observation, $\mathbf{X}_{n_1 \dots n_d}$ be the $(n_1 \dots n_d) \times p$ -dimensional design matrix, $\beta := (\beta_1, \dots, \beta_p)^\top$ be p -dimensional vector of unknown parameters, and $\mathcal{E}_{n_1 \dots n_d}$ be the $n_1 \dots n_d$ -dimensional random errors. Then the model under H_0 can also be written as

$$\mathbf{Y}_{n_1 \dots n_d} = \mathbf{X}_{n_1 \dots n_d} \beta + \mathcal{E}_{n_1 \dots n_d}, \quad (3)$$

where $\mathbf{Y}_{n_1 \dots n_d}$ has the $n_1 \dots n_d$ dimensional normal distribution with mean $\mathbf{X}_{n_1 \dots n_d} \beta$ and variance-covariance matrix given by $\sigma^2 \Sigma$, say, defined by

$$\Sigma := \begin{pmatrix} \sigma(\mathbf{t}_{1 \dots 1} - \mathbf{t}_{1 \dots 1}) & \sigma(\mathbf{t}_{1 \dots 1} - \mathbf{t}_{21 \dots 1}) & \cdots & \sigma(\mathbf{t}_{1 \dots 1} - \mathbf{t}_{n_1 \dots n_d}) \\ \sigma(\mathbf{t}_{21 \dots 1} - \mathbf{t}_{1 \dots 1}) & \sigma(\mathbf{t}_{21 \dots 1} - \mathbf{t}_{21 \dots 1}) & \cdots & \sigma(\mathbf{t}_{21 \dots 1} - \mathbf{t}_{n_1 \dots n_d}) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma(\mathbf{t}_{n_1 \dots n_d} - \mathbf{t}_{1 \dots 1}) & \sigma(\mathbf{t}_{n_1 \dots n_d} - \mathbf{t}_{21 \dots 1}) & \cdots & \sigma(\mathbf{t}_{n_1 \dots n_d} - \mathbf{t}_{n_1 \dots n_d}) \end{pmatrix}$$

which is considered as an unknown $(n_1 \dots n_d) \times (n_1 \dots n_d)$ -dimensional matrix. In the case of spatial isotropy, we have

$$\Sigma := \begin{pmatrix} \sigma(\|\mathbf{t}_{1 \dots 1} - \mathbf{t}_{1 \dots 1}\|) & \sigma(\|\mathbf{t}_{1 \dots 1} - \mathbf{t}_{21 \dots 1}\|) & \cdots & \sigma(\|\mathbf{t}_{1 \dots 1} - \mathbf{t}_{n_1 \dots n_d}\|) \\ \sigma(\|\mathbf{t}_{21 \dots 1} - \mathbf{t}_{1 \dots 1}\|) & \sigma(\|\mathbf{t}_{21 \dots 1} - \mathbf{t}_{21 \dots 1}\|) & \cdots & \sigma(\|\mathbf{t}_{21 \dots 1} - \mathbf{t}_{n_1 \dots n_d}\|) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma(\|\mathbf{t}_{n_1 \dots n_d} - \mathbf{t}_{1 \dots 1}\|) & \sigma(\|\mathbf{t}_{n_1 \dots n_d} - \mathbf{t}_{21 \dots 1}\|) & \cdots & \sigma(\|\mathbf{t}_{n_1 \dots n_d} - \mathbf{t}_{n_1 \dots n_d}\|) \end{pmatrix}.$$

For a positive definite matrix Σ the root square $\Sigma^{1/2}$ exists, where $\Sigma^{1/2}$ is a matrix with the property $\Sigma^{1/2} \Sigma^{1/2} = \Sigma$. For $i = 1, \dots, (n_1 \dots n_d)$, let λ_i be the i -th eigenvalue of Σ associated with the normalized eigenvector $\mathbf{v}_i$, and let $\mathbf{D} := (\mathbf{v}_1 \dots \mathbf{v}_{n_1 \dots n_d})$ and $\mathbf{\Lambda} := \text{diag}(\lambda_1 \dots \lambda_{n_1 \dots n_d})$. Then we have a decomposition $\Sigma = \mathbf{D} \mathbf{\Lambda} \mathbf{D}^\top$, where $\mathbf{D}$ is an orthogonal matrix. Hence, $\Sigma^{1/2} = \mathbf{D} \mathbf{\Lambda}^{1/2} \mathbf{D}^\top$, where $\mathbf{\Lambda}^{1/2} = \text{diag}(\sqrt{\lambda_1} \dots \sqrt{\lambda_{n_1 \dots n_d}})$. Furthermore we define $\Sigma^{-1/2}$ as a matrix with the property $\Sigma^{-1/2} := (\Sigma^{1/2})^{-1} = (\Sigma^{-1})^{1/2}$. Therefore we have

$$\Sigma^{-1/2} = ((\mathbf{D} \mathbf{\Lambda} \mathbf{D}^\top)^{1/2})^{-1} = (\mathbf{D} \mathbf{\Lambda}^{1/2} \mathbf{D}^\top)^{-1} = \mathbf{D} \mathbf{\Lambda}^{-1/2} \mathbf{D}^\top,$$

where $\mathbf{\Lambda}^{-1/2} = (\mathbf{\Lambda}^{1/2})^{-1} = \text{diag}(\frac{1}{\sqrt{\lambda_1}} \dots \frac{1}{\sqrt{\lambda_{n_1 \dots n_d}}})$.

Upon transforming Model 3 using the linear mapping represented by the matrix $\Sigma^{-1/2}$ we get a new model with independent and identically distributed errors

$$\mathbf{Y}_{n_1 \dots n_d}^* = \mathbf{X}_{n_1 \dots n_d}^* \beta + \mathcal{E}_{n_1 \dots n_d}^* \quad (4)$$

where $\mathbf{Y}_{n_1 \dots n_d}^* := \Sigma^{-1/2} \mathbf{Y}_{n_1 \dots n_d}$, $\mathbf{X}_{n_1 \dots n_d}^* := \Sigma^{-1/2} \mathbf{X}_{n_1 \dots n_d}$, and $\mathcal{E}_{n_1 \dots n_d}^* = \Sigma^{-1/2} \mathcal{E}_{n_1 \dots n_d}$. For this transformed model the new vector of observations $\mathbf{Y}_{n_1 \dots n_d}^*$ is $n_1 \dots n_d$ -dimensional normally distributed with $\mathbf{E}(\mathbf{Y}_{n_1 \dots n_d}^*) = \mathbf{X}_{n_1 \dots n_d}^* \beta$ and the variance-covariance matrix given by

$$Cov(\mathbf{Y}_{n_1 \dots n_d}^*) = Cov(\mathcal{E}_{n_1 \dots n_d}^*) = \sigma^2 \Sigma^{-1/2} \Sigma \Sigma^{-1/2} = \sigma^2 \mathbf{I}.$$

This means that the components of $\mathbf{Y}_{n_1 \dots n_d}^*$ are uncorrelated so that under the normality condition they are independent and identically distributed with unit variance.

In the estimation of the parameter we consider only the transformed model. By the well-known Gauss-Marcov theorem (cf. Arnold [2]) the best unbiased estimator for the parameter β is given by the least squares estimator

$$\hat{\beta} = (\mathbf{X}_{n_1 \dots n_d}^{*\top} \mathbf{X}_{n_1 \dots n_d}^*)^{-1} \mathbf{X}_{n_1 \dots n_d}^{*\top} \mathbf{Y}_{n_1 \dots n_d}^* = (\mathbf{X}_{n_1 \dots n_d}^\top \Sigma^{-1} \mathbf{X}_{n_1 \dots n_d})^{-1} \mathbf{X}_{n_1 \dots n_d}^\top \Sigma^{-1} \mathbf{Y}_{n_1 \dots n_d}.$$

To test Hypothesis 2 we only need to consider the vector of the residuals of the transformed observations given by

$$\begin{aligned} \hat{\mathbf{R}}_{n_1 \dots n_d}^* &:= (r_{j_1 \dots j_d}^*)_{j_1=1, \dots, j_d=1}^{n_1, \dots, n_d} = [\mathbf{I} - \mathbf{X}_{n_1 \dots n_d}^* (\mathbf{X}_{n_1 \dots n_d}^{*\top} \mathbf{X}_{n_1 \dots n_d}^*)^{-1} \mathbf{X}_{n_1 \dots n_d}^{*\top}] \mathbf{Y}_{n_1 \dots n_d}^* \\ &= [\mathbf{I} - pr \mathbf{X}_{n_1 \dots n_d}^*] \mathcal{E}_{n_1 \dots n_d}^*, \end{aligned}$$

where $pr \mathbf{X}_{n_1 \dots n_d}^*$ is the orthogonal projector onto the column space of $\mathbf{X}_{n_1 \dots n_d}^*$. This array of residuals is next embedded to a stochastic process having sample paths in the space $\mathcal{C}(\mathcal{A})$ by means of the set-indexed partial sums operator $\mathbf{V}_{n_1 \dots n_d} : \mathcal{R}^{n_1 \times \dots \times n_d} \rightarrow \mathcal{C}(\mathcal{A})$ defined by

$$\mathbf{V}_{n_1 \dots n_d}(\hat{\mathbf{R}}_{n_1 \dots n_d}^*)(B) = \sum_{j_1=1}^{n_1} \dots \sum_{j_d=1}^{n_d} \sqrt{n_1 \dots n_d} \lambda^d(B \cap C_{j_1 \dots j_d}) r_{j_1 \dots j_d}^*,$$

where $\mathcal{A}$ is the family of compact subsets of $[\mathbf{a}, \mathbf{b}]$, and $C_{j_1 \dots j_d} := \prod_{k=1}^d [t_{n_k j_k - 1}, t_{n_k j_k}]$ with $\lambda^d(C_{j_1 \dots j_d}) = 1/(n_1 \dots n_d)$, for $1 \leq j_k \leq n_k$, $k = 1, \dots, d$. We refer interested reader to Pyke [13], Alexander and Pyke [1], and recently to Somayasa [17] for more detailed discussion regarding this process.

We define reasonable statistics for testing (2) which are Kolmogorov-Smirnov (KS) and Cramér-von Misses (CvM) functionals of the process. That are

$$\begin{aligned} KS_{n_1 \dots n_d; \mathcal{A}} &:= \sup_{B \in \mathcal{A}} \left| \frac{1}{\sigma} \mathbf{V}_{n_1 \dots n_d}(\hat{\mathbf{R}}_{n_1 \dots n_d}^*)(B) \right| \\ CvM_{n_1 \dots n_d; \mathcal{A}} &:= \frac{1}{n_1 \dots n_d} \sum_{B \in \mathcal{A}} \left(\frac{1}{\sigma} \mathbf{V}_{n_1 \dots n_d}(\hat{\mathbf{R}}_{n_1 \dots n_d}^*)(B) \right)^2. \end{aligned}$$

Asymptotic test using the KS test will reject H_0 at a level of significance α if and only if $KS_{n_1 \dots n_d; \mathcal{A}} \geq q_{1-\alpha}$, where $q_{1-\alpha}$ is the $(1 - \alpha)$ quantile of the limit distribution of $KS_{n_1 \dots n_d; \mathcal{A}}$. Similarly, using the CvM test, H_0 will be rejected at a level of significance α if and only if $CvM_{n_1 \dots n_d; \mathcal{A}} \geq c_{1-\alpha}$, where $c_{1-\alpha}$ is the $(1 - \alpha)$ quantile of the limit distribution of $CvM_{n_1 \dots n_d; \mathcal{A}}$.

2.2. Extension to multivariate case

In the multivariate case we observe a vector of spatial processes $\{\mathbf{Y}(\mathbf{t}) := (Y_i(\mathbf{t}))_{i=1}^p : \mathbf{t} \in [\mathbf{a}, \mathbf{b}]\}$ on the experimental design $\Xi_{n_1 \dots n_d}$ to test the hypothesis defined as follows

$$H_0 : \mathbf{g} = (g_i)_{i=1}^p \in \prod_{i=1}^p \mathbf{W}_i \text{ against } H_1 : \mathbf{g} = (g_i)_{i=1}^p \notin \prod_{i=1}^p \mathbf{W}_i, \quad (5)$$

where for every $\mathbf{t} \in [\mathbf{a}, \mathbf{b}]$, $\mathbf{Y}(\mathbf{t})$ allows a decomposition $\mathbf{Y}(\mathbf{t}) = \mathbf{g}(\mathbf{t}) + \zeta(\mathbf{t})$ for some vector of random errors $\zeta := (\zeta_1, \dots, \zeta_p)^\top$ which follows a p -variate normal distribution with mean vector $\mathbf{0}$ and the variance-covariance matrix defined below. In this case the existence of correlation among the components of $\mathbf{Y}$ and among the observations is inherent. That is for any $\mathbf{t}_\ell, \mathbf{t}_k \in [\mathbf{a}, \mathbf{b}]$,

$Cov(\mathbf{Y}(\mathbf{t}_\ell), \mathbf{Y}(\mathbf{t}_k)) = \mathbf{\Sigma}(\mathbf{t}_\ell - \mathbf{t}_k)$, where $\mathbf{\Sigma}(\mathbf{t}_\ell - \mathbf{t}_k)$ is a $p \times p$ -dimensional matrix of functions of the difference $(\mathbf{t}_\ell - \mathbf{t}_k)$ taking the form

$$\mathbf{\Sigma}(\mathbf{t}_\ell - \mathbf{t}_k) = \begin{pmatrix} \sigma_{11}(\mathbf{t}_\ell - \mathbf{t}_k) & \sigma_{12}(\mathbf{t}_\ell - \mathbf{t}_k) & \cdots & \sigma_{1p}(\mathbf{t}_\ell - \mathbf{t}_k) \\ \sigma_{21}(\mathbf{t}_\ell - \mathbf{t}_k) & \sigma_{22}(\mathbf{t}_\ell - \mathbf{t}_k) & \cdots & \sigma_{2p}(\mathbf{t}_\ell - \mathbf{t}_k) \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1}(\mathbf{t}_\ell - \mathbf{t}_k) & \sigma_{p2}(\mathbf{t}_\ell - \mathbf{t}_k) & \cdots & \sigma_{pp}(\mathbf{t}_\ell - \mathbf{t}_k) \end{pmatrix}.$$

To simplify the notation and also for computational reason, we assume that $\mathbf{W}_1 = \cdots = \mathbf{W}_p$. Hence we get a more simple model

$$\mathbf{Y}(\ell/n, k/n) = \mathbf{g}(\ell/n, k/n) + \zeta(\ell/n, k/n), \quad 1 \leq \ell, k \leq n.$$

For a fixed $n \geq 1$, let $\mathbf{Y}_{nnp} := (vec^\top(Y_1^\top(\Xi_n)), vec^\top(Y_2^\top(\Xi_n)), \dots, vec^\top(Y_p^\top(\Xi_n)))^\top$ be pn^2 -dimensional vector of random observations, where $Y_i(\Xi_n) := (Y_i(\ell/n, k/n))_{k=1, \ell=1}^{n, n}$, and "vec" is the well-known vec operator (cf. Harville [9]). Then the model under H_0 can be represented as

$$\mathbf{Y}_{nnp} = \mathbf{A}_{pn^2 \times p^2} \beta + \zeta_{nnp} \quad (6)$$

where $\beta := (\beta_1^\top, \dots, \beta_p^\top)^\top$ is the p^2 -dimensional vector of parameters, with $\beta_i := (\beta_{1i}, \dots, \beta_{pi})^\top$, $\mathbf{A}_{pn^2 \times p^2} := diag(\mathbf{X}_{nn}, \dots, \mathbf{X}_{nn})$ is the $pn^2 \times p^2$ -dimensional block matrix, and ζ_{nnp} is the associated pn^2 -dimensional normally distributed vector of random errors with $\mathbf{E}(\zeta_{nnp}) = \mathbf{0}$ and $Cov(\zeta_{nnp}) = \mathbf{\Sigma}_{nnp}$, where $\mathbf{\Sigma}_{nnp}$ is the $pn^2 \times pn^2$ -dimensional unknown matrix. For further simplification of the problem we assume the covariance matrix $\mathbf{\Sigma}_{nnp}$ has a structure $\mathbf{\Sigma}_{nnp} = \mathbf{\Sigma}_p \otimes \Psi_{nn}$, where $\otimes$ stands for the well-known Kronecker product defined e.g. in Harville [9] given by

$$\mathbf{\Sigma}_p \otimes \Psi_{nn} = \begin{pmatrix} \sigma_{11}\Psi_{nn} & \sigma_{12}\Psi_{nn} & \cdots & \sigma_{1p}\Psi_{nn} \\ \sigma_{21}\Psi_{nn} & \sigma_{22}\Psi_{nn} & \cdots & \sigma_{2p}\Psi_{nn} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{p1}\Psi_{nn} & \sigma_{p2}\Psi_{nn} & \cdots & \sigma_{pp}\Psi_{nn} \end{pmatrix},$$

σ_{ij} represents the covariance between Y_i and Y_j , and Ψ_{nn} is the $n^2 \times n^2$ -dimensional matrix of covariance functions taking the form

$$\Psi_{nn} = \begin{pmatrix} \psi(\mathbf{t}_{n11} - \mathbf{t}_{n11}) & \cdots & \psi(\mathbf{t}_{n11} - \mathbf{t}_{nn1}) & \cdots & \psi(\mathbf{t}_{n11} - \mathbf{t}_{nnn}) & \cdots & \psi(\mathbf{t}_{n11} - \mathbf{t}_{nnn}) \\ \vdots & & \vdots & & \vdots & & \vdots \\ \psi(\mathbf{t}_{n1k} - \mathbf{t}_{n11}) & \cdots & \psi(\mathbf{t}_{n1k} - \mathbf{t}_{nn1}) & \cdots & \psi(\mathbf{t}_{n1k} - \mathbf{t}_{nnk}) & \cdots & \psi(\mathbf{t}_{n1k} - \mathbf{t}_{nnn}) \\ \vdots & & \vdots & & \vdots & & \vdots \\ \psi(\mathbf{t}_{nnn} - \mathbf{t}_{n11}) & \cdots & \psi(\mathbf{t}_{nnn} - \mathbf{t}_{nn1}) & \cdots & \psi(\mathbf{t}_{nnn} - \mathbf{t}_{nnn}) & \cdots & \psi(\mathbf{t}_{nnn} - \mathbf{t}_{nnn}) \end{pmatrix},$$

where for $1 \leq \ell, k \leq n$, $\mathbf{t}_{n\ell k} := (\ell/n, k/n)$. Note that in the case of spatial isotropy one has

$$\Psi_{nn} = (\psi(\|\mathbf{t}_{n\ell k} - \mathbf{t}_{n\ell' k'}\|))_{\ell=1, k=1, \ell'=1, k'=1}^{n, n, n, n} \in \mathcal{R}^{n^2 \times n^2}.$$

As in the univariate case, under the assumption that $\mathbf{\Sigma}_p \otimes \Psi_{nn}$ is positive definite, upon pre-multiplying Model 6 by $(\mathbf{\Sigma}_p \otimes \Psi_{nn})^{-1/2}$, we get the transformed model as

$$(\mathbf{\Sigma}_p \otimes \Psi_{nn})^{-1/2} \mathbf{Y}_{nnp} = (\mathbf{\Sigma}_p \otimes \Psi_{nn})^{-1/2} \mathbf{A}_{pn^2 \times p^2} \beta + (\mathbf{\Sigma}_p \otimes \Psi_{nn})^{-1/2} \zeta_{nnp}. \quad (7)$$

By using such transformation, the new model $\mathbf{Y}_{nnp}^* = \mathbf{A}_{pn^2 \times p^2}^* \beta + \zeta_{nnp}^*$ consists of the vector of random errors ζ_{nnp}^* owning the variance-covariance matrix $\mathbf{I}_{pn^2 \times pn^2}$. Hence, by the Gauss-Markov theorem the least squares estimator of β , denoted by $\hat{\beta}$ will become the BLUE of β which is given by

$$\hat{\beta} = (\mathbf{A}_{pn^2 \times p^2}^\top (\boldsymbol{\Sigma}_p \otimes \boldsymbol{\Psi}_{nn})^{-1} \mathbf{A}_{pn^2 \times p^2})^{-1} \mathbf{A}_{pn^2 \times p^2}^\top (\boldsymbol{\Sigma}_p \otimes \boldsymbol{\Psi}_{nn})^{-1} \mathbf{Y}_{nnp}.$$

Similarly, we get the corresponding vector of the least squares residuals as

$$\hat{\mathbf{R}}_{nnp}^* := [\mathbf{I}_{pn^2 \times pn^2} - \mathbf{A}_{pn^2 \times p^2}^* (\mathbf{A}_{pn^2 \times p^2}^{*\top} \mathbf{A}_{pn^2 \times p^2}^*)^{-1} \mathbf{A}_{pn^2 \times p^2}^{*\top}] \zeta_{nnp}^*.$$

The transformation defined above also implies that the array $\{\zeta^*(\ell/n, k/n) = (\zeta_i^*(\ell/n, k/n))_{i=1}^p : 1 \leq \ell, k \leq n\}$ consists of independent and identically normally distributed random vectors with $\mathbf{E}(\zeta^*(\ell/n, k/n)) = \mathbf{0}$ and $\text{Cov}(\zeta^*(\ell/n, k/n)) = \mathbf{I}_{p \times p}$.

The p -dimensional set-indexed partial sums operator is denoted by $\mathbf{V}_{nnp}$ defined on the space $\prod_{i=1}^p \mathcal{R}^{n^2}$ taking values in the space $\prod_{i=1}^p \mathcal{C}(\mathcal{A})$ given by

$$\mathbf{V}_{nnp}(\hat{\mathbf{R}}_{nnp}^*)(B) = \sum_{k=1}^n \cdots \sum_{\ell=1}^n n \lambda^2(B \cap C_{\ell k}) \mathbf{r}_{\ell k}^*, \quad (8)$$

where $C_{\ell k} := ((\ell-1)/n, \ell/n] \times [(k-1)/n, k/n]$, and $\mathcal{A}$ is the family of compact subsets of $[0, 1] \times [0, 1]$. Interested reader is referred to Somayasa [18] for the definition of (8). Note that $\mathbf{r}_{\ell k}^*$ is the vector of the residual in the point $(\ell/n, k/n)$ associated to the transformed model. The Kolmogorov-Smirnov (KS) and Cramér-von Misses (CvM) functionals of the p -dimensional process are defined as follows

$$KS_{nnp;\mathcal{A}} := \sup_{B \in \mathcal{A}} \|\mathbf{V}_{nnp}(\hat{\mathbf{R}}_{nnp}^*)(B)\| \text{ and } CvM_{nnp;\mathcal{A}} := \frac{1}{n^2} \sum_{B \in \mathcal{A}} \|\mathbf{V}_{nnp}(\hat{\mathbf{R}}_{nnp}^*)(B)\|^2.$$

Similarly, asymptotic test using the KS test will reject H_0 at a level of significance α if and only if $KS_{nnp;\mathcal{A}} \geq t_{1-\alpha}$, where $t_{1-\alpha}$ is the $(1-\alpha)$ quantile of the limit distribution of $KS_{nnp;\mathcal{A}}$. Using the CvM test, H_0 will be rejected at a level of significance α if and only if $CvM_{nnp;\mathcal{A}} \geq w_{1-\alpha}$, where $w_{1-\alpha}$ is the $(1-\alpha)$ quantile of the limit distribution of $CvM_{nnp;\mathcal{A}}$.

3. Limiting behavior of the tests

In this section we investigate the limit process of the sequence of set-indexed least squares residual partial sums process defined in the preceding section. Although in one hand the distribution model of the sample is known to be normal, but on the other hand we insists on the derivation of the asymptotic behavior of the KS and CvM tests.

Firstly let us consider the univariate case represented by Model 4 and without loss of the generality we consider the unit rectangle $[0, 1] \times [0, 1]$ and Ξ_n as the experimental region and design, respectively. Let $\mathbf{W}_{nn}^* := [f_1^*(\Xi_{nn}), \dots, f_p^*(\Xi_{nn})]$, where $f_i^*(\Xi_{nn}) := \boldsymbol{\Sigma}^{-1/2} f_i(\Xi_{nn})$ is a vector in $\mathcal{R}^{n^2}$, $i = 1, \dots, p$ which are assumed to be orthonormal. Then we have

$$\hat{\mathbf{R}}_{nn}^* = \mathcal{E}_{nn}^* - pr \mathbf{W}_{nn}^* \mathcal{E}_{nn}^* = \mathcal{E}_{nn}^* - \sum_{i=1}^p \langle f_i^*(\Xi_{nn}), \mathcal{E}_{nn}^* \rangle_{\mathcal{R}^{n^2}} f_i^*(\Xi_{nn}).$$

Hence by the linearity of $\mathbf{V}_{nn}$ and by referring to Lemma 5.1 in Somayasa [17] we further get

$$\begin{aligned} \mathbf{V}_{nn}(\hat{\mathbf{R}}_{nn}^*)(B) &= \mathbf{V}_{nn}(\mathcal{E}_{nn}^*)(B) - \sum_{i=1}^p \langle \mathbf{V}_{nn}(f_i^*(\Xi_{nn})), \mathbf{V}_{nn}(\mathcal{E}_{nn}^*) \rangle_{\mathcal{H}_Z} \mathbf{V}_{nn}(f_i^*(\Xi_{nn}))(B) \\ &= \mathbf{V}_{nn}(\mathcal{E}_{nn}^*)(B) - (pr \mathbf{V}_{nn}(\mathbf{W}_{nn}^*) \mathbf{V}_{nn}(\mathcal{E}_{nn}^*))(B), \quad \forall B \in \mathcal{A}, \end{aligned}$$

where $\mathbf{V}_{nn}(\mathbf{W}_{nn}^*) := [\mathbf{V}_{nn}(f_1^*(\Xi_{nn})), \dots, \mathbf{V}_{nn}(f_p^*(\Xi_{nn}))] \subseteq \mathcal{H}_Z \subset \mathcal{C}(\mathcal{A})$. For $i = 1, \dots, p$, let $s_i^{(n)}$ be a step function on $[0, 1] \times [0, 1]$ associated to $f_i^*(\Xi_{nn})$, defined by

$$s_i^{(n)}(t, s) := \sum_{\ell=1}^n \sum_{k=1}^n f_i^*(\ell/n, k/n) \mathbf{1}_{C_{\ell k}}(t, s),$$

then we have $h_{s_i^{(n)}}(B) := \int_B s_i^{(n)}(t, s) \lambda^2(dt, ds) = \frac{1}{n} \mathbf{V}_{nn}(f_i^*(\Xi_{nn}))$. Hence the set $\{h_{s_1^{(n)}}, \dots, h_{s_p^{(n)}}\}$ also builds a basis for the space $\mathbf{V}_{nn}(\mathbf{W}_{nn}^*)$. Now we are ready to state our main result which can be proved in the similar way as in the result of Somayasa [17].

Theorem 3.1 Let $f_1^*, \dots, f_p^*$ be orthonormal as functions in $L_2(\lambda^2, [0, 1] \times [0, 1])$, where f_i^* is obtained in such a way that $\|s_i^{(n)} - f_i^*\|_\infty$ converges to zero. If $f_1, \dots, f_p$ are continuous and have bounded variation on $[0, 1] \times [0, 1]$ and Σ is positive definite, then

$$\frac{1}{\sigma} \mathbf{V}_{nn}(\widehat{\mathbf{R}}_{nn}^*)(\cdot) \Rightarrow Z_{\mathbf{f}^*}(\cdot) := Z(\cdot) - \sum_{i=1}^p \left(\int_{[0,1] \times [0,1]}^{(R)} f_i^*(t, s) dZ(t, s) \right) h_{f_i^*}(\cdot),$$

where $Z(\cdot)$ is the centered Gaussian process indexed by $\mathcal{A}$, see [1, 13] for the definition of $Z(\cdot)$.

The following result is an immediate consequence of the well-known continuous mapping theorem, see e.g. Theorem 5.1 in Billingsley [4].

Corollary 3.2 Under the condition of Theorem 3.1 it holds

$$KS_{nn;\mathcal{A}} \Rightarrow \sup_{B \in \mathcal{A}} |Z_{\mathbf{f}^*}(B)| \text{ and } CvM_{nn} \Rightarrow \int_{[0,1] \times [0,1]} Z_{\mathbf{f}^*}^2(B) dB.$$

Corollary 3.2 suggests us to approximate the quantiles of $KS_{nn;\mathcal{A}}$ and CvM_{nn} by those of the distributions of the statistics $\sup_{B \in \mathcal{A}} |Z_{\mathbf{f}^*}(B)|$ and $\int_{[0,1] \times [0,1]} Z_{\mathbf{f}^*}^2(B) dB$ which can be directly derived from that of the set-indexed Brownian sheet. However in this work the finite sample quantiles of the statistics $KS_{nn;\mathcal{A}}$ and CvM_{nn} will be approximated by simulation using computer.

Next we extend our consideration to the multivariate model. The p -dimensional set-indexed partial sums operator on $\prod_{i=1}^p \mathcal{R}^{n^2}$ is defined component wise as in Equation 8. Hence the limit processes can also directly derived by using the convergence component wise.

Theorem 3.3 Let $f_1^*, \dots, f_p^*$ be orthonormal as functions in $L_2(\lambda^2, [0, 1] \times [0, 1])$, where f_i^* is obtained in such a way that $\|s_i^{(n)} - f_i^*\|_\infty$ converges to zero. If $f_1, \dots, f_p$ are continuous and have bounded variation on $[0, 1] \times [0, 1]$ and $\Sigma_p \otimes \Psi_{nn}$ is positive definite, then

$$\mathbf{V}_{nnp}(\widehat{\mathbf{R}}_{nnp}^*)(\cdot) \Rightarrow \mathbf{Z}_{\mathbf{f}^*}^{(p)}(\cdot) := \mathbf{Z}^{(p)}(\cdot) - \left(\sum_{j=1}^p \left(\int_{[0,1] \times [0,1]}^{(R)} f_j^*(t, s) dZ^{(i)}(t, s) \right) h_{f_j^*}(\cdot) \right)_{i=1}^p,$$

where $\mathbf{Z}^{(p)}(\cdot) := (Z^{(1)}, \dots, Z^{(p)})^\top$ is the p -dimensional centered Gaussian process indexed by $\mathcal{A}$, defined e.g. in Somayasa and Wibawa [18]. We notice that the convergence $\mathbf{V}_{nnp}(\widehat{\mathbf{R}}_{nnp}^*)(\cdot) \Rightarrow \mathbf{Z}^{(p)}(\cdot)$ can be shown by extending the uniform central limit theorem studied in [1] and [13] to the vectorial analog.

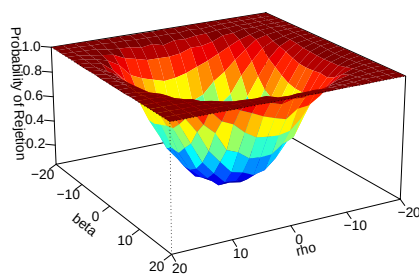
As a direct consequence of the continuous mapping theorem, we get the following corollary.

Corollary 3.4 Under the condition of Theorem 3.1 it holds

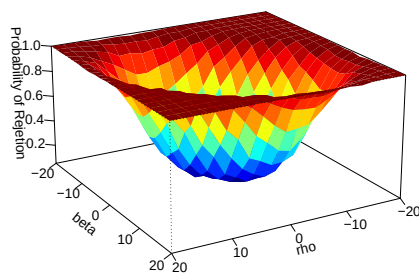
$$KS_{nnp;A} := \sup_{B \in \mathcal{A}} \|\mathbf{V}_{nnp}(\hat{\mathbf{R}}_{nnp}^*)(B)\| \Rightarrow \sup_{B \in \mathcal{A}} \|\mathbf{Z}_{\mathbf{f}^*}^{(p)}(B)\|$$

$$CvM_{nnp;A} := \frac{1}{n^2} \sum_{B \in \mathcal{A}} \|\mathbf{V}_{nnp}(\hat{\mathbf{R}}_{nnp}^*)(B)\|^2 \Rightarrow \int_{[0,1] \times [0,1]} \|\mathbf{Z}_{\mathbf{f}^*}(B)\|^2 dB.$$

Corollary 3.4 give us important way in approximating the sampling distributions of the multivariate KS and CvM statistics. In the practice the tests are realized by computing the quantiles of the limiting statistics presented in the right-hand side of Corollary 3.4. Unfortunately they are mathematically not tractable. By this reason in this work we develop simulation for approximating the quantiles of the statistics.



(a) Power function of KS test



(a) Power function of CvM test

Figure 1. The empirical power functions of the KS and CvM tests for univariate constant model.

4. Simulation study

In this section we study the finite sample properties of both tests by using Monte Carlo simulation using software package R. However by technical and computational reasons we only consider univariate and bivariate constant models, although theoretically we have more options for our choice such as any polynomial of finite degree.

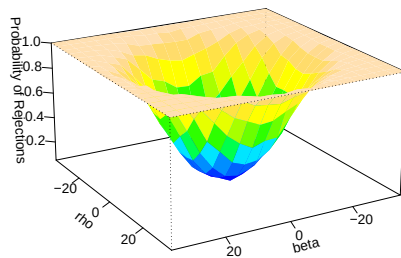
4.1. Simulation 1

We considering under H_0 a univariate constant model $Y(t, s) = 5 + \varepsilon(t, s)$, while under H_1 we assume a first order model $Y(t, s) = 5 + \rho t - \beta s + \varepsilon(t, s)$, for $(t, s) \in [0, 1] \times [0, 1]$, where $\varepsilon(t, s)$ is assumed to be normally distributed and isotropic having a covariance function

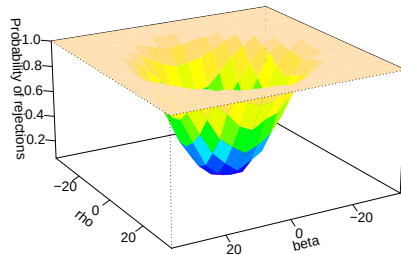
$$Cov(\varepsilon(t, s), \varepsilon(t', s')) = \sigma^2 \exp\{-\|(t, s) - (t', s')\|\}, \quad (t, s), (t', s') \in [0, 1] \times [0, 1].$$

Thereby σ^2 is an unknown parameter which is estimated by a consistent estimator defined in Arnold [2]. The samples are generated from a localized model $Y(\ell/n, k/n) = (5 + \rho\ell/n + \beta k/n)/n + \varepsilon(\ell/n, k/n)$, for $1 \leq \ell, k \leq n$, under $n \times n$ regular lattice Ξ_n . It is clear that the model is from H_0 if and only if ρ and β are simultaneously zero, otherwise the model is from H_1 . This means for a predetermined α , the power functions must attain α when both ρ and β are zero. Conversely, when either α or β are not zero, the power functions must take values larger than α .

The three dimensional empirical power functions of the KS and CvM tests are presented in Figure 1. The graphs are generated by drawing a surface through the points $(\rho, \beta, P(\rho, \beta))$ using three dimensional perspective plot of R, where for a fixed (ρ, β) , $P(\rho, \beta)$ is the relative number of the values of $KS_{nn;\mathcal{A}}$ (respectively $CvM_{nn;\mathcal{A}}$) exceeding the 0.95 quantile of $KS_{nn;\mathcal{A}}$ (respectively $CvM_{nn;\mathcal{A}}$). Figure 1 shows that both tests attain the level of significance $\alpha = 0.05$ as ρ and β are fixed to zero. When either ρ or β moves away from zero, the probability of rejection of H_0 gets larger. In general, for the constant model the KS test shows larger power than the CvM test.



(a). Power function of KS test



(b). Power function of CvM test

Figure 2. The empirical power functions of the KS and CvM tests for bivariate constant model.

4.2. Simulation 2

In the second case we simulate a multivariate constant model with two responses. The sample from localized bivariate model $\mathbf{Y}(\ell/n, k/n) = (5 + \rho k/n + \rho\ell/n, 3 + \beta k/n + \beta\ell/n)/n + \zeta(\ell/n, k/n)$. To simplify the computation, we assume that $\Psi_{nn} = \mathbf{I}_{n^2 \times n^2}$ and the random vector $\zeta = (\zeta_1, \zeta_2)^\top$ is generated from a bivariate normal distribution with mean zero and variance-covariance matrix

$$\Sigma := \begin{pmatrix} +6.26 & -0.50 \\ -0.50 & +6.26 \end{pmatrix}.$$

However Σ is assumed to be unknown and for that it is estimated by a consistent estimator proposed in Arnold [3].

Figure 2 exhibits the three-dimensional graphics of the empirical power function of the tests, where the first graph (a) represents the power function of the KS test and the second (b) represents that of the CvM test. Both power functions seem to achieve the pre signed level of significance $\alpha = 0.05$. The power of the tests get larger as the the model moves away from H_0 . Among the two tests the CvM test tends to have larger power.

5. Conclusion

We established asymptotic test procedures based on the Kolmogorov-Smirnov and Cramér-von Mises functionals of the partial sums processes of the least squares residuals. We obtain the limit processes as function of the set-indexed Brownian sheet for univariate case, whereas for multivariate case the limit process is a functional of the p -dimensional Brownian sheet indexed by sets. Observing constant model the finite sample size behavior of the tests are that in the univariate case the KS test tend to have a larger power than that of CvM test. Conversely in multivariate case the CvM test is slightly more powerful then the KS test. By this reason in the application we suggest the user to use the KS test rather than the CvM test for univariate constant model. However for bivariate constant model we propose to adopt the CvM test.

Acknowledgement

This work was supported in parts by SAME program 2015. The authors wish to thank the Institute for Stochastic Karlsruhe Institute of Technology (KIT) for hospitality.

References

- [1] Alexander K S and Pyke R 1986 *The Annals of Probability* **14** 582–97.
- [2] Arnold S F 1981 *The Theory of Linear Models and Multivariate Analysis* (New York: John Wiley & Sons, Inc.)
- [3] Arnold S F 1984 *Journal of the American Statistical Association* **75** (372) 890–94.
- [4] Billingsley P 1968 *Convergence of Probability Measures* (New York: John Wiley & Sons, Inc.)
- [5] Bischoff W 2002 *Theory of Stochastic Processes* **2** 23–8
- [6] Bischoff W and Somayasa W 2009 *J. Multivariate Analysis* **100** 2167–77
- [7] Christensen R 1991 *Linear Models for Multivariate, Time Series, and Spatial Data* (New York:Springer)
- [8] Cressie Noel A C 1991, *Statistics for Spatial Data* (New York:John Wiley & Sons)
- [9] Harville D A 1997 *Matrix Algebra from a Statistician's Perspective* (New York: Springer-Verlag)
- [10] Graybill F A 2000 *Theory and Application of Linear Model* (London:Duxbury press)
- [11] Johnson R A and Wichern D W 2007 *Applied Multivariate Statistical Analysis* (New York: Prentice-Hall)
- [12] MacNeill I B and Jandhyalla V K 1993 *Multivariate Environmental Statistics* ed G P Patil and C R Rao (Amsterdam: Elsevier Science Publishers B.V.) 298–306.
- [13] Pyke R 1983 *Probability, Statistics and Analysis* ed J F C Kingman and G E H Reuter (London: Math. Soc. Lect. Notes Series) **79** 219–40.
- [14] Schabenberger S and Gotway C A 2005 *Statistical Methods for Spatial Data Analysis* (London:Chapman)
- [15] Seber G A F and Lee A J 2003 *Linear Regression Analysis* (2nd ed.) (New York: John Wiley & Sons)
- [16] Somayasa W 2014 *J. Indones. Math.Soc.* **20**(2) 77–94.
- [17] Somayasa W, Ruslan, Cahyono E and Engkoimani L O 2015 *Fareast J. Math. Sciences* **96**(8) 933–66
- [18] Somayasa W, and Wibawa G N A 2015 *Fareast J. Math. Sciences* **98**(5) 613–39
- [19] Stute W 1997 *The Annals of Statistics* **25**(2) 613–41
- [20] Stute W, Mantaiga W G and Quindimil M P 1997 *J. Amer. Statistical Association* **93**(441) 141–9
- [21] Xie L and MacNeill I B 2006 *South African Statist. J.* **4** 33–3
- [22] Zellner A 1962 *J. Amer. Statistical Association* **57** (298) 348–68