S©MR ISSN 1813-3304

СИБИРСКИЕ ЭЛЕКТРОННЫЕ
МАТЕМАТИЧЕСКИЕ ИЗВЕСТИЯ

Siberian Electronic Mathematical Reports
http://semr.math.nsc.ru

Том 18, №2, стр. 1482–1492 (2021) УДК 519.233
DOI 10.33048/semi.2021.18.111 MSC 62F03

ASYMPTOTICS OF SUMS OF REGRESSION RESIDUALS UNDER MULTIPLE ORDERING OF REGRESSORS

M.G. CHEBUNIN, A.P. KOVALEVSKII

Abstract. We prove theorems about the Gaussian asymptotics of an empirical bridge built from residuals of a linear model under multiple regressor orderings. We study the testing of the hypothesis of a linear model for the components of a random vector: one of the components is a linear combination of the others up to an error that does not depend on the other components of the random vector. The independent copies of the random vector are sequentially ordered in ascending order of several of its components. The result is a sequence of vectors of higher dimension, consisting of induced order statistics (concomitants) corresponding to different orderings. For this sequence of vectors, without the assumption of a linear model for the components, we prove a lemma of weak convergence of the distributions of an appropriately centered and normalized process to a centered Gaussian process with almost surely continuous trajectories. Assuming a linear relationship of the components, standard least squares estimates are used to compute regression residuals, that is, the differences between response values and the predicted ones by the linear model. We prove a theorem of weak convergence of the process of sums of of regression residuals under the necessary normalization to a centered Gaussian process.

Keywords: Concomitants, copula, weak convergence, regression residuals.

Chebunin, M.G., Kovalevskii, A.P., Asymptotics of sums of regression residuals under multiple ordering of regressors.

The work is supported by Mathematical Center in Akademgorodok under agreement No. 075-15-2019-1675 with the Ministry of Science and Higher Education of the Russian Federation.

Received June, 14, 2021, published December, 2, 2021.

1. Introduction

An extremely useful method for analyzing multivariate statistics is the study of linear relationships between components. This analysis allows you to build a linear prediction of one variable based on the others. The very existence of dependencies is verified by calculating sample correlations and developing tests based on them. This class of tests is the subject of correlation analysis.

The construction of models for the linear dependence of one variable (response) on other variables (regressors), the estimation of the parameters of the linear dependence, and testing of their significance are the subject of regression analysis. However, standard regression analysis methods do not include methods of detecting that the proposed linear model is incorrect entirely. If the model is incorrect, then it must either be completely discarded or substantially modified.

Methods for testing linear models, as a rule, use functionals as statistics from random processes built according to the sequence of observations. If the observations are ordered by one of the regressors, then such statistical tests are often called tests discord detection. In the papers of Rao (1950), Page (1954), observations are ordered by time, and the alternative hypothesis is that the distribution changes at some time (the change point). In this case, the distribution before the change is assumed to be known, and the tests are focused on the fastest detection of the change. Moustakides (1986) proved that the CUSUM procedure proposed by Page (1954) is optimal in terms of Lorden (1971). Shiryaev (1996) generalized this result to a continuous-time analogue of CUSUM. Brodsky and Darkhovsky (2005) proved the asymptotic optimality of the adaptive CUSUM test for compact sets of unknown distribution parameters.

In situations where the distribution before the disorder is not described by a set of parameters from a compact set, and also for more complex linear models, the process of sums of regression residuals is used, see Shorack and Wellner (1986). MacNeill (1978) proposed such a test for time series, and Bishoff (1998) significantly relaxed the assumptions of MacNeill. An analysis of results in this direction can be found in Csorgo and Horváth (1997, Chapters 2 and 3) and MacNeill et al. (2020). Kovalevskii and Shatalin (2015, 2016), Kovalevskii (2020) proposed tests for matching of regression models using data ordering by one of the regressors. We offer a statistical test that uses multiple ordering of data across multiple regressors.

The rest of the work is organized as follows. We prove a lemma extending the functional central limit theorem by Davydov and Egorov (2000) for the multidimensional case (convergence to a Gaussian field) in Section 2. This lemma is based on the general result of Ossiander (1987) and allows one to obtain a limit theorem for multiple ordering. Section 3 contains this limit theorem for a general linear regression model. Proofs are in Section 4, a discussion is in Section 5.

2. Induced order statistics

Let $(\mathbf{X}_i, \mathbf{Y}_i)$, $i = 1, 2 \ldots$, be the independent copies of a random vector $(\mathbf{X}, \mathbf{Y})$ such that $\mathbf{X} = (X^{(1)}, \ldots, X^{(d_1)})$ takes values in $[0,1]^{d_1}$, $\mathbf{Y}$ takes values in $\mathbb{R}^{d_2}$. The distribution function (copula) of $\mathbf{X}$ is $C(\mathbf{u}) = \mathbf{P}(\mathbf{X} \leq \mathbf{u}) = \mathbf{P}\left(X^{(1)} \leq u^{(1)}, \ldots, X^{(d_1)} \leq u^{(d_1)}\right)$, $\mathbf{u} = (u^{(1)}, \ldots, u^{(d_1)}) \in [0,1]^{d_1}$.

We assume that there is copula density $c(\mathbf{u})$, that is,

$$C(\mathbf{u}) = \int_{\mathbf{v}\leq\mathbf{u}} c(\mathbf{v})d\mathbf{v}, \quad \mathbf{v} \in [0,1]^{d_1}. \tag{1}$$

Denote $X_{n,1}^{(k)} \leq X_{n,2}^{(k)} \leq \cdots \leq X_{n,n}^{(k)}$, $1 \leq k \leq d_1$, the order statistics of the k-th column of matrix X, and $\mathbf{Y}_{n,1}^{(k)}, \mathbf{Y}_{n,2}^{(k)}, \ldots, \mathbf{Y}_{n,n}^{(k)}$ the corresponding values of the vectors $\mathbf{Y}_i$. The random vectors $\left(\mathbf{Y}_{n,i}^{(k)}, i \leq n\right)$ are called induced order statistics (concomitants).

We study the asymptotic behavior of the random field

$$\mathbf{Q}_n(\mathbf{u}) = \sum_{j=1}^{n} \mathbf{Y}_j \mathbf{1}\left(\mathbf{X}_j \leq \mathbf{u}\right) = \sum_{j=1}^{n} \mathbf{Y}_j \mathbf{1}\left(X_j^{(1)} \leq u^{(1)}, \ldots, X_j^{(d_1)} \leq u^{(d_1)}\right), \ \mathbf{u} \in [0,1]^{d_1}.$$

Using the asymptotics of $\mathbf{Q}_n(\mathbf{u})$, we study the asymptotics of $d_1 \times d_2$-dimensional process of sums of induced order statistics under different orderings

$$\mathbf{Z}_n(t) = \left(\sum_{j=1}^{[nt]} \mathbf{Y}_{n,j}^{(1)}, \sum_{j=1}^{[nt]} \mathbf{Y}_{n,j}^{(2)}, \ldots, \sum_{j=1}^{[nt]} \mathbf{Y}_{n,j}^{(d_1)}\right), \ t \in [0,1].$$

Let $\mathbf{m}(\mathbf{u}) = \mathbf{E}(\mathbf{Y} \mid \mathbf{X} = \mathbf{u})$, $\mathbf{u} \in [0,1]^{d_1}$, and $\mathbf{f}(\mathbf{u}) = \int_{\mathbf{0}}^{\mathbf{u}} \mathbf{m}(\mathbf{v})c(\mathbf{v})d\mathbf{v}$.

Let

$$\sigma^2(\mathbf{u}) = \mathbf{E}\left\{(\mathbf{Y} - \mathbf{m}(\mathbf{X}))^T(\mathbf{Y} - \mathbf{m}(\mathbf{X})) \mid \mathbf{X} = \mathbf{u}\right\}$$

be the conditional covariance matrix of $\mathbf{Y}$ and $\sigma(\mathbf{u})$ be the positive definite matrix such that $\sigma(\mathbf{u})^T\sigma(\mathbf{u}) = \sigma^2(\mathbf{u})$.

Let $\mathbf{e}_{k,t} = (1, \ldots, 1, t, 1, \ldots, 1)$ the vector in $[0,1]^{d_1}$ with k-th coordinate being t and other coordinates being 1.

All our limit fields and processes are continuous, so we use the uniform metric. Let $\|\cdot\|$ denote the Euclidean norm in the corresponding space. Our random field $\mathbf{Q}_n$ takes values in the space $B([0,\,1]^{d_1};\,\mathbf{R}^{d_2})$ of bounded measurable functions with the Borel σ-algebra $\mathcal{B}$. This space is not separable, so we note that the random field $\mathbf{Q}_n$ takes values in its subset D with the smaller σ-algebra $\mathcal{D}$. This σ-algebra is generated by the d_1-dimensional analog of Skorohod metrics, see Straf (1972). So, let D be the uniform closure, in the space $B([0,\,1]^{d_1};\,\mathbf{R}^{d_2})$, of the vector subspace of simple functions (that is, linear combinations of step functions).

For $\mathbf{x}, \mathbf{y} \in D$, let the "Skorohod" distance be

$$d(\mathbf{x}, \mathbf{y}) = \inf\{\min(||\mathbf{x} - \mathbf{y}\lambda||_S, ||\lambda||_S) : \ \lambda \in \Lambda\},$$

where

$$||\mathbf{x} - \mathbf{y}\lambda||_S = \sup\{||\mathbf{x}(\mathbf{t}) - \mathbf{y}(\lambda(\mathbf{t}))||, \ \mathbf{t} \in [0,\,1]^{d_1}\},$$

$||\lambda||_S = \sup\{||\lambda(\mathbf{t}) - \mathbf{t}||, \ \mathbf{t} \in [0,\,1]^{d_1}\}$, Λ is the group of all transformations λ of the form $\lambda(t_1, \ldots, t_{d_1}) = (\lambda_1(t_1), \ldots, \lambda_{d_1}(t_{d_1}))$, where each $\lambda_i : [0,1] \to [0,1]$ is continuous, strictly increasing, and fixes 0 and 1.

With respect to the corresponding metric topology, D is complete and separable, and it's Borel σ-algebra $\mathcal{D}$ coincides with the σ-algebra generated by coordinate mappings, see Section 3 of Bickel & Wichura (1971) for details.

The process $\mathbf{Q}_n$ takes values in D. So, it is $\mathcal{D}$-measurable.

We define the weak convergence as follows (cf. Dudley, 1967): $\lim_{n\to\infty} \mathbf{E}f(\mathbf{Q}_n) = \mathbf{E}f(\mathbf{Q})$ for every bounded, continuous, and $\mathcal{D}$-measurable function $f : D \to \mathbf{R}^{d_2}$.

Davydov & Zitikis (2008) in their Proposition 1 have proved that the limitation of the class of functions f to the uniform continuity (instead of the continuity) gives the same definition of the weak convergence (that is, these definitions are equivalent).

We use the symbol $\Rightarrow$ to denote the weak convergence of random fields in the sense that has been mentioned above. We use the same symbol for the weak convergence of random variables and the weak convergence of stochastic processes in the uniform topology.

The following Lemma 1 generalizes the result of the first part of Theorem 2.1(1) by Davydov and Egorov (2000) to random fields.

Lemma 1. *If* $\mathbf{E}\|\mathbf{Y}\|^2 < \infty$ *then* $\widetilde{\mathbf{Q}}_n = \frac{\mathbf{Q}_n - \mathbf{f}}{\sqrt{n}} \Rightarrow \mathbf{Q}$, *a centered Gaussian field with covariance*

$$K(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{E}\mathbf{Q}^T(\mathbf{u}_1)\mathbf{Q}(\mathbf{u}_2) = \int_{\mathbf{0}}^{\min(\mathbf{u}_1,\mathbf{u}_2)} \sigma^2(\mathbf{v})c(\mathbf{v})d\mathbf{v}$$

$$+ \int_{\mathbf{0}}^{\min(\mathbf{u}_1,\mathbf{u}_2)} \mathbf{m}^T(\mathbf{v})\mathbf{m}(\mathbf{v})c(\mathbf{v})d\mathbf{v} - \int_{\mathbf{0}}^{\mathbf{u}_1} \mathbf{m}^T(\mathbf{v})c(\mathbf{v})d\mathbf{v} \int_{\mathbf{0}}^{\mathbf{u}_2} \mathbf{m}(\mathbf{v})c(\mathbf{v})d\mathbf{v},$$

$\mathbf{u}_1,\ \mathbf{u}_2 \in [0,1]^{d_1}$.

Lemma 2 generalizes the result of Theorem 2.1(2) by Davydov and Egorov (2000) to multiple ordering but under additional assumption $\mathbf{m} \equiv \mathbf{0}$.

Lemma 2. *If* $\mathbf{E}\|\mathbf{Y}\|^2 < \infty$, $\mathbf{m} \equiv \mathbf{0}$ *then* $\widetilde{\mathbf{Z}}_n = \frac{\mathbf{Z}_n}{\sqrt{n}} \Rightarrow \mathbf{Z}$, *a centered Gaussian* $(d_1 \times d_2)$*-dimensional process with covariance matrix function* $\mathbf{E}\mathbf{Z}^T(t_1)\mathbf{Z}(t_2) = (K(\mathbf{e}_{k_1,t_1}, \mathbf{e}_{k_2,t_2}))_{k_1,k_2=1}^{d_1}$,

$$K(\mathbf{e}_{k_1,t_1}, \mathbf{e}_{k_2,t_2}) = \mathbf{E}\mathbf{Q}^T(\mathbf{e}_{k_1,t_1})\mathbf{Q}(\mathbf{e}_{k_2,t_2}) = \int_{\mathbf{0}}^{\min(\mathbf{e}_{k_1,t_1},\mathbf{e}_{k_2,t_2})} \sigma^2(\mathbf{v})c(\mathbf{v})d\mathbf{v}.$$

3. Main result

Let $(\mathbf{X}_i, \xi_i, \eta_i) = (X_{i1}, \ldots, X_{id_1}, \xi_{i1}, \ldots, \xi_{i,d_2-1}, \eta_i)$ be independent and identically distributed random vector rows, $i = 1, \ldots, n$. All components of a raw can be dependent and $X_{i1}, \ldots, X_{id_1}$ have copula (so their marginal distributions are uniform on [0, 1]) and (1) is true.

Rows $(\mathbf{X}_i, \xi_i, \eta_i)$ form matrix (X, ξ, η).

We assume a linear regression hypothesis H_0:

$$\eta_i = \xi_i\theta + \varepsilon_i = \sum_{j=1}^{d_2-1} \xi_{ij}\theta_j + \varepsilon_i, \tag{2}$$

$\{\varepsilon_i\}_{i=1}^n$ and $\{(\mathbf{X}_i, \xi_i)\}_{i=1}^n$ are independent, $\mathbf{E}\,\varepsilon_1 = 0$, $\mathbf{Var}\,\varepsilon_1 > 0$.

Vector $\theta = (\theta_1, \ldots, \theta_{d_2-1})^T$ and constant $\mathbf{Var}\,\varepsilon_1$ are unknown. We consider d_1 orderings of rows of the matrix (X, ξ, η) in acsending order of columns of X.

The result of d_1 orderings is a sequence of d_1 matrices $(X^{(j)}, \xi^{(j)}, \eta^{(j)})$ with rows $(\mathbf{X}_i^{(j)}, \xi_i^{(j)}, \eta_i^{(j)}) = (X_{i1}^{(j)}, \ldots, X_{id_1}^{(j)}, \xi_{i1}^{(j)}, \ldots, \xi_{i,d_2-1}^{(j)}, \eta_i^{(j)})$, $j = 1, \ldots, d_1$.

Let $\widehat{\theta}$ be LSE:

$$\widehat{\theta} = (\xi^T\xi)^{-1}\xi^T\eta.$$

It does not depend on the order of rows.

Let $h^{(j)}(x) = \mathbf{E}\{\xi_1 | X_{1j} = x\}$ be conditional expectations, $L^{(j)}(x) = \int\limits_0^x h^{(j)}(s)\, ds$ be *induced theoretical generalised Lorentz curves* (see Davydov and Egorov (2000)),

$$b_j^2(x) = \mathbf{E}\left((\xi_1 - h^{(j)}(x))^T(\xi_1 - h^{(j)}(x)) \mid X_{1j} = x\right)$$

be matrices of conditional covariances.

Let $G = \mathbf{E}\xi_1^T\xi_1$. Then

$$\int_0^1 \left(b_j^2(x) + (h^{(j)}(x))^T h^{(j)}(x)\right) dx = G$$

for any $j = 1, \ldots, d_2 - 1$.

Let $\widehat{\varepsilon}_i^{(j)} = \eta_i^{(j)} - \xi_i^{(j)}\widehat{\theta}$ be regression residuals, $\widehat{\Delta}_k^{(j)} = \sum\limits_{i=1}^k \widehat{\varepsilon}_i^{(j)}$ be its partial sums, $\widehat{\Delta}_0^{(j)} = 0$.

Let $\widehat{Z}_n^{(j)} = \{\widehat{Z}_n^{(j)}(t),\, 0 \le t \le 1\}$ be a piecewise linear random function with nodes

$$\left(\frac{k}{n},\ \frac{\widehat{\Delta}_k^{(j)}}{\sqrt{n\mathbf{Var}\varepsilon_1}}\right), \quad k = 0,\ 1, \ldots,\ n.$$

From Theorem 1 (Kovalevskii, 2020) we have

Theorem 1. *If matrix G exists and is non-degenerate and H_0 is true then $\widehat{Z}_n^{(j)} \implies \widehat{Z}^{(j)}$ for any $j = 1, \ldots, d_1$. Here $\widehat{Z}^{(j)}$ is a centered Gaussian process with continuous a.s. sample paths and covariance function*

$$\widehat{K}_{jj}(s,t) = \min(s,t) - L^{(j)}(s)G^{-1}(L^{(j)}(t))^T, \quad s,t \in [0,1].$$

We prove that the d_1-dimensional process $\widehat{Z}_n = (\widehat{Z}_n^{(j)}, j = 1, \ldots, d_1)$ has a Gaussian limit.

Theorem 2. *If matrix G exists and is non-degenerate and H_0 is true then $\widehat{Z}_n \implies \widehat{Z}$. Here $\widehat{Z}$ is a centered d_1-dimensional Gaussian process with continuous a.s. sample paths and covariance matrix function* $\widehat{K}(s,t) = \left(\widehat{K}_{ij}(s,t)\right)_{i,j=1}^{d_1}$,

$$\widehat{K}_{ij}(s,t) = \mathbf{P}(X_{1i} \le s, X_{1j} \le t) - L^{(i)}(s)G^{-1}(L^{(j)}(t))^T, \quad s,t \in [0,1].$$

4. Proofs

Proof of Lemma 1.

For simplicity, we consider the case $d_1 = 2$ since the construction of the proof given below can be easily extended to the case $d_1 > 2$. Now let $d_2 = 1$, we will generalize it to $d_2 \ge 1$ using the Cramer-Wold theorem.

Thus, we consider a random field

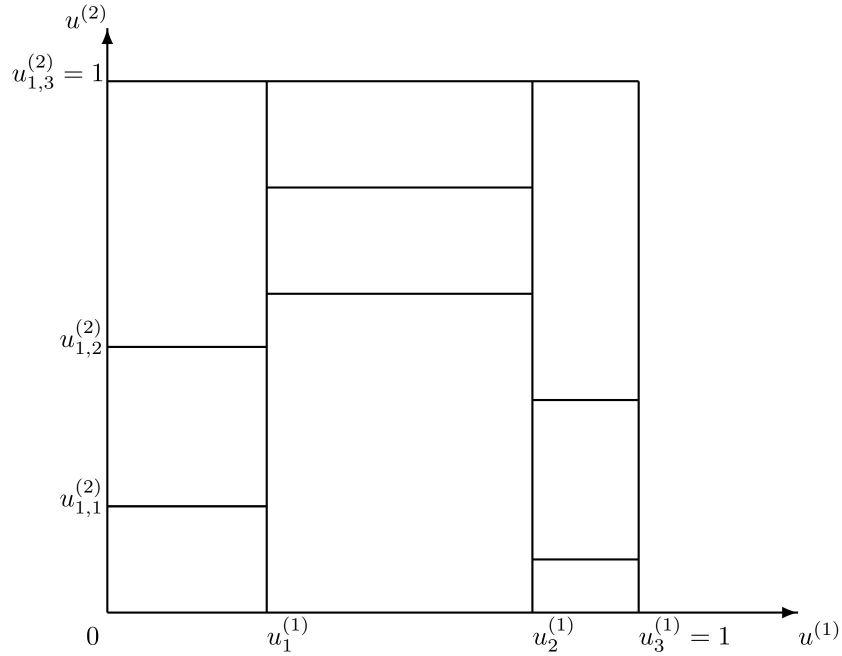
$$Q_n(\mathbf{u}) = \sum_{j=1}^n Y_j \mathbf{1}\left(X_j^{(1)} \le u^{(1)}, X_j^{(2)} \le u^{(2)}\right),\ \mathbf{u} \in [0,1]^2.$$

Let us define the partition of the unit square $[0,1]^2$ into N^2 parts as follows. Let $u_0^{(1)} = 0 < u_1^{(1)} < u_2^{(1)} < \ldots < u_N^{(1)} = 1$ be a partition of the interval [0,1], such that

$$\int_{(u_{i-1}^{(1)},0)}^{(u_i^{(1)},1)} \left(\mathbf{E}\left(Y^2 \mid \mathbf{X} = \mathbf{v}\right) + \mathbf{E}Y^2\right) c(\mathbf{v}) d\mathbf{v} = 2\mathbf{E}Y^2/N,\ i = 1, 2, \ldots, N,$$

and for any fixed $i = 1, 2, \ldots, N$ let $u^{(2)}_{i,0} = 0 < u^{(2)}_{i,1} < u^{(2)}_{i,2} < \ldots < u^{(2)}_{i,N} = 1$ be another partition of the interval [0,1] (see Pic.1) such that

$$\int_{(u^{(1)}_{i-1}, u^{(2)}_{i,j-1})}^{(u^{(1)}_{i}, u^{(2)}_{i,j})} \left(\mathbf{E}\left(Y^2 \mid \mathbf{X} = \mathbf{v}\right) + \mathbf{E}Y^2\right) c(\mathbf{v}) d\mathbf{v} = 2\mathbf{E}Y^2/N^2, \ j = 1, 2, \ldots, N.$$



Pic. 1. An example of the partition of $[0,1]^2$ for $N = 3$.

So we have $N+1$ points $u^{(1)}_i$ in the first coordinate and not greater then $(N-1)^2 + 1$ different points $u^{(2)}_{i,j}$ in the second coordinate. For any $\mathbf{u} = (u^{(1)}, u^{(2)}) \in [0,1]^2$, there are indexes $i^{(1)}, i^{(2)}_1, i^{(2)}_2, j^{(2)}_1, j^{(2)}_2 \in \{0, 1, \ldots, N\}$ such that $u^{(1)}_{i^{(1)}-1} \leq u^{(1)} \leq u^{(1)}_{i^{(1)}}$ and

$$u^{(2)}_{i^{(2)}_1, j^{(2)}_1} = \max_{i,j} \{u^{(2)}_{i,j} \leq u^{(2)}\},$$

$$u^{(2)}_{i^{(2)}_2, j^{(2)}_2} = \min_{i,j} \{u^{(2)}_{i,j} \geq u^{(2)}\},$$

so $u^{(2)}_{i^{(2)}_1, j^{(2)}_1} \leq u^{(2)} \leq u^{(2)}_{i^{(2)}_2, j^{(2)}_2}$.

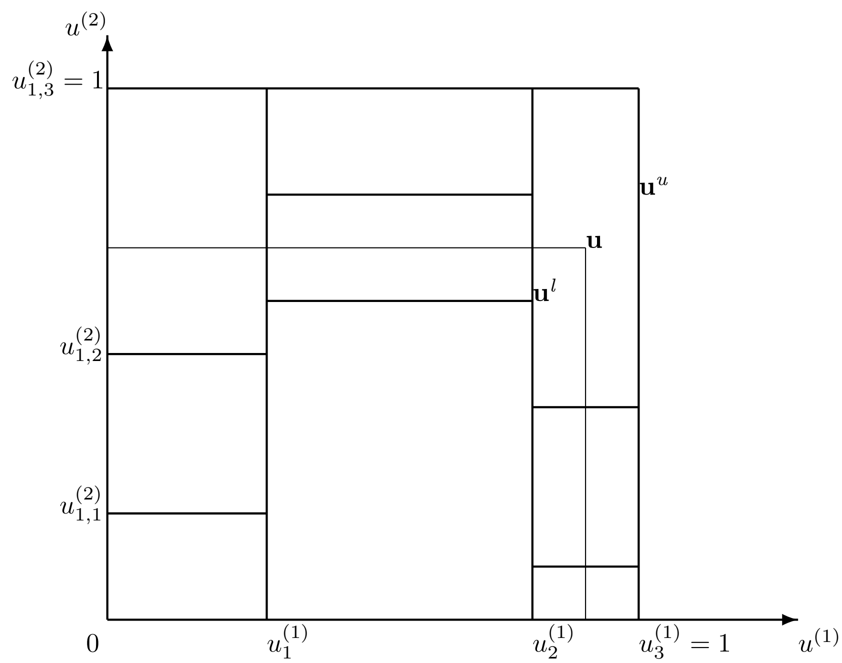
Denote $\mathbf{u}^l = (u^{(1)}_{i^{(1)}-1}, u^{(2)}_{i^{(2)}_1, j^{(2)}_1})$ and $\mathbf{u}^u = (u^{(1)}_{i^{(1)}}, u^{(2)}_{i^{(2)}_2, j^{(2)}_2})$ (Pic. 2).

Define the metric entropy with bracketing for the special separable pseudometric space (S, ρ), where

$$S = \left\{h_{\mathbf{u}}(\mathbf{x}, y) = y\mathbf{1}(\mathbf{x} \leq \mathbf{u}), \quad \mathbf{u} \in [0,1]^2\right\}$$
$$\rho^2\left(h_{\mathbf{u_1}}, h_{\mathbf{u_2}}\right) = \mathbf{E}\left(h_{\mathbf{u_1}}(\mathbf{X}, Y) - h_{\mathbf{u_2}}(\mathbf{X}, Y)\right)^2$$

Let $S(\delta) = \{s_1, s_2, \ldots, s_M\} \subseteq S$ be such that for some random variables $f^l(s_i)$ and $f^u(s_i), i \le M$, the following conditions are valid. For any $s \in S$ there exists $s_i \in S(\delta)$ such that

$$\rho(s, s_i) \le \delta,$$
$$f^l(s_i) \le f(s, \mathbf{X}, Y) \le f^u(s_i) \quad a.s.,$$
$$\rho\left(f^u(s_i), f^l(s_i)\right) \le \delta, \quad i \le n.$$



Pic. 2. An example of upper and lower points for $N = 3$.

Then $H^B(\delta, S, \rho) = \min\{M : S(\delta) \subseteq S\}$ is called the metric entropy with bracketing.

According to Ossiander's (1987) theorem, to prove Lemma for $\widetilde{Q}_n$ we must show that

$$\int_0^1 \left(H^B(t, S, \rho)\right)^{1/2} dt < \infty.$$

To prove it we show that

$$H^B(\delta, S, \rho) \le C \log\left(\frac{1}{\delta}\right).$$

First notice that $f(h_{\mathbf{u}}, \mathbf{X}, Y) = h_{\mathbf{u}}(\mathbf{X}, Y) - Eh_{\mathbf{u}}(\mathbf{X}, Y), \mathbf{u} \in [0,1]^2$ is a separable random process satisfying the conditions $(2.1)-(2.3)$ of Ossiander's work, moreover $\widetilde{Q}_n = n^{-1/2}\sum_{i=1}^n f(h_{\mathbf{u}}, \mathbf{X}_i, Y_i)$. Then write $f(h_{\mathbf{u}}, \mathbf{x}, y)$ as follows

$$f(h_{\mathbf{u}}, \mathbf{x}, y) = y_+\mathbf{1}(\mathbf{x} \le \mathbf{u}) - y_-\mathbf{1}(\mathbf{x} \le \mathbf{u}) - \int_{\mathbf{0}}^{\mathbf{u}} m_+(\mathbf{v})c(\mathbf{v})d\mathbf{v} + \int_{\mathbf{0}}^{\mathbf{u}} m_-(\mathbf{v})c(\mathbf{v})d\mathbf{v}$$

Let

$$f^u = Y_+\mathbf{1}(\mathbf{X} \le \mathbf{u}^u) - Y_-\mathbf{1}(\mathbf{X} \le \mathbf{u}^l) - \int_{\mathbf{0}}^{\mathbf{u}^l} m_+(\mathbf{v})c(\mathbf{v})d\mathbf{v} + \int_{\mathbf{0}}^{\mathbf{u}^u} m_-(\mathbf{v})c(\mathbf{v})d\mathbf{v},$$

$$f^l = Y_+\mathbf{1}(\mathbf{X} \le \mathbf{u}^l) - Y_-\mathbf{1}(\mathbf{X} \le \mathbf{u}^u) - \int_{\mathbf{0}}^{\mathbf{u}^u} m_+(\mathbf{v})c(\mathbf{v})d\mathbf{v} + \int_{\mathbf{0}}^{\mathbf{u}^l} m_-(\mathbf{v})c(\mathbf{v})d\mathbf{v}.$$

Then

$$f^u - f^l = |Y|\mathbf{1}(\mathbf{u}^l \le \mathbf{X} \le \mathbf{u}^u) + \int_{\mathbf{u}^l}^{\mathbf{u}^u} |m(\mathbf{v})|c(\mathbf{v})d\mathbf{v}.$$

We use the Cauchy-Bunyakovsky and Jensen inequalities, as well as simple algebraic inequalities, and obtain that (see Davydov and Egorov (2000) for details)

$$\left(E\left(f^u - f^l\right)^2\right)^{1/2} \le \left(\int_{\mathbf{u}^l}^{\mathbf{u}^u} E\left(Y^2 \mid \mathbf{X} = \mathbf{v}\right)c(\mathbf{v})d\mathbf{v}\right)^{1/2}$$

$$+\left(EY^2\int_{\mathbf{u}^l}^{\mathbf{u}^u} c(\mathbf{v})d\mathbf{v}\right)^{1/2} \le 2\left(\int_{\mathbf{u}^l}^{\mathbf{u}^u} \left(E\left(Y^2 \mid \mathbf{X} = \mathbf{v}\right) + EY^2\right)c(\mathbf{v})d\mathbf{v}\right)^{1/2}$$

$$\le 2\sqrt{2N \cdot 2EY^2/N^2} = 4\sqrt{EY^2/N} = \delta.$$

In the last inequality, we used the fact that the region of integration is included in $2N$ rectangles from the constructed partition of the unit square.

Summing up these equalities we get

$$M < (N+1)^3 \le \left(\left[\frac{16EY^2}{\delta^2}\right] + 2\right)^3.$$

Hence by Ossiander's theorem

$$\widetilde{Q}_n \Rightarrow Q$$

where Q is the Gaussian field, $EQ(\mathbf{u}) = 0, \mathbf{E}(Q(\mathbf{u}_1), Q(\mathbf{u}_2)) = K(\mathbf{u}_1, \mathbf{u}_2)$. Elementary calculations show that

$$K(\mathbf{u}_1, \mathbf{u}_2) = \mathbf{E}(Y^2\mathbf{1}(\mathbf{X} \le \mathbf{u}_1, \mathbf{X} \le \mathbf{u}_2)) - \int_{\mathbf{0}}^{\mathbf{u}_1} m(\mathbf{v})c(\mathbf{v})d\mathbf{v}\int_{\mathbf{0}}^{\mathbf{u}_2} m(\mathbf{v})c(\mathbf{v})d\mathbf{v}.$$

This construction of the proof can be easily extended to the case $d_1 > 2$ by splitting the corresponding integral into pieces of size $2\mathbf{E}Y^2/N^{d_1}$. So we use Theorem 3.1 of Ossiander (1987) and get the weak convergence $Q_n \Rightarrow Q$ for $d_2 = 1$.

From the tightness of Q_n for $d_2 = 1$ we get the tightness of $\mathbf{Q}_n$ because the tightness is coordinatewise. According to the Cramer-Wold theorem since $\mathbf{Q}_n(\mathbf{u})$ is linear, we obtain the convergence of finite-dimensional distributions for any $d_2 \ge 1$.

The proof is complete.

Proof of Lemma 2.

Note that

$$\mathbf{Z}_n(t) = \left(Q_n\left(\mathbf{e}_{1,X^{(1)}_{n,[nt]+1}}\right), Q_n\left(\mathbf{e}_{2,X^{(2)}_{n,[nt]+1}}\right), \ldots, Q_n\left(\mathbf{e}_{d_1,X^{(d_1)}_{n,[nt]+1}}\right)\right),\ t \in [0,1].$$

Hence, due to Lemma 4.1 by Davydov and Egorov (2000), for all $i = 1, \ldots, d_1$

$$\sup_{0<t<1}\left|\widetilde{Q}_n\left(\mathbf{e}_{i,X^{(i)}_{n,[nt]+1}}\right) - \widetilde{Q}_n\left(\mathbf{e}_{i,t}\right)\right| \xrightarrow{p} 0.$$

We assume $m \equiv 0$, so

$$\frac{1}{\sqrt{n}} \sup_{0<t<1} \left| Q_n\left(\mathbf{e}_{i,X^{(i)}_{n,[nt]+1}}\right) - Q_n(\mathbf{e}_{i,t}) \right| \xrightarrow{p} 0.$$

Therefore the limiting process for $\widetilde{Z}_n$ is the same as for

$$\frac{1}{\sqrt{n}}\left(Q_n(\mathbf{e}_{1,t}), Q_n(\mathbf{e}_{2,t}), \ldots, Q_n(\mathbf{e}_{d_1,t})\right).$$

We have

$$K(\mathbf{e}_{k_1,t}, \mathbf{e}_{k_2,t}) = \mathbf{E}(Y^2 \mathbf{1}(\mathbf{X} \le \mathbf{e}_{k_1,t}, \mathbf{X} \le \mathbf{e}_{k_2,t})).$$

The proof is complete.

Proof of Theorem 2.

Let $\varepsilon_i^{(j)} = \eta_i^{(j)} - \xi_i^{(j)}\theta$ be regression mistakes. From (2) we have $\{\varepsilon_i^{(j)}\}_{i=1}^n$ are i.i.d. with ε_1 and independent with $\{(X_i^{(j)}, \xi_i^{(j)})\}$ for any $j = 1, \ldots, d_1$.

Let $\varepsilon^{(j)} = (\varepsilon_1^{(j)}, \ldots, \varepsilon_n^{(j)})^T$. Note that

$$\widehat{\Delta}_k^{(j)} = \sum_{i=1}^k (\eta_i^{(j)} - \xi_i^{(j)}\widehat{\theta}) = \sum_{i=1}^k (\xi_i^{(j)}(\theta - \widehat{\theta}) + \varepsilon_i^{(j)})$$

$$= \sum_{i=1}^k (\xi_i^{(j)}(\theta - (\xi^T\xi)^{-1}\xi^T\eta) + \varepsilon_i^{(j)})$$

$$= \sum_{i=1}^k \left(\xi_i^{(j)}\left(\theta - (\xi^T\xi)^{-1}\left(\xi^{(j)}\right)^T(\xi^{(j)}\theta + \varepsilon^{(j)})\right) + \varepsilon_i^{(j)}\right)$$

$$= \sum_{i=1}^k \left(\varepsilon_i^{(j)} - \xi_i^{(j)}(\xi^T\xi)^{-1}\left(\xi^{(j)}\right)^T \varepsilon^{(j)}\right).$$

Note that

$$\left\{\sum_{i=1}^{[nt]} \xi_i^{(j)}/n,\ t \in [0,1]\right\} \to L^{(j)}$$

a.s. uniformely, and $\xi^T\xi/n \to G$ a.s.

So we study process

$$\left\{\sum_{i=1}^{[nt]} \varepsilon_i^{(j)} - L^{(j)}(t)G^{-1}\left(\xi^{(j)}\right)^T \varepsilon^{(j)}, \quad 1 \le j \le d_1, \quad t \in [0,1]\right\}.$$

This process is a bounded linear functional of a $d_1 \times d_2$-dimensional process

$$\left\{\sum_{i=1}^{[nt]} (\xi_i^{(j)}\varepsilon_i^{(j)},\ \varepsilon_i^{(j)}), \quad 1 \le j \le d_1, \quad t \in [0,1]\right\}.$$

This is the process from Lemma 2 with $\mathbf{m} \equiv 0$. So we have convergence to a Gaussian process and calculate covariances using Lemma 2.

The proof is complete.

5. Discussion

We now describe the application of this result to testing the hypothesis of linear dependence. Let $d_2 - 1 = d_1$. Let $(\xi_i, \eta_i) = (\xi_{i1}, \dots, \xi_{i,d_1}, \eta_i)$ be independent and identically distributed random vector rows, $i = 1, \dots, n$. In addition, we assume that the column ξ_{i,d_1} consists of ones:

$$\xi_{i,d_1} \equiv 1. \tag{3}$$

We want to test the linear dependence (2). To do this we estimate the parameters θ and $\mathbf{Var}\,\varepsilon_1$, sort the data in ascending order of each of the first $d_1 - 1$ columns of the regressor and calculate processes of the sums of regression residuals. We use the quantile functions $F_{\xi_{1j}}^{-1}$ to apply Theorem 2. So we assume that $\xi_{ij} = F_{\xi_{1j}}^{-1}(X_{ij})$, $i = 1, \dots, n$, $j = 1, \dots, d_1 - 1$. If the matrix G exists and is non-degenerate then under the true hypothesis H_0 we are in the conditions of Theorem 2.

From (3) we have $\widehat{Z}_n(1) = \mathbf{0}$. So we can use a statistics of omega squared type and calculate its limiting distribution by lines of Chakrabarty et al. (2020):

$$\omega_n^2 = \sum_{j=1}^{d_1-1} \int_0^1 \left(\widehat{Z}_n^{(j)}(t)\right)^2 dt \Rightarrow \omega^2 = \sum_{j=1}^{d_1-1} \int_0^1 \left(\widehat{Z}^{(j)}(t)\right)^2 dt.$$

We estimate the covariance function in Theorem 2 from empirical data. An estimate for $\mathbf{P}(X_{1i} \le s,\ X_{1j} \le t)$ is

$$\frac{1}{n} \sum_{k=1}^{n} \mathbf{1}\{\xi_{ki} \le \xi_{[ns],i}^{(i)},\ \xi_{kj} \le \xi_{[nt],j}^{(j)}\}.$$

It converges to the probability uniformely on (s,t) in $[0,1]^2$. We estimate functions $L^{(j)}$ and matrix G by their empirical counterparts.

Thus we construct a statistical test for accurate analysis of the data correspondence to the linear regression model. This test allows one to use multiple ordering of the initial multidimensional data and, due to this, to find non-obvious differences of the investigated data from the model.

Acknowledgement

The work is supported by Mathematical Center in Akademgorodok under agreement No. 075-15-2019-1675 with the Ministry of Science and Higher Education of the Russian Federation. The authors like to thank an anonymous referee for helpful and constructive comments and suggestions.

References

[1] Bickel P. J., Wichura M. J., *Convergence Criteria for Multiparameter Stochastic Processes and Some Applications*, Ann. Math. Stat., **42**:5 (1971), 1656–70. MR0383482

[2] Bischoff W., 1998. *A functional central limit theorem for regression models*, Ann. Stat., **26** (1998), 1398–1410. MR1647677

[3] Brodsky B., Darkhovsky B., *Asymptotically optimal methods of change-point detection for composite hypotheses*, Journal of Statistical Planning and Inference, **133**:1 (2005), 123–138. MR2162571

[4] Chakrabarty A., Chebunin M., Kovalevskii A., Pupyshev I. M., Zakrevskaya N. S., Zhou Q., *A statistical test for correspondence of texts to the Zipf - Mandelbrot law*, Siberian Electronic Mathematical Reports, **17** (2020), 1959–1974. MR4239147

[5] Csorgo M., Horváth L., *Limit Theorems in Change-Point Analysis*, NY: Wiley, 1997. MR2743035

[6] Davydov Y., Egorov V., *Functional limit theorems for induced order statistics*, Mathematical Methods of Statistics, **9**:3 (2000), 297–313. MR1807096

[7] Davydov Y., Zitikis R., *On weak convergence of random fields*, Annals of the Institute of Statistical Mathematics, **60** (2008), 345–365. MR2403523

[8] Dudley, R. M., . *Measures on non-separable metric spaces*, Illinois Journal of Mathematics, **11** (1967), 449–453. MR0235087

[9] Kovalevskii A. P., Shatalin E. V., *Asymptotics of Sums of Residuals of One-Parameter Linear Regression on Order Statistics*, Theory of Probability and its Applications, **59**:3 (2015), 375–387. MR3415974

[10] Kovalevskii A., Shatalin E., *A limit process for a sequence of partial sums of residuals of a simple regression on order statistics*, Probability and Mathematical Statistics **36**:1 (2016), 113–120. MR3529343

[11] Kovalevskii A. P., *Asymptotics of an empirical bridge of regression on induced order statistics*, Siberian Electronic Mathematical Reports, **17** (2020), 954–963. MR4217195

[12] Lorden G., *Procedures for Reacting to a Change in Distribution*, Ann. Math. Statist., **42**:6 (1971), 1897–1908. MR0309251

[13] MacNeill I. B., *Limit processes for sequences of partial sums of regression residuals*, Ann. Prob., **6** (1978), 695–698. MR0494708

[14] MacNeill I.B., Jandhyala V.K., Kaul A., Fotopoulos S.B., *Multiple change-point models for time series*, Environmetrics, **31**:1 (2020), e2593. MR4061127

[15] Moustakides G. V., *Optimal Stopping Times for Detecting Changes in Distributions*, Ann. Statist., **14**:4 (1986), 1379–1387. MR0868306

[16] Ossiander M., *A Central limit theorem under metric entropy with L_2 bracketing*, Ann. Prob., **15** (1987), 897–919. MR0893905

[17] Page E. S., *Continuous inspection schemes*, Biometrika, **41**:1–2 (1954), 100–115. MR0088850

[18] Rao C., *Sequential Tests of Null Hypotheses*, Sankhya, **10**:4 (1950), 361–370. MR0039196

[19] Shiryaev A. N., *Minimax Optimality of the Method of Cumulative Sum (Cusum) in the Case of Continuous Time*, Russian Mathematical Surveys, **51** (1996), 750–751. MR1422244

[20] Shorack G., Wellner J., *Empirical processes with applications to statistics*, Wiley N.-Y., 1986 MR3396731

[21] Straf M.L., *Weak convergence of stochastic processes with several parameters*, Proc. Sixth Berkely Symp. Math. Statist. Prob., **2** (1972), 187–222. MR0402847

Mikhail Georgievich Chebunin
Karlsruhe Institute of Technology,
Institute of Stochastics,
Karlsruhe, 76131, Germany
Novosibirsk State University,
1, Pirogova str.,
Novosibirsk, 630090, Russia
Email address: chebuninmikhail@gmail.com

Artyom Pavlovich Kovalevskii
Novosibirsk State Technical University,
20, K. Marksa ave.,
Novosibirsk, 630073, Russia
Novosibirsk State University,
1, Pirogova str.,
Novosibirsk, 630090, Russia
Email address: artyom.kovalevskii@gmail.com