

## Research Article

# Simultaneous Model Change Detection in Multivariate Linear Regression With Application to Indonesian Economic Growth Data

Wayan Somayasa <sup>1</sup>, Muhammad Kabil Djafar <sup>1</sup>, Norma Muhtar <sup>1</sup>,  
and Desak Ketut Sutiari <sup>2</sup>

<sup>1</sup>Department of Mathematic, Halu Oleo University, Jl. H.E.A. Mokodompit, Kendari 93232, Indonesia

<sup>2</sup>Department of Medical Electrical Engineering, Mandala Waluya University, Anduonohu, Kendari 93232, Indonesia

Correspondence should be addressed to Wayan Somayasa; wayan.somayasa@uho.ac.id

Received 18 July 2023; Revised 24 March 2024; Accepted 26 March 2024; Published 10 May 2024

Academic Editor: Arvind Kumar Misra

Copyright © 2024 Wayan Somayasa et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

In this paper, we study asymptotic model change detection in multivariate linear regression by using the Kolmogorov–Smirnov function of the partial sum process of recursive residuals. We approximate the rejection region and also the power function of the test by establishing a functional central limit theorem for the sequence of the partial sum processes of the recursive residuals of the observations. When the assumed model is true, the limit process is given by the standard multivariate Brownian motion which does not depend on the regression functions. However, when the assumed model is not true (some models change), the limit process is represented by a vector of deterministic trend plus the standard multivariate Brownian motion. The finite sample size rejection region and the power of the test are investigated by means of Monte Carlo simulation. The simulation study shows evidence that the proposed test is consistent in the sense that it attains the power larger than the size of the test when the hypothesis is not true. We also demonstrate the application of the proposed test method to Indonesian economic growth data in which we test the adequacy of three-variate low-order polynomial model. The test result shows that the growth of the Indonesian economy is neither simultaneously constant nor linear. The test has successfully detect the appearance of a change in the model which is mainly caused by the COVID-19 pandemic in 2020.

**Keywords:** economic growth; functional central limit theorem; gross domestic product; Kolmogorov–Smirnov type test; model change detection; multivariate linear regression; multivariate standard Brownian motion; partial sum process; recursive residuals

## 1. Introduction

As stated in Fujikoshi [1], He et al. [2], Somayasa [3], and Somayasa, Ruslan, and Sutiari [4], multivariate linear regression has been applied in many fields of study including agriculture, economics, geology, biology, and chemistry, among others. One important part of statistical modelling using multivariate linear regression is checking whether each component of the mean of the response vector can be adequately represented by a set of regression functions or whether we need more functions to represent some components of the mean of the response vector. This intention can be realized in practice by conducting simultaneous detection whether there are some changes in

regression model. Cotos-Yéñez, Pérez-González, and González-Manteiga [5], Fujikoshi [1], and Zimmerman [6] called this step as model check. In other literatures, it also frequently called goodness of fit test for regression (see Das [7]).

Methods for change detection in multivariate linear regression has been intensively studied. They are commonly developed based on the residuals of the responses. Likelihood ratio test using modified Wilk's lambda statistic has been documented in the literatures of multivariate linear regression (cf. [1, 2]). This method is applicable only when the response vector is normally distributed. Somayasa et al. [8] developed asymptotic model check method for multivariate linear regression with spatial observations on a compact

rectangle based on the partial sum process of the vector of ordinary least squares residuals. The limit process obtained in [8] has been derived analytically by using some transformation theorems and the linear property of the partial sum operator. However, as it can be seen therein, the limit process appeared as an intricate function of the standard multivariate Brownian motion. It depends not only on the design of experiment but also on the assumed regression functions. By this reason, the quantiles of the Kolmogorov–Smirnov as well as the Cramér–von Mises type test statistics can not be computed analytically. This clearly restricted the application of the proposed method in practice. Although the limit process has a complex structure, nevertheless it is geometrically interpretable as a projection of the standard multivariate Brownian motion on its kernel space. This simple structure gives advantage in analyzing the power of the test when the assumed model is not true (see Somayasa [9]).

Partial sum method has been initially investigated for the purpose of quality monitoring in production process. The application of the method has been continuously developed to the problem of monitoring structural change for regression by defining a test based on the partial sum process of recursive residuals. Recent publications related to this context are Jiang and Kurozumi [10], Dao [11], and Otto and Breitung [12], among others. Groen, Kapetanios, and Price [13] pioneered the extension of the application of the method for monitoring structural change in multivariate linear regression based on the partial sum process of the vector of recursive residuals. However, their approaches required a normal distribution assumption attached to the vector of the errors, so that they are not purely asymptotic in the strictly sense. When the vector of random errors in the model is multivariate normally distributed, then the vector of recursive residuals builds a sequence of mutually independent random vectors, so that the limit distribution of the corresponding sequence of the partial sum processes of the vector of recursive residuals can be immediately obtained. This means that the application of the method in the practice must be started with a goodness of fit test or diagnostic check regarding the normality of the vector of recursive residuals.

In this paper, we study the application of the partial sum process of multivariate recursive residuals in asymptotic model check for multivariate linear regression defined on a closed and bounded experimental region. For that, we need to establish a functional central limit theorem for the partial sum process of recursive residuals obtained from multivariate linear regression under more general setting than that defined in Groen, Kapetanios, and Price [13]. It is worth mentioning that the results presented in [10–13] have been derived solely for univariate times series linear regression, where the experimental region is given by the set of positive integers which is not closed and bounded. In contrast to time series, when the experimental region is compact, we have triangular arrays of observations. For that, we need to establish a functional central limit theorem applied to triangular arrays of observations drawn from a multivariate linear regression model. To the knowledge of the authors, this result has not been yet documented elsewhere. It will require more effort on one hand by the existence of the correlation

among the components of the response vector, but on the other hand, linear regression is attractive statistical tool which is applied in many areas such as in response surface methodology, chemical industry, and mining industry, among others.

The paper is organized as follows. In Section 2, we discuss literature review. Multivariate regression model with univariate experimental region together with the definition of the test statistic for checking the adequacy of the model is defined in Section 3. The limit process of the sequence of the partial sum processes of recursive residuals is investigated in Section 4 for the situation under  $H_0$  as well as under  $H_1$ . The proofs of the results are postponed to the appendix. Section 5 discusses the results of numerical simulation. In Section 6, we demonstrate the application of the proposed test method to real data which is the Indonesian economic growth data. Some conclusions and remarks are presented in Section 7.

## 2. Literature Review

Model check or model change detection plays important role in every serious empirical model building for a random phenomenon. As in geostatistics, Bassani and Costa [14] informed that optimum prediction (kriging) of the spatial process in unobserved geographical positions depends on the adequateness of a proposed regression model. Myers, Montgomery, and Anderson-Cook [15] stated that in response surface methodology, the validity of an assumed regression model determines the level of accuracy of the optimum condition of a production process under study. This means that to be able to obtain accurate prediction result, we have to carry out model change detection before the assumed model is used in the prediction (see also Huang and He [16]).

As documented in Fahrmeir et al. [17] and Zimmerman [6], model check or model change detection in multivariate linear regression can be conducted in several various ways: graphical method, Akaike information criterion (AIC), and likelihood ratio test by making use of Wilk's lambda statistic. All methods have similar approach in that they are conducted by investigating the residuals of the observations. As a classical method, the decision obtained by graphical method is subjective so that this method is rarely used in the practice. AIC and likelihood ratio test are two inferential methods that have drawback in the application in that they are conducted under the assumption that the observations are normally distributed (see Fujikoshi [1] and He et al. [2]).

Bischoff and Gegg [18] proposed a purely asymptotic test method for detecting change in multivariate linear regression based on the partial sum process of ordinary least squares residuals. This innovative approach successfully incorporates the theory of high-dimensional stochastic process especially high-dimensional Gaussian process in the statistical inference. Many authors proposed test method based on the partial sums of recursive residuals instead of ordinary least squares residuals. Recently, Dao [11] studied the application of this approach for condition monitoring and fault diagnosis of wind turbines. Jiang and Kurozumi

[10] investigated the power properties of the test based on the modified partial sum process of recursive residuals. Otto and Breitung [12] applied the partial sum method for testing and monitoring structural change of COVID-19.

To the best knowledge of the authors, the only work that investigated the application of the partial sum process of recursive residuals obtained from multivariate linear regression observed over time is that written by Groen, Kapetanios, and Price [13]. As it has been already mentioned in Section 1, they derived the limit process under multivariate normally distributed random error which clearly restricted the application of the method. Motivated by [13], in this work, we propose asymptotic test method based on the partial sum process of recursive residuals with application to simultaneous model change detection in multivariate linear regression of Indonesian economic growth data.

As defined in Bonokeling et al. [19], economic growth is an increase in the production of economic goods and services in one period of time compared with a previous period. Economic growth of a country reflects and measures the ability of the government in developing the economy of the corresponding country. Economic growth is commonly measured in terms of the increase in aggregated market value of additional goods and services produced which is measured based on the gross domestic product (GDP).

According to the study documented in Sari [20] and Febriyanti [21], there are at least four variables that frequently influence the economic growth of a country, namely, investment, government expenditure, export, and import. Each variable has a different impact on economic growth. While investment, government spending, and export positively affect economic growth, import negatively affects economic growth.

As it has been quoted in [19], other factor that can cause negative influence to the economic growth is disaster, such as COVID-19 outbreaks. The COVID-19 has destroyed world economy in 2020 which has brought the world to worst economic recession. Indonesia is one of more than 210 countries in the world that has been hit by the COVID-19 pandemic. Muhyiddin and Nugroho [22] reported that this situation has caused the Indonesian economic grew negatively in the second, third, and fourth quarters of 2020 after a positive growth achieved in the first quarter of 2020. In this work, we aim to check asymptotically that the COVID-19 causes change in the model of the Indonesian economic growth so that it can not be modelled anymore using low-degree polynomial over time.

### 3. Model Definition

We consider a nonparametric multivariate regression

$$\mathbf{Y}(x) = \mathbf{g}(x) + \mathcal{E}(x), x \in [a, b] \quad (1)$$

where  $\mathbf{Y} := (Y_1, \dots, Y_p)^\top$  is the random response vector,  $\mathbf{g} = (g_1, \dots, g_p)^\top$  is the true but unknown vector of regression functions whose components are assumed to be continuous and of bounded variation on  $\mathbf{D} := [a, b] \subset \mathcal{R}$ , and  $\mathcal{E} =$

$(\varepsilon_1, \dots, \varepsilon_p)^\top$  is the random error vector with  $\mathbf{E}(\mathcal{E}) = \mathbf{0}$  and  $\mathbf{Cov}(\mathcal{E}) = \mathbf{\Sigma} = (\sigma_{\ell k})_{\ell=1, k=1}^{p,p}$ . We assume throughout that  $\mathbf{\Sigma}$  is a positive definite matrix. Let  $s_1, \dots, s_m$  be linearly independent regression functions in  $L_2(P_0)$ , where  $P_0$  is the Lebesgue measure on the measure space  $(\mathbf{D}, \mathcal{B}(\mathbf{D}))$ . Our goal is to find an asymptotic simultaneous monitoring procedure to check whether or not there are some changes in the regression models. More specifically, we aim to investigate a test procedure for the hypotheses  $H_0 : g_i \in \mathbf{W} \forall i \in \{1, \dots, p\}$  versus  $H_1 : \exists i \in \{1, \dots, p\}$ , such that  $g_i \notin \mathbf{W}$ , where  $\mathbf{W}$  is a linear subspace in  $L_2(P_0)$  generated by the regression functions  $\{s_1, \dots, s_m\}$ . In contrast to the classical inference method for multivariate linear regression, in this work, we do not need normal assumption for the error vector.

Suppose Model 1 is observed independently over an equidistant experimental design on  $\mathbf{D}$ , with triangular array of design points, given by

$$\gamma_n := \left\{ x_{nj} = a + (b-a) \frac{j}{n}, j = 1 \dots, n \right\}, \quad n \geq 1$$

Correspondingly, let  $\mathbf{Y}_{nj} := \mathbf{Y}(x_{nj})$ ,  $\mathbf{g}_{nj} := \mathbf{g}(x_{nj})$ , and  $\mathcal{E}_{nj} := \mathcal{E}(x_{nj})$ . Then, the triangular array of observations  $\mathbf{Y}_{nj}$  satisfies the model

$$\mathbf{Y}_{nj} = \mathbf{g}_{nj} + \mathcal{E}_{nj}, j = 1, \dots, n \quad (2)$$

Next, for  $j = m+1, \dots, n$ , let  $\gamma_{nj}$  be the subset of  $\gamma_n$  consisting of the first  $j$  design points. Associated with  $\gamma_{nj}$ , we define the following  $j$ -dimensional vectors as follows:

$$\begin{aligned} \mathbf{Y}^{(j)}(\gamma_{nj}) &:= (Y_{i1}, \dots, Y_{ij})^\top \\ \mathbf{g}^{(j)}(\gamma_{nj}) &:= (g_i(x_{n1}), \dots, g_i(x_{nj}))^\top \\ \boldsymbol{\varepsilon}^{(j)}(\gamma_{nj}) &:= (\varepsilon_{i1}, \dots, \varepsilon_{ij})^\top \end{aligned}$$

with  $\mathbf{E}(\boldsymbol{\varepsilon}^{(j)}(\gamma_{nj})) = \mathbf{0} \in \mathcal{R}^j$  and  $\mathbf{Cov}(\boldsymbol{\varepsilon}^{(j)}(\gamma_{nj})) = \sigma_{ii} \mathbf{I}_j$ , where  $\mathbf{I}_j$  is the  $j \times j$  identity matrix. The realization of Model 1 as well as Model 2 when observed on  $\gamma_{nj}$  can be written as

$$\mathbf{Y}_{j \times p} = \mathbf{G}_{j \times p} + \mathcal{E}_{j \times p}, j = m+1, \dots, n \quad (3)$$

where

$$\begin{aligned} \mathbf{Y}_{j \times p} &= \left( \mathbf{Y}^{(1)}(\gamma_{nj}), \dots, \mathbf{Y}^{(p)}(\gamma_{nj}) \right) \\ \mathbf{G}_{j \times p} &= \left( \mathbf{g}^{(1)}(\gamma_{nj}), \dots, \mathbf{g}^{(p)}(\gamma_{nj}) \right) \\ \mathcal{E}_{j \times p} &= \left( \boldsymbol{\varepsilon}^{(1)}(\gamma_{nj}), \dots, \boldsymbol{\varepsilon}^{(p)}(\gamma_{nj}) \right) \end{aligned}$$

with  $\mathbf{E}(\mathcal{E}_{j \times p}) = \mathbf{O}_{j \times p}$  and  $\mathbf{Cov}(\mathbf{vec}(\mathcal{E}_{j \times p})) = \mathbf{\Sigma} \otimes \mathbf{I}_j$ .

Let  $\mathbf{W}_{j \times m}$  be the design matrix of Model 3. That is,  $\mathbf{W}_{j \times m}$  is a  $j \times m$  matrix, defined by

$$\mathbf{W}_{j \times m} = \left( s_1(\gamma_{nj}), \dots, s_m(\gamma_{nj}) \right)$$

whose  $k$ -th column is given by  $s_k(\gamma_{nj})$ , where we define  $s_k(\gamma_{nj}) := (s_k(x_{n1}), \dots, s_k(x_{nj}))^\top \in \mathcal{R}^j$ , for  $k = 1, \dots, m$ . If  $H_0$  is true, then we have the following multivariate linear regression model

$$\mathbf{Y}_{j \times p} = \mathbf{W}_{j \times m} \mathbf{B} + \mathcal{E}_{j \times p}, j = m + 1, \dots, n \quad (4)$$

where  $\mathbf{B}$  is the  $m \times p$  matrix of unknown parameters, defined by

$$\mathbf{B} = \begin{pmatrix} \beta_{11} & \dots & \beta_{p1} \\ \vdots & \ddots & \vdots \\ \beta_{1m} & \dots & \beta_{pm} \end{pmatrix} = \left( \beta^{(1)}, \dots, \beta^{(p)} \right)$$

The least squares estimator of  $\mathbf{B}$  in Model 4 based on the first  $j$  vector of observations can be computed by the following formula:

$$\widehat{\mathbf{B}}_j = \left( \mathbf{W}_{j \times m}^\top \mathbf{W}_{j \times m} \right)^{-1} \mathbf{W}_{j \times m}^\top \mathbf{Y}_{j \times p} \quad (5)$$

Hence, by the definition of the component-wise projection, we have for  $i = 1, \dots, p$ ,

$$\widehat{\beta}_j^{(i)} = \left( \mathbf{W}_{j \times m}^\top \mathbf{W}_{j \times m} \right)^{-1} \mathbf{W}_{j \times m}^\top \mathbf{Y}^{(i)}(\gamma_{nj}) \quad (6)$$

We note that index  $j$  in  $\widehat{\mathbf{B}}_j$  as well as in  $\widehat{\beta}_j^{(i)}$  presented in Equations (5) and (6) means that the estimation is based on the first  $j$  observations. By following the univariate case, we define the  $p$ -dimensional recursive residual of Model 4 as

$$\mathbf{u}_{nj} = \left( u_{nj}^{(1)}, \dots, u_{nj}^{(p)} \right)^\top = \frac{\mathbf{Y}_{nj} - \widehat{\mathbf{B}}_{j-1}^\top \mathbf{s}(x_{nj})}{\sqrt{e_{jj}}} \quad (7)$$

where

$$e_{jj} := 1 + \mathbf{s}^\top(x_{nj}) \left( \mathbf{W}_{(j-1) \times m}^\top \mathbf{W}_{(j-1) \times m} \right)^{-1} \mathbf{s}(x_{nj})$$

with  $\mathbf{s}(x_{nj}) := (s_1(x_{nj}), \dots, s_m(x_{nj}))^\top \in \mathcal{R}^m$ .

For the purpose of testing the hypotheses defined above, we investigate the sequence of the partial sum processes of the  $p$ -dimensional recursive residuals (Equation 7) by transforming the random matrix

$$\mathbf{U}_{p \times (n-m)} = \begin{pmatrix} \mathbf{u}_{n,m+1}^\top \\ \vdots \\ \mathbf{u}_{nn}^\top \end{pmatrix}^\top$$

into a sequence of  $p$ -dimensional stochastic processes  $\{\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})(x) : x \in \mathbf{D}\}$ , defined by

$$\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})(x) := \sum_{j=1}^{\lfloor (n(x-a))/(b-a) \rfloor} \mathbf{u}_{nj} + \left( \frac{n(x-a)}{b-a} - \left\lfloor \frac{n(x-a)}{b-a} \right\rfloor \right) \mathbf{u}_{n, \lfloor (n(x-a))/(b-a) \rfloor + 1} \quad (8)$$

where  $\lfloor t \rfloor := \max \{z \in \mathcal{Z} | z \leq t\}$  and  $\mathbf{u}_{nj} = \mathbf{0} \in \mathcal{R}^p$ , for  $j = 1, \dots, m$ . Let us call  $\{\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})(x) : x \in \mathbf{D}\}$  throughout the paper  $p$ -dimensional recursive residual partial sum process (RRPSP). By the definition, for every  $n \geq 1$ ,  $\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})$  builds a stochastic process with sample path in the space of  $p$ -dimensional vector of continuous functions  $\mathcal{C}^p(\mathbf{D}) := \times_{i=1}^p \mathcal{C}(\mathbf{D})$ . We define a test using the Kolmogorov–Smirnov functional, given by

$$\text{KS}_p^{(n)} := \max_{1 \leq j \leq n} \left\| \boldsymbol{\Sigma}^{-1/2} \mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})(x_{nj}) \right\| \quad (9)$$

It is clear that the wider the dispersion of the assumed model to the true-unknown model, the larger the value of  $\text{KS}_p^{(n)}$  will be. This means that the statistic  $\text{KS}_p^{(n)}$  in Equation (9) measures the discrepancy between the true and the assumed model. By this reason,  $\text{KS}_p^{(n)}$  can be reasonably used as a test statistic in detecting the occurrence of some changes in the model. We will reject  $H_0$  when  $\text{KS}_p^{(n)}$  is large. For that, the limit distribution of  $\text{KS}_p^{(n)}$  under  $H_0$  as well as under  $H_1$  needs to be investigated.

We notice that the partial sum process (Equation 8) differs to those defined in [13] in that they did not include the rest term of Equation (8) in their definition. Consequently, the process defined in [13] has sample path in the space of  $p$ -dimensional right continuous functions on  $\mathbf{D}$  with left limit, denoted by  $\mathcal{D}^p(\mathbf{D})$ , instead of the space  $\mathcal{C}^p(\mathbf{D})$ .

#### 4. Approximation to the Test Statistics

Since the exact distribution of  $\text{KS}_p^{(n)}$  is mathematically not tractable, we investigate their limit distribution. We firstly obtain the limit process of  $\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})$  by applying Theorem 7.5 of Billingsley [23] (see also Theorem 1.5.4 of Van der Vaart and Wellner [24]).

By substituting

$$\widehat{\mathbf{B}}_{j-1} = \left( \mathbf{W}_{(j-1) \times m}^\top \mathbf{W}_{(j-1) \times m} \right)^{-1} \mathbf{W}_{(j-1) \times m}^\top \mathbf{Y}_{(j-1) \times p}$$

into Equation (7), the vector of the recursive residuals can be expressed as follows:

$$\mathbf{u}_{nj} = \frac{\mathbf{Y}_{nj}}{\sqrt{e_{jj}}} - \frac{\mathbf{Y}_{(j-1) \times p}^T \mathbf{W}_{(j-1) \times m} \left( \mathbf{W}_{(j-1) \times m}^T \mathbf{W}_{(j-1) \times m} \right)^{-1} \mathbf{s}(x_{nj})}{\sqrt{e_{jj}}}$$

Since we have  $\mathbf{Y}_{(j-1) \times p} = \mathbf{W}_{(j-1) \times m} \mathbf{B} + \mathcal{E}_{(j-1) \times p}$  and  $\mathbf{Y}_{nj} = \mathbf{B}^T \mathbf{s}(x_{nj}) + \mathcal{E}_{nj}$ , then by substituting these two equations into the preceding one, we get

$$\mathbf{u}_{nj} = \frac{\mathcal{E}_{nj}}{\sqrt{e_{jj}}} - \frac{\mathcal{E}_{(j-1) \times p}^T \mathbf{W}_{(j-1) \times m} \left( \mathbf{W}_{(j-1) \times m}^T \mathbf{W}_{(j-1) \times m} \right)^{-1} \mathbf{s}(x_{nj})}{\sqrt{e_{jj}}}$$

Hence, for  $j = m + 1, \dots, n$ , it holds

$$\mathbf{u}_{nj}^T = \frac{\left( -\mathbf{s}^T(x_{nj}) \left( \mathbf{W}_{(j-1) \times m}^T \mathbf{W}_{(j-1) \times m} \right)^{-1} \right)}{\sqrt{e_{jj}}} \times \mathbf{W}_{(j-1) \times m}^T \mathbf{1} \mathcal{E}_{j \times p} \quad (10)$$

Expression (10) shows that for every  $j = m + 1, \dots, n$ , there exists a column vector  $\mathbf{c}_{nj} \in \mathcal{R}^n$ , defined by

$$\mathbf{c}_{nj}^T := (\mathbf{b}_{nj} | \mathbf{O}_{n-j})$$

where

$$\mathbf{b}_{nj} := \frac{\left( -\mathbf{s}^T(x_{nj}) \left( \mathbf{W}_{j-1, m}^T \mathbf{W}_{j-1, m} \right)^{-1} \mathbf{W}_{j-1, m}^T \mathbf{1} \right)^T}{\sqrt{e_{jj}}} \in \mathcal{R}^j$$

$$\mathbf{O}_{n-j} := (0, 0, \dots, 0)^T \in \mathcal{R}^{n-j}$$

The column vector  $\mathbf{c}_{nj}^T$  satisfies

$$\mathbf{c}_{nj}^T \mathbf{c}_{nj'} = \begin{cases} 1; & \text{if } j = j' \\ 0; & \text{if } j \neq j' \end{cases} \quad (11)$$

So the  $p$ -dimensional vector of recursive residuals  $\mathbf{u}_{nj}$  can be written as

$$\mathbf{u}_{nj}^T = \mathbf{c}_{nj}^T \mathcal{E}_{n \times p} \quad (12)$$

By Equation (12), we get  $\mathbf{E}(\mathbf{u}_{nj}) = \mathbf{0}$ , and by Equation (11), it holds

$$\begin{aligned} \text{Cov}(\mathbf{u}_{nj}, \mathbf{u}_{nj'}) &= \begin{pmatrix} \sigma_{11} \mathbf{c}_{nj}^T \mathbf{c}_{nj'} & \cdots & \sigma_{1p} \mathbf{c}_{nj}^T \mathbf{c}_{nj'} \\ \vdots & \ddots & \vdots \\ \sigma_{p1} \mathbf{c}_{nj}^T \mathbf{c}_{nj'} & \cdots & \sigma_{pp} \mathbf{c}_{nj}^T \mathbf{c}_{nj'} \end{pmatrix} \\ &= \begin{cases} \boldsymbol{\Sigma}; & \text{if } j = j' \\ \mathbf{O}; & \text{if } j \neq j' \end{cases} \end{aligned}$$

Thus,  $\{\mathbf{u}_{nj} = (u_{nj}^{(1)}, \dots, u_{nj}^{(p)})^T : j \geq m + 1\}$  builds a sequence of uncorrelated  $p$ -dimensional random vectors with  $\mathbf{E}(\mathbf{u}_{nj}) = \mathbf{0} \in \mathcal{R}^p$  and  $\text{Cov}(\mathbf{u}_{nj}) = \boldsymbol{\Sigma} \in \mathcal{R}^{p \times p}$ . If the random error vector  $\mathcal{E}_{nj}$  is distributed as  $N_p(\mathbf{0}, \boldsymbol{\Sigma})$ , then  $\mathbf{u}_{nj}$  is also distributed as  $N_p(\mathbf{0}, \boldsymbol{\Sigma})$ . Hence, under such condition, the set  $\{\mathbf{u}_{nj} = (u_{nj}^{(1)}, \dots, u_{nj}^{(p)})^T : j \geq m + 1\}$  constitutes a sequence of independent and identically distributed random vectors.

Furthermore, by Equation (12), there exists an  $(n - m) \times n$  dimensional lower triangular matrix  $\mathbf{C}$ , say, where

$$\mathbf{C} = \begin{pmatrix} \mathbf{c}_{n, m+1}^T \\ \vdots \\ \mathbf{c}_{n, n}^T \end{pmatrix} = \begin{pmatrix} \mathbf{b}_{n, m+1}^T | \mathbf{O}_{n-m-1}^T \\ \vdots \\ \mathbf{b}_{nn}^T \end{pmatrix}$$

with the properties  $\mathbf{C}\mathbf{C}^T = \mathbf{I}_{n-m}$ , such that

$$\begin{pmatrix} \mathbf{u}_{n, m+1}^T \\ \vdots \\ \mathbf{u}_{nn}^T \end{pmatrix} = \mathbf{C} \mathcal{E}_{n \times p} \quad (13)$$

For  $i = 1, \dots, n - m$  and  $k = 1, \dots, n$ , let  $c_{ik}$  be the entry of  $\mathbf{C}$  in the  $i$ -th row and  $k$ -th column. Then, by the definition,  $c_{ik}$  can be concretely written as

$$c_{ik} = \frac{-\mathbf{s}^T(x_{n, m+i}) \left( \mathbf{W}_{(m+i-1) \times m}^T \mathbf{W}_{(m+i-1) \times m} \right)^{-1}}{\sqrt{e_{m+i, m+i}}} \times \mathbf{s}(x_{n, k}) \quad (14)$$

for  $k = 1, \dots, m + i$ , and  $c_{ik} = 0$ , for  $k = m + i + 1, \dots, n$ .

Now we are in the position to state the limit distribution of  $(1/\sqrt{n-m}) \boldsymbol{\Sigma}^{-1/2} \mathbf{Q}_{n-m} (\mathbf{U}_{p \times (n-m)})$  under  $H_0$ . The proof is given in the appendix.

**Theorem 1.** Let the regression functions  $s_1, \dots, s_m$  be linearly independent in  $L_2(\mathbf{P}, \mathbf{D})$ , continuous, and have bounded

variation on  $\mathbf{D}$ . Suppose that  $\mathbf{E}(u_{nj}^{(i)})^4 < \infty$ , for  $i = 1, \dots, p$ ,  $n \geq 1$  and  $j = m + 1, \dots, n$ . If  $H_0$  is true, then for  $n \rightarrow \infty$ ,  $(1/\sqrt{n-m})\boldsymbol{\Sigma}^{-1/2}\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})$  converges in distribution to  $\mathbf{B}_p$ . Thereby,  $\mathbf{B}_p$  is the standard  $p$ -variate Brownian motion on  $\mathbf{D}$ . That is, a centered  $p$ -variate Gaussian process with the covariance function given by

$$\mathcal{E}(x_1, x_2) = \frac{(\min\{x_1, x_2\} - a)}{(b - a)} \mathbf{I}_p, (x_1, x_2) \in \mathbf{D}$$

We consult the reader to Durrett [25] for the definition of the standard  $p$ -variate Brownian motion  $\mathbf{B}_p$ .

By Theorem 1, it is clear for arbitrary fixed  $x \in \mathbf{D}$ , and the sequence of random vectors

$$\frac{1}{\sqrt{n-m}} \boldsymbol{\Sigma}^{-1/2} \mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})(x)$$

converges in distribution to a  $p$ -variate normal distribution  $N_p(\mathbf{0}, \mathbf{A}_x)$ , where

$$\mathbf{A}_x := \frac{x-a}{b-a} \mathbf{I}_p$$

So that by applying Theorem 2.7 in [23] or Theorem 1.3.6 in [24] (continuous mapping theorem), the quadratic form defined by Equation (15)

$$\frac{1}{n-m} \left\| \mathbf{A}_x^{-1/2} \boldsymbol{\Sigma}^{-1/2} \mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})(x) \right\|^2 \quad (15)$$

converges in distribution to a chi-square distribution with  $p$  degrees of freedoms, denoted by  $\chi^2(p)$ . Furthermore, Theorem 1 gives us an approximation to the probability distribution of the Kolmogorov–Smirnov type test statistics  $\text{KS}_p^{(n)}$  when  $H_0$  is true allowing us in approximating the rejection region of the test. It is constructed based on the probability distribution of the statistic  $\sup_{x \in \mathbf{D}} \|\mathbf{B}_p(x)\|$ . For  $\alpha \in (0, 1)$ , an asymptotic size  $\alpha$ -test will reject  $H_0$ , if and only if  $\text{KS}_p^{(n)} \geq v_{1-\alpha}$ , where  $v_{1-\alpha}$  is a positive constant that satisfies the condition  $\mathbf{P}\{\sup_{x \in \mathbf{D}} \|\mathbf{B}_p(x)\| \geq v_{1-\alpha}\} = \alpha$ . In practice,  $\boldsymbol{\Sigma}$  is usually unknown. It is estimated under  $H_0$  by a consistent estimator  $\widehat{\boldsymbol{\Sigma}}_n$ , defined by

$$\widehat{\boldsymbol{\Sigma}}_n := \frac{1}{n} \begin{pmatrix} \mathbf{r}_{n1}^\top \\ \vdots \\ \mathbf{r}_{nn}^\top \end{pmatrix}^\top \begin{pmatrix} \mathbf{r}_{n1}^\top \\ \vdots \\ \mathbf{r}_{nn}^\top \end{pmatrix} \quad (16)$$

where for  $j = 1, \dots, n$ ,

$$\mathbf{r}_{nj} = \mathbf{Y}_{nj} - \mathbf{Y}_{n \times p}^\top \mathbf{W}_{n \times m} (\mathbf{W}_{n \times m}^\top \mathbf{W}_{n \times m})^{-1} \mathbf{s}(x_{nj})$$

This means that Equation (16) is computed based on the  $p$ -dimensional ordinary least squares residuals.

To be able to assess the power of the test, we need to find out the limit distribution of the test statistic when  $H_0$  is not true. For that, we consider the scaled version of Model 2, defined by

$$\mathbf{Y}_{nj} = \frac{1}{\sqrt{n-m}} \mathbf{g}_{nj} + \mathcal{E}_{nj}, j = 1, \dots, n \quad (17)$$

When  $H_0$  is not true, the model can be written as

$$\mathbf{Y}_{j \times p} = \frac{1}{\sqrt{n-m}} \mathbf{G}_{j \times p} + \mathcal{E}_{j \times p}, j = m+1, \dots, n \quad (18)$$

Let  $\mathbf{u}_{nj}^*$  be the vector of the recursive residuals when  $H_0$  is not true. Then, by recalling Equations (10) and (17), we have

$$\mathbf{u}_{nj}^* = \frac{(1/\sqrt{n-m})\mathbf{g}_{nj} + \mathcal{E}_{nj}}{\sqrt{e_{jj}}} - \frac{\mathbf{Y}_{(j-1) \times p}^\top \mathbf{W}_{(j-1) \times m} (\mathbf{W}_{(j-1) \times m}^\top \mathbf{W}_{(j-1) \times m})^{-1} \mathbf{s}(x_{nj})}{\sqrt{e_{jj}}}$$

By substituting

$$\mathbf{Y}_{(j-1) \times p}^\top = \frac{1}{\sqrt{n-m}} \mathbf{G}_{(j-1) \times p}^\top + \mathcal{E}_{(j-1) \times p}^\top$$

obtained from Equation (18) into the preceding equation, we get

$$\mathbf{u}_{nj}^* = \frac{\mathbf{g}_{nj}}{\sqrt{e_{jj}(n-m)}} - \frac{\mathbf{G}_{(j-1) \times p}^\top \mathbf{W}_{(j-1) \times m} (\mathbf{W}_{(j-1) \times m}^\top \mathbf{W}_{(j-1) \times m})^{-1}}{\sqrt{e_{jj}(n-m)}} \times \mathbf{s}(x_{nj}) + \mathbf{u}_{nj} \quad (19)$$

where  $\mathbf{u}_{nj}$  is the recursive residuals under  $H_0$ . It is clear that when  $H_0$  is true,  $\mathbf{u}_{nj}^*$  defined in Equation (19) coincides with  $\mathbf{u}_{nj}$ . This can be obtained by replacing the terms  $\mathbf{g}_{nj}/\sqrt{n-m}$  and  $\mathbf{G}_{(j-1) \times p}^\top/\sqrt{n-m}$  by  $\mathbf{B}^\top \mathbf{s}(x_{nj})$  and  $\mathbf{W}_{(j-1) \times m}^\top$ , respectively.

The limit process of  $\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)}^*)$ , for  $\mathbf{U}_{p \times (n-m)}^* := (\mathbf{u}_{n,m+1}^*, \dots, \mathbf{u}_{n,n}^*)$  when  $H_0$  is not true, is summarized in Theorem 2. The proof is postponed to the appendix.

**Theorem 2.** Let the regression function  $s_1, \dots, s_m$  be linearly independent in  $L_2(\mathbf{P}_0, \mathbf{D})$ , continuous, and have bounded variation on  $\mathbf{D}$ . Suppose that the model

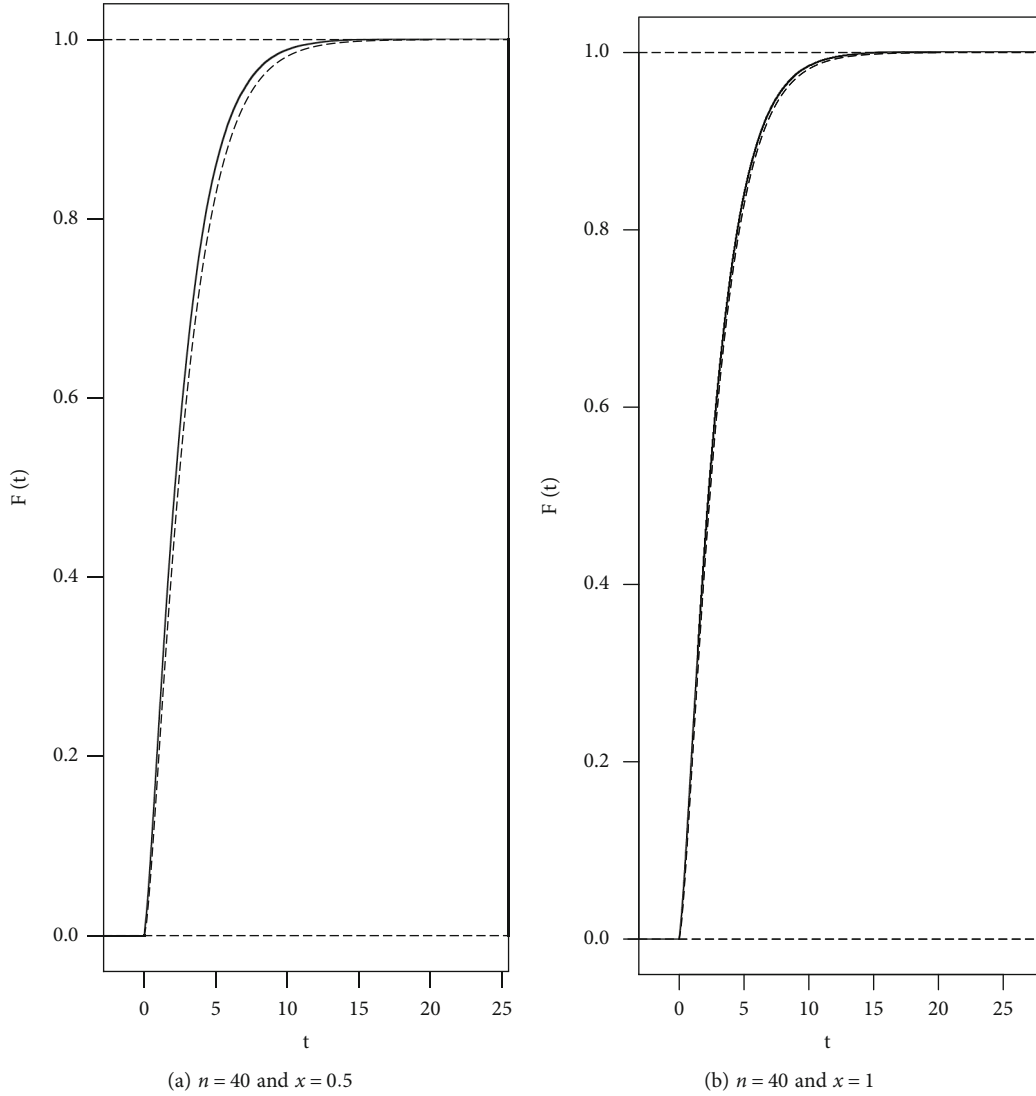


FIGURE 1: The graphs of the ECDF of the quadratic form of the three-variate RRPSP for first-order model (step line) and the CDF of  $\chi^2(3)$  (dotted line).

$$\mathbf{Y}_{nj} = \frac{1}{\sqrt{n-m}} \mathbf{g}(x_{nj}) + \mathcal{E}_{nj}$$

is observed over  $\gamma_n$ . When  $H_0$  is not true, then under the condition of Theorem 1,  $(1/\sqrt{n-m})\Sigma^{-1/2}\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)}^*)$  converges in distribution to the process  $\mathbf{h}_G + \mathbf{B}_p$ , as  $n \rightarrow \infty$ , where for every  $x \in \mathbf{D}$ ,

$$\begin{aligned} \mathbf{h}_G(x) := & \Sigma^{-1/2} \left[ \int_{[a,x]} \tilde{\mathbf{g}}(z) \mathbf{P}_0(dz) - \int_{[a,x]} \left( \int_{[a,z]} \mathbf{g}(v) \mathbf{s}^\top(v) \mathbf{P}_0(dv) \right) \right. \\ & \left. \times \left( \int_{[a,z]} \mathbf{s}(v) \mathbf{s}^\top(v) \mathbf{P}_0(dv) \right)^{-1} \mathbf{s}(z) \mathbf{P}_0(dz) \right] \end{aligned}$$

The convergence result presented in Theorem 2 provides an approximation to the power function of the test.

Let  $\Psi_{\text{KS}_p^{(n)}}(\mathbf{g})$  be the power of the test evaluated in a regression function vector  $\mathbf{g}$ . That is,

$$\Psi_{\text{KS}_p^{(n)}}(\mathbf{g}) := \mathbf{P} \left\{ \text{KS}_p^{(n)} > \nu_{1-\alpha} | \mathbf{g} \right\} \quad (20)$$

Then, by Theorem 2 and the well-known continuous mapping theorem,  $\Psi_{\text{KS}_p^{(n)}}(\mathbf{g})$  can be approximated by the following boundary crossing probability:

$$\Psi(\mathbf{g}) := \mathbf{P} \left\{ \sup_{x \in \mathbf{D}} \|\mathbf{h}_G(x) + \mathbf{B}_p(x)\| > \nu_{1-\alpha} \right\} \quad (21)$$

When  $H_0$  is true, the power computed in Equations (20) and (21) will reduce to the probability of type I error of the test of size  $\alpha$ . Conversely, when  $H_0$  is not true, both determine the probability of the rejection of  $H_0$  provided

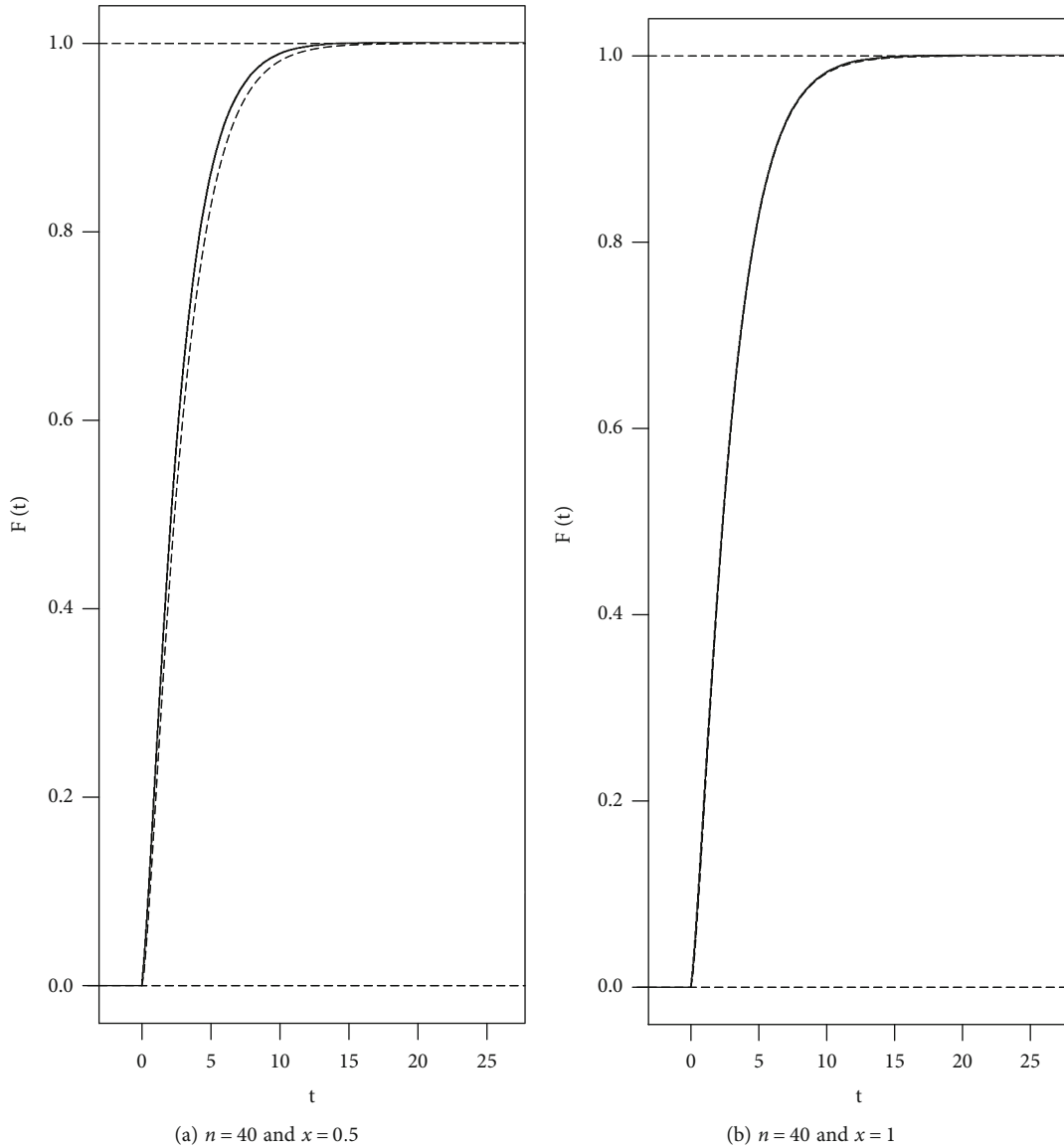


FIGURE 2: The graphs of the ECDF of the quadratic form of the three-variate RRPSP for second-order model (step line) and the CDF of  $\chi^2(3)$  (dotted line).

that  $\mathbf{g}$  is a true regression function vector. In other words,  $\Psi_{\text{KS}_p^{(n)}}(\mathbf{g})$  and  $\Psi(\mathbf{g})$  supply information regarding the ability of the test in detecting the existence of model change. A good test should own the general property stated in Lehmann and Romano [26] and Rasch and Schott [27]. That is, the larger the power under  $H_1$  is, the better is the test.

### 5. Numerical Simulation

In this section, we report on and discuss numerical simulation to study the finite sample size performance of the convergence result and the behavior of the test investigated in the preceding section. To be more specific, we consider  $p$ -variate polynomial model defined by  $s_k(x) := x^{k-1}$ ,

for  $k = 1, \dots, m$ , with experimental region restricted to the unit interval  $[0,1]$  and equidistance design points of size  $n$ . So that the design matrix of the  $j$ -th part of the model for  $j = m + 1, \dots, n$  is given by

$$\mathbf{W}_{j \times m} := \begin{pmatrix} 1 & \frac{1}{n} & \left(\frac{1}{n}\right)^2 & \dots & \left(\frac{1}{n}\right)^{m-1} \\ 1 & \frac{2}{n} & \left(\frac{2}{n}\right)^2 & \dots & \left(\frac{2}{n}\right)^{m-1} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ 1 & \frac{j}{n} & \left(\frac{j}{n}\right)^2 & \dots & \left(\frac{j}{n}\right)^{m-1} \end{pmatrix}$$



TABLE 1: The approximated lower quantiles of  $KS_p^{(n)}$ . The simulation results are based on 100000 runs.

$n$	$\widehat{v}_{0.75}$	$\widehat{v}_{0.85}$	$\widehat{v}_{0.90}$	$\widehat{v}_{0.95}$	$\widehat{v}_{0.99}$
$p = 2$					
25	1.9229	2.1770	2.3591	2.6438	3.2097
30	1.9263	2.1794	2.3583	2.6389	3.2043
35	1.9270	2.1839	2.3680	2.6518	3.2108
40	1.9305	2.1822	2.3639	2.6415	3.1978
45	1.9362	2.1909	2.3712	2.6423	3.1923
50	1.9282	2.1829	2.3618	2.6444	3.1902
$p = 3$					
25	2.2615	2.5203	2.6999	2.9738	3.5337
29	2.2611	2.5176	2.6971	2.9763	3.5235
35	2.2693	2.5232	2.7050	2.9893	3.5511
40	2.2643	2.5208	2.7018	2.9770	3.5278
45	2.2674	2.5241	2.7038	2.9840	3.5250
50	2.2690	2.5218	2.6988	2.9785	3.5211
$p = 4$					
25	2.4970	2.7477	2.9298	3.2032	3.7301
30	2.4979	2.7499	2.9302	3.1978	3.7332
35	2.5054	2.7545	2.9303	3.2013	3.7448
40	2.5075	2.7605	2.9392	3.2172	3.7514
45	2.5121	2.7672	2.9456	3.2205	3.7544
50	2.5174	2.7699	2.9490	3.2157	3.7573

Polynomial model clearly satisfies the condition of Theorem 1.

5.1. *Simulation Under  $H_0$ .* We simulate the convergence result formulated in Theorem 1 by demonstrating the finite sample size attitude of Equation (15) based on two scenarios under  $H_0$ . In the first scenario, we generate the samples from the three-variate polynomial model of degree 1 ( $m = 2$ ), that is, three-variate straight line regression model. The three-dimensional error vectors are generated independently from the three-variate centered normal distribution with the covariance matrix

$$\Sigma = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

The simulation result for  $n = 40$  is devoted in Figure 1, where Figure 1(a) is for  $x = 0.5$  and Figure 1(b) is for  $x = 1$ . The graph of the empirical cumulative distribution function (ECDF) of the quadratic form of the three-variate RRPSP is scattered using the step line. The curve indicated by the dotted line is for the cumulative distribution function (CDF) of  $\chi^2(3)$ . All graphs are generated using R under 10000 runs. In the second scenario, we simulate three-

TABLE 2: Simulated rejection probabilities of size  $\alpha = 0.05$  test for three-variate first-order model computed for several varied values of  $\rho$ ,  $\gamma$ , and  $\delta$  with  $n = 30$ .

$\rho$	$\gamma$	$\delta$	$\widehat{\Psi}(\widehat{\mathbf{g}})$	$\rho$	$\gamma$	$\delta$	$\widehat{\Psi}(\widehat{\mathbf{g}})$
$n = 30$							
0.0	0.0	0.0	0.0497	0.0	0.0	0.5	0.0495
4.5	0.0	0.0	0.0576	0.0	0.0	2.5	0.0609
10.5	0.0	0.0	0.0875	0.0	0.0	5.5	0.1091
12.5	0.0	0.0	0.1109	0.0	0.0	8.5	0.2079
14.5	0.0	0.0	0.1340	0.0	0.0	12.5	0.4353
16.5	0.0	0.0	0.1622	0.0	0.0	15.0	0.6061
20.0	0.0	0.0	0.2259	0.0	0.0	20.0	0.8722
25.0	0.0	0.0	0.3434	0.0	0.0	25.0	0.9801
30.0	0.0	0.0	0.4911	0.0	0.0	30.0	0.9985
40.0	0.0	0.0	0.7722	0.5	0.5	0.5	0.0526
0.0	0.5	0.0	0.0516	2.5	2.5	2.5	0.0948
0.0	4.0	0.0	0.0735	3.5	3.5	3.5	0.1405
0.0	10.5	0.0	0.2512	5.5	5.5	5.5	0.3228
0.0	15.0	0.0	0.4995	6.5	6.5	6.5	0.4430
0.0	20.0	0.0	0.7810	8.5	8.5	8.5	0.7035
0.0	30.0	0.0	0.9917	10.0	10.0	10.0	0.8540
0.0	35.0	0.0	0.9994	12.5	12.5	12.5	0.9725
0.0	40.0	0.0	0.9999	15.0	15.0	15.0	0.9973

TABLE 3: Simulated rejection probabilities of size  $\alpha = 0.05$  test for three-variate first-order model computed for several varied values of  $\rho$ ,  $\gamma$ , and  $\delta$  with  $n = 40$ .

$\rho$	$\gamma$	$\delta$	$\widehat{\Psi}(\widehat{\mathbf{g}})$	$\rho$	$\gamma$	$\delta$	$\widehat{\Psi}(\widehat{\mathbf{g}})$
$n = 40$							
0.0	0.0	0.0	0.0534	0.0	0.0	0.5	0.0497
4.5	0.0	0.0	0.0576	0.0	0.0	2.5	0.0590
10.5	0.0	0.0	0.0913	0.0	0.0	5.5	0.1047
12.5	0.0	0.0	0.1063	0.0	0.0	8.5	0.1995
14.5	0.0	0.0	0.1290	0.0	0.0	12.5	0.4215
16.5	0.0	0.0	0.1586	0.0	0.0	15.0	0.5840
20.0	0.0	0.0	0.2153	0.0	0.0	20.0	0.8565
25.0	0.0	0.0	0.3366	0.0	0.0	25.0	0.9754
30.0	0.0	0.0	0.4747	0.0	0.0	30.0	0.9973
40.0	0.0	0.0	0.7576	0.5	0.5	0.5	0.0533
0.0	0.5	0.0	0.0493	2.5	2.5	2.5	0.0958
0.0	4.0	0.0	0.0742	3.5	3.5	3.5	0.1450
0.0	10.5	0.0	0.2431	5.5	5.5	5.5	0.3167
0.0	15.0	0.0	0.4744	6.5	6.5	6.5	0.4339
0.0	20.0	0.0	0.7723	8.5	8.5	8.5	0.6818
0.0	24.0	0.0	0.9244	10.5	10.5	10.5	0.8393
0.0	30.0	0.0	0.9912	12.5	12.5	12.5	0.9681
0.0	35.0	0.0	0.9988	15.0	15.0	15.0	0.9963

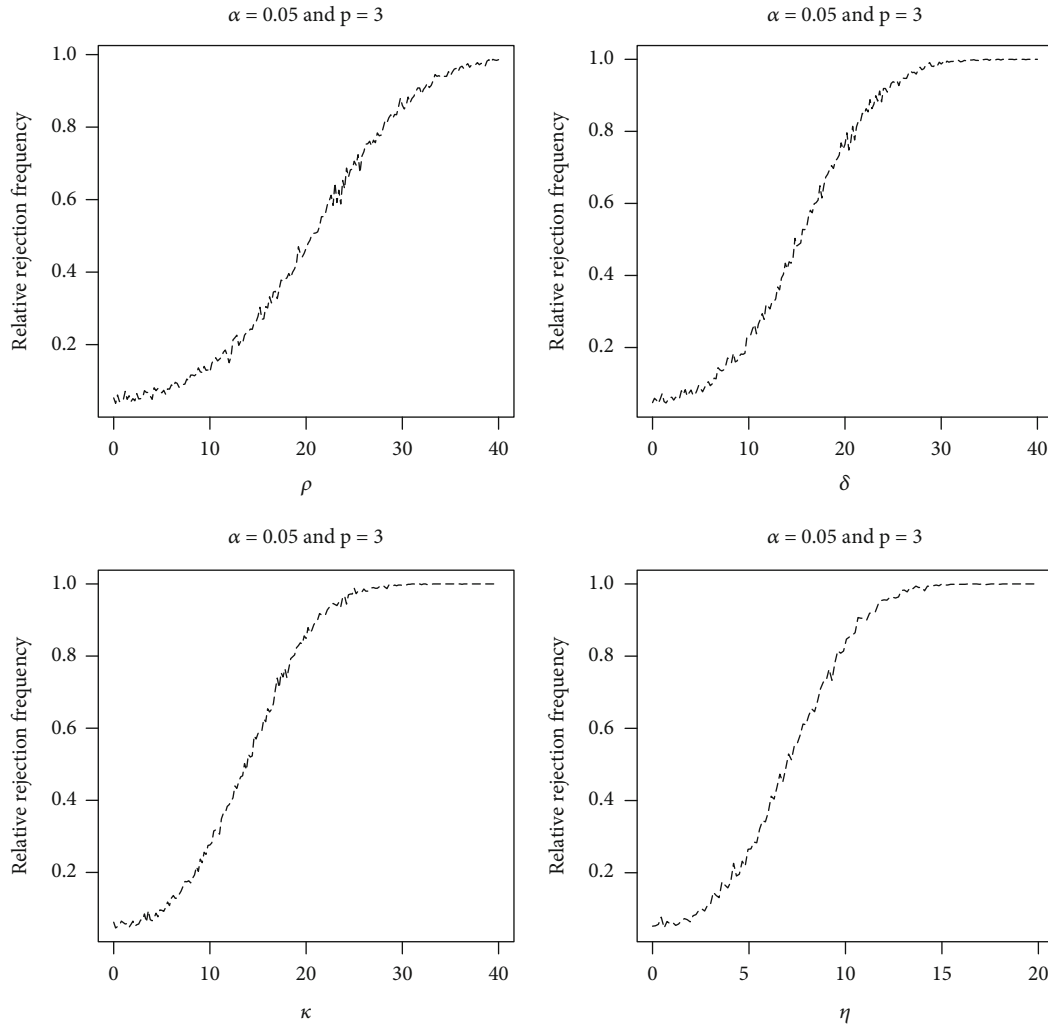


FIGURE 3: The graphs of simulated rejection probabilities of the test for three-variate first-order model using  $n = 40$  sample points. The graphs are generated under 10000 runs.

variate quadratic model ( $m = 3$ ) defined on the unit interval  $[0, 1]$ . The error vectors are generated independently from the three-variate centered normal distribution having the same covariance matrix as in the first case. The graphs of the ECDF of the quadratic form of the RRPSP simulated for  $n = 40$ ,  $x = 0.5$ , and  $x = 1$  together with the graphs of the CDF of  $\chi^2(3)$  are presented in Figures 2(a) and 2(b), respectively.

The simulation results show that independent to the proposed regression model and to the chosen value of  $x \in \{0.5, 1\}$ ,  $\chi^2(3)$  gives a good approximation to the distribution of the quadratic form of the three-variate RRPSP.

Next we approximate the finite sample size lower quantile of the Kolmogorov–Smirnov type statistic  $KS_p^{(n)}$  under  $H_0$  by Monte Carlo simulation. For this purpose, we generate the samples based on the  $p$ -variate polynomial model of degree one ( $m = 2$ ) and two ( $m = 3$ ), for  $p = 2, 3, 4$ , where the error vectors are generated independently from the  $p$ -variate normal distribution  $N_p(\mathbf{0}, \Sigma_p)$ , with the following corresponding covariance matrix:

$$\Sigma_2 = \begin{pmatrix} 2 & -1 \\ -1 & 2 \end{pmatrix}$$

$$\Sigma_3 = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

$$\Sigma_4 = \begin{pmatrix} 9 & 3 & -6 & 12 \\ 3 & 26 & -7 & -11 \\ -6 & -7 & 9 & 7 \\ 12 & -11 & 7 & 65 \end{pmatrix}$$

Table 1 consists of the simulated lower quantiles of  $KS_p^{(n)}$  for  $\alpha = 65\%$ ,  $75\%$ ,  $85\%$ ,  $90\%$ ,  $95\%$ , and  $99\%$ , simulated for  $n = 25, 30, 35, 40, 45$ , and  $50$ , under 100000 runs. These

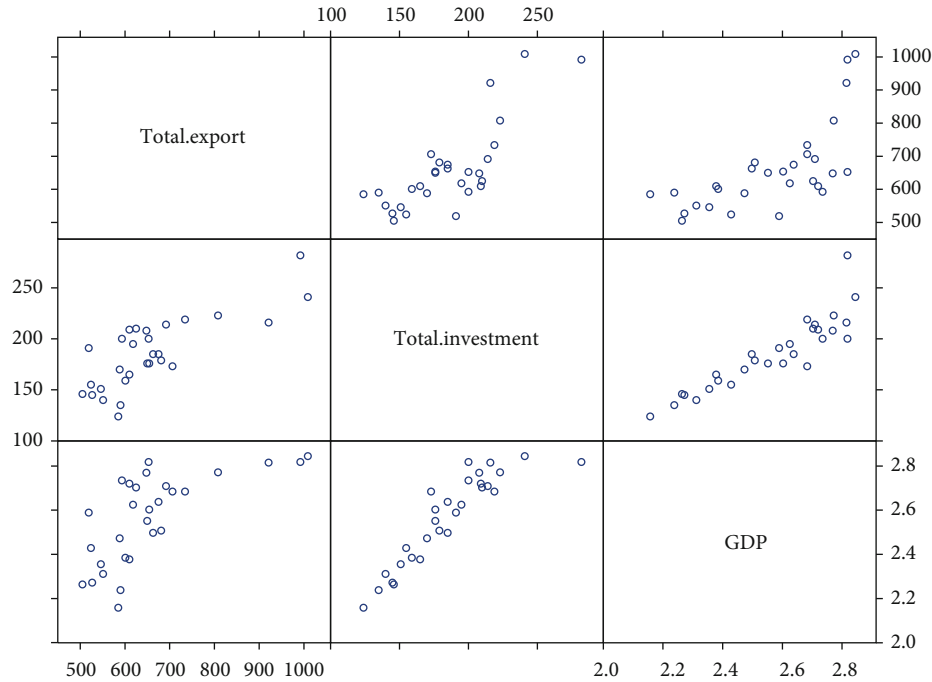


FIGURE 4: The scatter matrix between the total value of export, total value of investment, and the GDP.

quantiles are used in constructing the finite sample size rejection region of the test. The R coding for generating the graphs and the lower tail quantiles can be obtained by request to the authors.

5.2. *Simulation Under  $H_1$ .* We simulate the case of testing  $H_0 : g_i \in \mathbf{W}$ , for all  $i \in \{1, 2, 3\}$ , versus  $g_i \in \mathbf{W}$ , for some  $i \in \{1, 2, 3\}$ , where  $\mathbf{W} = [s_1, s_2]$ , with  $s_k(x) = x^{k-1}$ , for  $x \in [0, 1]$  and  $k = 1, 2$ . This means that under  $H_0$ , we assume a three-variate first-order model on the unit interval  $[0, 1]$ . In this simulation, the samples are generated based on the following three-variate scaled regression model

$$\begin{pmatrix} Y_1 \\ Y_2 \\ Y_3 \end{pmatrix} = \begin{pmatrix} \frac{2 + 0.5(k/n) + 2\rho \exp \{k/n\}}{\sqrt{n-2}} \\ \frac{-0.2 + 0.25(k/n) + 3.5\gamma \exp \{k/n\}}{\sqrt{n-2}} \\ \frac{0.25 + 2.5(k/n) + 4.5\delta \exp \{k/n\}}{\sqrt{n-2}} \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \end{pmatrix}$$

where the error vector is generated independently from the three-variate normal distribution  $N_3(\mathbf{0}, \Sigma)$ , with  $\Sigma$  is given by

$$\Sigma = \begin{pmatrix} 2 & -1 & 0 \\ -1 & 2 & -1 \\ 0 & -1 & 2 \end{pmatrix}$$

When the constants  $\rho, \gamma$ , and  $\delta$  are simultaneously set to zero, the condition under  $H_0$  is fulfilled which means that

there are no changes in the model. In such a case, the power of the test should be equal to the size of the test 5%. However, when at least one of these constants are non-zero, then there exists some  $i \in \{1, 2, 3\}$  such that  $g_i \in \mathbf{W}$ . In other words, there exist some models that change simultaneously.

Some numerical empirical powers of size 5% test simulated for  $n = 30$  and  $n = 40$  with several chosen values of  $\rho, \gamma$ , and  $\delta$  under 10000 runs are given in Tables 2 and 3, respectively. The tables present the empirical power of four types of alternative:

1.  $H_{11} : g_1 \in \mathbf{W}$  (when  $\rho \neq 0, \gamma = 0$ , and  $\delta = 0$ )
2.  $H_{12} : g_2 \in \mathbf{W}$  (when  $\rho = 0, \gamma \neq 0$ , and  $\delta = 0$ )
3.  $H_{13} : g_3 \in \mathbf{W}$  (when  $\rho = 0, \gamma = 0$ , and  $\delta \neq 0$ )
4.  $H_{14} : \mathbf{g}_i \in \mathbf{W}$  for all  $i = 1, 2, 3$  attained when  $\rho \neq 0, \gamma \neq 0$ , and  $\delta \neq 0$ . It can be seen therein that when  $\rho = 0, \gamma = 0$ , and  $\delta = 0$ , the power of the test attains the value 0.04967 for  $n = 30$  and attains the value 0.05340 for  $n = 40$ , which are approximately equal to 5%. So the size of the test is achieved. The graphs of the simulated power function of size 5% test associated with the alternatives  $H_{11}, H_{12}, H_{13}$ , and  $H_{14}$  are scattered in Figure 3 simulated for  $n = 40$  under 10000 runs. The graphs indicate increasing power functions. They have a common feature in that the larger the values of  $\rho, \gamma$ , and  $\delta$  are, the greater the powers. All powers reach the size of the test at the starting points, that is when  $\rho, \gamma$ , and  $\delta$  are simultaneously fixed to zero. Tables 2 and 3 and Figure 3 show that the power of the test for such alternatives gets large as the

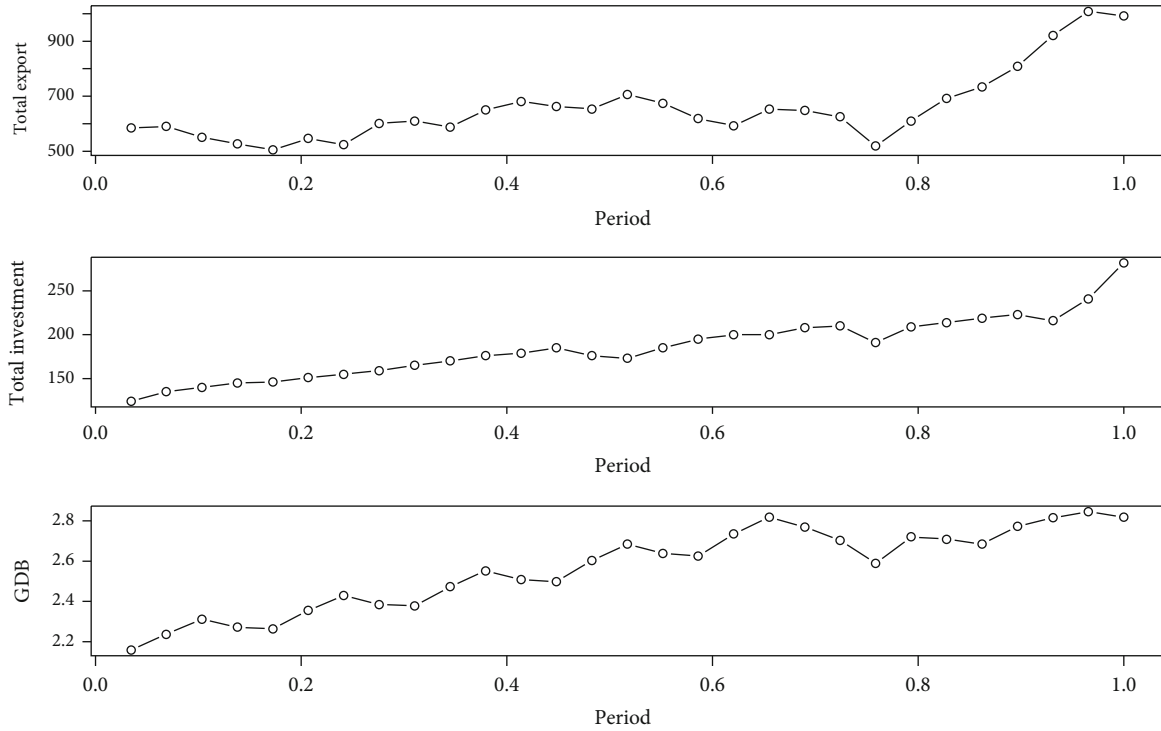


FIGURE 5: The scatter plot of 29 total value of exports, total value of investments, and the GDB of Indonesia observed on the unit interval  $[0, 1]$ .

model moves away from  $H_0$ . This means that the test has good power in detecting the existence of some changes in the model when some changes exist. By referring to Ghosh, Delampady, and Samanta [28], it can be concluded based on the simulation results that the test is unbiased

### 6. Application

In this section, we consider Indonesian economic growth data provided by the Central Bureau of Statistics of the Republic of Indonesia (Badan Pusat Statistik (BPS)) [29]. The data consists of quarterly simultaneous measurements of the total value of export, total value of investment, and the GDP of Indonesia measured starting from the first quarter of 2015 until the first quarter of 2022. The first observation was recorded at the end of March 2015, the second one was at the end of June 2015, and so on. The last one which is the 29th observation was recorded at the end of March 2022. All variables are measured in IDR trillion (see BPS [29]). The data considered in this work can also be obtained by request to the authors. The sample coefficient correlation matrix of the three variables based on the data of size 29 is given by

$$\begin{pmatrix} 1.00000 & 0.78547 & 0.67968 \\ 0.78547 & 1.00000 & 0.90174 \\ 0.67968 & 0.90174 & 1.00000 \end{pmatrix}$$

TABLE 4: The critical values of KS type test for constant, first-order, second-order, and third-order polynomial model.

Model	Critical value
Constant	9.3802
First order	3.0876
Second order	1.8136
Third order	2.8870

The matrix indicates the existence of strong correlations among the three variables. These tendencies are also visualized graphically in Figure 4 which is the scatter matrix of the data. The strongest correlation appears between the total value of investment and the GDP, namely, 0.90174, whereas the weakest correlation is that between the total value of export and the GDP, that is, 0.67968. By reviewing this fact, the statistical analysis of the three variables must be handled simultaneously applying multivariate method.

We aim to model the data using three-variate polynomial regression model and conduct a test for checking the adequacy of a proposed model based on the partial sum process of the recursive residuals. For that we interpret, the data as a realization of 29 independent three-dimensional vector responses admitting three-variate regression model observed on equidistance design points  $\{1/29, 2/29, \dots, 1\}$  over the unit interval  $[0,1]$ . The scatter plot of the total value of investment, the total value of export, and the GDP are presented in Figure 5. The existence of deep peaks in the scatter

plot of the data appeared as the impact of the contraction of the Indonesia's economic growth during the COVID-19 pandemic in 2020. In the normal situation, the total values of export, import, and the GDP should smoothly increase so that they can be modelled by means of lower degree polynomial functions. We infer that the COVID-19 pandemic will change the model in which they need to be modelled by means of higher degree polynomial model. The existence of these changes will be detected by using the proposed test procedure. The computation result of the test statistic  $KS_p^{(n)}$  for several low-order three-variate polynomial regression models is presented in Table 4. The R coding for computing  $KS_p^{(n)}$  can be obtained by request to the author. Since the lower 95% quantile of the distribution of  $KS_p^{(n)}$  for  $n = 29$  and  $p = 3$  takes the value 2.9763 (see Table 1), then the asymptotic size 5% test will reject constant model and first-order model, whereas second-order model and third-order model will not be rejected. The test result seems to be synchronous with the scatter plot of the data in that constant and first-order models are not plausible for the Indonesia's economic growth data during the period of March 2020 until March 2022.

### 7. Conclusion

A limit theorem for the sequence of a random function defined by the RRPSP of multivariate linear regression has been established. The result can be applied in detecting the existence of changes in regression model. In the case of no change, we successfully obtain the limit. It has been given by the standard multivariate Brownian motion  $\{\mathbf{B}_p(x) : x \in \mathbf{D}\}$  which is a model free limit process which depended only on the dimension of the response vector. When there exists at least one change, the limit process has been obtained as a vector of trends plus the standard multivariate Brownian motion, that is,  $\{\mathbf{h}_G(x) + \mathbf{B}_p(x) : x \in \mathbf{D}\}$ . We have built simulation to approximate the finite sample size critical values as well as the power of the Kolmogorov-Smirnov type test. The simulation showed that the test based on the multivariate RRPSP leads to an unbiased test with good power. This test method can be implemented in computer using statistical package like R, so that the computation is quite fast.

### Appendices

#### Proof of Theorem 1

Without loss of generality, we assume for the rectangle  $\mathbf{D} = [a, b]$ , that  $a = 0$  and  $b = 1$ , with  $x_{nj} = j/n$ , for  $j = 1, \dots, n$ . According to the well-known Donsker theorem (cf. Billingsley [23] and Van der Vaart and Wellner [24]), we need to show that the finite dimensional distributions of  $(1/\sqrt{n-m})\Sigma^{-1/2}\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})$  converges to those of  $B_p$  and that  $(1/\sqrt{n-m})\Sigma^{-1/2}\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})$  is tight. For arbitrary  $q \geq 1$ , let  $0 \leq x_1 \leq \dots \leq x_q \leq 1$  be  $q$  different points in  $[0, 1]$  and  $\kappa_1, \dots, \kappa_q$  be nonzero constants. We show that  $\sum_{i=1}^q (\kappa_i/$

$\sqrt{n-m})\Sigma^{-1/2}\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})(x_i)$  converges in distribution to  $\sum_{i=1}^q \kappa_i \mathbf{B}_p(x_i)$  which follows a centered  $p$ -variate normal distribution with the covariance matrix  $\mathbf{I}_p \sum_{i=1}^q \sum_{r=1}^q \kappa_i \kappa_r \min\{x_i, x_r\}$ . By recalling Equation (12) and by defining a notation

$$q_{n;x_i} := nx_i - \lfloor nx_i \rfloor$$

we get the following expressions

$$\begin{aligned} \sum_{i=1}^q \kappa_i \mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})(x_i) &= \sum_{i=1}^q \kappa_i \left[ \sum_{j=1}^{\lfloor nx_i \rfloor} \mathbf{c}_{ij}^\top \mathcal{E}_{n \times p} + q_{n;x_i} \mathbf{c}_{n, \lfloor nx_i \rfloor + 1}^\top \mathcal{E}_{n \times p} \right] \\ &= \sum_{\ell=1}^n \sum_{i=1}^q \kappa_i \left[ \sum_{j=1}^{\lfloor nx_i \rfloor} c_{j\ell} \mathcal{E}_{n\ell} + q_{n;x_i} c_{\lfloor nx_i \rfloor + 1, \ell} \mathcal{E}_{n\ell} \right] \end{aligned}$$

Hence, we have

$$\sum_{i=1}^q \frac{\kappa_i \mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})(x_i)}{\sqrt{n-m}} = \sum_{\ell=1}^n a_{n\ell} \mathcal{E}_{n\ell} \tag{A.1}$$

where

$$a_{n\ell} := \sum_{i=1}^q \frac{\kappa_i}{\sqrt{n-m}} \left( \sum_{j=1}^{\lfloor nx_i \rfloor} c_{j\ell} + q_{n;x_i} c_{\lfloor nx_i \rfloor + 1, \ell} \right)$$

Thus, by Equation (A.1), the problem now reduces to that of showing  $\sum_{\ell=1}^n \Sigma^{-1/2} a_{n\ell} \mathcal{E}_{n\ell}$  converges in distribution to  $\sum_{i=1}^q \kappa_i \mathbf{B}_p(x_i)$ . Since  $\{a_{n\ell} \mathcal{E}_{n\ell} : \ell = 1, \dots, n\}$  are independent, with  $\mathbf{E}(a_{n\ell} \mathcal{E}_{n\ell}) = \mathbf{0}$  and  $\mathbf{Cov}(a_{n\ell} \mathcal{E}_{n\ell}) = a_{n\ell}^2 \Sigma$ , by the well-known Lindeberg-Feller multivariate central limit theorem, it is suffices to show that the covariance of  $\sum_{\ell=1}^n \Sigma^{-1/2} a_{n\ell} \mathcal{E}_{n\ell}$  converges to that of  $\sum_{i=1}^q \kappa_i \mathbf{B}_p(x_i)$  and it satisfies Lindeberg-Feller condition. That is, for every  $\varepsilon > 0$ ,

$$\lim_{n \rightarrow \infty} \frac{1}{\sqrt{n-m}} \sum_{\ell=1}^n \mathbf{E} \left( \|a_{n\ell} \Sigma^{-1/2} \mathcal{E}_{n\ell}\|_{\mathcal{R}^p}^2 \times \mathbf{1}_{\{\|a_{n\ell} \Sigma^{-1/2} \mathcal{E}_{n\ell}\|_{\mathcal{R}^p} > \varepsilon \sqrt{n-m}\}} \right) = 0$$

It is clear that  $\mathbf{E}(\sum_{\ell=1}^n \Sigma^{-1/2} a_{n\ell} \mathcal{E}_{n\ell}) = \mathbf{0}$  and

$$\mathbf{Cov} \left( \sum_{\ell=1}^n \Sigma^{-1/2} a_{n\ell} \mathcal{E}_{n\ell} \right) = \sum_{\ell=1}^n a_{n\ell}^2 \mathbf{I}_p$$

where

$$\begin{aligned}
\sum_{\ell=1}^n a_{n\ell}^2 &= \sum_{i=1}^q \sum_{r=1}^q \frac{\kappa_i \kappa_r}{n-m} \sum_{j=1}^{\lfloor nx_i \rfloor} \sum_{s=1}^{\lfloor nx_r \rfloor} \sum_{\ell=1}^n c_{j\ell} c_{s\ell} \\
&+ \sum_{i=1}^q \sum_{r=1}^q \kappa_i \kappa_r \frac{q_{n;x_i}}{n-m} \sum_{s=1}^{\lfloor nx_r \rfloor} \sum_{\ell=1}^n c_{\lfloor nx_i \rfloor + 1, \ell} c_{s\ell} \\
&+ \sum_{i=1}^q \sum_{r=1}^q \kappa_i \kappa_r \frac{q_{n;x_r}}{n-m} \sum_{j=1}^{\lfloor nx_i \rfloor} \sum_{\ell=1}^n c_{j\ell} c_{\lfloor n(x_r - a)/(b-a) \rfloor + 1, \ell} \\
&+ \sum_{i=1}^q \sum_{r=1}^q \frac{\kappa_i \kappa_r q_{n;x_i} q_{n;x_r}}{n-m} \sum_{\ell=1}^n c_{\lfloor nx_i \rfloor + 1, \ell} c_{\lfloor nx_r \rfloor + 1, \ell} \\
&= \sum_{i=1}^q \sum_{r=1}^q \frac{\kappa_i \kappa_r}{n-m} \sum_{j=1}^{\lfloor nx_i \rfloor} \sum_{s=1}^{\lfloor nx_r \rfloor} \mathbf{c}_{nj}^\top \mathbf{c}_{ns} \\
&+ \sum_{i=1}^q \sum_{r=1}^q \kappa_i \kappa_r \frac{q_{n;x_i}}{n-m} \sum_{s=1}^{\lfloor nx_r \rfloor} \mathbf{c}_{n, \lfloor nx_i \rfloor + 1}^\top \mathbf{c}_{ns} \\
&+ \sum_{i=1}^q \sum_{r=1}^q \kappa_i \kappa_r \frac{q_{n;x_r}}{n-m} \sum_{j=1}^{\lfloor nx_i \rfloor} \mathbf{c}_{nj}^\top \mathbf{c}_{n, \lfloor nx_r \rfloor + 1} \\
&+ \sum_{i=1}^q \sum_{r=1}^q \kappa_i \kappa_r \frac{q_{n;x_i} q_{n;x_r}}{n-m} \mathbf{c}_{n, \lfloor nx_i \rfloor + 1}^\top \mathbf{c}_{n, \lfloor nx_r \rfloor + 1}
\end{aligned}$$

By recalling Equation (11) and the fact that  $q_{n;x_i}/(n-m)$  and  $q_{n;x_r}/(n-m)$  converge to zero, then the right-hand side of the last equation converges as  $n \rightarrow \infty$  to

$$\sum_{i=1}^q \sum_{r=1}^q \kappa_i \kappa_r \min \{x_i, x_r\} \quad (\text{A.2})$$

Expression (A.2) is the formula for the covariance function of  $\sum_{i=1}^q \kappa_i \mathbf{B}_p(x_i)$ . Next, let  $\varepsilon > 0$  be arbitrary small number and let  $M := \max_{m+1 \leq j \leq n} \max_{1 \leq \ell \leq n} |nc_{j\ell}|$ . By the definition, we have

$$\begin{aligned}
|a_{n\ell}| &\leq \frac{\sum_{i=1}^q |\kappa_i|}{\sqrt{n-m}} \left( \sum_{j=1}^{\lfloor nx_i \rfloor} |c_{j\ell}| + |q_{n;x_i}| |c_{n, \lfloor nx_i \rfloor + 1, \ell}| \right) \\
&= \sum_{i=1}^q \frac{|\kappa_i|}{\sqrt{n-m}} \left( M \frac{1}{n} \lfloor nx_i \rfloor + \frac{1}{n} |q_{n;x_i}| M \right) \\
&\leq \sum_{i=1}^q \frac{2M |\kappa_i|}{\sqrt{n-m}} \leq \frac{2Mq}{\sqrt{n-m}}
\end{aligned}$$

Then, we get by applying the well-known bounded convergence theorem

$$\begin{aligned}
\lim_{n \rightarrow \infty} \frac{1}{\sqrt{n-m}} \sum_{\ell=1}^n \mathbf{E} \left( \|a_{n\ell} \Sigma^{-1/2} \mathcal{E}_{n\ell}\|_{\mathcal{E}^p}^2 \times \mathbf{1}_{\{\|a_{n\ell} \Sigma^{-1/2} \mathcal{E}_{n\ell}\| > \varepsilon \sqrt{n-m}\}} \right) \\
\leq \lim_{n \rightarrow \infty} \frac{n4M^2 q^2}{(n-m)} \\
\times \frac{\mathbf{E} \left( \|\Sigma^{-1/2} \mathcal{E}_{n1}\|^2 \times \mathbf{1}_{\{\|\Sigma^{-1/2} \mathcal{E}_{n1}\| > (\varepsilon(n-m)/2Mq)\}} \right)}{\sqrt{n-m}} = 0
\end{aligned}$$

Next we show that the process  $\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})$  is tight. Since the modulus of continuity of the sequence  $\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})$  satisfies

$$\varphi \left( \mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)}); \delta \right) \leq \sum_{i=1}^p \varphi \left( \mathbf{Q}_{n-m}(\mathbf{u}_{(n-m)}^{(i)}); \delta \right)$$

the process  $\mathbf{Q}_{n-m}(\mathbf{U}_{p \times (n-m)})$  is tight only if  $\mathbf{Q}_{n-m}(\mathbf{u}_{(n-m)}^{(i)})$  is tight, for all  $i = 1, \dots, p$ . By some characterizations of tightness in the space  $\mathcal{E}^p(\mathbf{D})$ , we only need to show that  $\mathbf{E} \left( \mathbf{Q}_{n-m}(\mathbf{u}_{(n-m)}^{(i)}) \right)^4 < \infty$ . The result follows by the assumption that  $\mathbf{E}(u_{nj}^{(i)})^4 < \infty$ , for  $i = 1, \dots, p$ ,  $n \geq 1$ , and  $j = m+1, \dots, n$ . The proof finishes.

## Proof of Theorem 2

By recalling the definition of the operator  $\mathbf{Q}_{n-m}$ , we have

$$\begin{aligned}
\frac{1}{\sqrt{n-m}} \mathbf{Q}_{n-m} \left( \mathbf{U}_{p \times (n-m)}^* \right) (x) \\
= \sum_{j=1}^{\lfloor nx \rfloor} \frac{\mathbf{g}_{nj}}{(n-m)\sqrt{e_{jj}}} + \frac{q_{n;x} \mathbf{g}_{n, \lfloor nx \rfloor + 1}}{(n-m)\sqrt{e_{jj}}} - \sum_{j=1}^{\lfloor nx \rfloor} \frac{\mathbf{H}_{j;m} \mathbf{V}_{m;j}^{-1} \mathbf{s}(x_{nj})}{(n-m)\sqrt{e_{jj}}} \\
+ \frac{q_{n;x} \mathbf{H}_{\lfloor nx \rfloor + 1, m} \mathbf{V}_{\lfloor nx \rfloor, m}^{-1} \mathbf{s}(x_{n, \lfloor nx \rfloor + 1})}{(n-m)\sqrt{e_{\lfloor nx \rfloor + 1, \lfloor nx \rfloor + 1}}} \\
+ \frac{\mathbf{Q}_{n-m} \left( \mathbf{U}_{p \times (n-m)} \right) (x)}{\sqrt{n-m}}, x \in \mathbf{D}
\end{aligned} \quad (\text{A.3})$$

where for  $j = 1, \dots, \lfloor nx \rfloor$ , we define

$$\mathbf{V}_{j;m} := \mathbf{W}_{(j-1) \times m}^\top \mathbf{W}_{(j-1) \times m}$$

$$\mathbf{H}_{j;m} := \mathbf{G}_{(j-1) \times p}^\top \mathbf{W}_{(j-1) \times m}$$

By conducting a little algebraic manipulation, Equation (A.3) can be further written as

$$\begin{aligned}
 & \frac{1}{\sqrt{n-m}} \mathbf{Q}_{n-m} \left( \mathbf{U}_{p \times (n-m)}^* \right) (x) \\
 &= \frac{n}{n-m} \left( \sum_{j=1}^{\lfloor nx \rfloor} \frac{\mathbf{g}_{nj}}{n \sqrt{e_{jj}}} + \frac{q_{n;x}}{n} \frac{\mathbf{g}(x_{n, \lfloor nx \rfloor + 1})}{\sqrt{e_{jj}}} \right) \\
 & \quad - \frac{n}{(n-m) \sqrt{e_{jj}}} \sum_{j=1}^{\lfloor nx \rfloor} \left( \frac{1}{n} \sum_{\ell=1}^{j-1} \mathbf{g}(x_{n\ell}) \mathbf{s}^\top(x_{n\ell}) \right) \\
 & \quad \times \left( \frac{1}{n} \sum_{\ell=1}^{j-1} \mathbf{s}(x_{n\ell}) \mathbf{s}^\top(x_{n\ell}) \right)^{-1} \frac{1}{n} \mathbf{s}(x_{nj}) \tag{A.4} \\
 & \quad + \frac{n}{(n-m)} \frac{q_{n;x}}{n} \frac{\left( (1/n) \sum_{\ell=1}^{\lfloor nx \rfloor} \mathbf{g}(x_{n\ell}) \mathbf{s}^\top(x_{n\ell}) \right)}{\sqrt{e_{[\lfloor nx \rfloor + 1, \lfloor nx \rfloor + 1]}}} \\
 & \quad \times \left( \frac{1}{n} \sum_{\ell=1}^{\lfloor nx \rfloor} \mathbf{s}(x_{n\ell}) \mathbf{s}^\top(x_{n\ell}) \right)^{-1} \mathbf{s}(x_{n, \lfloor nx \rfloor + 1}) \\
 & \quad + \frac{1}{\sqrt{n-m}} \mathbf{Q}_{n-m} \left( \mathbf{U}_{p \times (n-m)} \right) (x)
 \end{aligned}$$

Let  $\mathbf{P}_n$ ,  $n \geq 1$  be a sequence of discrete probability measure defined on the  $\sigma$ -field  $\mathcal{B}(\mathbf{D})$  associated with the sequence of the experimental design  $\gamma_n$ , given by

$$\mathbf{P}_n(B) := \frac{1}{n} \sum_{j=1}^n \mathbf{1}_B(x_{nj}), B \in \mathcal{B}(\mathbf{D}), n \geq 1$$

Then, we can write Equation (A.4) in terms of the integrals with respect to  $\mathbf{P}_n$  as follows:

$$\begin{aligned}
 & \frac{1}{\sqrt{n-m}} \mathbf{Q}_{n-m} \left( \mathbf{U}_{p \times (n-m)}^* \right) (x) \\
 &= \frac{n}{(n-m)} \int_{[0,x]} \frac{\mathbf{g}(v) \mathbf{P}_n(dv)}{\sqrt{e_{jj}}} + o_p(1) \\
 & \quad - \frac{n}{\sqrt{(n-m) e_{jj}}} \int_{[0,x]} \left( \int_{[0,z]} \mathbf{g}(u) \mathbf{s}^\top(u) \mathbf{P}_n(du) \right) \\
 & \quad \times \left( \int_{[0,z]} \mathbf{s}(u) \mathbf{s}^\top(u) \mathbf{P}_n(du) \right)^{-1} \mathbf{s}(v) \mathbf{P}_n(dv) \tag{A.5} \\
 & \quad + o_p(1) \left( \int_{[0,x]} \mathbf{g}(u) \mathbf{s}^\top(u) \mathbf{P}_n(du) \right) \\
 & \quad \times \left( \int_{[0,x]} \mathbf{s}(u) \mathbf{s}^\top(u) \mathbf{P}_n(du) \right)^{-1} \mathbf{s}(x_{n, \lfloor nx \rfloor + 1}) \\
 & \quad + \frac{1}{\sqrt{n-m}} \mathbf{Q}_{n-m} \left( \mathbf{U}_{p \times (n-m)} \right) (x)
 \end{aligned}$$

It is clear that  $\mathbf{P}_n$  converges in distribution to the Lebesgue measure  $\mathbf{P}_0$  mentioned in Section 2, which is defined by  $\mathbf{P}_0((0, x]) := x$ . Moreover, since the components of  $\mathbf{g}$  and  $\mathbf{s}$  are bounded and continuous on  $\mathbf{D}$ ,  $q_{n;x}/n$  converges to zero and  $e_{jj}$  converges to one, as  $n \rightarrow \infty$ ; then, by applying

either Theorem 2.1 in [23] (Portmanteau theorem) or Theorem 1.3.4 in Van der Vaart and Wellner [24], all integrals presented in the right-hand side of Equation (A.5) converge as  $n \rightarrow \infty$  to the integral with respect to  $\mathbf{P}_0$ . So by recalling Theorem 1, we get

$$\begin{aligned}
 & \frac{\Sigma^{-1/2}}{\sqrt{n-m}} \mathbf{Q}_{n-m} \left( \mathbf{U}_{p \times (n-m)}^* \right) (x) \\
 & \Rightarrow \Sigma^{-1/2} \int_{[0,x]} \mathbf{g}(v) \mathbf{P}_0(dv) - \Sigma^{-1/2} \int_{[0,x]} \left( \int_{[a,z]} \mathbf{g}(u) \mathbf{s}^\top(u) \mathbf{P}_0(du) \right) \\
 & \quad \times \left( \int_{[0,z]} \mathbf{s}(u) \mathbf{s}^\top(u) \mathbf{P}_0(du) \right)^{-1} \mathbf{s}(v) \mathbf{P}_0(dv) + \mathbf{B}_p(x)
 \end{aligned}$$

We notice that the  $p \times p$  symmetric matrix  $\int_{[0,z]} \mathbf{s}(u) \mathbf{s}^\top(u) \mathbf{P}_0(du)$  is invertible since the columns are linearly independent (see also Somayasa and et al. [8]), establishing the proof.

### Data Availability Statement

The PDF data used to support the findings of this study were supplied by BPS under license and so cannot be made freely available. Requests for access to these data should be made to Wayan Somayasa (wayan.somayasa@uho.ac.id).

### Conflicts of Interest

The authors declare no conflicts of interest.

### Funding

This work was supported in part by the Indonesian Ministry of Research, Technology and Higher Education through the KLN and Publikasi Internasional Research Grant 2019.

### References

- [1] Y. Fujikoshi, "Likelihood ratio tests in multivariate linear model," *Applied Linear Algebra in Action*, vol. 139, 2016.
- [2] Y. He, T. Jiang, T. Wen, and G. Xu, "Likelihood ratio test in multivariate linear regression: from low to high dimension," *Statistica Sinica*, vol. 31, pp. 1215–1238, 2021.
- [3] W. Somayasa, "Approximated optimum condition of second order response surface model with correlated observations," *Journal of Physics: Conference Series*, vol. 725, article 012002, pp. 1–10, 2016.
- [4] W. Somayasa, R. Ruslan, and D. K. Sutiari, "Assessing the optimum condition of multivariate second order response surface model through the asymptotic inference of the eigenvalues," *AIP Conference Proceedings*, vol. 2668, article 070005, 2022.
- [5] T. R. Cotos-Yéñez, A. Pérez-González, and W. González-Manteiga, "Model checks for nonparametric regression with missing data: a comparative study," *Journal of Statistical Computation and Simulation*, vol. 86, no. 16, pp. 3188–3204, 2016.
- [6] D. L. Zimmerman, *Linear Model Theory*, Springer Cham, Berlin, 2020.

- [7] P. Das, "Linear regression model: goodness of fit and testing of hypothesis," in *Econometrics in Theory and Practice*, Springer, Singapore, 2019.
- [8] W. Somayasa, G. N. A. Wibawa, L. Hamimu, and L. O. Ngkoimani, "Asymptotic theory in model diagnostic for general multivariate spatial regression," *International Journal of Mathematics and Mathematical Sciences*, vol. 2016, Article ID 2601601, 16 pages, 2016.
- [9] W. Somayasa, "Assessing the power of tests based on set-indexed partial sums of multivariate regression residuals," *Journal of Applied Mathematics*, vol. 2018, Article ID 2071861, 13 pages, 2018.
- [10] P. Jiang and E. Kurozumi, "Power properties of the modified CUSUM tests," *Communications in Statistics - Theory and Methods*, vol. 48, no. 12, pp. 2962–2981, 2019.
- [11] P. B. Dao, "A CUSUM-based approach for condition monitoring and fault diagnosis of wind turbines," *Energies*, vol. 14, no. 11, p. 3236, 2021.
- [12] S. Otto and J. Breitung, "Backward CUSUM for testing and monitoring structural change with an application to COVID-19 pandemic data," *Econometric Theory*, vol. 39, no. 4, pp. 659–692, 2023.
- [13] J. Groen, G. Kapetanios, and S. Price, "Multivariate methods for monitoring structural change," *Journal of Applied Econometrics*, vol. 28, no. 2, pp. 250–274, 2013.
- [14] M. A. A. Bassani and J. F. L. Costa, *Geostatistics with Data of Different Support Applied to Mining Engineering*, Springer Cham, Berlin, 2022.
- [15] R. H. Myers, D. C. Montgomery, and C. M. Anderson-Cook, *Response Surface Methodology: Process and Product Optimization Using Designed Experiments*, John Wiley & Sons, New York, 4th edition, 2016.
- [16] H. Huang, Z. He, and College of Management and Economics, Tianjin University, Tianjin 300072, China, "A global optimization method for multiple response optimization problems," *Journal of Industrial and Management Optimization*, vol. 19, no. 3, pp. 1755–1769, 2023.
- [17] L. Fahrmeir, T. Kneib, S. Lang, and B. D. Marx, *Regression: Models, Methods and Applications*, Springer Cham, Berlin, 2021.
- [18] W. Bischoff and A. Gegg, "Partial sums process to check regression models with multiple correlated response: with an application for testing a change-point in profile data," *Journal of Multivariate Analysis*, vol. 102, pp. 281–291, 2013.
- [19] D. E. Bonokeling, M. Sholeh, and Mispandi, "The effect of investment, national government expenditure, exports, and imports on Indonesia economic growth," *Jurnal Ekonomi dan Pembangunan*, vol. 30, no. 1, pp. 56–58, 2022.
- [20] M. A. Sari, "Impact of investment, labor, and infrastructure on Java Island economic growth 2011–2017," *Efficient: Indonesian Journal of Development Economics*, vol. 1, no. 3, pp. 230–241, 2018.
- [21] D. F. Febriyanti, "Effect of export and import of gross domestic product in Indonesia 2008–2017," *Ecoplan*, vol. 2, no. 1, pp. 10–20, 2019.
- [22] Muhyiddin and H. Nugroho, "Indonesia development update a year of Covid-19: a long road to recovery and acceleration of Indonesia's development," *The Indonesian Journal of Development Planning*, vol. 5, no. 1, pp. 1–19, 2021.
- [23] P. Billingsley, *Convergence of Probability Measures*, John Wiley & Sons, New York, 2nd edition, 1999.
- [24] A. W. Van der Vaart and J. A. Wellner, *Weak Convergence and Empirical Processes with Applications to Statistics*, Springer, Berlin, 2023.
- [25] R. Durrett, *Probability: Theory and Examples*, Cambridge University Press, London, 2019.
- [26] E. L. Lehmann and J. P. Romano, *Testing Statistical Hypotheses*, Springer Cham, New York, 4th edition, 2022.
- [27] D. Rasch and D. Schott, *Mathematical Statistics*, John Wiley & Sons, New Jersey, 2018.
- [28] J. K. Ghosh, M. Delampady, and T. Samanta, *An Introduction to Bayesian Analysis: Theory and Methods*, Springer Science Media, Berlin, 2006.
- [29] BPS, *Produk Domestik Bruto Indonesia Triwulanan*, Badan Pusat Statistik, Jakarta, 2022.