Journal of Mathematical Research with Applications
Jul., 2015, Vol. 35, No. 4, pp. 448–462
DOI:10.3770/j.issn:2095-2651.2015.04.010
Http://jmre.dlut.edu.cn

Monitoring Change in the Mean Vector of Multivariate Normal Distribution

Weiqi LIU[1,2], Jinfang LIU[2], Ruibing QIN[2,*]

1. *Institute of Management and Decision-making, Shanxi University, Shanxi* 030006*, P. R. China;*
2. *School of Mathematical Sciences, Shanxi University, Shanxi* 030006*, P. R. China*

Abstract In this paper, we propose two monitoring schemes to monitor change in the mean vector of independent multivariate process after a period of size m. The first procedure is based on the CUSUM of residuals, and the second procedure employs the CUSUM of recursive residuals. The corresponding asymptotic distributions of the statistics are derived. Simulations show that the proposed monitoring procedures perform well. The empirical application illustrates the practicability and effectiveness of the procedures.

Keywords change-point; residuals; recursive residuals; online monitoring

MR(2010) Subject Classification 62F12; 60F17; 62H15

1. Introduction

Since the seminal work of Page [1], change point problem attracts increasing attention. Bai [2] estimated the change point in a linear process by the method of least squares. Bai [3] proposed a likelihood approach related to multiple structural changes in regression models. Perron and Qu [4] considered the multiple structural changes in a linear regression model where restrictions are imposed on the parameters. In the last several years, considerable efforts have been made to the change point of multivariate time series. Horváth et al. [5] proposed several statistics to detect the changes in the mean of multivariate dependent stationary processes. Qu and Perron [6] considered a model with changes in the covariance matrix of the errors and in the regression coefficients at some unknown time. Boutahar [7] presented a nonparametric CUSUM test for structural change in the mean of multivariate time series with varying covariance. For a general review of the change point analysis, the readers are referred to Csörgő and Horváth [8], Perron [9].

In practice, new data arrive steadily. Given the costs of failing to detect the change points, it is desirable to detect them as rapidly as possible. Originated with Chu et al. [10], the online change point monitoring arouses the interest of economists and statisticians. Chu et al. [10] proposed a method on the basis of a historical data of fixed size, which can monitor the new

Received April 29, 2014; Accepted May 27, 2015
Supported by Humanities and Social Science Research Project Fund of Ministry of Education (Grant No. 14JA790034) and the National Natural Science Foundation of Tianyuan Fund (Grant No. 11226217).
* Corresponding author
E-mail address: liuwq@sxu.edu.cn (Weiqi LIU); liujf.2009@163.com (Jinfang LIU); rbqin@hotmail.com (Ruibing QIN)

observed samples and give an alarm quickly when a change point occurs. According to this idea, Carsoule and Franses [11] proposed a sequential testing approach to monitor change point in variance. Horváth et al. [12] proposed two CUSUM monitoring schemes respectively by employing residuals and recursive residuals for linear regression models. Aue et al. [13] developed the asymptotic theory for the two monitoring schemes in a linear regression model by martingale difference errors. Additionally, Andreou and Ghysels [14], Zeileis et al. [15] dealt with change point monitoring problem for dynamic econometric models.

However, there is little literature to focus on the multivariate change point monitoring. Considering this and motivated by Horváth et al. [5], we propose two monitoring procedures in the mean vector of multivariate normal distribution in this paper. In this way, we can monitor the factors of a product simultaneously, and alarming when any one of the factors changed. Thus we do not need to monitor these factors one by one. The first procedure is based on the residuals from a regression on a constant, and the second one is based on the recursive residuals. Under the null hypothesis we derive the limiting distribution, and tabulate some critical values, we also prove its consistency under the alternative hypothesis.

The rest of the paper is organized as follows: Section 2 introduces all the necessary assumptions. In Section 3, we discuss the monitoring procedures and analyze their asymptotic properties. In Section 4, we evaluate the performances of the two monitoring procedures by Monte Carlo simulations. Additionally, an empirical application is also given in this section. All proofs of the theorems are collected in Section 5.

2. Models and assumptions

We consider the following model:

$$\boldsymbol{X}_i = \boldsymbol{\mu}_i + \boldsymbol{e}_i, \quad 1 \leqslant i < \infty,$$

where $\boldsymbol{X}_i$ is a $d \times 1$ dimensional normal random vector, $\boldsymbol{\mu}_i$ is a $d \times 1$ vector and $\{\boldsymbol{e}_i\}$ is a $d \times 1$ dimensional error sequence.

First, we state the assumptions which are needed to prove the asymptotic properties of our procedure.

Assumption 2.1 The components of vector $\boldsymbol{X}_i$ are independent of each other, and we denote

$$\overline{\boldsymbol{X}}_k = \frac{1}{k} \sum_{1 \leqslant i \leqslant k} \boldsymbol{X}_i.$$

Assumption 2.2 The error sequences satisfy:

$$\begin{aligned} E\boldsymbol{e}_{i,j} = 0,\ 1 \leqslant i < \infty \ \text{ and } \ 1 \leqslant j \leqslant d, \\ \{\boldsymbol{e}_i, 1 \leqslant i < \infty\} \ \text{ is an independent process,} \\ E\|\boldsymbol{e}_i\|^{\nu} < \infty \ \text{ for some } \nu > 2, \end{aligned} \tag{2.1}$$

and there exists an integer m satisfying

$$\sigma\{\boldsymbol{e}_i, 1 \leqslant i < k\} \ \text{ and } \ \sigma\{\boldsymbol{e}_i, l \leqslant i < \infty\} \tag{2.2}$$

are independent for each l and k satisfying $l-k \geqslant m$.

Assumption 2.3 If Assumption 2.2 holds, then there exists a d dimensional diagonal matrix $\boldsymbol{D}$ such that

$$\lim_{n\to\infty} \frac{1}{n} E\Big(\sum_{1\leqslant i\leqslant n} \boldsymbol{e}_i\Big)\Big(\sum_{1\leqslant i\leqslant n} \boldsymbol{e}_i\Big)^{\mathrm{T}} = \boldsymbol{D}. \tag{2.3}$$

Assumption 2.4 There is no change in the historical data set of size m, i.e.,

$$\boldsymbol{\mu}_i = \boldsymbol{\mu}_0, \quad 1 \leqslant i \leqslant m. \tag{2.4}$$

Remark 2.5 Assumption 2.1 is for the model. Assumptions 2.2 and 2.3 are same as Horváth et al.[5], which are necessary to prove the limit distribution of the statistics in this paper. Assumption 2.4 is "non-contamination assumption" for monitoring the mean change.

Now we observe new incoming data sequentially and monitor if a change occurs in the mean vector. Namely, we want to test the null hypothesis

$$\mathrm{H}_0 : \boldsymbol{\mu}_i = \boldsymbol{\mu}_0, \quad i = m+1, m+2, \ldots, \tag{2.5}$$

against the alternative hypothesis

$$\begin{aligned} \mathrm{H}_1 :&\text{there is a } k^* \geqslant 1 \text{ such that } \boldsymbol{\mu}_i = \boldsymbol{\mu}_0, \quad i = m+1, \ldots, m+k^*, \\ &\text{but } \boldsymbol{\mu}_i = \boldsymbol{\mu}_1, \quad i = m+k^*+1, m+k^*+2, \ldots, \text{with } \boldsymbol{\mu}_0 \neq \boldsymbol{\mu}_1. \end{aligned} \tag{2.6}$$

The mean vectors $\boldsymbol{\mu}_0, \boldsymbol{\mu}_1$ and the so called change point k^* are assumed unknown.

3. Monitoring procedures and main results

In this section we define the two monitoring procedures and state their asymptotic properties under both the null and the alternative.

3.1. The CUSUM of residuals

The CUSUM of residuals statistic we consider is given by

$$T(k) = \sum_{m<i\leqslant m+k} (\boldsymbol{X}_i - \overline{\boldsymbol{X}}_m)^{\mathrm{T}} \boldsymbol{D}^{-1} \sum_{m<i\leqslant m+k} (\boldsymbol{X}_i - \overline{\boldsymbol{X}}_m)$$

and

$$g(m,k) = cm^{1/2}(1+\frac{k}{m})(\frac{k}{k+m})^{\gamma}$$

is chosen as the boundary function, where $c = c(\alpha)$ (α is the significance level) and

$$0 \leqslant \gamma < \frac{1}{2}.$$

Next, we introduce two asymptotic results under the null hypothesis and alternative hypothesis, respectively.

Theorem 3.1 *Assume that the previous assumptions (2.1)–(2.4) hold. Then under the null hypothesis (H_0), we have*

$$\lim_{m\to\infty} P\Big\{\sup_{1\leqslant k\leqslant\infty} T(k)/g_1^2(m,k)\leqslant c\Big\} = P\Big\{\sup_{0<t\leqslant 1}\sum_{1\leqslant i\leqslant d} W_i^2(t)/t^{2\gamma}\leqslant c\Big\}, \tag{3.1}$$

where $g_1(m,k)=m^{1/2}(1+\frac{k}{m})(\frac{k}{k+m})^{\gamma}$, $\{W_i(t), 0\leqslant t<\infty\}$ *is a Wiener process.*

Theorem 3.2 *Assume that the previous assumptions (2.1)–(2.4) hold. Then under the null hypothesis (H_1), we have*

$$\sup_{1\leqslant k\leqslant\infty} T(k)/g_1^2(m,k)\xrightarrow{p}\infty,\quad m\to\infty. \tag{3.2}$$

3.2. The CUSUM of recursion residuals

In this section, we construct the test detector based on the recursion residuals:

$$\widetilde{T}(k)=\sum_{m<i\leqslant m+k}(\boldsymbol{X}_i-\overline{\boldsymbol{X}}_{i-1})^{\mathrm{T}}\boldsymbol{D}^{-1}\sum_{m<i\leqslant m+k}(\boldsymbol{X}_i-\overline{\boldsymbol{X}}_{i-1})$$

and employ $g(m,k)=cm^{1/2}h(\frac{k}{m})$ as the boundary function with

$$\lim_{t\to 0}\frac{t^{\gamma}}{h(t)}=0\ \text{ with some } 0\leqslant\gamma<\frac{1}{2},$$
$$\limsup_{t\to\infty}\frac{(t\log\log t)^{1/2}}{h(t)}<\infty \tag{3.3}$$

where $h(t)$ is positive and continuous on $(0,\infty)$.

Similarly to Theorems 3.1 and 3.2, we have the following two asymptotic results:

Theorem 3.3 *Assume that the previous assumptions (2.1)–(2.4) hold. Then under null hypothesis (H_0), we have*

$$\lim_{m\to\infty} P\Big\{\sup_{1\leqslant k\leqslant\infty}\widetilde{T}(k)/g_1^2(m,k)\leqslant c\Big\} = P\Big\{\sup_{0<t\leqslant 1}\sum_{1\leqslant i\leqslant d} W_i^2(t)/h^2(t)\leqslant c\Big\}, \tag{3.4}$$

where $g_1(m,k)=m^{1/2}h(\frac{k}{m})$, $\{W_i(t), 0\leqslant t<\infty\}$ *is a Wiener process.*

Theorem 3.4 *Assume that the previous assumptions (2.1)–(2.4) hold. Then under the alternative hypothesis (H_1), we have*

$$\sup_{1\leqslant k<\infty}\widetilde{T}(k)/g_1^2(m,k)\xrightarrow{p}\infty,\quad m\to\infty. \tag{3.5}$$

4. Simulations and an empirical application

4.1. Simulations

In this section, we evaluate the performance of the previous monitoring procedures through Monte Carlo simulations.

In order to simulate the results, we replace the matrix $\boldsymbol{D}$ with an appropriate estimator $\boldsymbol{D}_n$

satisfying:

$$\| \boldsymbol{D}_n - \boldsymbol{D} \| = o_p((\log\log n)^{-1/2}), \tag{4.1}$$

which is based on the sample $\{\boldsymbol{X}_i, 1 \leqslant i \leqslant n\}$. If the previous assumptions hold, we can estimate $\boldsymbol{D}$ by

$$\boldsymbol{D}_n = \frac{1}{n-1} \sum_{1\leqslant i\leqslant n} (\boldsymbol{X}_i - \overline{\boldsymbol{X}}_n)(\boldsymbol{X}_i - \overline{\boldsymbol{X}}_n)^{\mathrm{T}}$$

and it can be verified that the previous theorems still hold.

For convenience, we define following notations firstly:

$$\boldsymbol{D}_\gamma = T(k)/(m(1+\frac{k}{m})^2(\frac{k}{k+m})^{2\gamma}), \quad \boldsymbol{D}_a = \widetilde{T}(k)/(mh^2(k/m)),$$

where $h_a(m,k) = [(t+1)(a^2 + \log(1+t))]^{1/2}$ and a satisfies $\exp(-a^2/2) = \alpha$.

γ	α				
	0.010	0.025	0.050	0.100	0.250
0.00	12.5688	10.6249	9.0864	7.5673	5.3846
0.15	12.7989	10.9510	9.4475	7.9320	5.7459
0.25	13.3873	11.3953	9.8468	8.2786	6.1145
0.35	14.0561	12.0574	10.5146	8.9395	6.8383
0.45	16.0328	14.0248	12.4084	10.8205	8.6183
0.49	18.2926	16.2669	14.7174	13.0148	10.6133

Table 1 Critical values

Table 1 gives the critical values we obtain based on 50000 replications of

$$\sup_{0<t\leqslant 1} \sum_{1\leqslant i\leqslant d} W_i^2(t)/t^{2\gamma}$$

for $d = 3$ from a Monte Carlo simulation.

Table 2 reports the empirical size of the monitoring procedures based on Theorems 3.1 and 3.3. We generate data from i.i.d. $\boldsymbol{N}_3(0, I)$ random vectors and simulate the false alarm rate under the null hypothesis with the historical sample size of m =25, 100 and 300. For each m, the processes are monitored from time $m+1$ until q which is set two, four, six and nine times the historical sample size. The reported results are based on 2500 replications. From the table we conclude: The false alarm rate increases with γ increasing at the same levels when m is 25, 100 and 300 respectively. If, however, the monitoring takes place over an interval of fixed length, then increasing m visibly reduces the probability of a false rejection; Moreover, the crossing probability approaches to the theoretical values 10% and 5% when m is large enough meanwhile q raises towards infinity. From the table, we can see the monitoring based on the recursive residuals is more sensitive than that based on D_γ. Though it is impossible setting q and m to extend to infinity in the simulations, the results are appropriate except when $m = 25$.

	q	$\gamma = 0.00$		$\gamma = 0.25$		$\gamma = 0.45$		Recurisive	
		10%	5%	10%	5%	10%	5%	10%	5%
$m = 25$	$2m$	1.92	1.24	6.56	3.56	12.78	8.40	16.04	8.96
	$4m$	8.44	5.10	14.16	8.42	16.34	11.30	23.78	12.98
	$6m$	12.04	7.12	15.74	10.06	18.36	11.38	25.66	13.36
	$9m$	13.70	9.26	17.34	11.60	19.02	13.34	25.76	13.78
$m = 100$	$2m$	0.64	0.24	2.90	1.44	7.34	3.88	11.30	4.86
	$4m$	4.06	1.90	7.36	3.44	9.48	5.16	15.46	6.52
	$6m$	6.34	3.30	8.76	4.16	9.78	5.96	15.68	6.68
	$9m$	8.48	4.16	10.62	5.52	10.08	6.06	16.56	6.86
$m = 300$	$2m$	0.36	0.10	2.30	0.82	6.48	3.20	9.68	4.54
	$4m$	3.38	1.62	5.90	2.84	8.54	4.70	14.30	5.12
	$6m$	5.50	2.50	8.08	3.86	9.34	4.78	14.46	5.70
	$9m$	6.98	2.96	9.24	4.40	9.60	5.04	15.02	5.86

Table 2 Empirical sizes

	$m = 25, k^* = 1$					$m = 100, k^* = 1$			
	$D_{0.00}$	$D_{0.25}$	$D_{0.45}$	$\widetilde{D}_a$		$D_{0.00}$	$D_{0.25}$	$D_{0.45}$	$\widetilde{D}_a$
min	3	1	1	1	min	9	3	1	3
Q_1	8	5	2	3	Q_1	16	8	3	6
med	10	6	4	4	med	18	10	5	8
Q_3	13	9	6	6	Q_3	21	12	7	9
max	47	38	91	19	max	44	30	17	17
	$m = 50, k^* = 5$					$m = 100, k^* = 100$			
	$D_{0.00}$	$D_{0.25}$	$D_{0.45}$	$\widetilde{D}_a$		$D_{0.00}$	$D_{0.25}$	$D_{0.45}$	$\widetilde{D}_a$
min	9	5	1	3	min	40	17	1	8
Q_1	16	11	9	9	Q_1	125	120	119	111
med	19	13	11	11	med	133	128	129	117
Q_3	22	16	13	12	Q_3	142	137	138	122
max	50	38	36	26	max	206	179	207	144

Table 3 Five number summary for the detect time distribution

Table 3 reports the summary statistics for the distribution of the detection time. The mean vector is $(0,\ 0,\ 0)'$ before the change point and $(1,\ 1,\ 1)'$ after it. As the table shows: If the historical sample was fixed, then the procedures are relatively sensitive when the change point occurs at the early beginning of the monitoring. Otherwise the detecting time is a little longer. If the location of the change point stays the same, then the detecting time decreases as the historical sample increases.

Figures 1 − 4 show the estimated densities of the stopping time. All the density curves are asymptotic normality. Just as in Table 2, m is the length of the training period, and $m + k^*$ is the point at which a change in the mean occurred. If the change point occurs at the early stage of the monitoring, then the detector with $\gamma = 0.45$ has the shortest detection delay time, see Figures 1 and 3. But we do not recommend when the change point occurs after a longer period of the monitoring because of the high false alarm rate, see Figures 2 and 4. The truth of $\widetilde{D}_a$ is almost exactly opposite to the detects $\widetilde{D}_\gamma$, that is to say, $\widetilde{D}_a$ is worse than the detects D_γ when k^* is very small and outclass D_γ when k^* is large. The detector with $\gamma = 0.25$ is relatively moderate no matter the change point occurs in the early or the later. Additionally, the distributions of the stoping time do not vary much with the change of m and are only associated with the location of the change point.

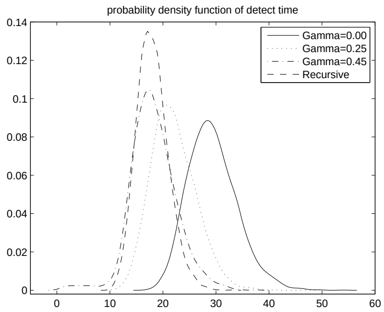

Figure 1 $m = 50, k^* = 5$

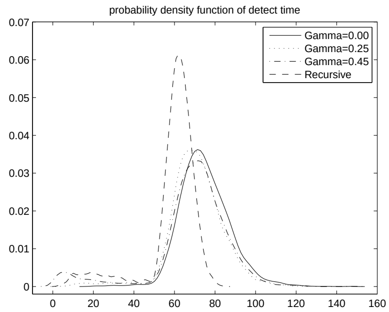

Figure 2 $m = 50, k^* = 50$

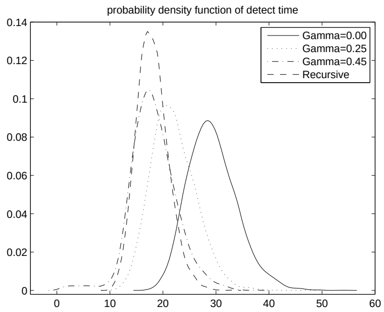

Figure 3 $m = 100, k^* = 10$

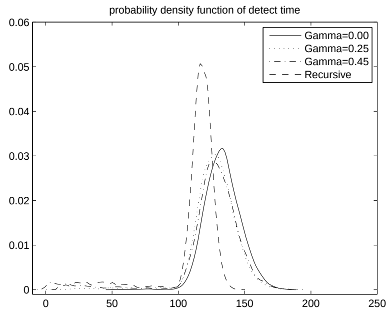

Figure 4 $m = 100, k^* = 100$

Table 4 reports the ARL (Average Run Length) of the procedures. As the table shows, the ARL decreases as the γ increases under the same significance level. Additionally, if the change point occurs at the early stage of the monitoring, then the delay time is relatively short. If, however, the change point occurs after a longer monitoring time, then the delay time increases significantly.

Table 5 shows the empirical powers of the procedures. From the table we can see: It grows as the q increases when m , k^* and γ were fixed, that is to say the longer the monitoring time is, the easier to detect the change point. And it increases as γ grows when m , k^* and q were constants, which suggests that the larger the γ is chosen, the more sensitive this procedure is, but the limit of γ is 1/2 and not equal to 1/2.

	k^*	$\gamma = 0.00$		$\gamma = 0.25$		$\gamma = 0.45$		Recurisive	
		10%	5%	10%	5%	10%	5%	10%	5%
$m = 50$	5	19.3916	21.2674	14.0490	15.5298	11.7140	12.7088	10.7130	10.9746
	50	75.0274	78.9274	70.9194	74.8330	71.8250	76.6748	58.4420	62.4592
$m = 100$	10	29.4708	31.8708	22.0812	23.5934	18.2630	19.5556	18.0388	18.2312
	100	133.6872	138.2840	127.8190	132.7320	127.7902	134.0710	111.1418	117.0960
$m = 300$	30	61.8152	65.6110	79.0064	81.4256	72.6190	84.2880	73.4454	73.4506
	300	354.9752	361.9664	344.7468	352.6002	333.3176	344.6128	312.5988	325.0098

Table 4 The ARLs of the procedures

	k^*	q	$\gamma = 0.00$		$\gamma = 0.25$		$\gamma = 0.45$		Recurisive	
			10%	5%	10%	5%	10%	5%	10%	5%
$m = 50$	5	$2m$	0.9068	0.9972	0.9798	0.9400	0.9824	0.9512	1	1
		$4m$	1	1	1	1	1	1	1	1
		$6m$	1	1	1	1	1	1	1	1
	50	$2m$	0.3690	0.2182	0.4820	0.3706	0.3888	0.2572	1	1
		$4m$	0.9968	0.9908	0.9974	0.9928	0.9930	0.8666	1	1
		$6m$	1	0.9998	1	0.9994	0.9998	0.9938	1	1
$m = 100$	10	$2m$	1	1	1	1	1	1	1	1
		$4m$	1	1	1	1	1	1	1	1
		$6m$	1	1	1	1	1	1	1	1
	100	$3m$	0.9230	0.8216	0.9570	0.8948	0.9324	0.8636	1	1
		$4m$	1	1	1	1	1	1	1	1
		$6m$	1	1	1	1	1	1	1	1
$m = 300$	30	$2m$	1	1	1	1	1	1	1	1
		$4m$	1	1	1	1	1	1	1	1
		$6m$	1	1	1	1	1	1	1	1
	300	$3m$	1	1	1	1	1	1	1	1
		$4m$	1	1	1	1	1	1	1	1
		$6m$	1	1	1	1	1	1	1	1

Table 5 The Empirical powers of the procedures

4.2. Empirical application

In this section, we illustrate our previous monitoring schemes by a real data example of a white wine production process from May 2004 to February 2007. The data contains totally 4898 observations, and is publicly available in the "Wine Quality Data Set" of the UCI Machine Learning Repository (http://archive.ics.uci.edu/ml/datasets/Wine+Quality). The data were recorded by a computerized system and contains eleven variables, indicating the quality of the wine, including fixed acidity, volatile acidity, citric acid, residual sugar, chlorides, free sulfur dioxide, total sulfur dioxide, density, pH, sulphates and alcohol(denoted by $y_1, y_2, \ldots, y_{11}$, respectively). Another categorical quality variable between 0 (very bad) and 10 (very excellent) is also provided based on sensory analysis. For more detail about this example we refer to Cortez et al. [16].

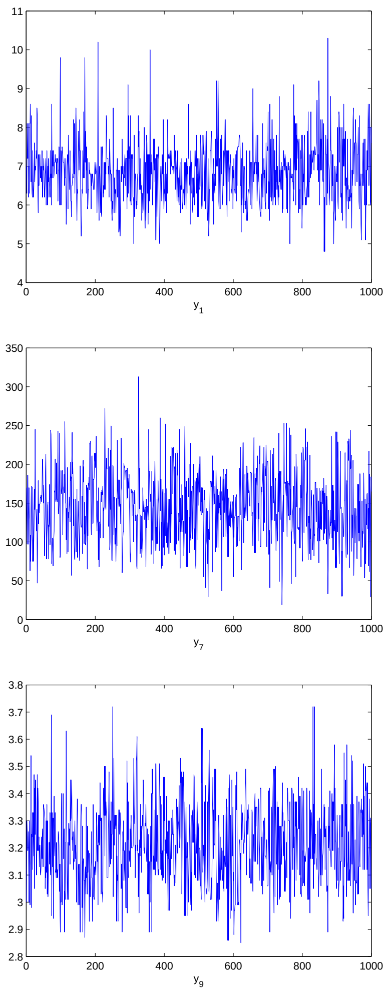

Figure 5 Line chart of y_1, y_7 and y_9 respectively

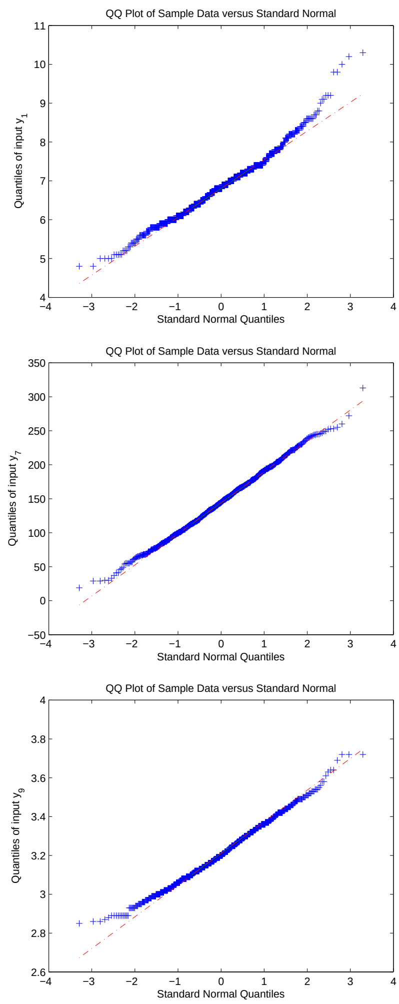

Figure 6 Q-Q plot of y_1, y_7 and y_9 respectively

The goal is to monitor the production process of Vinho Verde wine to guarantee its quality. For convenience, we just select three variables here, fixed acidity(y_1), total sulfur dioxide(y_7), PH(y_9), based on the first 1000 samples. The sample correlation coefficient matrix of this data (not reported here) implies that the variables are not correlated. Figure 5 shows the Line Chart of the raw data for the three variables. The Q-Q plots in Figure 6 indicates that the marginal distributions of these variables are close to normal distributions. And the Chi-square test of this three variables are 7.4252e-016, 0.59 and 7.3715e-094 respectively, which indicates that the original data obey the normal distribution. Furthermore, Figure 7 also illustrates that the distribution of raw data is close to a multinormal distribution.

Then, we begin to monitor the change point. Let initial 150 data be the historical samples and we monitor from the 151^{st} data at the level $\alpha = 0.1$ and $\gamma = 0.25$. As a result, the stoping time is 191 which indicates that there occurs the change point before 191, which coincides with the real change point at 165 based on the sensory analysis.

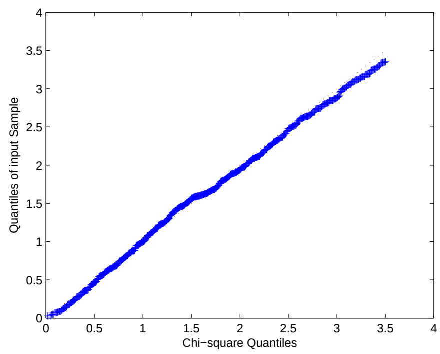

Figure 7 Q-Q plot of y_1, y_7 and y_9 jointly

5. Proofs of theorems

Let

$$\boldsymbol{U}_m(k) = \sum_{m<i\leqslant m+k} \boldsymbol{e}_i - \frac{k}{m}\sum_{1\leqslant i\leqslant m} \boldsymbol{e}_i, \quad \boldsymbol{W}_m(k) = \boldsymbol{W}_{1,m}(k) - \frac{k}{m}\boldsymbol{W}_{2,m}(m).$$

Lemma 5.1 *If the conditions in Theorem 3.1 hold, then there exist two independent multivariate Wiener processes* $\{\boldsymbol{W}_{1,m}(t), 0 \leqslant t < \infty\}$ *and* $\{\boldsymbol{W}_{2,m}(t), 0 \leqslant t < \infty\}$ *such that*

$$\sup_{1\leqslant k<\infty} |\boldsymbol{U}_m^{\mathrm{T}}(k)\boldsymbol{D}^{-1}\boldsymbol{U}_m(k) - \boldsymbol{W}_m^{\mathrm{T}}(k)\boldsymbol{W}_m(k)|/g_1^2(m,k) = o_p(1), \tag{5.1}$$

where $g_1(m,k) = m^{1/2}(1+\frac{k}{m})(\frac{k}{k+m})^\gamma$.

Proof We note that

$$\boldsymbol{U}_m(k) = (U_{m1}(k), \dots, U_{md}(k)),$$

where $U_{mj}(k) = \sum_{m<i\leqslant m+k} e_{ij} - \frac{k}{m}\sum_{1\leqslant i\leqslant m} e_{ij}$ with $1 \leqslant j \leqslant d$.

Let d_{jj} be the j^{th} row of the j^{th} column of $\boldsymbol{D}$ for $1 \leqslant j \leqslant d$. By Lemma 5.3 in Horváth et al. [12], we have

$$\begin{aligned}
&\sup_{1\leqslant k<\infty} \big|U_{mj}(k) - d_{jj}(W_{1,mj}(k) - \frac{k}{m}W_{2,mj}(m))\big|/g_1(m,k) \\
&= O_p(1)\sup_{1\leqslant k<\infty}\{k^{1/\nu} + \frac{k}{m}m^{1/\nu}\}/\{m^{1/2}(1+\frac{k}{m})(\frac{k}{m+k})^\gamma\} \\
&= O_p(1)I.
\end{aligned} \tag{5.2}$$

We denote $I = \sup_{1\leqslant k<\infty}\{k^{1/\nu} + \frac{k}{m}m^{1/\nu}\}/\{m^{1/2}(1+\frac{k}{m})(\frac{k}{m+k})^\gamma\}$.

When $\gamma < \dfrac{1}{2}$ and $\nu > 2$, we have

$$I \leqslant 2^\gamma\{m^{\gamma-1/2}\max_{1\leqslant k\leqslant m} k^{1/\nu-\gamma} + m^{1/\nu-1/2}\} = o(1), \quad \text{as} \quad 1 \leqslant k \leqslant m \tag{5.3}$$

and

$$I \leqslant 2^{\gamma+1}m^{1/\nu-1/2} = o(1), \quad \text{as} \quad m \leqslant k < \infty \tag{5.4}$$

Hence

$$\begin{aligned}
&\sup_{1\leqslant k<\infty} |\boldsymbol{U}_m^{\mathrm{T}}(k)\boldsymbol{D}^{-1}\mathbf{U}_m(k) - \boldsymbol{W}_m^{\mathrm{T}}(k)\,\boldsymbol{W}_m(k)|/g_1^2(m,k)\\
&= |(O_{p1}(1)I,\ldots,O_{pd}(1)I)^{\mathrm{T}}\boldsymbol{D}^{-1}(O_{p1}(1)I,\ldots,O_{pd}(1)I) - \boldsymbol{W}_m^{\mathrm{T}}(k)\,\boldsymbol{W}_m(k)|/g_1^2(m,k)\\
&\leqslant O_p(1)|(I,\ldots,I)^{\mathrm{T}}\mathrm{D}^{-1}(I,\ldots,I) - \boldsymbol{W}_m^{\mathrm{T}}(k)\,\boldsymbol{W}_m(k)|\\
&\leqslant O_p(1)\sum_{1\leqslant i\leqslant d}\sup_{1\leqslant k<\infty}(k^{1/\nu}+\frac{k}{m}m^{1/\nu})^2/g_1^2(m,k)\\
&= O_p(1)I^2.
\end{aligned} \tag{5.5}$$

Putting together (5.2)–(5.5), we prove the Lemma 5.1. □

Proof of Theorem 3.1 Horváth et al. [12] has proved that

$$\sup_{1\leqslant k<\infty}\frac{W_{1,mj}(k)-\frac{k}{m}W_{2,mj}(m)}{g_1(m,k)} \xrightarrow{\mathcal{D}} \sup_{0<t<\infty}\frac{W_{1j}(t)-tW_{2j}(t)}{(1+t)(t/(1+t))^{\gamma}} \overset{\mathcal{D}}{=} \sup_{0<t\leqslant 1}\frac{W_j(t)}{t^{\gamma}}.$$

Considering that the distribution of $\{\boldsymbol{W}_{1,m}(t), \boldsymbol{W}_{2,m}(t), 0\leqslant t<\infty\}$ does not depend on m, we can prove that

$$\begin{aligned}
&\sup_{1\leqslant k<\infty}\frac{\boldsymbol{W}_m^{\mathrm{T}}(k)\,\boldsymbol{W}_m(k)}{g_1^2(m,k)}\\
&= \sup_{1\leqslant k<\infty}\frac{(\boldsymbol{W}_{1,m}(k)-\frac{k}{m}\boldsymbol{W}_{2,m}(m))^{\mathrm{T}}(\boldsymbol{W}_{1,m}(k)-\frac{k}{m}\boldsymbol{W}_{2,m}(m))}{g_1^2(m,k)}\\
&\xrightarrow{\mathcal{D}} \sup_{0<t<\infty}\frac{(\boldsymbol{W}_1(t)-t\,\boldsymbol{W}_2(1))^{\mathrm{T}}(\boldsymbol{W}_1(t)-t\,\boldsymbol{W}_2(1))}{(1+t)^2(\frac{t}{1+t})^{2\gamma}}\\
&\overset{\mathcal{D}}{=} \sup_{0<t\leqslant 1}\frac{\boldsymbol{W}^{\mathrm{T}}(t)\,\boldsymbol{W}(t)}{t^{2\gamma}},
\end{aligned} \tag{5.6}$$

where $\boldsymbol{W}(t)$ is a d dimensional Wiener process.

Under the null hypothesis, it is obvious that

$$T(k) = \boldsymbol{U}_m^{\mathrm{T}}(k)\boldsymbol{D}^{-1}\,\boldsymbol{U}_m(k). \tag{5.7}$$

Hence, Theorem 3.1 follows from Lemma 5.1 and (5.6)–(5.7). □

Proof of Theorem 3.2 Let $\tilde{k} = k^* + m$. By the alternative hypothesis we have

$$\begin{aligned}
T(\tilde{k}) &= \sum_{m<i\leqslant m+\tilde{k}}(\boldsymbol{X}_i-\bar{\boldsymbol{X}}_m)^{\mathrm{T}}\boldsymbol{D}^{-1}\sum_{m<i\leqslant m+\tilde{k}}(\boldsymbol{X}_i-\bar{\boldsymbol{X}}_m)\\
&=\Big(\sum_{i=m+k^*}^{m+\tilde{k}}(\boldsymbol{\mu}_*-\boldsymbol{\mu}_0)+\sum_{i=m+1}^{m+\tilde{k}}(\mathbf{e}_i-\frac{1}{m}\sum_{1\leqslant i\leqslant m}\boldsymbol{e}_i)\Big)^{\mathrm{T}}\boldsymbol{D}^{-1}\\
&\quad\Big(\sum_{i=m+k^*}^{m+\tilde{k}}(\boldsymbol{\mu}_*-\boldsymbol{\mu}_0)+\sum_{i=m+1}^{m+\tilde{k}}(\boldsymbol{e}_i-\frac{1}{m}\sum_{1\leqslant i\leqslant m}\boldsymbol{e}_i)\Big).
\end{aligned} \tag{5.8}$$

Let

$$\boldsymbol{V}_m(k) = \sum_{i=m+1}^{m+k} (\boldsymbol{X}_i - \overline{\boldsymbol{X}}_m). \tag{5.9}$$

Then

$$\boldsymbol{V}_m(\tilde{k}) = \sum_{i=m+k^*}^{m+\tilde{k}} (\boldsymbol{\mu}_* - \boldsymbol{\mu}_0) + \boldsymbol{U}_m(\tilde{k}). \tag{5.10}$$

Putting (5.8)–(5.10) together, we get

$$\begin{aligned} T(\tilde{k}) = &\sum_{i=m+k^*}^{m+\tilde{k}} (\boldsymbol{\mu}_* - \boldsymbol{\mu}_0)^{\mathrm{T}} \boldsymbol{D}^{-1} \sum_{i=m+k^*}^{m+\tilde{k}} (\boldsymbol{\mu}_* - \boldsymbol{\mu}_0) + \sum_{i=m+k^*}^{m+\tilde{k}} (\boldsymbol{\mu}_* - \boldsymbol{\mu}_0)^{\mathrm{T}} \boldsymbol{D}^{-1} \boldsymbol{U}_m(\tilde{k}) + \\ &\boldsymbol{U}_m^{\mathrm{T}}(\tilde{k}) \boldsymbol{D}^{-1} \sum_{i=m+k^*}^{m+\tilde{k}} (\boldsymbol{\mu}_* - \boldsymbol{\mu}_0) + \boldsymbol{U}_m^{\mathrm{T}}(\tilde{k}) \boldsymbol{D}^{-1} \boldsymbol{U}_m(\tilde{k}). \end{aligned} \tag{5.11}$$

Moreover, we have

$$\begin{aligned} &\boldsymbol{U}_m^{\mathrm{T}}(\tilde{k}) \boldsymbol{D}^{-1} \sum_{i=m+k^*}^{m+\tilde{k}} (\boldsymbol{\mu}_* - \boldsymbol{\mu}_0)/g_1^2(m,k) \\ &= \sum_{i=m+k^*}^{m+\tilde{k}} (\boldsymbol{\mu}_* - \boldsymbol{\mu}_0)^{\mathrm{T}} \boldsymbol{D}^{-1} \boldsymbol{U}_m(\tilde{k})/g_1^2(m,k) \\ &= (\boldsymbol{\mu}_* - \boldsymbol{\mu}_0)^{\mathrm{T}} \boldsymbol{D}^{-1} \sum_{i=m+k^*}^{m+\tilde{k}} \boldsymbol{U}_m(\tilde{k})/g_1^2(m,k) = o_p(1), \end{aligned} \tag{5.12}$$

and

$$\sum_{i=m+k^*}^{m+\tilde{k}} (\boldsymbol{\mu}_* - \boldsymbol{\mu}_0)^{\mathrm{T}} \boldsymbol{D}^{-1} \sum_{i=m+k^*}^{m+\tilde{k}} (\boldsymbol{\mu}_* - \boldsymbol{\mu}_0)/g_1^2(m,k) = O_p(m). \tag{5.13}$$

Using the previous proof, we can obtain

$$\boldsymbol{U}_m^{\mathrm{T}}(\tilde{k}) \boldsymbol{D}^{-1} \boldsymbol{U}_m(\tilde{k})/g_1^2(m,k) = O_p(1). \tag{5.14}$$

Putting (5.12)–(5.14), we have $\liminf_{m\to\infty} T(\tilde{k})/g_1^2(m,k) > 0$, i.e.,

$$\sup_{1\leqslant k\leqslant\infty} T(k)/g_1^2(m,k) \xrightarrow{p} \infty,$$

as $m \to \infty$, and the proof of Theorem 3.2 is completed. □

In order to prove Theorems 3.3 and 3.4, let

$$\begin{aligned} \tilde{\boldsymbol{U}}_m(k) &= \sum_{m<i\leqslant m+k} \boldsymbol{e}_i - \frac{k}{i-1} \sum_{1\leqslant j\leqslant i-1} \boldsymbol{e}_j, \\ \boldsymbol{W}(k) &= \boldsymbol{W}_{1,m}(k) - \sum_{i=m+1}^{m+k} \frac{i}{m} \boldsymbol{W}_{1,m}(i-m) - \log(1+\frac{k}{m}) \boldsymbol{W}_{2,m}(m), \end{aligned}$$

and denote that

$$\tilde{\boldsymbol{U}}_m(k) = (\tilde{U}_{m1}(k), \ldots, \tilde{U}_{md}(k)), \quad \boldsymbol{W}(k) = (W_1(k), \ldots, W_d(k)),$$

where $\tilde{U}_{ml}(k)=\sum_{i=m+1}^{m+k}e_{il}-\frac{k}{i-1}\sum_{1\leqslant j\leqslant i-1}e_{jl}$, $W_l(k)=W_{1,ml}(k)-\sum_{i=m+1}^{m+k}\frac{i}{m}W_{1,ml}(i-m)-\log(1+\frac{k}{m})W_{2,ml}(m)$, with $1\leqslant l\leqslant d$.

Lemma 5.2 *If the conditions of Theorem 3.3 are satisfied, then there exist two independent Wiener processes $\{\boldsymbol{W}_{1,m}(t),0\leqslant t<\infty\}$ and $\{\boldsymbol{W}_{2,m}(t),0\leqslant t<\infty\}$ such that*

$$\sup_{1\leqslant m<\infty}|\tilde{\boldsymbol{U}}_m^{\mathrm{T}}(k)\boldsymbol{D}^{-1}\tilde{\boldsymbol{U}}_m(k)-\boldsymbol{W}(k)^{\mathrm{T}}\boldsymbol{W}(k)|/(mh^2(k/m))=o_p(1).$$

Proof Let d_{jj} be the j^{th} row of the j^{th} column of $\boldsymbol{D}$ for $1\leqslant l\leqslant d$. By Lemma 6.4 in Horváth et al. [12], we have

$$\sup_{1\leqslant k<\infty}|U_{ml}(k)-d_{jj}W_l(k)|/mh(k/m)=o_p(1). \tag{5.15}$$

Hence, we can prove

$$\sup_{1\leqslant k<\infty}|\tilde{\boldsymbol{U}}_m^{\mathrm{T}}(k)\boldsymbol{D}^{-1}\tilde{\boldsymbol{U}}_m(k)-\boldsymbol{W}_m^{\mathrm{T}}(k)\boldsymbol{W}_m(k)|/mh^2(k/m)=o_p(1). \tag{5.16}$$

So the proof of Lemma 5.2 is completed. □

Lemma 5.3 *If the conditions of Theorem 3.3 hold, then*

$$\sup_{1\leqslant k<\infty}\boldsymbol{W}_m^{\mathrm{T}}(k)\boldsymbol{W}_m(k)/mh^2(k/m)\xrightarrow{\mathcal{D}}\sup_{0<t<\infty}\frac{\Gamma(t)^2}{h(t)^2}.$$

Proof By Lemma 6.5 in Horváth et al. [12]. For any $1\leqslant l\leqslant d$, we have

$$\sup_{1\leqslant k<\infty}W_l(k)\xrightarrow{\mathcal{D}}\sup_{0<t<\infty}\frac{\Gamma(t)}{h(t)},$$

so we can deduce

$$\begin{aligned}\sup_{1\leqslant k<\infty}\boldsymbol{W}_m^{\mathrm{T}}(k)\boldsymbol{W}_m(k)&=\sup_{1\leqslant k<\infty}\sum_{1\leqslant l\leqslant d}W_l(k)^2\leqslant\sum_{1\leqslant l\leqslant d}\sup_{1\leqslant k<\infty}W_l(k)^2\\&\xrightarrow{\mathcal{D}}\sum_{1\leqslant l\leqslant d}\sup_{0<t<\infty}\frac{\Gamma(t)^2}{h(t)^2}=\sup_{0<t<\infty}\sum_{1\leqslant l\leqslant d}\frac{\Gamma(t)^2}{h(t)^2}.\end{aligned}$$

The proof of Lemma 5.3 is completed. □

Proof of Theorem 3.3 Computing the expectation and covariance functions, we can obtain

$$E\Gamma(t)=0\ \text{ and }\ E\Gamma(t)\Gamma(s)=\min(t,s),$$

and

$$\{\Gamma(t),0\leqslant t<\infty\}\overset{\mathcal{D}}{=}\{W(t),0\leqslant t<\infty\},$$

where $\{W(t),0\leqslant t<\infty\}$ is a Wiener process. Then Theorem 3.3 follows from Lemmas 5.2 and 5.3. □

Proof of Theorem 3.4 Let $\tilde{k}=k^*+m$. By the alternative hypothesis we have

$$T(\tilde{k})=\sum_{m<i\leqslant m+\tilde{k}}(\boldsymbol{X}_i-\bar{\boldsymbol{X}}_{i-1})^{\mathrm{T}}\boldsymbol{D}^{-1}\sum_{m<i\leqslant m+\tilde{k}}(\boldsymbol{X}_i-\bar{\boldsymbol{X}}_{i-1})$$

$$=\Big(\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)+\sum_{i=m+1}^{m+\tilde{k}}\Big(\boldsymbol{e}_i-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{e}_j\Big)\Big)^{\mathrm{T}}\boldsymbol{D}^{-1}$$

$$\Big(\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)+\sum_{i=m+1}^{m+\tilde{k}}\Big(\boldsymbol{e}_i-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{e}_j\Big)\Big). \tag{5.17}$$

Let

$$\tilde{\boldsymbol{V}}_m(k)=\sum_{i=m+1}^{m+k}(\boldsymbol{X}_i-\overline{\boldsymbol{X}}_{i-1}). \tag{5.18}$$

Then

$$\tilde{\boldsymbol{V}}_m(\tilde{k})=\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)+\tilde{\boldsymbol{U}}_m(\tilde{k}). \tag{5.19}$$

Putting (5.17)–(5.19) together, we get

$$\begin{aligned}T(\tilde{k})=&\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)^{\mathrm{T}}\boldsymbol{D}^{-1}\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)+\\&\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)^{\mathrm{T}}\boldsymbol{D}^{-1}\tilde{\boldsymbol{U}}_m(\tilde{k})+\\&\tilde{\boldsymbol{U}}_m^{\mathrm{T}}(\tilde{k})\boldsymbol{D}^{-1}\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)+\tilde{\boldsymbol{U}}_m^{\mathrm{T}}(\tilde{k})\boldsymbol{D}^{-1}\tilde{\boldsymbol{U}}_m(\tilde{k}).\end{aligned} \tag{5.20}$$

Moreover, we have

$$\begin{aligned}&\tilde{\boldsymbol{U}}_m^{\mathrm{T}}(\tilde{k})\boldsymbol{D}^{-1}\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)/mh^2(k/m)\\&=\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)^{\mathrm{T}}\boldsymbol{D}^{-1}\tilde{\boldsymbol{U}}_m(\tilde{k})/mh^2(k/m)\\&=o_p(1),\end{aligned} \tag{5.21}$$

and

$$\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)^{\mathrm{T}}\boldsymbol{D}^{-1}\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)=O_p(m). \tag{5.22}$$

From the previous derivation $\tilde{\boldsymbol{U}}_m^{\mathrm{T}}(\tilde{k})\boldsymbol{D}^{-1}\tilde{\boldsymbol{U}}(\tilde{k})/mh^2(k/m)=O_p(1)$, that is to say

$$\tilde{\boldsymbol{V}}_m(\tilde{k})/mh^2(k/m)=\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)+O_p(g_1(m,k)),$$

however,

$$\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)^{\mathrm{T}}\boldsymbol{D}^{-1}\sum_{i=m+k^*}^{m+\tilde{k}}\Big(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j\Big)>0,$$

so we get

$$\liminf_{m\to\infty}\frac{\sum_{i=m+k^*}^{m+\tilde{k}}(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j)^{\mathrm{T}}\boldsymbol{D}^{-1}\sum_{i=m+k^*}^{m+\tilde{k}}(\boldsymbol{\mu}_*-\frac{1}{i-1}\sum_{1\leqslant j\leqslant i-1}\boldsymbol{\mu}_j)}{mh^2(m,k)}>0.$$

The proof of Theorem 3.4 is completed. □

References

[1] E. S. PAGE. *A test for a change in a parameter occurring at an unknown point*. Biometrika, 1955, **42**(3): 523–527.

[2] Jushan BAI. *Least squares estimation of a shift in linear processes*. J. Time Ser. Anal., 1994, **15**(5): 453–472.

[3] Jushan BAI. *Likelihood ratio tests for multiple structural changes*. J. Econometrics, 1999, **91**(2): 299–323.

[4] P. PERRON, Zhongjun QU. *Estimating restricted structural change models*. J. Econometrics, 2006, **134**(2): 373–399.

[5 L. HORVÁTH, P. KOKOSZKA, J. STEINEBACH. *Testing for changes in multivariate dependent observations with an application to temperature changes*. J. Multivariate Anal., 1999, **68**(1): 96–119.

[6] Zhongjun QU, P. PERRON. *Estimating and testing structural changes in multivariate regressions*. Econometrica, 2007, **75**(2): 459–502.

[7] M. BOUTAHAR. *Testing for change in mean of independent multivariate observations with time varying covariance*. J. Probab. Stat. 2012, 1–17.

[8] M. CSÖRGŐ, L. HORVÁTH. *Limit Theorems in Change-point Analysis*. John Wiley & Sons, Ltd., Chichester, 1997.

[9] P. PERRON. *Dealing with Structural Breaks*. New York, Palgrave Macmillan, 2006.

[10] C. S. J. CHU, M. STINCHCOMBE, H. WHITE. *Monitoring structural change*. Econometrica, 1996, **64**(5): 1045–1063.

[11] F. CARSOULE, P. H. FRANSES. *Monitoring structural change in variance, with an application to european nomianl exchange rate volatility*. Econometric, 1999, **6**: 1–33.

[12] L. HORVÁTH, M. HUŠKOVÁ, P. KOKOSZKA, et al. *Monitoring changes in linear models*. J. Statist. Plann. Inference, 2004, **126**(1): 225–251.

[13] A. AUE, L. HORVÁTH, M. HUŠKOVÁ, et al. *Change-point monitoring in linear models*. Econom. J., 2006, **9**(3): 373–403.

[14] E. ANDREOU, E. GHYSELS. *Monitoring disruptions in finacial markets*. J. Econometrics, 2006, **135**(1): 77–124.

[15] A.ZEILEIS, F. LEISCH, C. KLEIBER, et al. *Monitoring structural change in dynamic econometric models*. J. Appl. Econometrics, 2005, **20**(1): 99–121.

[16] P. CORTEZ, A. CERDEIRA, F. ALMEIDA, et al. *Modeling wine preferences by data mining from physicochemical properties*. Decision Support Systems, 2009, **47**(4): 547–533.