Research Article

Testing for Change in Mean of Independent Multivariate Observations with Time Varying Covariance

Mohamed Boutahar

Institute of Mathematics of Luminy, 163 Avenue de Luminy, 13288 Marseille Cedex 9, France

Correspondence should be addressed to Mohamed Boutahar, mohammed.boutahar@univ-amu.fr

We consider a nonparametric CUSUM test for change in the mean of multivariate time series with time varying covariance. We prove that under the null, the test statistic has a Kolmogorov limiting distribution. The asymptotic consistency of the test against a large class of alternatives which contains abrupt, smooth and continuous changes is established. We also perform a simulation study to analyze the size distortion and the power of the proposed test.

1. Introduction

In the statistical literature there is a vast amount of works on testing for change in the mean of univariate time series. Sen and Srivastava [1, 2], Hawkins [3], Worsley [4], and James et al. [5] considered tests for mean shifts of normal i.i.d. sequences. Extension to dependent univariate time series has been studied by many authors, see Tang and MacNeill [6], Antoch et al. [7], Shao and Zhang [8], and the references therein. Since the paper of Srivastava and Worsley [9] there are a few works on testing for change in the mean of multivariate time series. In their paper they considered the likelihood ratio tests for change in the multivariate i.i.d. normal mean. Tests for change in mean with dependent but stationary error terms have been considered by Horváth et al. [10]. In a more general context of regression, Qu and Perron [11] considered a model where changes in the covariance matrix of the errors occur at the same time as changes in the regression coefficients, and hence the covariance matrix of the errors is a step-function of time. To our knowledge there are no results testing for change in the mean of multivariate models when the covariance matrix of the errors is time varying with

unknown form. The main objective of this paper is to handle this problem. More precisely we consider the d-dimensional model

$$Y_t = \mu_t + \Gamma_t \varepsilon_t, \quad t = 1, \ldots, n, \tag{1.1}$$

where (ε_t) is an i.i.d. sequence of random vectors (not necessary normal) with zero mean and covariance I_d, the identity matrix. The sequence of matrices (Γ_t) is deterministic with unknown form. The null and the alternative hypotheses are as follows:

$$H_0 : \mu_t = \mu \quad \forall t \geq 1$$

against   (1.2)

$$H_1 : \text{There exist } t \neq s \text{ such that } \mu_t \neq \mu_s.$$

In practice, some particular models of (1.1) have been considered in many areas. For instance, in the univariate case ($d = 1$), Starica and Granger [12] show that an appropriate model for the logarithm of the absolute returns of the S&P500 index is given by (1.1) where μ_t and Σ_t are step functions, that is,

$$\mu_t = \mu_{(j)} \quad \text{if } t = n_{j-1} + 1, \ldots, n_j, n_j = [\lambda_j n], \quad 0 < \lambda_1 < \cdots < \lambda_{m_1} < 1,$$
$$\sum\nolimits_t = \sigma_{(j)} \quad \text{if } t = t_{j-1} + 1, \ldots, t_j, t_j = [\tau_j n], \quad 0 < \tau_1 < \cdots < \tau_{m_2} < 1, \tag{1.3}$$

for some integers m_1 and m_2. They also show that model (1.1) and (1.3) gives forecasts superior to those based on a stationary GARCH(1,1) model. In the multivariate case ($d > 1$), Horváth et al. [10] considered the model (1.1) where μ_t is subject to change and $\Sigma_t = \Sigma$ is constant; they applied such model to temperature data to provide evidence for the global warming theory. For financial data, it is well known that assets' returns have a time varying covariance. Therefore, for example, in portfolio management, our test can be used to indicate if the mean of one or more assets returns are subject to change. If so, then taking into account such a change is very useful in computing the portfolio risk measures such as the value at Risk (VaR) or the expected shortfall (ES) (see Artzner et al. [13] and Holton [14] for more details).

2. The Test Statistic and the Assumptions

In order to construct the test statistic let

$$B_n(\tau) = \widehat{\Gamma}^{-1} \frac{1}{\sqrt{n}} \sum_{t=1}^{[n\tau]} (Y_t - \overline{Y}) \quad \forall \tau \in [0, 1], \tag{2.1}$$

where $\widehat{\Gamma}$ is a square root of $\widehat{\Sigma}$, that is, $\widehat{\Sigma} = \widehat{\Gamma}\widehat{\Gamma}'$,

$$\widehat{\Sigma} = \frac{1}{n} \sum_{t=1}^{n} (Y_t - \overline{Y})(Y_t - \overline{Y})', \quad \overline{Y} = \frac{1}{n} \sum_{t=1}^{n} Y_t \tag{2.2}$$

are the empirical covariance and mean of the sample $(Y_1, \ldots, Y_n)'$, respectively, $[x]$ is the integer part of x, and X' is the transpose of X.

The CUSUM test statistic we will consider is given by

$$\mathcal{B}_n = \sup_{\tau \in [0,1]} \|B_n(\tau)\|_\infty, \tag{2.3}$$

where

$$\|X\|_\infty = \max_{1 \le i \le d} \left| X^{(i)} \right| \quad \text{if } X = \left(X^{(1)}, \ldots, X^{(d)} \right)'. \tag{2.4}$$

Assumption 1. The sequence of matrices (Γ_t) is bounded and satisfies

$$\frac{1}{n} \sum_{t=1}^{n} \Gamma_t \Gamma_t' \longrightarrow \Sigma > 0 \quad \text{as } n \longrightarrow \infty. \tag{2.5}$$

Assumption 2. There exists $\delta > 0$ such that $E(\|\varepsilon_1\|^{2+\delta}) < \infty$, where $\|X\|$ denotes the Euclidian norm of X.

3. Limiting Distribution of $\mathcal{B}_n$ under the Null

Theorem 3.1. *Suppose that Assumptions 1 and 2 hold. Then, under H_0,*

$$\mathcal{B}_n \xrightarrow{\mathcal{L}} \mathcal{B}_\infty = \sup_{\tau \in [0,1]} \|B(\tau)\|_\infty, \tag{3.1}$$

$\xrightarrow{\mathcal{L}}$ *denotes the convergence in distribution and $B(\tau)$ is a multivariate Brownian Bridge with independent components.*

Moreover, the cumulative distribution function of B_∞ is given by

$$F_{B_\infty}(z) = \left(1 + 2 \sum_{k=1}^{\infty} (-1)^k \exp\left\{ -2k^2 z^2 \right\} \right)^d. \tag{3.2}$$

To prove Theorem 3.1 we will establish first a functional central limit theorem for random sequences with time varying covariance. Such a theorem is of independent interest. Let $D = D[0,1]$ be the space of random functions that are right-continuous and have left limits, endowed with the Skorohod topology. For a given $d \in \mathbb{N}$, let $D^d = D^d[0,1]$ be the product space. The weak convergence of a sequence of random elements X_n in D^d to a random element X in D^d will be denoted by $X_n \Rightarrow X$.

For two random vectors X and Y, $X \stackrel{\text{law}}{=} Y$ means that X has the same distribution as Y.

Consider an i.i.d. sequence (ε_t) of random vectors such that $E(\varepsilon_t) = 0$ and $\mathrm{var}(\varepsilon_t) = I_d$. Let (Γ_t) satisfy (2.5) and set

$$W_n(\tau) = \frac{\Gamma^{-1}}{\sqrt{n}} \sum_{t=1}^{[n\tau]} \Gamma_t \varepsilon_t, \quad \tau \in [0,1], \tag{3.3}$$

where Γ is a square root of Σ, that is, $\Sigma = \Gamma\Gamma'$. Many functional central limit theorems were established for covariance stationary random sequences, see Boutahar [15] and the references therein. Note that the sequence $(\Gamma_t\varepsilon_t)$ we consider here is not covariance stationary.

There are two sufficient conditions to prove that $W_n \Rightarrow W$ (see Billingsley [16] and Iglehart [17]), namely,

(i) the finite-dimensional distributions of W_n converge to the finite-dimensional distributions of W,

(ii) $W_n^{(i)}$ is tight for all $1 \le i \le d$, if $W_n = (W_n^{(1)}, \ldots, W_n^{(d)})'$.

Theorem 3.2. *Assume that (ε_t) is an i.i.d. sequence of random vectors such that $E(\varepsilon_t) = 0$, $\mathrm{var}(\varepsilon_t) = I_d$ and that Assumptions 1 and 2 hold. Then*

$$W_n \implies W, \tag{3.4}$$

where W is a standard multivariate Brownian motion.

Proof. Write $F_t = \Gamma^{-1}\Gamma_t$, $F_t(i,j)$ the (i,j)-th entry of the matrix F_t, $\varepsilon_t = (\varepsilon_t^{(1)}, \ldots, \varepsilon_t^{(d)})'$. To prove that the finite-dimensional distributions of W_n converge to those of W it is sufficient to show that for all integer $r \ge 1$, for all $0 \le \tau_1 < \cdots < \tau_r \le 1$, and for all $\alpha_i \in \mathbb{R}^d$, $1 \le i \le r$,

$$Z_n = \sum_{i=1}^{r} \alpha_i' W_n(\tau_i) \xrightarrow{\mathcal{D}} Z = \sum_{i=1}^{r} \alpha_i' W(\tau_i). \tag{3.5}$$

Denote by $\Phi_{Z_n}(u) = E(\exp(iuZ_n))$ the characteristic function of Z_n and by C a generic positive constant, not necessarily the same at each occurrence. We have

$$\begin{aligned}
\Phi_{Z_n}(u) &= E\left[\exp\left(\frac{iu}{\sqrt{n}} \sum_{k=1}^{r} \alpha_k' \sum_{t=1}^{[n\tau_k]} F_t \varepsilon_t\right)\right] \\
&= E\left[\exp\left(\frac{iu}{\sqrt{n}} \sum_{k=1}^{r} \left(\sum_{j=k}^{r} \alpha_j'\right) \sum_{t=[n\tau_{k-1}]+1}^{[n\tau_k]} F_t \varepsilon_t\right)\right], \quad \tau_0 = 0 \\
&= \prod_{k=1}^{r} \Phi_{k,n}(u),
\end{aligned} \tag{3.6}$$

where

$$\Phi_{k,n}(u) = E\left[\exp\left(\frac{iu}{\sqrt{n}}\left(\sum_{j=k}^{r}\alpha'_j\right)\sum_{t=[n\tau_{k-1}]+1}^{[n\tau_k]}F_t\varepsilon_t\right)\right]. \tag{3.7}$$

Since (ε_t) is an i.i.d. sequence of random vectors we have

$$\left(\varepsilon_{[n\tau_{k-1}]+1},\ldots,\varepsilon_{[n\tau_k]}\right)\overset{\text{law}}{=}\left(\varepsilon_1,\ldots,\varepsilon_{[n\tau_k]-[n\tau_{k-1}]}\right). \tag{3.8}$$

Hence

$$\Phi_{k,n}(u) = E\left[\exp\left(\frac{iu}{\sqrt{n}}\left(\sum_{j=k}^{r}\alpha'_j\right)\sum_{t=1}^{[n\tau_k]-[n\tau_{k-1}]}F_{[n\tau_{k-1}]+t}\varepsilon_t\right)\right]. \tag{3.9}$$

Let $\mathbf{I}(A) = 1$ if the argument A is true and 0 otherwise, $k_n = [n\tau_k] - [n\tau_{k-1}]$,

$$\xi_{n,i} = \frac{1}{\sqrt{n}}\left(\sum_{j=k}^{r}\alpha'_j\right)F_{[n\tau_{k-1}]+i}\varepsilon_i, \qquad M_{n,k_n} = \sum_{i=1}^{k_n}\xi_{n,i}, \tag{3.10}$$

$\mathcal{F}_{n,t} = \sigma(\varepsilon_1,\ldots,\varepsilon_t, t \le k_n)$ the filtration spanned by $\varepsilon_1,\ldots,\varepsilon_t$.

Then $(M_{n,i}, \mathcal{F}_{n,i}, 1 \le i \le k_n, n \ge 1)$ is a zero-mean square-integrable martingale array with differences $\xi_{n,i}$. Observe that

$$\sum_{i=1}^{k_n}E\left(\xi_{n,i}^2 \mid \mathcal{F}_{n,i-1}\right) = \frac{1}{n}\left(\sum_{j=k}^{r}\alpha'_j\right)\sum_{i=1}^{k_n}F_{[n\tau_{k-1}]+i}F'_{[n\tau_{k-1}]+i}\left(\sum_{j=k}^{r}\alpha_j\right)$$

$$\longrightarrow \sigma_k^2 = (\tau_k - \tau_{k-1})\left\|\sum_{j=k}^{r}\alpha_j\right\|^2 \quad \text{as } n \longrightarrow \infty. \tag{3.11}$$

Now using Assumption 1 we obtain that $\|\Gamma_t\| < K$ uniformly on t for some positive constant K, hence Assumption 2 implies that for all $\varepsilon > 0$,

$$\sum_{i=1}^{k_n}E\left(\xi_{n,i}^2\mathbf{I}(|\xi_{n,i}| > \varepsilon) \mid \mathcal{F}_{n,i-1}\right) \le \frac{1}{\varepsilon^\delta}\sum_{i=1}^{k_n}E\left(|\xi_{n,i}|^{2+\delta} \mid \mathcal{F}_{n,i-1}\right)$$

$$\le \frac{Ck_n}{n^{1+\delta/2}} \tag{3.12}$$

$$\longrightarrow 0 \quad \text{as } n \longrightarrow \infty,$$

where

$$C = \frac{E\left(\|\varepsilon_1\|^{2+\delta}\right)}{\varepsilon^\delta}\left(K\|\Gamma^{-1}\|\left\|\sum_{j=k}^{r}\alpha_j\right\|\right)^{2+\delta},$$

(3.13)

consequently (see Hall and Heyde [18], Theorem 3.2)

$$M_{n,k_n} \xrightarrow{\mathscr{Q}} Z_k,$$

(3.14)

where Z_k is a normal random variable with zero mean and variance σ_k^2. Therefore

$$\Phi_{k,n}(u) = E(\exp(iuM_{n,k_n})) \longrightarrow \exp\left(-\frac{1}{2}\sigma_k^2 u^2\right) \quad \text{as } n \longrightarrow \infty,$$

(3.15)

which together with (3.6) implies that

$$\Phi_{Z_n}(u) \longrightarrow \exp\left(-\frac{1}{2}\sum_{k=1}^{r}\sigma_k^2 u^2\right) = \Phi_Z(u) \quad \text{as } n \longrightarrow \infty,$$

(3.16)

the last equality holds since, with $\tau_0 = 0$,

$$\sum_{k=1}^{r}(\tau_k - \tau_{k-1})\left\|\sum_{j=k}^{r}\alpha_j\right\|^2 = \sum_{1\le i,\,j\le r}\alpha_i'\alpha_j\min(\tau_i,\tau_j).$$

(3.17)

For $1 \le i \le d$, fixed, in order to obtain the tightness of $W_n^{(i)}$ it suffices to show the following inequality (Billingsley [16], Theorem 15.6):

$$E\left(\left|W_n^{(i)}(\tau) - W_n^{(i)}(\tau_1)\right|^\gamma \left|W_n^{(i)}(\tau_2) - W_n^{(i)}(\tau)\right|^\gamma\right) \le (F(\tau_2) - F(\tau_1))^\alpha,$$

(3.18)

for some $\gamma > 0$, $\alpha > 1$, where F is a nondecreasing continuous function on $[0,1]$ and $0 < \tau_1 < \tau < \tau_2 < 1$.

We have

$$E\left(\left|W_n^{(i)}(\tau) - W_n^{(i)}(\tau_1)\right|^2 \left|W_n^{(i)}(\tau_2) - W_n^{(i)}(\tau)\right|^2\right) = T_1 T_2,$$

(3.19)

where

$$T_1 = \frac{1}{n} E\left(\left|\sum_{t=[n\tau_1]+1}^{[n\tau]}\sum_{j=1}^{d}F_t(i,j)\varepsilon_t^{(j)}\right|^2\right),$$

$$T_2 = \frac{1}{n} E\left(\left|\sum_{t=[n\tau]+1}^{[n\tau_2]}\sum_{j=1}^{d}F_t(i,j)\varepsilon_t^{(j)}\right|^2\right).$$

(3.20)

Now observe that

$$T_1 = \frac{1}{n}\sum_{t,s}\mathrm{cov}\left(\sum_{j=1}^{d}F_t(i,j)\varepsilon_t^{(j)},\sum_{j=1}^{d}F_s(i,j)\varepsilon_s^{(j)}\right)$$

$$= \frac{1}{n}\sum_{t=[n\tau_1]+1}^{[n\tau]}\sum_{j=1}^{d}(F_t(i,j))^2$$

(3.21)

$$\leq C(\tau - \tau_1) \quad \text{for some constant } C > 0.$$

Likewise $T_2 \leq C(\tau_2 - \tau)$. Since $(\tau - \tau_1)(\tau_2 - \tau) \leq (\tau_2 - \tau_1)^2/2$, the inequality (3.18) holds with $\gamma = \alpha = 2$, $F(t) = Ct/\sqrt{2}$. $\qquad\square$

In order to prove Theorem 3.1 we need also the following lemma.

Lemma 3.3. *Assume that (Y_t) is given by (1.1), where (ε_t) is an i.i.d sequence of random vectors such that $E(\varepsilon_t) = 0$, $\mathrm{var}(''_t) = I_d$ and that (Γ_t) satisfies (2.5). Then under the null H_0, the empirical covariance of Y_t satisfies*

$$\widehat{\Sigma} \overset{a.s.}{\longrightarrow} \Sigma,$$

(3.22)

where $\overset{a.s.}{\longrightarrow}$ denotes the almost sure convergence.

Proof. Let $W_t = \Gamma_t \varepsilon_t$, $\mathcal{F}_t = \sigma(\varepsilon_1, \ldots, \varepsilon_t)$ and for i, j fixed, $1 \leq i \leq d$, $1 \leq j \leq d$, $e_t = W_t^{(i)}W_t^{(j)} - E(W_t^{(i)}W_t^{(j)} \mid \mathcal{F}_{t-1})$.

Then (e_t) is a martingale difference sequence with respect to $\mathcal{F}_t$. Since $e_t = W_t^{(i)}W_t^{(j)} - \sum_{k=1}^{d}\Gamma_t(i,k)\Gamma_t(j,k)$ and the matrix Γ_t is bounded, it follows that

$$E\left(|e_t|^{(2+\delta)/2}\right) \leq C + E\left(\left|W_t^{(i)}W_t^{(j)}\right|^{(2+\delta)/2}\right)$$

$$\leq E\left(\left|W_t^{(i)}\right|^{2+\delta}\right)^{1/2}E\left(\left|W_t^{(j)}\right|^{2+\delta}\right)^{1/2}$$

(3.23)

$$\leq C,$$

since by using Assumptions 1 and 2 we get

$$
E\left(\left|W_t^{(i)}\right|^{2+\delta}\right) \leq \left(\sum_{k=1}^{d} E\left(\left|\Gamma_t(i,k)\varepsilon_t^{(k)}\right|^{2+\delta}\right)^{1/(2+\delta)}\right)^{2+\delta}
$$

$$
\leq C.
$$

(3.24)

Therefore, Theorem 5 of Chow [19] implies that

$$
\sum_{t=1}^{n} e_t = o(n) \text{ almost surely}
$$

(3.25)

or

$$
\frac{1}{n}\sum_{t=1}^{n} W_t^{(i)} W_t^{(j)} = \frac{1}{n}\sum_{t=1}^{n} (\Gamma_t \Gamma_t')(i,j) + o(1) \text{ almost surely,}
$$

(3.26)

where $(\Gamma_t \Gamma_t')(i,j)$ denotes the (i,j)-th entry of the matrix $\Gamma_t \Gamma_t'$. Hence

$$
\frac{1}{n}\sum_{t=1}^{n} W_t W_t' \xrightarrow{\text{a.s.}} \sum.
$$

(3.27)

$\square$

Lemma 2 of Lai and Wei [20], page 157, implies that with probability one

$$
\sum_{t=1}^{n} \Gamma_t(i,k)\varepsilon_t^{(k)} = o\left(\sum_{t=1}^{n} (\Gamma_t(i,k))^2\right) + O(1)
$$

$$
= o(n) + O(1) \quad \forall 1 \leq i \leq d,
$$

(3.28)

or

$$
\frac{1}{n}\sum_{t=1}^{n} \Gamma_t(i,k)\varepsilon_t^{(k)} = o(1) + O\left(\frac{1}{n}\right) \text{ almost surely,}
$$

(3.29)

which implies that

$$
\frac{1}{n}\sum_{t=1}^{n} W_t = \sum_{k=1}^{d} \left(\frac{1}{n}\sum_{t=1}^{n} \Gamma_t(1,k)\varepsilon_t^{(k)}, \ldots, \frac{1}{n}\sum_{t=1}^{n} \Gamma_t(d,k)\varepsilon_t^{(k)}\right)' \xrightarrow{\text{a.s.}} 0.
$$

(3.30)

Note that $Y_t = \mu + W_t$, hence combining (3.27) and (3.30) we obtain

$$
\begin{aligned}
\widehat{\Sigma} &= \frac{1}{n}\sum_{t=1}^{n} Y_t Y_t' - \overline{Y}\overline{Y}' \\
&= \mu\frac{1}{n}\sum_{t=1}^{n}W_t + \frac{1}{n}\sum_{t=1}^{n}W_t\mu' + \mu\mu' + \frac{1}{n}\sum_{t=1}^{n}W_t W_t' - \overline{Y}\overline{Y}' \\
&\xrightarrow{\text{a.s.}} \Sigma.
\end{aligned}
\tag{3.31}
$$

Proof of Theorem 3.1. Under the null H_0 we have $Y_t = \mu + \Gamma_t \varepsilon_t$, thus recalling (3.3) we can write

$$
\begin{aligned}
B_n(\tau) &= \widehat{\Gamma}^{-1}\frac{1}{\sqrt{n}}\sum_{t=1}^{[n\tau]}\left(Y_t - \overline{Y}\right) \\
&= \widehat{\Gamma}^{-1}\Gamma\frac{1}{\sqrt{n}}\Gamma^{-1}\sum_{t=1}^{[n\tau]}\left[(Y_t - \mu) - \left(\overline{Y} - \mu\right)\right] \\
&= \widehat{\Gamma}^{-1}\Gamma\left(W_n(\tau) - \frac{[n\tau]}{n}W_n(1)\right).
\end{aligned}
\tag{3.32}
$$

Therefore the result (3.1) holds by applying Theorem 3.2, Lemma 3.3, and the continuous mapping theorem. $\qquad\square$

4. Consistency of $\mathcal{B}_n$

We assume that under the alternative H_1 the means (μ_t) are bounded and satisfy the following.

Assumption H1. There exists a function U from $[0,1]$ into $\mathbb{R}^d$ such that

$$
\forall \tau \in [0,1], \quad \frac{1}{n}\sum_{t=1}^{[n\tau]}\mu_t \longrightarrow U(\tau) \quad \text{as } n \longrightarrow \infty.
\tag{4.1}
$$

Assumption H2. There exists $\tau^* \in (0,1)$ such that

$$
\mathcal{U}(\tau^*) = U(\tau^*) - \tau^* U(1) \neq 0.
\tag{4.2}
$$

Assumption H3. There exists $\sum_\mu$ such that

$$
\frac{1}{n}\sum_{t=1}^{n}(\mu_t - \overline{\mu})(\mu_t - \overline{\mu})' \longrightarrow \sum_\mu \quad \text{as } n \longrightarrow \infty,
\tag{4.3}
$$

where $\overline{\mu} = (1/n)\sum_{t=1}^{n}\mu_t$.

Theorem 4.1. *Suppose that Assumptions 1 and 2 hold. If (Y_t) is given by (1.1) and the means (μ_t) satisfy the Assumptions H1, H2, and H3, then the test based on $\mathcal{B}_n$ is consistent against H_1, that is,*

$$\mathcal{B}_n \xrightarrow{P} +\infty, \tag{4.4}$$

where $\xrightarrow{P}$ denotes the convergence in probability.

Proof. We have

$$B_n(\tau) = B_n^0(\tau) + B_n^1(\tau), \tag{4.5}$$

where

$$B_n^0(\tau) = \frac{\widehat{\Gamma}^{-1}}{\sqrt{n}} \sum_{t=1}^{[n\tau]} \left(W_t - \overline{W} \right), \quad W_t = \Gamma_t \varepsilon_t, \ \overline{W} = \frac{1}{n} \sum_{t=1}^{n} W_t,$$

$$B_n^1(\tau) = \frac{\widehat{\Gamma}^{-1}}{\sqrt{n}} \sum_{t=1}^{[n\tau]} \left(\mu_t - \overline{\mu} \right). \tag{4.6}$$

Straightforward computation leads to

$$\widehat{\sum} \xrightarrow{\text{a.s.}} \sum_* = \sum + \sum_\mu. \tag{4.7}$$

Therefore

$$B_n^0(\tau) \xrightarrow{\mathscr{L}} \Gamma_*^{-1} \Gamma B(\tau), \tag{4.8}$$

where Γ_* is a square root of Σ_*, that is, $\Sigma_* = \Gamma_* \Gamma_*'$, and

$$\frac{B_n^1(\tau)}{\sqrt{n}} \xrightarrow{\text{a.s.}} \Gamma_*^{-1} \mathcal{U}(\tau). \tag{4.9}$$

Hence

$$\|B_n(\tau^*)\|_\infty \xrightarrow{P} +\infty, \tag{4.10}$$

which implies that

$$\mathcal{B}_n \xrightarrow{P} +\infty. \tag{4.11}$$

$\square$

4.1. Consistency of $\mathcal{B}_n$ against Abrupt Change

Without loss of generality we assume that under the alternative hypothesis H_1 there is a single break date, that is, (Y_t) is given by (1.1) where

$$
\mu_t = \begin{cases} \mu_{(1)} & \text{if } 1 \leq t \leq [n\tau_1] \\ \mu_{(2)} & \text{if } [n\tau_1] + 1 \leq t \leq n \end{cases} \quad \text{for some } \tau_1 \in (0,1) \text{ and } \mu_{(1)} \neq \mu_{(2)}. \tag{4.12}
$$

Corollary 4.2. *Suppose that Assumptions 1 and 2 hold. If (Y_t) is given by (1.1) and the means (μ_t) satisfy (4.12), then the test based on $\mathcal{B}_n$ is consistent against H_1.*

Proof. It is easy to show that (4.1)–(4.3) are satisfied with

$$
\mathcal{U}(\tau) = \begin{cases} \tau(1 - \tau_1)(\mu_{(1)} - \mu_{(2)}) & \text{if } \tau \leq \tau_1 \\ \tau_1(1 - \tau_1)(\mu_{(1)} - \mu_{(2)}) & \text{if } \tau > \tau_1, \end{cases}
$$

$$
\sum_{\mu} = \tau_1(\mu_{(1)} - \mu_{(2)})(\mu_{(1)} - \mu_{(2)})'. \tag{4.13}
$$

Note that (4.2) is satisfied for all $0 < \tau^* < \tau_1$ since $\mu_{(1)} \neq \mu_{(2)}$. □

Remark 4.3. The result of Corollary 4.2 remains valid if under the alternative hypothesis there are multiple breaks in the mean.

4.2. Consistency of $\mathcal{B}_n$ against Smooth Change

In this subsection we assume that the break in the mean does not happen suddenly but the transition from one value to another is continuous with slow variation. A well-known dynamic is the smooth threshold model (see Teräsvirta [21]), in which the mean μ_t is time varying as follows

$$
\mu_t = \mu_{(1)} + (\mu_{(2)} - \mu_{(1)})F\left(\frac{t}{n}, \tau_1, \gamma\right), \quad 1 \leq t \leq n, \ \mu_{(1)} \neq \mu_{(2)}, \tag{4.14}
$$

where $F(x, \tau_1, \gamma)$ is a the smooth transition function assumed to be continuous from $[0,1]$ into $[0,1]$, $\mu_{(1)}$ and $\mu_{(2)}$ are the values of the mean in the two extreme regimes, that is, when $F \to 0$ and $F \to 1$. The slope parameter γ indicates how rapid the transition between two extreme regimes is. The parameter τ_1 is the location parameter.

Two choices for the function F are frequently evoked, the logistic function given by

$$
F_L(x, \tau_1, \gamma) = [1 + \exp(-\gamma(x - \tau_1))]^{-1}, \tag{4.15}
$$

and the exponential one

$$
F_e(x, \tau_1, \gamma) = 1 - \exp\left(-\gamma(x - \tau_1)^2\right). \tag{4.16}
$$

For example, for the logistic function with $\gamma > 0$, the extreme regimes are obtained as follows:

(i) if $x \to 0$ and γ large then $F \to 0$ and thus $\mu_t = \mu_{(1)}$,

(ii) if $x \to 1$ and γ large then $F \to 1$ and thus $\mu_t = \mu_{(2)}$.

This means that at the beginning of the sample μ_t is close to $\mu_{(1)}$ and then moves towards $\mu_{(2)}$ and becomes close to it at the end of the sample.

Corollary 4.4. *Suppose that Assumptions 1 and 2 hold. If (Y_t) is given by (1.1) and the means (μ_t) satisfy (4.14), then the test based on $\mathcal{B}_n$ is consistent against H_1.*

Proof. The assumptions (4.1) and (4.3) are satisfied with

$$\mathcal{U}(\tau) = \left(\mu_{(2)} - \mu_{(1)}\right)T(\tau), \tag{4.17}$$

where

$$T(\tau) = \int_0^\tau F(x, \tau_1, \gamma)dx - \tau \int_0^1 F(x, \tau_1, \gamma)dx,$$

$$\sum_{\mu} = \left(\mu_{(2)} - \mu_{(1)}\right)\left(\mu_{(2)} - \mu_{(1)}\right)'\left\{\int_0^1 F^2(x, \tau_1, \gamma)dx - \left(\int_0^1 F(x, \tau_1, \gamma)dx\right)^2\right\}. \tag{4.18}$$

Since $\mu_{(2)} - \mu_{(1)} \neq 0$, to prove (4.2), it suffices to show that there exists τ^* such $T(\tau^*) \neq 0$. Assume that $T(\tau) = 0$ for all $\tau \in (0, 1)$ then

$$\frac{dT(\tau)}{d\tau} = F(\tau, \tau_1, \gamma) - \int_0^1 F(x, \tau_1, \gamma)dx = 0 \quad \forall \tau \in (0, 1), \tag{4.19}$$

which implies that $F(\tau, \tau_1, \gamma) = \int_0^1 F(x, \tau_1, \gamma)dx = C$ for all $\tau \in (0, 1)$ or

$$\mu_t = \mu_{(1)} + \left(\mu_{(2)} - \mu_{(1)}\right)C = \mu \quad \forall t \geq 1, \tag{4.20}$$

and this contradicts the alternative hypothesis H_1. $\qquad\square$

4.3. Consistency of $\mathcal{B}_n$ against Continuous Change

In this subsection we will examine the behaviour $\mathcal{B}_n$ under the alternative where the mean (μ_t) varies at each time, and hence can take an infinite number of values. As an example we consider a polynomial evolution for μ_t:

$$\mu_t = \left(P_1\left(\frac{t}{n}\right), \ldots, P_d\left(\frac{t}{n}\right)\right)', \qquad P_j(x) = \sum_{k=0}^{p_j} \alpha_{j,k}x^k, \quad 1 \leq j \leq d. \tag{4.21}$$

Corollary 4.5. *Suppose that Assumptions 1 and 2 hold. If (Y_t) is given by (1.1) and the means (μ_t) satisfy (4.21), then the test based on $\mathcal{B}_n$ is consistent against H_1.*

Proof. The assumptions H1–H3 are satisfied with

$$
\mathcal{U}(\tau) = \left(\sum_{k=0}^{p_1} \frac{\alpha_{1,k}}{k+1} \left(\tau^{k+1} - \tau \right), \dots, \sum_{k=0}^{p_d} \frac{\alpha_{d,k}}{k+1} \left(\tau^{k+1} - \tau \right) \right)',
$$

$$
\sum_{\mu}(i,j) = \sum_{k=0}^{p_i} \sum_{l=0}^{p_j} \frac{1}{l+k+1} \alpha_{i,k} \alpha_{j,l} - \left(\sum_{k=0}^{p_i} \frac{\alpha_{i,k}}{k+1} \right) \left(\sum_{k=0}^{p_j} \frac{\alpha_{j,k}}{k+1} \right).
$$

(4.22)

Note that (4.2) is satisfied for all $0 < \tau^* < 1$, provided that there exist i, $1 \leq i \leq d$ and k, $1 \leq k \leq p_i$ such that $\alpha_{i,k} \neq 0$. □

5. Finite Sample Performance

All models are driven from an i.i.d. sequences $\varepsilon_t = (\varepsilon_t^{(1)}, \dots, \varepsilon_t^{(d)})'$, where each $\varepsilon_t^{(j)}$, $1 \leq j \leq d$, has a $t(3)$ distribution, a Student distribution with 3 degrees of freedom, and $\varepsilon_t^{(i)}$ and $\varepsilon_t^{(j)}$ are independent for all $i \neq j$. Simulations were performed using the software R. We carry out an experiment of 1000 samples for seven models and we use three different sample sizes, $n = 30$, $n = 100$, and $n = 500$. The empirical sizes and powers are calculated at the nominal levels $\alpha = 1\%$, 5%, and 10%, in both cases.

5.1. Study of the Size

In order to evaluate the size distortion of the test statistic $\mathcal{B}_n$ we consider two bivariate models $Y_t = \mu_t + \Gamma_t \varepsilon_t$ with the following.

Model 1 (constant covariance).

$$
\mu_t = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \Gamma_t = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}.
$$

(5.1)

Model 2 (time varying covariance).

$$
\mu_t = \begin{pmatrix} 1 \\ 1 \end{pmatrix}, \qquad \Gamma_t = \begin{pmatrix} 2\sin(t\omega) & -1 \\ -1 & 2\cos(t\omega) \end{pmatrix}, \qquad \omega = \frac{\pi}{4}.
$$

(5.2)

From Table 1, we observe that for small sample size ($n = 30$) the test statistic $\mathcal{B}_n$ has a severe size distortion. But as the sample size n increases, the distortion decreases. The empirical size becomes closer to (but always lower than) the nominal level. The distortion in the nonstationary Model 2 (time varying covariance) is a somewhat greater than the one in the stationary Model 1 (constant covariance). However the test seems to be conservative in both cases.

Table 1: Empirical sizes (%).

	α	$n = 30$	$n = 100$	$n = 500$
	1%	0.2	0.3	0.4
Model 1	5%	2.1	2.9	4
	10%	4.9	7.3	8.9
	1%	0.0	0.2	0.3
Model 2	5%	1.1	2.7	3.0
	10%	2.9	6.4	7.3

5.2. Study of the Power

In order to see the power of the test statistic $\mathcal{B}_n$ we consider five bivariate models $Y_t = \mu_t + \Gamma_t \varepsilon_t$ with the following.

5.2.1. Abrupt Change in the Mean

Model 3 (constant covariance).

$$\mu_t = \begin{cases} \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \text{if } 1 \leq t \leq \left[\frac{n}{2}\right] \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \text{if } \left[\frac{n}{2}\right] + 1 \leq t \leq n, \end{cases} \qquad \Gamma_t = \begin{pmatrix} 2 & 1 \\ 1 & 2 \end{pmatrix}. \tag{5.3}$$

Model 4. In this model the mean and the covariance are subject to an abrupt change at the same time:

$$\mu_t = \begin{cases} \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \text{if } 1 \leq t \leq \left[\frac{n}{2}\right] \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \text{if } \left[\frac{n}{2}\right] + 1 \leq t \leq n \end{cases} \qquad \Gamma_t = \begin{cases} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} & \text{if } 1 \leq t \leq \left[\frac{n}{2}\right] \\ \begin{pmatrix} 2 & 1 \\ 0 & 2 \end{pmatrix} & \text{if } \left[\frac{n}{2}\right] + 1 \leq t \leq n \end{cases}. \tag{5.4}$$

Model 5. The mean is subject to an abrupt change and the covariance is time varying (see Figure 1):

$$\mu_t = \begin{cases} \begin{pmatrix} 0 \\ 1 \end{pmatrix} & \text{if } 1 \leq t \leq \left[\frac{n}{2}\right] \\ \begin{pmatrix} 1 \\ 0 \end{pmatrix} & \text{if } \left[\frac{n}{2}\right] + 1 \leq t \leq n, \end{cases} \qquad \Gamma_t = \begin{pmatrix} 2\sin(t\omega) & -1 \\ -1 & 2\cos(t\omega) \end{pmatrix}, \quad \omega = \frac{\pi}{4}. \tag{5.5}$$
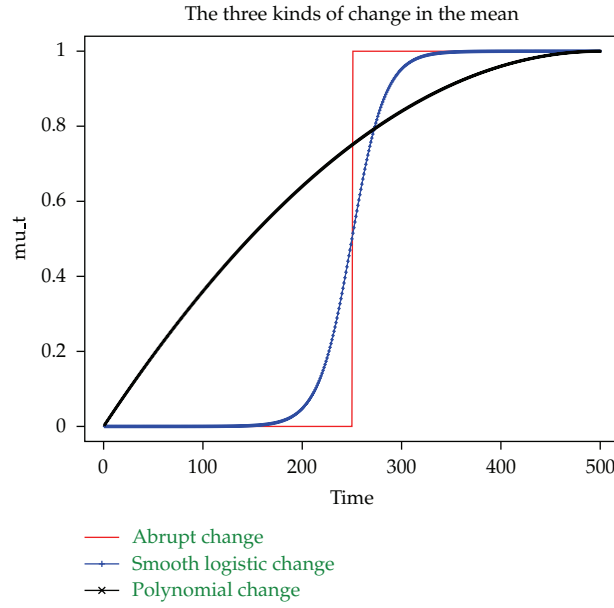
Figure 1: The three kinds of change in the mean.

5.2.2. Smooth Change in the Mean

Model 6. We consider a logistic smooth transition for the mean and a time varying covariance (see Figure 1):

$$\mu_t = \frac{\mu_{(1)} + \left(\mu_{(2)} - \mu_{(1)}\right)}{\left(1 + \exp(-30(t/n - 1/2))\right)}, \quad 1 \le t \le n,$$

$$\mu_{(1)} = \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \quad \mu_{(2)} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \Gamma_t = \begin{pmatrix} 2\sin(t\omega) & -1 \\ -1 & 2\cos(t\omega) \end{pmatrix}, \quad \omega = \frac{\pi}{4}. \tag{5.6}$$

5.2.3. Continuous Change in the Mean

Model 7. In this model the mean is a polynomial of order two and the covariance matrix is also time varying as in the preceding Models 5 and 6 (see Figure 1):

$$\mu_t = \begin{pmatrix} \dfrac{t}{n}\left(2 - \dfrac{t}{n}\right) \\ \dfrac{t}{n}\left(2 - \dfrac{t}{n}\right) \end{pmatrix}, \quad \Gamma_t = \begin{pmatrix} 2\sin(t\omega) & -1 \\ -1 & 2\cos(t\omega) \end{pmatrix}, \quad \omega = \frac{\pi}{4}. \tag{5.7}$$

From Table 2, we observe that for small sample size ($n = 30$), the test statistic $\mathcal{B}_n$ has a low power. However, for the five models, the power becomes good as the sample size n

Table 2: Empirical powers (%).

	α	$n = 30$	$n = 100$	$n = 500$
	1%	11.4	81.8	100
Model 3	5%	34.1	92.1	100
	10%	46.9	95.1	100
	1%	3.9	49.8	99.9
Model 4	5%	15.7	71.4	99.9
	10%	29	80.4	100
	1%	1.4	22.7	95.9
Model 5	5%	8.3	44.7	98.4
	10%	15.9	56.5	99.3
	1%	1.4	17.2	94
Model 6	5%	8.5	38.5	97.7
	10%	16.3	51.9	98.7
	1%	0.1	5.3	44.2
Model 7	5%	2.2	16.1	70.0
	10%	5.9	25.5	79.4

increases. The powers in nonstationary models are always smaller than those of stationary models. This is not surprising since, from Table 1, the test statistic $\mathcal{B}_n$ is more conservative in nonstationary models. We observe also that the power is almost the same in abrupt and logistic smooth changes (compare Models 5 and 6). However, for the polynomial change (Model 7) the power is lower than those of Models 5 and 6. To explain this underperformance we can see, in Figure 1, that in the polynomial change, the time intervals where the mean stays near the extreme values 0 and 1 are very short compared to those in abrupt and smooth changes. We have simulated other continuous changes, linear and cubic polynomial, trigonometric, and many other functions. Like in Model 7, changes are hardly detected for small values of n, and the test $\mathcal{B}_n$ has a good performance only in large samples.

Acknowledgment

References

[1] A. Sen and M. S. Srivastava, "On tests for detecting change in mean," *The Annals of Statistics*, vol. 3, pp. 98–108, 1975.

[2] A. Sen and M. S. Srivastava, "Some one-sided tests for change in level," *Technometrics*, vol. 17, pp. 61–64, 1975.

[3] D. M. Hawkins, "Testing a sequence of observations for a shift in location," *Journal of the American Statistical Association*, vol. 72, no. 357, pp. 180–186, 1977.

[4] K. J. Worsley, "On the likelihood ratio test for a shift in location of normal populations," *Journal of the American Statistical Association*, vol. 74, no. 366, pp. 365–367, 1979.

[5] B. James, K. L. James, and D. Siegmund, "Tests for a change-point," *Biometrika*, vol. 74, no. 1, pp. 71–83, 1987.

[6] S. M. Tang and I. B. MacNeill, "The effect of serial correlation on tests for parameter change at unknown time," *The Annals of Statistics*, vol. 21, no. 1, pp. 552–575, 1993.

[7] J. Antoch, M. Hušková, and Z. Prášková, "Effect of dependence on statistics for determination of change," *Journal of Statistical Planning and Inference*, vol. 60, no. 2, pp. 291–310, 1997.

[8] X. Shao and X. Zhang, "Testing for change points in time series," *Journal of the American Statistical Association*, vol. 105, no. 491, pp. 1228–1240, 2010.

[9] M. S. Srivastava and K. J. Worsley, "Likelihood ratio tests for a change in the multivariate normal mean," *Journal of the American Statistical Association*, vol. 81, no. 393, pp. 199–204, 1986.

[10] L. Horváth, P. Kokoszka, and J. Steinebach, "Testing for changes in multivariate dependent observations with an application to temperature changes," *Journal of Multivariate Analysis*, vol. 68, no. 1, pp. 96–119, 1999.

[11] Z. Qu and P. Perron, "Estimating and testing structural changes in multivariate regressions," *Econometrica*, vol. 75, no. 2, pp. 459–502, 2007.

[12] C. Starica and C. W. J. Granger, "Nonstationarities in stock returns," *The Review of Economics and Statistics*, vol. 87, no. 3, pp. 503–522, 2005.

[13] P. Artzner, F. Delbaen, J.-M. Eber, and D. Heath, "Coherent measures of risk," *Mathematical Finance*, vol. 9, no. 3, pp. 203–228, 1999.

[14] A. G. Holton, *Value-at-Risk: Theory and Practice*, Academic Press, 2003.

[15] M. Boutahar, "Identification of persistent cycles in non-Gaussian long-memory time series," *Journal of Time Series Analysis*, vol. 29, no. 4, pp. 653–672, 2008.

[16] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, NY, USA, 1968.

[17] D. L. Iglehart, "Weak convergence of probability measures on product spaces with application to sums of random vectors," Technical Report, Stanford University, Department of Statistics, Stanford, Calif, USA, 1968.

[18] P. Hall and C. C. Heyde, *Martingale Limit Theory and Its Application*, Academic Press, New York, NY, USA, 1980.

[19] Y. S. Chow, "Local convergence of martingales and the law of large numbers," *Annals of Mathematical Statistics*, vol. 36, no. 2, pp. 552–558, 1965.

[20] T. L. Lai and C. Z. Wei, "Least squares estimates in stochastic regression models with applications to identification and control of dynamic systems," *The Annals of Statistics*, vol. 10, no. 1, pp. 154–166, 1982.

[21] T. Teräsvirta, "Specification, estimation, and evaluation of smooth transition autoregressive models," *Journal of American Statistical Association*, vol. 89, pp. 208–218, 1994.