TESTING ADDITIVE ASSUMPTIONS ON MEANS OF REGULAR MONITORING DATA: A MULTIVARIATE NONSTATIONARY TIME SERIES APPROACH

Ting Zhang

Boston University

Supplementary Material

S1. Appendix: Technical Proofs

Let $\boldsymbol{\Xi}_n(t) = \boldsymbol{\Sigma}(t)^{-1/2} \sum_{j=1}^n w_{j,n}(t)\boldsymbol{A}\boldsymbol{e}_j$, $t \in [0,1]$. We shall here provide the technical proofs of our main results.

Proof. (Theorem 3.1) By properties of local linear weights and Lemma A.4 of Zhang and Wu (2012), we have

$$(nb_n)\sum_{j=1}^n \{w_{j,n}(t)\}^2 \boldsymbol{\Lambda}(j/n) = (nb_n)\sum_{j=1}^n \{w_{j,n}(t)\}^2 \boldsymbol{\Lambda}(t) + o(1) = K_2\boldsymbol{\Lambda}(t) + o(1).$$

Since $\hat{\boldsymbol{\mu}}_n(t) - E\{\hat{\boldsymbol{\mu}}_n(t)\} = \sum_{j=1}^n w_{j,n}(t)\boldsymbol{e}_j$, by the proof of Lemma A1 in Zhang and Wu (2011) and the m-dependence approximation, we obtain the asymptotic normality

$$(nb_n)^{1/2}[\hat{\boldsymbol{\mu}}_n(t) - E\{\hat{\boldsymbol{\mu}}_n(t)\}] \Rightarrow N\{0, K_2\boldsymbol{\Lambda}(t)\}.$$

Note that $E\{\hat{\boldsymbol{\mu}}_n(t)\} = \boldsymbol{\mu}(t) + b_n^2\kappa_2\boldsymbol{\mu}''(t)/2 + O(b_n^3)$ by properties of local linear estimators, the first claim follows. By (3.4) and Lemma A.1 of Zhang and Wu (2012),

$$\begin{aligned}
\hat{\boldsymbol{a}}_n - E(\hat{\boldsymbol{a}}_n) &= \sum_{j=1}^n \left\{\int_0^1 w_{j,n}(t)dt\right\} \boldsymbol{A}\boldsymbol{e}_j \\
&= n^{-1}\sum_{j=1}^n \boldsymbol{A}\boldsymbol{e}_j + O_p\{n^{1/2}(nb_n)^{-2}b_n + n^{-1}(nb_n)^{1/2}\}.
\end{aligned}$$

By the proof of Lemma A1 in Zhang and Wu (2011) and the m-dependence approximation, we have

$$n^{-1/2}\sum_{j=1}^n \boldsymbol{A}\boldsymbol{e}_j \Rightarrow N\left\{0, \int_0^1 \boldsymbol{\Sigma}(t)dt\right\},$$

the second claim follows by noting that $nb_n \to \infty$ and $b_n \to 0$. Under the null hypothesis (3.1), we have $\boldsymbol{A}\boldsymbol{\mu}(t) = \boldsymbol{a}$, $t \in [0,1]$, and thus

$$E(\hat{\boldsymbol{a}}_n) = \int_0^1 \left\{ \sum_{j=1}^n w_{j,n}(t) \boldsymbol{A}\boldsymbol{\mu}(j/n) \right\} dt = \boldsymbol{a},$$

by observing that $\sum_{j=1}^n w_{j,n}(t) = 1$, $t \in [0,1]$, for local linear weights. □

Proof. (Theorem 3.2) Under the null hypothesis (3.1), by properties of local linear weights,

$$E\{\boldsymbol{A}\hat{\boldsymbol{\mu}}_n(t)\} = \sum_{j=1}^n w_{j,n}(t)\boldsymbol{a} = \boldsymbol{a}.$$

By Theorem 3.1,

$$\sup_{t\in[b_n,1-b_n]} |\boldsymbol{D}_n(t) - \boldsymbol{\Xi}_n(t)| = \sup_{t\in[b_n,1-b_n]} |\boldsymbol{\Sigma}(t)^{-1/2}\{E(\hat{\boldsymbol{a}}_n) - \hat{\boldsymbol{a}}_n\}| = O_p(n^{-1/2}).$$

Since $(b_n \log n)^{1/2} \to 0$ under the specified conditions, it suffices to prove that

$$\mathrm{pr}\left\{ \left(\frac{2nb_n \log m_n}{K_2} \right)^{1/2} \sup_{t\in[b_n,1-b_n]} |\boldsymbol{\Xi}_n(t)| - B_K(m_n) \le u \right\} \to \exp\{-2\exp(-u)\}. \tag{S1.1}$$

Let $\boldsymbol{S}_{\boldsymbol{A},k} = \sum_{j=1}^k \boldsymbol{A}\boldsymbol{e}_j$, $k = 1, \ldots, n$, be the partial sum process. By Theorem 2 of Zhou and Wu (2010), on a richer probability space, there exists a process $(\boldsymbol{S}^\diamond_{\boldsymbol{A},k})_{k=1}^n$ and iid r-dimensional standard multivariate normal random vectors $\boldsymbol{Z}_1, \ldots, \boldsymbol{Z}_n$ such that $(\boldsymbol{S}^\diamond_{\boldsymbol{A},k})_{k=1}^n$ has the same joint distribution as $(\boldsymbol{S}_{\boldsymbol{A},k})_{k=1}^n$ and

$$\max_{1\le k\le n} \left| \boldsymbol{S}^\diamond_{\boldsymbol{A},k} - \sum_{j=1}^k \boldsymbol{\Sigma}(j/n)^{1/2}\boldsymbol{Z}_j \right| = o_p(n^{3/10}\log n).$$

By the technique of summation by parts, we have

$$\sup_{t\in[b_n,1-b_n]} \left| \sum_{j=1}^n w_{j,n}(t)(\boldsymbol{S}^\diamond_{\boldsymbol{A},j} - \boldsymbol{S}^\diamond_{\boldsymbol{A},j-1}) - \sum_{j=1}^n w_{j,n}(t)\boldsymbol{\Sigma}(j/n)^{1/2}\boldsymbol{Z}_j \right| = o_p\left(\frac{n^{3/10}\log n}{nb_n} \right).$$

Since $(nb_n)^{-1/2} n^{3/10} (\log n)^{3/2} \to 0$ under the specified conditions, it suffices to prove that (S1.1) holds with $\boldsymbol{\Xi}_n(t)$ therein replaced by $\boldsymbol{\Sigma}(t)^{-1/2} \sum_{j=1}^n w_{j,n}(t)\boldsymbol{\Sigma}(j/n)^{1/2}\boldsymbol{Z}_j$. By the proof of Lemmas A.3 and A.4 of Zhang and Wu (2012),

$$\sup_{t\in\mathcal{T}_n} \left| \sum_{j=1}^n w_{j,n}(t) \left\{ \boldsymbol{\Sigma}(j/n)^{1/2} - \boldsymbol{\Sigma}(t)^{1/2} \right\} \boldsymbol{Z}_j \right| = O_p\{(nb_n)^{-1/2}(\log n)^{1/2} b_n^{1/2}\},$$

and

$$\sup_{t\in\mathcal{T}_n}\left|\sum_{j=1}^{n}[w_{j,n}(t)-(nb_n)^{-1}K\{(j/n-t)/b_n\}]\boldsymbol{Z}_j\right| = O_p\{(nb_n)^{-3/2}(\log n)^{1/2}\}.$$

Theorem 3.2 follows by Lemma 1 of Zhou and Wu (2010). □

Proof. (Proposition 3.1) Under the alternative hypothesis (3.7),

$$\boldsymbol{A}\hat{\boldsymbol{\mu}}_n(t)=\sum_{j=1}^{n}w_{j,n}(t)\boldsymbol{A}\boldsymbol{e}_j+\boldsymbol{a}+d_n\boldsymbol{f}(t)+O(d_nb_n^2) \tag{S1.2}$$

uniformly over $t\in[0,1]$ and

$$\hat{\boldsymbol{a}}_n=\{\hat{\boldsymbol{a}}_n-E(\hat{\boldsymbol{a}}_n)\}+\boldsymbol{a}+d_n\int_0^1\boldsymbol{f}(t)dt+O(d_nb_n^2). \tag{S1.3}$$

Let $\boldsymbol{\Delta}(t)=\boldsymbol{\Sigma}(t)^{-1/2}\{\boldsymbol{f}(t)-\int_0^1\boldsymbol{f}(t)dt\}$, by Theorem 3.1,

$$\begin{aligned}\sup_{t\in[b_n,1-b_n]}|\boldsymbol{D}_n(t)| &= \sup_{t\in[b_n,1-b_n]}|\boldsymbol{\Xi}_n(t)+d_n\boldsymbol{\Delta}(t)|+O_p(d_nb_n^2+n^{-1/2})\\ &\geq d_n\sup_{t\in[b_n,1-b_n]}|\boldsymbol{\Delta}(t)|-\sup_{t\in[b_n,1-b_n]}|\boldsymbol{\Xi}_n(t)|+O_p(d_nb_n^2+n^{-1/2}).\end{aligned}$$

Since $B_K(m_n)=O(\log m_n)$, by (S1.1) we have

$$\sup_{t\in[b_n,1-b_n]}|\boldsymbol{\Xi}_n(t)|=O_p\{(nb_n)^{-1/2}(\log m_n)^{1/2}\}.$$

Note that $(2nb_n\log m_n)^{1/2}d_n/B_K(m_n)\to\infty$ and $(2b_n\log m_n)^{1/2}\to 0$ under the specified conditions on the bandwidth sequence, Proposition 3.1 follows. □

Proof. (Theorem 3.3) Since $E\{\boldsymbol{A}\hat{\boldsymbol{\mu}}_n(t)\}=\hat{\boldsymbol{a}}_n=\boldsymbol{a}$ for all $t\in[0,1]$ under the null hypothesis (3.1), we have

$$\boldsymbol{A}\hat{\boldsymbol{\mu}}_n(t)-\hat{\boldsymbol{a}}_n=\sum_{j=1}^{n}w_{j,n}(t)\boldsymbol{A}\boldsymbol{e}_j-\{\hat{\boldsymbol{a}}_n-E(\hat{\boldsymbol{a}}_n)\},\quad t\in[0,1],$$

and thus

$$|\boldsymbol{D}_n(t)|^2=|\boldsymbol{\Xi}_n(t)|^2+|\boldsymbol{\Sigma}(t)^{-1/2}\{\hat{\boldsymbol{a}}_n-E(\hat{\boldsymbol{a}}_n)\}|^2-2\{\hat{\boldsymbol{a}}_n-E(\hat{\boldsymbol{a}}_n)\}^\top\boldsymbol{\Sigma}(t)^{-1}\sum_{j=1}^{n}w_{j,n}(t)\boldsymbol{A}\boldsymbol{e}_j$$

holds for all $t\in[0,1]$. By the proof of Theorem 3.1, we have

$$\int_0^1|\boldsymbol{D}_n(t)|^2dt=\int_0^1|\boldsymbol{\Xi}_n(t)|^2dt+O_p(n^{-1}).$$

By Lemmas A.2 and A.5 of Zhang and Wu (2012),

$$nb_n^{1/2}\left\{\int_0^1 |\boldsymbol{\Xi}_n(t)|^2 dt - (nb_n)^{-1} r K^*(0)\right\} \Rightarrow N(0, 4rK_2^*), \tag{S1.4}$$

Theorem 3.3 follows. □

Proof. (Proposition 3.2) Let $\boldsymbol{\Delta}(t) = \boldsymbol{\Sigma}(t)^{-1/2}\{\boldsymbol{f}(t) - \int_0^1 \boldsymbol{f}(t)dt\}$ be as in the proof of Proposition 3.1, and $\bar{\Delta}^2 = \int_0^1 |\boldsymbol{\Delta}(t)|^2 dt$. Then under the alternative hypothesis (3.7), by (S1.2) and (S1.3),

$$\boldsymbol{D}_n(t) = \boldsymbol{\Xi}_n(t) - \boldsymbol{\Sigma}(t)^{-1/2}\{\hat{\boldsymbol{a}}_n - E(\hat{\boldsymbol{a}}_n)\} + d_n\{\boldsymbol{\Delta}(t) + O(b_n^2)\}$$

uniformly over $t \in [0,1]$. Hence, by Lemma A.1 of Zhang and Wu (2012),

$$\int_0^1 |\boldsymbol{D}_n(t)|^2 dt = \int_0^1 |\boldsymbol{\Xi}_n(t)|^2 dt + d_n^2 \int_0^1 |\boldsymbol{\Delta}(t)|^2 dt + O_p(n^{-1} + n^{-1/2} d_n + d_n^2 b_n^2).$$

Since $n^{-1/2} = o(d_n)$, we have for sufficiently large n,

$$\begin{aligned}
&\mathrm{pr}\left\{\int_0^1 |\boldsymbol{D}_n(t)|^2 dt > (nb_n)^{-1} r K^*(0) + n^{-1} b_n^{-1/2} (4rK_2^*)^{1/2} q_{1-\alpha}\right\} \\
&\quad \geq \mathrm{pr}\left[nb_n^{1/2}\left\{\int_0^1 |\boldsymbol{\Xi}_n(t)|^2 dt - (nb_n)^{-1} r K^*(0)\right\} > (4rK_2^*)^{1/2} q_{1-\alpha} - \frac{nb_n^{1/2} d_n^2}{2}\bar{\Delta}^2\right].
\end{aligned}$$

Since $nb_n^{1/2} d_n^2 \to \infty$, Proposition 3.2 follows by (S1.4). □

Proof. (Corollary 5.1) Let $\boldsymbol{\vartheta}_k(t) = \boldsymbol{G}(t; \boldsymbol{\mathcal{F}}_k) - E\{\boldsymbol{G}(t; \boldsymbol{\mathcal{F}}_k)\}$ and $\boldsymbol{\eta}_k(t) = \boldsymbol{A}\boldsymbol{\vartheta}_k(t)$, then $\boldsymbol{e}_j = \boldsymbol{\vartheta}_j(j/n)$, $j = 1, \ldots, n$, and under the null hypothesis (3.1) the difference $\boldsymbol{A}\hat{\boldsymbol{\mu}}_n(t) - \boldsymbol{a} = \sum_{j=1}^n \boldsymbol{\eta}_j(j/n) w_{j,n}(t)$. Recall that $\boldsymbol{A} = (a_{ij})$ is a prespecified matrix determined by the goal of the test (for example testing the mean equivalence of two particular time series), whose nonzero entries are bounded away from zero, namely there exists a constant $\varepsilon > 0$ such that $|a_{ij}| \geq \varepsilon$ if $a_{ij} \neq 0$. Let $\vartheta_{i,k}(t)$ be the i-th component of the vector $\boldsymbol{\vartheta}_k(t)$, $i = 1, \ldots, p$. Then, by the Cauchy-Schwarz inequality, we have for any $t_1, t_2 \in [0,1]$,

$$\begin{aligned}
\|\boldsymbol{\eta}_k(t_1) - \boldsymbol{\eta}_k(t_2)\|^2 &= E\left(\sum_{i=1}^r \left[\sum_{j=1}^p a_{ij}\{\vartheta_{j,k}(t_1) - \vartheta_{j,k}(t_2)\}\right]^2\right) \\
&\leq \sum_{i=1}^r \left[\left(\sum_{j=1}^p a_{ij}^2\right) E\left\{\sum_{l=1}^p |\vartheta_{l,k}(t_1) - \vartheta_{l,k}(t_2)|^2 \left|\frac{a_{il}}{\varepsilon}\right|^2\right\}\right] \\
&\leq \|\boldsymbol{A}\|_F^2 \left\{\max_{1 \leq l \leq p} \|\vartheta_{l,k}(t_1) - \vartheta_{l,k}(t_2)\|^2\right\} \frac{\|\boldsymbol{A}\|_F^2}{\varepsilon^2} \leq \frac{c_F^4 (c_0^\sharp)^2}{\varepsilon^2} |t_1 - t_2|^2,
\end{aligned}$$

and thus $\|\boldsymbol{\eta}_k(t_1)-\boldsymbol{\eta}_k(t_2)\| \le \varepsilon^{-1}c_F^2 c_0^\sharp |t_1-t_2|$. Let $\boldsymbol{\vartheta}_k^\star(t)=\boldsymbol{G}(t;\boldsymbol{\mathcal{F}}_k^\star)-E\{\boldsymbol{G}(t;\boldsymbol{\mathcal{F}}_k^\star)\}$ be the coupled version of $\boldsymbol{\vartheta}_k(t)$ and similarly define $\boldsymbol{\eta}_k^\star(t)=\boldsymbol{A}\boldsymbol{\vartheta}_k^\star(t)$, then for any $q\ge 2$,

$$
\begin{aligned}
\|\boldsymbol{\eta}_k(t)-\boldsymbol{\eta}_k^\star(t)\|_q^q &= E\left\{\left(\sum_{i=1}^r\left[\sum_{j=1}^p a_{ij}\{\vartheta_{j,k}(t)-\vartheta_{j,k}^\star(t)\}\right]^2\right)^{q/2}\right\} \\
&\le E\left\{\left(\sum_{i=1}^r\left[\left(\sum_{j=1}^p a_{ij}^2\right)\left\{\sum_{l=1}^p|\vartheta_{l,k}(t)-\vartheta_{l,k}^\star(t)|^2\left|\frac{a_{il}}{\varepsilon}\right|^2\right\}\right]\right)^{q/2}\right\} \\
&\le \|\boldsymbol{A}\|_F^q\cdot E\left[\left\{\sum_{i=1}^r\sum_{l=1}^p|\vartheta_{l,k}(t)-\vartheta_{l,k}^\star(t)|^2\left|\frac{a_{il}}{\varepsilon}\right|^2\right\}^{q/2}\right].
\end{aligned}
$$

By taking $q=2$ in the above equation, we obtain that $\|\boldsymbol{\eta}_k(t)-\boldsymbol{\eta}_k^\star(t)\|\le\varepsilon^{-1}\|\boldsymbol{A}\|_F^2\theta_{k,2}^\sharp$ for all $t\in[0,1]$. On the other hand, if $q>2$, then by the Hölder inequality,

$$
\begin{aligned}
\sum_{i=1}^r\sum_{l=1}^p|\vartheta_{l,k}(t)-\vartheta_{l,k}^\star(t)|^2\left|\frac{a_{il}}{\varepsilon}\right|^2 &\le \left\{\sum_{i=1}^r\sum_{l=1}^p|\vartheta_{l,k}(t)-\vartheta_{l,k}^\star(t)|^q\left|\frac{a_{il}}{\varepsilon}\right|^2\right\}^{2/q} \\
&\quad\times\left\{\sum_{i=1}^r\sum_{l=1}^p\left|\frac{a_{il}}{\varepsilon}\right|^2\right\}^{(q-2)/q},
\end{aligned}
$$

and as a result,

$$
\|\boldsymbol{\eta}_k(t)-\boldsymbol{\eta}_k^\star(t)\|_q^q\le\|\boldsymbol{A}\|_F^q\left\{\max_{1\le l\le p}\|\vartheta_{l,k}(t)-\vartheta_{l,k}^\star(t)\|_q^q\right\}\left(\frac{\|\boldsymbol{A}\|_F^2}{\varepsilon^2}\right)^{q/2}\le\left(\frac{\|\boldsymbol{A}\|_F^2\theta_{k,q}^\sharp}{\varepsilon}\right)^q.
$$

Therefore, we have $\|\boldsymbol{\eta}_k(t)-\boldsymbol{\eta}_k^\star(t)\|_q\le\varepsilon^{-1}c_F^2\theta_{k,q}^\sharp$ for all $q\ge2$ and $t\in[0,1]$. Note that $|\boldsymbol{\eta}_k(t)|\in\mathbb{R}$, the proof follows by a similar argument as in Theorems 3.2 and 3.3. □

References

Zhang, T. and Wu, W. B. (2011). Testing parametric assumptions of trends of a nonstationary time series. *Biometrika* **98**, 599–614.

Zhang, T. and Wu, W. B. (2012). Inference of time-varying regression models. *Ann. Statist.* **40**, 1376–1402.

Zhou, Z. and Wu, W. B. (2010). Simultaneous inference of linear models with time varying coefficients. *J. Roy. Statist. Soc. B* **72**, 513–531.