Corrections on "Multiscale and multilevel technique for consistent segmentation of nonstationary time series" [1]

Haeran Cho * Piotr Fryzlewicz †

January 28, 2015

We correct Lemma 2 in the supplementary document of [1] as below.

Lemma 2' Suppose (1.2) holds and let $\eta \equiv \eta_{p_0+r} \in [s,e]$ for some $r \in \{1,\ldots,q\}$, denote a true change-point. Then there exists $c_0 \in (0,\infty)$ such that for b satisfying $|\widetilde{\mathbb{S}}^b_{s,e}| < |\widetilde{\mathbb{S}}^\eta_{s,e}|$ and $|\eta - b| \ge c_0\epsilon_T$ with $\epsilon_T = T^{5/2-2\Theta}\log T$, we have $|\widetilde{\mathbb{S}}^\eta_{s,e}| - |\widetilde{\mathbb{S}}^b_{s,e}| \ge 2\log T$.

Proof. Let both $\widetilde{\mathbb{S}}^\eta_{s,e}$, $\widetilde{\mathbb{S}}^b_{s,e} \ge 0$ without loss of generality.
The proof follows directly from the proof of Lemma 2.6 in [2]. We only consider Case 2 of Lemma 2.6, since adapting the proof of Case 1 (when there is a single change-point within $[s,e]$) to that of the current lemma takes analogous arguments.
Using the notations therein, it is shown that the term E_{1l} is dominant over E_{2l} and E_{3l} in $\widetilde{\mathbb{S}}^\eta_{s,e} - \widetilde{\mathbb{S}}^b_{s,e}$, where $l = c_0\epsilon_T$. Noting further that $i = \eta - s + 1$, $h = \delta_T$, $j = e - \eta - h$ and $a = \sum_{t=s}^{\eta}\sigma^2(t/T) - (e-s+1)^{-1}\sum_{t=s}^{e}\sigma^2(t/T)$, and that $h \ge 2l$,

$$\begin{aligned} E_{1l} &= \frac{la\sqrt{i+j+h}}{\sqrt{i}\sqrt{j+h}} \cdot \frac{h-l}{\sqrt{i+l}\sqrt{j+h-l}\{\sqrt{(i+l)(j+h-l)} + \sqrt{i(j+l)}\}} \\ &\ge \widetilde{\mathbb{S}}^\eta_{s,e} \cdot C\epsilon_T\delta_T T^{-2} \ge 2\log T \end{aligned}$$

for large T. □

With Lemma 2', we derive Theorem 1' under Assumption 2', which correct Theorem 1 and Assumption 2 of [1], respectively. Theorem 2 of [1] is also updated with regards to the rate given to the bias term ϵ_T in Theorem 1' below.

Assumption 2' For $\Theta \in (7/8, 1]$ and $\theta \in (5/4 - \Theta, \Theta - 1/2)$, the length of each segment in $\sigma^2(t/T)$ is bounded from below by $\delta_T = CT^\Theta$. Further, there exists some constant $c \in (0,\infty)$ such that,

$$\max_{1\le p\le N}\left\{\sqrt{\frac{\eta_p - \eta_{p-1}}{\eta_{p+1}-\eta_p}}, \sqrt{\frac{\eta_{p+1}-\eta_p}{\eta_p - \eta_{p-1}}}\right\} \le c.$$

*School of Mathematics, University of Bristol, UK. haeran.cho@bristol.ac.uk
†Department of Statistics, London School of Econimics, UK.

Theorem 1' Suppose that $\{Y_{t,T}\}_{t=0}^{T-1}$ follows the model (3). Assume that there exist $M, m > 0$ such that $\sup_t |\sigma^2(t/T)| \le M$ and $\inf_{1\le i\le N} |\sigma^2((\eta_i+1)/T) - \sigma^2(\eta_i/T)| \ge m$. Under Assumption 2', the number and locations of the detected breakpoints are consistent. That is, $\mathbb{P}\left\{\widehat{N} = N; |\widehat{\eta}_p - \eta_p| \le C\epsilon_T, 1 \le p \le N\right\} \to 1$ as $T \to \infty$, where $\widehat{\eta}_p$, $p = 1, \ldots, \widehat{N}$ are detected breakpoints and $\epsilon_T = T^{5/2-2\Theta}\log T$. (Interpreting this in the rescaled time interval $[0,1]$, $\epsilon_T/T = T^{3/2-2\Theta}\log T \to 0$ as $T \to 0$.)

References

[1] Cho, H. and Fryzlewicz, P. (2012), "Multiscale and multilevel technique for consistent segmentation of nonstationary time series," *Statistica Sinica*, 22, 207–229.

[2] Venkatraman, E. S. (1992), "Consistency results in multiple change-point problems," *Technical Report No. 24, Department of Statistics, Stanford University.*