Discussion of “Multiscale change point inference” by Frick, Munk and Sieling

Piotr Fryzlewicz
Department of Statistics
London School of Economics, UK
p.fryzlewicz@lse.ac.uk

October 1, 2013

I would like to congratulate the authors for the thought-provoking article.

Revisiting the results of Table 1 from the paper, it is interesting to note that the authors exhibit the performance of SMUCE for two particular values of the tuning parameter, $\alpha = 0.45$ and $\alpha = 0.6$, the latter presumably chosen in order to improve the performance in the high variance setting of $\sigma = 0.3$. In this note, we apply the Wild Binary Segmentation (WBS) technique of Fryzlewicz (2012) to the same example, including in a version that requires no choice of any tuning parameters on the part of the user.

More specifically, we apply the WBS algorithm to the trend-free signal from Table 1 of the paper, with $\sigma = 0.1$ and $\sigma = 0.3$. The WBS algorithm uses the default value of 5000 random draws (each execution took the average of 1.20 seconds on a standard PC). In the thresholding stopping rule, we use the threshold $c\,\hat{\sigma}\sqrt{2\log n}$, where $\hat{\sigma}$ is the Median Absolute Deviation estimator of σ suitable for iid Gaussian noise, n is the sample size, and the constant c is selected manually. In the fully automatic stopping rule based on an information criterion, we propose a new criterion termed “strengthened BIC” (sBIC), which can be shown to be consistent for the number and locations of change-points when coupled with WBS, and works by replacing the BIC penalty of $\log(n)\times$ number of change-points by $\log^{\gamma}(n)\times$ number of change-points, for $\gamma > 1$, where we use the default value of $\gamma = 1.01$ in order to remain close to the BIC.

Table 1 in this note shows the results. It is clear that SMUCE, even with an optimally chosen tuning parameter, struggles in the higher variance setting, where it is comfortably outperformed by WBS. WBS sBIC performs very well in both settings, and we emphasise again that it is a completely automatic procedure in which no tuning parameters need to be chosen. For the case $\sigma = 0.3$, WBS sBIC would have come on top also in Table 1 of the paper, and for the case $\sigma = 0.1$, very close to the top.

Finally, we note that the Unbalanced Haar (UH) method of Fryzlewicz (2007), included by the authors in the same simulation study, is not a consistent change-point detector. Out of

Method	σ	Results ≤ 4	5	6	7	≥ 8
SMUCE $\alpha = 0.45$	0.1	0	0	99	1	0
WBS $c = 1.3$	0.1	0	0	97	2	1
WBS $c = 1.4$	0.1	0	0	99	1	0
WBS $c \in [1.5, 2]$	0.1	0	0	100	0	0
WBS sBIC	0.1	0	0	97	2	1
SMUCE $\alpha = 0.45$	0.3	3	34	62	1	0
SMUCE $\alpha = 0.6$	0.3	0	10	80	9	0
WBS $c = 1.3$	0.3	0	3	94	2	1
WBS $c = 1.4$	0.3	2	7	90	1	0
WBS sBIC	0.3	0	1	95	3	1

Table 1: Distributions of the estimated numbers of change-points for SMUCE and WBS expressed as percentages. Results for SMUCE taken from the paper and averaged to the nearest integer. Results for WBS based on 100 simulation runs, with the random seed in R set to the arbitrary value of 1 before the first simulation run, for reproducibility.

the three methods studied (SMUCE, UH, WBS), users will be best-off applying the WBS method in this context.

References

P. Fryzlewicz. Unbalanced Haar technique for nonparametric function estimation. *J. Amer. Stat. Assoc.*, 102:1318–1327, 2007.

P. Fryzlewicz. Wild Binary Segmentation for multiple change-point detection. *Preprint*, 2012.