Multiscale analysis of multiple change points using unbalanced wavelets

Maarten Jansen

TU Eindhoven, Department of Mathematics,
PO Box 513, 5600 MB Eindhoven, the Netherlands
and K.U.Leuven, Department of Computer Science,
Celestijnenlaan 200A, 3001 Leuven, Belgium
mjansen@win.tue.nl

Abstract – We present a multiscale analysis of piecewise constant signals subject to noise. The objective is to find the locations of the change points. To this end, we first apply a continuous wavelet analysis and construct lines of local wavelet maxima, in a way similar to the well-known construction by Mallat and Hwang. The second stage of our approach is an extension towards unbalanced wavelet analyses in order to improve the statistical power of our detection.

1 Introduction, problem description

Wavelet thresholding [1] provides an interesting tool in smoothing piecewise smooth signals subject to noise. The idea is that singularities in the underlying (i.e., noise-free) signal give rise to wavelet coefficients whose magnitudes are significantly higher than the magnitudes of coefficients that are not related to any of the singularities. The technique can be proven to be statistically optimal in several aspects. Moreover, on the practical side, it can be extended and refined using nondecimated wavelet transforms, resolution level dependent thresholds, tree structured coefficient selection, block thresholding, hard- or soft-thresholding or any intermediate operation, such as thresholds in a Bayesian framework [2].

In spite of this success, wavelet smoothing is sometimes criticised for its outputs which often contain spurious features, typically spikes, due to falsely selected wavelet coefficients. Every coefficient selection method inevitably shows at least a few of these false discoveries. Moreover, the true significant coefficients also carry some proportion of noise. Although these spurious features in general have little impact on the signal-to-noise ratio of the output, they often affect its visual quality. The output becomes smoother if the adopted selection is more conservative (i.e., if less coefficients pass the selection), such as in the universal threshold procedure, but the price to pay is inaccuracy, Gibbs phenomena near the jumps.

Wavelet thresholding uses the wavelet coefficients both for locating the positions of the singularities and for smoothing the intervals between those singularities. Both tasks are performed simultaneously: the coefficients locate the singularities in an implicit, passive way. The idea of this paper is to actively estimate the precise locations of singularities first, thereby separating the detection of singularities from the actual smoothing.

In this paper we consider piecewise constant functions only. Smoothing given estimates for the jump locations is then of course straightforward by taking the averages of the observations between the jumps.

The problem of locating the jumps is known in statistical literature as *change point detection*. In the presented algorithm, the number of change point is not a priori bounded. The algorithm is refines a technique based on the analysis of local maxima in a continuous wavelet transform [3, 4], further explained below. It also extends this local maxima analysis with towards an unbalanced transform, which is a sort of so-called second generation wavelet. As explained below, this unbalanced transform allows to find for each change point, the specific scales (i.e., ranges) of the two intervals of smooth (in this paper: constant) behaviour on the left- and right hand side of the change point. All together, unlike the wavelet threshold approach, the proposed algorithm does not *operate* on the wavelet coefficients directly. It rather uses the coefficients as a tool for a fast search for the precise location and left- and right hand scales of each change point. A full search on these three parameters would require $\mathcal{O}(N^3)$ computations for N samples. The presented algorithm finds them with a computational complexity of $\mathcal{O}(N \log N)$.

2 The algorithm

Suppose we are given N noisy samples $y_i, i = 1, \ldots, N$ of a piecewise constant signal $\mu_i, i = 1, \ldots, N$. Consecutive observations have the same intensity, except at some transition points:

$$\mu_k = \mu_{\tau_r}, \text{ for } k = \tau_r, \ldots, \tau_{r+1} - 1,$$

where $0 < \tau_0 < \ldots < \tau_r < \tau_{r+1} < \ldots < \tau_R \leq n-1$ is a sequence of R change points, and $0 \leq R \leq n-1$ is unknown. The change points are specified by the (integer) index τ_r of the first observation from the segment with a certain intensity. The proposed algorithm to estimate the τ_r's proceeds as follows:

1. Compute a (discretised) continuous wavelet transform $\boldsymbol{w}$ of $\boldsymbol{y}$, Let J be the number of discretised scales, then $\boldsymbol{w}$ is a $J \times N$ matrix.

2. For each scale $j = 1, \ldots J$, find the local maxima. If this maximum is sufficiently large (say, if its absolute value is larger than 3), the corresponding location is considered as a candidate change point. In the presence of noise, the fine scales of the wavelet transform have a lot of local maxima. In order to save computations, it is interesting to smooth the wavelet transform within the scale and compute the local maxima of that smoothed version first. The obtained values serve as provisional estimates of the local maxima. In a second step, we compute the global maxima of the original, non-smoothed transform, on the intervals between every pair of provisional maxima and we replace the provisional values by their corresponding new values. Let $\mathcal{M}_j$ denote the set of indices corresponding to these selected maxima at scale j, i.e.,

$$\mathcal{M}_j = \{k = 0, \ldots, n-1 \big| |w_{j,k}| \geq |w_{j,k\pm1}|\}.$$

3. Link local maxima at successive scales. Two maxima at successive scales are linked if both are the closest maximum to the other one. More precisely, a maximum at scale j, location k_j is connected to a maximum at scale $j+1$, location k_{j+1} if and only if

$$\begin{aligned} k_j &= \arg\min_{l \in \mathcal{M}_j} |l - k_{j+1}|, \text{ and} \\ k_{j+1} &= \arg\min_{l \in \mathcal{M}_{j+1}} |l - k_j|. \end{aligned}$$

4. Merge lines with overlapping locations into a single line: some different lines show up at the same location, but different scales, for instance if there is a gap between scales of local maxima. The algorithm starts from the longest existing lines. If such a line does not continue all the scales down, we check if a bridge can be constructed from its end point to another line at a neighboring location. If there is more than one candidate, take the shortest bridge, where the length of the bridge is defined based on the shift in location and scale to jump into the new maxima line. If two candidate lines can be reached by bridges of equal lengths, we select the line whose average location over all scales is closest to the average of the original line that we want to extend. As soon as a candidate line of maxima is selected, the original line is completed by filling in the locations of the secondary line at scales where the original line had no maxima. The secondary line is then removed from the set of maxima lines.

5. For each line of maxima, select the scale j on which the coefficient has the largest magnitude.

6. Make the basis functions unbalanced. Introduce two scale variables, j_l and j_r, left and right from the location k of the line of maxima at the scale j selected in the previous step. The extension to unbalanced wavelets is trivial in the Haar case, leading to the unbalanced Haar wavelet transform. Chose the values of j_l and j_r such that the resulting wavelet coefficient is maximised.

7. Starting with the largest coefficients, select the locations of significant change points. Once a location is selected, recompute the remaining coefficients such that the corresponding basis function lie entirely on one side of the previously selected change points.

An example of the output of the algorithm is given in figure 1. The noise in this example is generated by a Poisson process.

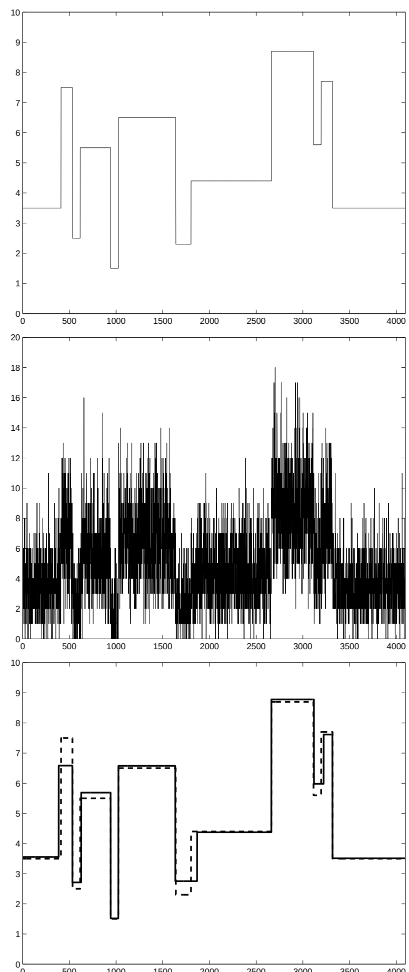

Fig. 1: A simulated example of Poisson data with time varying intensities. On the left the plot of the intensity curve. This is a scaled and vertically translated version of the well-known ‘Blocks’ test example [1]. In the middle a random realization. On the right the estimation from that realization, using the procedure proposed in this paper.

References

[1] D. L. Donoho and I. M. Johnstone. Ideal spatial adaptation via wavelet shrinkage. *Biometrika*, 81:425–455, 1994.

[2] I. M. Johnstone and B. W. Silverman. Needles and straw in haystacks: empirical bayes estimates of possibly sparse sequences. *Annals of Statistics*, 32(4):1594–1649, 2004.

[3] S. Mallat and W. L. Hwang. Singularity detection and processing with wavelets. *IEEE Transactions on Information Theory*, 38(2):617–643, 1992.

[4] S. Mallat and S. Zhong. Characterization of signals from multiscale edges. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 14:710–732, 1992.