Iterative Filtered Derivative with t-Value for Change Point Detection

Pierre Raphaël BERTRAND*†,
Doha HADOUNI*†

13 avril 2017

Résumé: Notre étude concerne la détection de rupture *a posteriori* en utilisant la méthode de Dérivée Filtrée avec t-Value Itérative (I-FDtV). Il s'agit d'un algorithme en deux étapes. Dans la première, on utilise la fonction dérivée filtrée (FD), basée sur deux extra-paramètres : le seuil de détection et la taille de la fenêtre, pour sélectionner un ensemble de points de rupture potentiels. Dans la seconde étape, on calcule la *t-value* pour chaque point de rupture potentiel que l'on compare avec une t-value critique itérativement incrémentée. Ainsi, on retient les vrais positifs (avec probabilité $1-\alpha_1$) et on rejette les faux positifs (avec probabilité $1-\alpha_2$). De plus, nous présentons les résultats théoriques et les choix pratiques pour les extra-paramètres de la méthode.

Mots clès: Détection de point de rupture ; La fonction Dérivée Filtrée ; La méthode Dérivée Filtrée avec *t-Value* Iterative ; Extra-paramètres de la Dérivée Filtrée.

Abstract: The study deals with *off-line* change point detection using the Iterative Filtered Derivative with *t-Value* method. The i-FDtV method is a two-step procedure for change point analysis. The first step is based on the Filtered Derivative function (FD) to select a set of potential change points, using its extra-parameters - namely the threshold for detection and the sliding window size. The second step is an iteration set of eliminations with an increasing *t-value* threshold in order to discard the change points with a *t-value* lower than the threshold, called false alarms (with probability $1-\alpha_2$), and keep the true positives (with probability $1-\alpha_1$) once the stopping condition is checked. Furthermore, we give the theoretical results and the practical choices of the extra-parameters.

Keywords: Change points detection ; Filtered Derivative function ; Iterative Filtered Derivative with *t-Value* ; Filtered Derivative extra-parameters.

Introduction

La détection de rupture *a posteriori* d'une série chronologique est une méthode statistique pertinente pour les applications en finance [5, 11], médecine [7], neuro-physiologie [9, 6]. La méthode FDtV itérative a une complexité en $\mathcal{O}(n)$ en temps de calcul et en espace mémoire, où n est la taille de la série [4, 10, 8]. Comme toute méthode, la méthode I-FDtV dépend d'extra-paramètres. L'objectif est donc de trouver ces extra-paramètres tels que le risque de zero non-detection soit supérieur à $1-\alpha_1$ et le risque de fausse alarme inférieur à α_2. On pourra ainsi détecter tous les vrais positifs et la méthode I-FDtV sera exacte. Autrement dit, grâce à cette méthode, nous détectons tous les points de rupture réels au risque α_1 de non-détection et au risque α_2 de fausses alarmes.

*Université Clermont Auvergne, CNRS, LMBP, F-63000 Clermont–Ferrand, France.
†Supporté par le projet ANR- 12-BS01-0016-01 intitulé *"Do Well B"*.

1 La méthode Dérivée Filtrée avec t-Value Itérative

1.1 Modèle

Soit $\mathbf{X} = (X_1, X_2, \ldots, X_n)$ une série indexée par le temps $\mathbf{t} = 1, 2, \ldots, n$ telle que :

- $t \longmapsto \mu(t) = \mu_k$ une fonction constante par morceaux pour tout $t \in (\tau_k, \tau_{k+1}]$;
- $\tau = (\tau_1, \ldots, \tau_K)$ une configuration de K points de rupture, complétée par convention $\tau_0 = 0$ et $\tau_{K+1} = n$;
- $L_0 = \inf\{|\tau_{k+1} - \tau_k|, pour\ k = 0, \ldots, K\}$ la distance minimale entre deux points de rupture consécutifs ;
- $\mu = (\mu_0, \ldots, \mu_K)$ une configuration de K moyennes ;
- $\delta = (\delta_1, \ldots, \delta_K)$ une configuration des décalages où $\delta_k = \mu_k - \mu_{k-1}$, pour $k = 1, \ldots, K$ tel que $\delta_0 = \inf\{|\delta_k|,\ k = 1, \ldots, K\}$ la valeur absolue minimale des décalages ;
- $X_t \in \mathcal{N}(\mu_k, \sigma^2)$ pour $t \in (\tau_k, \tau_{k+1}]$ avec $k = 0, \ldots, K$,
- Le ratio signal-bruit $SNR = \delta_0/\sigma$.

1.2 Définition de la méthode I-FDtV

La méthode I-FDtV est basée sur deux étapes. La première consiste à utiliser la fonction Dérivée Filtrée (FD). La seconde étape consiste à éliminer les faux positifs et garder les vrais positifs de manière itérative afin de ne pas avoir de non-détection. Pour plus de précision, la méthode I-FDtV est définie comme suit :

Etape 1 **: Dérivée Filtrée**

La première étape consiste à utiliser la fonction Dérivée Filtrée (FD) pour sélectionner les points de ruptures potentiels $T_1 = \{\tau_1^\star, \ldots, \tau_{K^\star}^\star\}$. Elle dépend de deux paramètres : la taille de la fenêtre A et le seuil de séléction C_1.

1. **Calcul de la fonction Dérivée Filtrée :**

   Définition 1.1 *La fonction Dérivée Filtrée est définie par la formule suivante :*

   $$FD(t, A) = \widehat{\mu}(X, [t+1, t+A]) - \widehat{\mu}(X, [t-A+1, t]), \quad pour\ A < t < n - A, \tag{1.1}$$

   avec $\widehat{\mu}(X, [u, v]) := \frac{\sum_{t=u}^{v} X_t}{(v-u+1)}$ *la moyenne empirique des variables* X_t *avec* $t \in [u, v]$.

   Cette méthode consiste à filtrer les données en calculant les estimateurs du paramètres μ avant d'appliquer une dérivation discrète. Ce qui explique le nom *"méthode de la Dérivée Filtrée "* [2, 1]. La quantité $A.FD(t, A)$ peut être calculée de manière itérative en utilisant

   $$A.FD(t+1, A) = A.FD(t, A) + X(t+1+A) - 2X(t+1) + X(t-A+1).$$

2. **Détermination des points de rupture potentiels**

   Définition 1.2 *Les points de rupture potentiels notés* τ_k^**, pour* $k = 1, \ldots, K^*$ *sont les maximums locaux de la valeur absolue de la fonction dérivée filtrée* $|FD(t, A)|$.

   Pour la réalisation de la détection, on suit l'algorithme suivant :

   (a) Sélectionner le point de rupture potentiels τ_k^* qui est le maximum global de la fonction $|FD_k(t, A)|$,

(b) Définir les valeurs de la fonction FD_{k+1} à zero afin que la largeur de l'amplitude du point de rupture τ_k^* soit égale à $2A$.

(c) Itérer cet algorithme tant que $|FD_k(\tau_k^*, A)| > C_1$.

Etape 2 : Elimination des faux positifs

Un point de rupture potentiel peut être une fausse alarme ou une estimation d'un point de rupture réel. Le but est donc d'éliminer les fausses alarmes sans enlever les vrais positifs. Pour cela, à l'étape $(j+1)$, on teste

$$(H_{0,k}) : \widehat{\mu}_k^{(j)} = \widehat{\mu}_{k+1}^{(j)} \quad \text{versus} \quad (H_{1,k}) : \widehat{\mu}_k^{(j)} \neq \widehat{\mu}_{k+1}^{(j)}$$

où $\widehat{\mu}_k^{(j)}$ est la moyenne empirique calculée sur l'intervalle $I = [\tau_k^{(j)} + \varepsilon_0, \tau_{k+1}^{(j)} - \varepsilon_0]$ où ε_0 désigne l'incertitude de la localisation de la rupture réelle $\tau_k^{(j)}$, il est choisit selon la formule 2.1 et la figure 1 ci-dessous.
On applique alors le test de Student :

$$t_k^{(j)} = \frac{\widehat{\mu}_k^{(j)} - \widehat{\mu}_{k-1}^{(j)}}{\sqrt{\frac{\left(S_{k-1}^{(j)}\right)^2}{N_{k-1}^{(j)}} + \frac{\left(S_k^{(j)}\right)^2}{N_k^{(j)}}}}, \tag{1.2}$$

où $\widehat{\mu}_k^{(j)} := \widehat{\mu}(X, [\tau_k^{(j)} + \varepsilon_0, \tau_{k+1}^{(j)} - \varepsilon_0])$, $N_k^{(j)} := \tau_{k+1}^{(j)} - \tau_k^{(j)} - 2\varepsilon_0$, et l'écart type calculé sur l'intervalle I

$$S_k^{(j)} = \sqrt{\left(\frac{1}{N_k^{(j)}} \sum_{t=\tau_k^{(j)}+\varepsilon_0}^{\tau_{k+1}^{(j)}-\varepsilon_0} X_t^2\right) - \left(\widehat{\mu}_k^{(j)}\right)^2}.$$

Ensuite, nous la comparons à une *t-value* critique fixée t_{c_j}. Si $|t_k^{(j)}| < t_{c_j}$ alors on élimine $\tau_k^{(j)}$ ainsi on obtient une sous-famille de points de rupture $T_{j+1} = \left\{\tau_k^{(j)} \in T_k \text{ tel que } |t_k^{(j)}| > t_{c_j}\right\}$ tel que $T_{j+1} \subset T_j$. A la première itération, on utilise l'ensemble T_1 obtenu à l'étape 1 avec une t-value critique $t_{c_0} = 1$. Ensuite, nous l'augmentons à chaque itération $t_{c_{j+1}} = t_{c_j} + 0.1$. On arrête l'itération dès qu'on a $t_{c_j} > t_{FA}$, où t_{FA} est l'estimation de la t-value maximale des fausses alarmes au risque α_2.

2 Automatisation du choix des extra-paramètres

2.1 Résultats théoriques

1. **Choix du paramètre de l'incertitude de la localisation ε_0**
Les simulations de Monte-Carlo montrent que

$$\varepsilon_0 = \nu(SNR) \times \left(\frac{\sigma}{\delta_0}\right)^2 \tag{2.1}$$

où ν est tel que $\nu(SNR) \leq 10$ ([3], Remark. 3, p.225) avec $SNR = \delta_0/\sigma$, voir Figure 1 ci-dessous

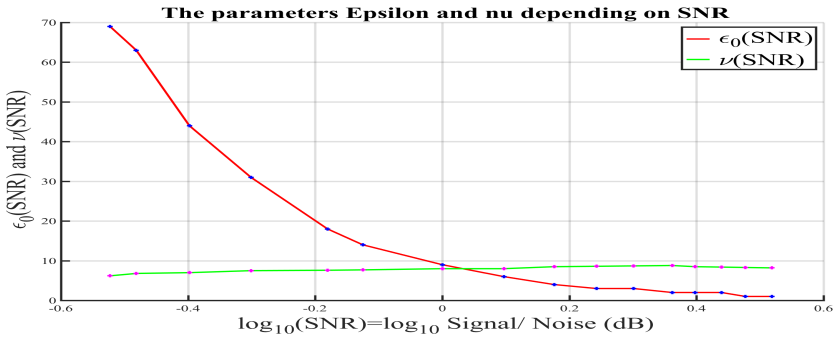

FIGURE 1 – *L'incertitude sur la localisation du vrai positif en fonction du ratio SNR et* $\nu(SNR)$*, pour* $\alpha_{ND} = 0.01$ *et* $\alpha_{FA} = 0.05$.

2. **Choix de la t-value critique t_c**

Proposition 2.1 *Soit* $\tau_k^{(j)} \in T_j$ *un point de rupture potential selectionné à l'étape* j *and* $t_k^{(j)}$ *sa t-value calculée avec la formule (1.2).*
i) Si $\tau_k^{(j)}$ *est une fausse alarme alors* $t_k^{(j)} \sim St_{N_k+N_{k+1}-2}$
ii) Si $\tau_k^{(j)}$ *est un point de rupture réel alors* $t_k^{(j)} \sim \Delta_k + St_{N_k^{(j)}+N_{k+1}^{(j)}-2}$
avec St_ν *la loi de Student de degré* ν*, et*

$$\Delta_k^{(j)} = \frac{\delta_k}{\sqrt{S_k^2/N_k^{(j)} + S_{k+1}^2/N_{k+1}^{(j)}}} \tag{2.2}$$

Proposition 2.2 *Supposons* $A < L_0$*. Soit* $(\alpha_1, \alpha_2) \in (0,1)^2$ *les risques de non-détection et de fausse alarme respectivement.*
i) Si

$$t_c \geq \Psi^{-1}\left(\frac{1-(1-\alpha_2)^{1/NFA}}{2}\right),$$

où NFA *est le nombre de fausses alarmes. Alors*

$$\mathbb{P}(|t_k^{(j)}| \geq t_c, \text{ pour tout } \tau_k^{(j)} \text{ fausse alarme}) = \alpha_2$$

ii) Soit $N_0^{(j)} = \min\{N_k^{(j)}, pour\ k = 1, \ldots, \widehat{K}^{(j)}\}$ *et* $\Delta_0^{(j)} := \left(\frac{\delta_0}{\sigma}\right) \times \sqrt{\frac{N_0^{(j)}}{2}}$ *pout tout* $k = 1, \ldots, \widehat{K}^{(j)}$*. Si*

$$\Delta_0^{(j)} \geq t_c + \Phi^{-1}\left((1-\alpha_1)^{1/K}\right),$$

où K *est le nombre réel de points de rupture. Alors*

$$\mathbb{P}(|t_k^{(j)}| \geq t_c, \text{ pour tout } \tau_k^{(j)} \text{ vrai positif}) = (1-\alpha_1), \tag{2.3}$$

2.2 Choix Pratiques

Proposition 2.3 *Supposons que le ratio signal-bruit* SNR *et* n *connus. Soit* $(\alpha_1, \alpha_2) \in (0,1)^2$ *le risque de non-detection et de fausse alarme respectivement. On définit*

$$
\begin{aligned}
x_1(SNR) &= \Psi^{-1}\left(\frac{1-(1-\alpha_2)^{2/Kmax}}{2}\right)+\Phi^{-1}\left((1-\frac{\alpha_1}{2})^{2/Kmax}\right),\\
x_2(SNR) &= t_{c_0}+\Phi^{-1}\left((1-\frac{\alpha_1}{2})^{2/Kmax}\right) \qquad par\ exemple \quad t_{c_0}=1,\\
A_2(SNR) &:= y+2\nu(SNR)\times(SNR)^{-2} \qquad avec \quad y=32,\\
Kmax(SNR) &= \left\lfloor \frac{n}{A_2(SNR)}-1\right\rfloor,\\
\lambda_1(SNR) &:= 1-\frac{\Phi^{-1}\left((1-\frac{\alpha_1}{2})^{1/Kmax}\right)}{\sqrt{x_1(SNR)^2+\nu\,(SNR)}},\\
\lambda_2(SNR) &:= 1-\frac{\Phi^{-1}\left((1-\frac{\alpha_1}{2})^{1/Kmax}\right)}{\sqrt{x_2(SNR)^2+\nu\,(SNR)}},\\
A_1(SNR) &:= \frac{2}{(1-\lambda_1)^2}\times\Phi^{-1}\left((1-\frac{\alpha_1}{2})^{1/Kmax}\right)^2\times(SNR)^{-2},\\
A_3(SNR) &:= \frac{2}{(1-\lambda_2)^2}\times\Phi^{-1}\left((1-\frac{\alpha_1}{2})^{1/Kmax}\right)^2\times(SNR)^{-2},\\
A_4(SNR) &:= \left[2x_2(SNR)^2+2\nu(SNR)\right]\times(SNR)^{-2},\\
C_1 &= \lambda(SNR)\times SNR\times\widehat{\sigma}_1,\\
\delta_1 &= SNR\times\widehat{\sigma}_1,
\end{aligned}
$$

où $\widehat{\sigma}_1=\delta_0/\widehat{SNR}$. *If*

$$A\geq A_1(SNR)\vee A_3(SNR)\vee A_4(SNR),$$

alors la méthode iterative FDtV est exacte au risque α_1,α_2, *pour toute résolution en temps* $L_0\geq A$ *et tout signal* $\delta_0\geq\delta_1$.

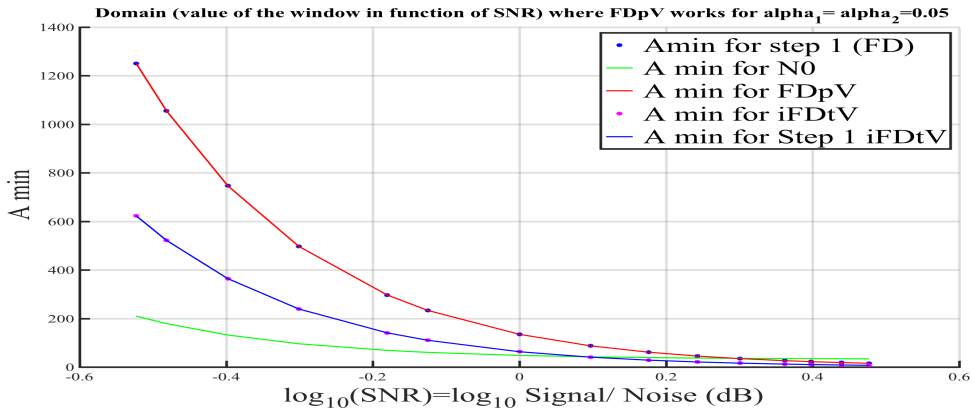

FIGURE 2 – A_1, A_2, A_3 *and* A_4 *en fonction du ratio* SNR.

La figure 2 ci dessus montre l'évolution de la taille de la fenêtre A en fonction du ratio SNR. On constate que la valeur minimale de la taille de la fenêtre A pour la méthode i-FDtV a été divisée par deux par rapport à celle de la méthode FDpV [4].

Conclusion

Notre analyse suggère une méthode générale pour optimiser les paramètres de la fonction Dérivée Filtrée ainsi que ceux des itérations de l'étape des éliminations. Les extra-paramètres des deux étapes de la méthode I-FDtV sont calculés de manière automatique. De plus le résultat est exact, donc la méthode est optimale. Autrement dit, à la première étape, on calcule la fonction Dérivée Filtrée telle que le risque de non-détection est de α_2 et celui de fausses alarmes est α_1. De plus, à la seconde étape, les éliminations se font de manière itérative en incrémentant le seuil de la t-value t_c.

Références

[1] M. Basseville and I. V. Nikiforov, *Detection of abrupt changes : theory and application*, Prentice Hall Information and System Sciences Series, Prentice Hall Inc., Englewood Cliffs, NJ, 1993. MR MR1210954 (95g :62153)

[2] A. Benveniste and M. Basseville, *Detection of abrupt changes in signals and dynamical systems : some statistical aspects*, Analysis and optimization of systems, Part 1 (Nice, (1984), Lecture Notes in Control and Inform. Sci., vol. 62, Springer, Berlin, 1984, pp. 145–155. MR MR876686

[3] P. R. Bertrand, *A local method for estimating change points : the "hat-function"*, Statistics **34** (2000), no. 3, 215–235. MR MR1802728 (2001j :62032)

[4] P. R. Bertrand, M. Fhima, and A Guillin, *Off-line detection of multiple change points by the filtered derivative with p-value method*, Sequential Analysis **30 (2)** (2011), 172–207.

[5] S. Bianchi, A. Pantanella, and A. Painese, *Efficient markets and behacioral finance : a comprehensive multifractal model*, Advances in Complex Systems **18** (2015).

[6] S. Grun, M. Diesmann, and A. Aertsen, *Unitary events in multiple single neuron spiking activity : I. detection and significance.*, Neural Comput. **14** (Jan, 2002).

[7] N. Khalfa, P.R. Bertrand, G. Boudet, A. Chamoux, and V. Billat, *Heart rate regulation processed through wavelet analysis and change detection. some case studies.*, Acta Biotheoretica. **60** (2012), 109–129.

[8] M. Messer, M. Kirchener, J. Shiemann, J. Roeper, R. Neininger, and G. Schneider, *A multiple filter test for the detection of rate changes in renewal processes with varying variance*, The Annals of Applied Statistics **8** (2015), no. 4, 2027–2067.

[9] G. Schneider, *Messages of oscillatory correlograms : A spike train model*, Neural Comput **20** (2008), 1211–1238.

[10] Y. S. Soh and V. Chandrasekaran, *High-dimensional change-point estimation : Combining filtering with convex optimization*, Article in Applied and Computational Harmonic Analysis (2015).

[11] W. Xiao, W. Zhang, and X. Zhang, *Parameter identification for drift fractional brownian motions with application to the chinese stock markets*, Communications in Statistics - Simulation and Computation **44** (2015), 2117–2136.