6 Détection de ruptures

La détection de *ruptures* dans les signaux et systèmes est un problème particulièrement important dans la mesure où, souvent, c'est précisément dans le changement brusque que se concentre une part prépondérante de l'information. Parmi quelques applications, on peut citer tous les problèmes de *segmentation* (reconnaissance de début et fin d'événement, par exemple en parole ou en bio-médical, tri en zones homgènes, détection de contours ligne par ligne en vidéo, ...) ainsi que ceux de *surveillance* (changement de modes vibratoires, apparition de chocs, ...).

Le problème de changement le plus simple est celui du saut de moyenne, mais on peut encore envisager le cas d'un saut de variance ou encore tout type de variation dans les paramètres d'un modèle.

6.1 Structure de détection

Une formulation "hors-ligne" du problème de détection de rupture est la suivante : on dispose de N observations $y_1, \ldots, y_N$ et on désire tester l'homogénéité de la séquence ou, en d'autres termes, détecter si un changement (au plus) a eu lieu en un échantillon de la séquence que l'on notera y_r.

On peut donc écrire :

$$\left\{ \begin{array}{lcl} H_0 & : & y_i \in p_0 \quad ; \quad 1 \leq i \leq N \\ H_1 & : & y_i \in p_0 \quad ; \quad 1 \leq i \leq r-1 \\ & & y_i \in p_1 \quad ; \quad r \leq i \leq N \end{array} \right.$$

en appelant p_0 la densité de probabilité des observations *avant* changement et p_1 celle *après* changement.

En supposant les échantillons y_i indépendants, on a :

$$p(y_1, \ldots, y_N | H_0) = \prod_{i=1}^{N} p_0(y_i)$$

et

$$p(y_1, \ldots, y_N | H_1) = \prod_{i=r-1}^{N} p_0(y_i). \prod_{i=r}^{N} p_1(y_i).$$

Le rapport de vraisemblance s'écrit par suite :

$$\Lambda(y_1, \ldots, y_N) = \prod_{i=r}^{N} \frac{p_1(y_i)}{p_0(y_i)},$$

d'où la règle de décision générale :

$$\prod_{i=r}^{N} \frac{p_1(y_i)}{p_0(y_i)} \underset{H_0}{\overset{H_1}{\gtrless}} \eta.$$

Si l'on suppose maintenant que l'on spécialise le problème au *saut de moyenne* (à variance inchangée) d'un processus *gaussien* en posant

$$p_k = \mathcal{N}(\mu_k, \sigma); k = 0, 1,$$

on en déduit que

$$\begin{aligned}
\Lambda(y_1, \ldots, y_N) &= \prod_{i=r}^{N} \exp\left\{ -\frac{1}{2\sigma^2} \left[(y_i - \mu_1)^2 - (y_i - \mu_0)^2 \right] \right\} \\
&= \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=r}^{N} \left[(y_i - \mu_1)^2 - (y_i - \mu_0)^2 \right] \right\},
\end{aligned}$$

d'où la log-vraisemblance :

$$\log \Lambda(y_1, \ldots, y_N) = \frac{\mu_1 - \mu_0}{\sigma^2} \sum_{i=r}^{N} \left(y_i - \frac{\mu_1 + \mu_0}{2} \right).$$

De façon à mettre l'accent sur le changement éventuel par rapport à la moyenne initiale μ_0, il est commode d'introduire la quantité :

$$S_k^l(\mu_0, \nu) := \nu \sum_{i=k}^{l} \left(y_i - \mu_0 - \frac{\nu}{2} \right)$$

selon laquelle :

$$\log \Lambda(y_1, \ldots, y_N) = \frac{1}{\sigma^2} S_r^N(\mu_0, \nu).$$

L'instant de changement r étant *a priori* inconnu, on peut le remplacer par son estimation au sens du maximum de vraisemblance :

$$\hat{r} = \arg \max_{1 \leq r \leq N} \Lambda(y_1, \ldots, y_N)$$

et donc à :

$$\hat{r} = \arg \max_{1 \leq r \leq N} S_r^N(\mu_0, \nu).$$

Ceci implique que le test de détection de saut de moyenne s'écrit :

$$g_N := \max_r S_r^N(\mu_0, \nu) \begin{array}{c} H_1 \\ > \\ < \\ H_0 \end{array} \eta.$$

Il s'agit dans ce cas d'un test *hors-ligne* et *rétrograde* relatif à l'écart entre les observations et la moyenne (demi-somme) de μ_0 et μ_1. Il est néanmoins possible de transformer ce test rétrograde en un test *progressif* en remarquant que :

$$\max_r S_r^N(\mu_0, \nu) = S_1^N(\mu_0, \nu) - \min_k S_1^k(\mu_0, \nu).$$

En remplaçant l'indice *fixe* N par un indice *courant* n, ceci permet de reformuler le test en une version *séquentielle* pouvant opérer *en ligne* (test dit "de Page-Hinkley").

La détetion de saut de moyenne a lieu lorsque

$$g_n = S_1^n(\mu_0, \nu) - \min_{1 \leq k \leq n} S_1^k(\mu_0, \nu) > \eta,$$

l'estimée de l'instant de rupture étant alors fournie par le *dernier indice* pour lequel la valeur minimum de $S_1^n(\mu_0, \nu)$ a été atteinte.

Lorsque le saut de moyenne attendu est inconnu (en valeur algébrique), une possibilité est :

1. de définir *a priori* un *saut minimum* d'amplitude ν_m ;

2. d'utiliser *deux tests* en parallèle :

   - l'un pour une *augmentation* :

   $$\begin{aligned} U_0 &= 0 \\ U_n &= \sum_{k=1}^{n} \left(y_k - \mu_0 - \frac{\nu_m}{2} \right) ; n \geq 1 \\ m_n &= \min_{0 \leq k \leq n} U_k, \end{aligned}$$

   avec détection si $U_n - m_n > \eta$.

- l'autre pour une *diminution* :

$$\begin{aligned} T_0 &= 0 \\ T_n &= \sum_{k=1}^{n} \left(y_k - \mu_0 + \frac{\nu_m}{2} \right) ; n \geq 1 \\ M_n &= \max_{0 \leq k \leq n} T_k, \end{aligned}$$

avec détection si $M_n - T_n > \eta$.

6.2 Performances

Dans le cas d'une détection séquentielle, deux types de performances sont recherchées :

1. la minimisation du *retard moyen à la détection* R, c'est-à-dire du temps qui sépare l'instant effecfif de changement de la prise de décision correspondante ;

2. la maximisation du *temps moyen entre fausses alarmes* F.

Ces deux exigences sont évidemment contradictoires et, en pratique, on cherchera par exemple à construire des tests optimaux en ce sens qu'ils minimisent R pour un F donné. Le test de Page-Hinkley est précisément optimal en ce sens. On montre alors que le comportement asymptotique de R est de la forme :

$$R \sim \frac{2\sigma^2}{\nu^2} \log F$$

lorsque $F \to \infty$.

6.3 Exemples

`mean_change.m` (test de Page-Hinkley) —