BRCP(Bayesian Regression Change Point) for Anomaly Detection

For the changepoint at time τ, we assume error function is independent with time, then:

$$y = \begin{cases} f_1(x_t) + \varepsilon_{1,t}, & \text{if } t \le \tau \\ f_2(x_t) + \varepsilon_{2,t}, & \text{if } t > \tau \end{cases} \qquad (1)$$

τ is the time for change point, $f_1(x_t)$ and $f_2(x_t)$ are polynomial fuction,y is obvious point value,t is time series index,ε_1、ε_2 are timeseries data with mean: 0 and variance: σ_1 σ_2.
The detail is below：

Initialization：

$$l = 0, \alpha^{(0)} = \left\{\alpha_1^{(0)}, \alpha_2^{(0)}, \cdots, \alpha_n^{(0)}\right\} = \{1.0/n, 1.0/n, \cdots, 1.0/n\}$$

l is iteration number, $\alpha_i^{(0)}$ stand for the probability that time point i is a change point, n is the size of the time series data.

Step 1：Calculate μ_{kj} with formula（2）;

$$\mu_{kj} = \begin{cases} \sum_{i=j}^{n-1} \alpha_i, & \text{if } k = 1 \text{ and } 1 \le j \le n-1 \\ 0, & \text{if } k = 1 \text{ and } j = n \\ 1 - \mu_{1j}, & \text{if } k = 2 \end{cases} \qquad (2)$$

$k = \{1,2\}$, $j = \{1,2,\cdots,n\}$, μ_{kj} denote the probability for $\tau \le j$.

Step 2：Calculate β_k with formula（3）;

$$\beta_k = (X^T U_k X)^{-1} X^T U_k Y \qquad (3)$$

$X = [x_1, x_2, \cdots, x_n]$, $U_k = \mathrm{diag}([\mu_{k1}, \mu_{k2}, \cdots, \mu_{kn}])$, $Y = [y_1, y_2, \cdots, y_n]$ and β_k need meet the condition below:

$$y_t = f_k(x_t) + \varepsilon_{k,t} = x_t \beta_k + \varepsilon_{k,t} \qquad (4)$$

Step 3：Calculate d_i, with formula (5)：

$$d_i = \sum_{j=1}^{i} \log\left(g_1(\varepsilon_{1,j})\right) + \sum_{j=i+1}^{n} \log\left(g_2(\varepsilon_{2,j})\right) \qquad (5)$$

g_1 and g_2 are the distribution function of the noise. If the distribution function is Ga ussian, the formula (5) can rewrite as：

$$d_i = \sum_{j=1}^{i} \left(\frac{{\varepsilon_{1,j}}^2}{sd(\varepsilon_1)^2} - 0.5\log\left(sd(\varepsilon_1)\right)\right) + \sum_{j=i+1}^{n} \frac{{\varepsilon_{2,j}}^2}{sd(\varepsilon_2)^2} - 0.5\log\left(sd(\varepsilon_2)\right) \quad (6)$$

when $f_k(x_t)$ is a good prediction, we have $\frac{{\varepsilon_{1,j}}^2}{sd(\varepsilon_1)^2} \approx {\varepsilon_{1,j}}^2$ and $\frac{{\varepsilon_{2,j}}^2}{sd(\varepsilon_2)^2} \approx {\varepsilon_{2,j}}^2$, then :

$$d_i \approx \sum_{j=1}^{i} \left({\varepsilon_{1,j}}^2 - 0.5\log\left(sd(\varepsilon_1)\right)\right) + \sum_{j=i+1}^{n} {\varepsilon_{2,j}}^2 - 0.5\log\left(sd(\varepsilon_2)\right) \quad (7)$$

Step 4：Calculate $\alpha^{(l+1)} = \left\{\alpha_1^{(l+1)}, \alpha_2^{(l+1)}, \cdots, \alpha_n^{(l+1)}\right\}$ with formula (8)：

$$\alpha_i^{(l+1)} = \frac{\exp(-{d_i}^2)}{\sum_{j=1}^{n-1} \exp(-{d_j}^2)} \qquad (8)$$

Step 5：If $\left\|\alpha^{(l+1)} - \alpha^{(l)}\right\| < \epsilon$,then stop the iteration,else come back to step 1.

Step 6：$CP = \underset{i}{\mathrm{argmax}}\, \alpha_i^{(l)}$.