veiledning

Per August Jarval Moen

4/5/2022

Setup

We observe p-dimensional Gaussian vectors

$$X_i \sim \mathrm{N}(\mu_i, I_p),$$

for $i = 1, ..., n$.

There are K change-points $\eta_1, \ldots, \eta_K$. At the v-th change-point η_v, denote the change by

$$\theta_v = \mu_{\eta_v+1} - \mu_{\eta_v}.$$

Minimax testing

Suppose $K = 1$. Let $(s, e] = [1, \ldots, \eta_1, \ldots, n]$. Let $\Delta_1 = \min(\eta_1, n - \eta_1)$.

Let $s = \|\theta_1\|_0$, i.e. the number of components that undergo a change at the change-point.

If $s \leq \sqrt{p \log(n)}$ we call the change sparse, and dense otherwise.

The minimax testing "rate" for a change-point in $(s, e]$ at any location is

$$\Delta_1 \|\theta_1\|^2 \gtrsim \begin{cases} s \log\left(\frac{ep \log(n)}{s^2}\right) \vee \log(n), & \text{if sparse} \\ \sqrt{p \log n} & \text{, if dense} \end{cases}$$

At $s = \sqrt{p \log n}$ the rates are the same.

Setup

For any segment $(s, e] = [s+1, e]$ and any $b \in [s+1, ..., e-1]$, let

$$\mathcal{C}^b_{(s,e]}(X_j) = \mathrm{CUSUM}^b_{(s,e]}(X_j).$$

(Cusum of j-th component, ie weighted difference in mean between $X_{j,s+1}, \dots, X_{j,b}$ and $X_{j,b+1}, \dots, X_{j,e}$).

If there is no change-point in $(s, e]$, then

$$\mathcal{C}^b_{(s,e]}(X_j) \overset{i.i.d.}{\sim} \mathrm{N}(0, 1),$$

for $j = 1, \dots, p$.

Test statistic

For any sparsity level t, let

$$a(t)^2 := \begin{cases} 4\log\left(\frac{ep\log(n)}{t^2}\right) & \text{, if } t < \sqrt{p\log(n)} \\ 0 & \text{, else} \end{cases}$$

$$T^b_{(s,e]}(t) := \sum_{j=1}^{p}\left(\mathcal{C}^b_{(s,e]}(X_j)^2 - \nu_{a(t)}\right) I_{\left\{\left|\mathcal{C}^b_{(s,e]}(X_j)\right| > a(t)\right\}},$$

where $\nu_a = \mathbb{E}\left(Z^2 \mid |Z| > a\right)$, $Z \sim \mathrm{N}(0,1)$.

Behavior

If there is no change, then with probability greater than

$$1 - \frac{1}{n},$$

for all b, and all sparsity levels t, we have

$$T^b_{(s,e]}(t) < \widetilde{C} \begin{cases} t \log\left(\frac{ep\log(n)}{t^2}\right) \vee \log(n), & \text{if } t < \sqrt{p \log n} \\ \sqrt{p \log n} & \text{, else} \end{cases}$$

Behavior

If there is a change at η_1, with sparsity level s, and the minimax rate is "satisfied", then

$$T^{\eta_1}_{(s,e]}(s) > \widetilde{C} \begin{cases} s \log\left(\frac{ep\log(n)}{s^2}\right) \vee \log(n), & \text{if } s < \sqrt{p \log n} \\ \sqrt{p \log n} & \text{, else,} \end{cases}$$

also with probability larger than $1 - \frac{1}{n}$.

Remark

For the dense case, as long as $p \geq \log(n)$, the minimum requirement for testing existence of a change-point is the same as the condition for consistenlty estimating the location.

Idea for change-point localization