Optimal Covariance Change Point Detection in High Dimension

Joint work with Daren Wang and Alessandro Rinaldo, CMU

Yi Yu
School of Mathematics, University of Bristol

- Review of change point detection problem
- Setup
- Main results
    - Two consistent change point detection algorithms
    - Phase transition - minimax optimality

World War II, Abraham Wald (SPRT, 1945)

CUmulative SUM (CUSUM, Page, 1954)

$$X_i \sim \begin{cases} F_1, & i = 1, \ldots, t_1 - 1, \\ F_2, & i = t_1, \ldots, t_2 - 1, \\ \ldots, & \\ F_K, & i = t_{K-1}, \ldots, n, \end{cases}$$

where $F_1, \ldots, F_K$, $K \geq 2$, are distribution functions, $F_{k-1} \neq F_k$, $k = 2, \ldots, K$, and $\{t_1, \ldots t_{K-1}\} \subset \{2, \ldots, n-1\}$ are unknown change point locations.

$\{X_i = f_i + \varepsilon_i\}_{i=1}^n$ is a univariate time series, and $\{f_i\}$ is piecewise constant univariate signal.

Least squares estimators (error rate $O_p(1)$, Yao and Au, 1989), wavelets, fused lasso, trend filtering, ...

Let

$$X_i = f_i + \varepsilon_i,\ i = 1, \ldots, n,$$

where $\{f_i\}$ is univariate piecewise constant. For an interval (s, e) with $0 \leq s < e - 1 < n$, the CUSUM statistic is defined as

$$\widetilde{X}_t^{s,e} = \sqrt{\frac{e-t}{(e-s)(t-s)}} \sum_{i=s+1}^{t} X_i - \sqrt{\frac{t-s}{(e-s)(e-t)}} \sum_{i=t+1}^{e} X_i,\ t = s+1, \ldots, e-1.$$

To find mean change points in $\{1, \ldots, n\}$, binary segmentation (BS, Vostrikova, 1981) searches for

$$a = \max_{t=1,\ldots,n-1} \left|\widetilde{X}_t^{0,n}\right| \quad \text{and} \quad b \in \operatorname*{argmax}_{t=1,\ldots,n-1} \left|\widetilde{X}_t^{0,n}\right|.$$

Wild binary segmentation (WBS, Fryzlewicz, 2014) randomly draws a collection of intervals (s_m, e_m), $m = 1, \ldots, M$, satisfying certain conditions, calculates their corresponding a_m and b_m, and finds

$$m^* = \operatorname*{argmax}_{m=1,\ldots,M} a_m.$$

Let

$$X_i = f_i + \varepsilon_i,\ i = 1, \ldots, n,$$

where $\{f_i\}$ is univariate piecewise constant. For an interval (s, e) with $0 \leq s < e - 1 < n$, the CUSUM statistic is defined as

$$\widetilde{X}_t^{s,e} = \sqrt{\frac{e-t}{(e-s)(t-s)}} \sum_{i=s+1}^{t} X_i - \sqrt{\frac{t-s}{(e-s)(e-t)}} \sum_{i=t+1}^{e} X_i,\ t = s+1, \ldots, e-1.$$

To find mean change points in $\{1, \ldots, n\}$, binary segmentation (BS, Vostrikova, 1981) searches for

$$a = \max_{t=1,\ldots,n-1} \left|\widetilde{X}_t^{0,n}\right| \quad \text{and} \quad b \in \operatorname*{argmax}_{t=1,\ldots,n-1} \left|\widetilde{X}_t^{0,n}\right|.$$

Wild binary segmentation (WBS, Fryzlewicz, 2014) randomly draws a collection of intervals (s_m, e_m), $m = 1, \ldots, M$, satisfying certain conditions, calculates their corresponding a_m and b_m, and finds

$$m^* = \operatorname*{argmax}_{m=1,\ldots,M} a_m.$$

Spacing parameter: Δ, the minimal distance between two consecutive change points.

In BS, $\Delta \geq Cn^{1-\beta}$, $0 \leq \beta < 1/8$; in WBS, $\Delta \geq \log(n)$.

Multi or high-dimensional settings

Covariance change point detection

Wald (1945), Picard (1985), Inclan and Tiao (1994), Gombay et al. (1996), Berkes et al. (2009), Aue et al. (2009), Avanesov and Buzun (2016), Barigozzi et al. (2016), *etc.*

Hypothesis testing for high dimensional covariance matrices .

Spacing parameter: Δ, the minimal distance between two consecutive change points.

In BS, $\Delta \geq Cn^{1-\beta}$, $0 \leq \beta < 1/8$; in WBS, $\Delta \geq \log(n)$.

Multi or high-dimensional settings

Covariance change point detection

Wald (1945), Picard (1985), Inclan and Tiao (1994), Gombay et al. (1996), Berkes et al. (2009), Aue et al. (2009), Avanesov and Buzun (2016), Barigozzi et al. (2016), *etc.*

Hypothesis testing for high dimensional covariance matrices .

Assumption 1 (Model) Let $X_1, \ldots, X_n \in \mathbb{R}^p$ be independent sub-Gaussian random vectors such that $\mathbb{E}(X_i) = 0$, $\mathbb{E}(X_i X_i^\top) = \Sigma_i$ and $\|X_i\|_{\psi_2} \leq B < \infty$. Let $\{\eta_k\}_{k=0}^{K+1} \subset \{0, \ldots, n\}$ be a collection of change points, such that $\eta_0 = 0$ and $\eta_{K+1} = n$ and that

$$\Sigma_{\eta_k+1} = \Sigma_{\eta_k+2} = \ldots = \Sigma_{\eta_{k+1}}, \text{ for any } k = 0, \ldots, K.$$

Assume the jump size $\kappa = \kappa(n)$ and the spacing $\Delta = \Delta(n)$ satisfy that

$$\inf_{k=1,\ldots,K+1} \{\eta_k - \eta_{k-1}\} \geq \Delta > 0,$$

and

$$\|\Sigma_{\eta_k} - \Sigma_{\eta_{k-1}}\|_{\mathrm{op}} = \kappa_k \geq \kappa > 0, \text{ for any } k = 1, \ldots, K+1.$$

The difficulty of change point localisation problem

- increases in p and B,
- decreases in Δ and κ.

Assumption 1 (Model) Let $X_1, \ldots, X_n \in \mathbb{R}^p$ be independent sub-Gaussian random vectors such that $\mathbb{E}(X_i) = 0$, $\mathbb{E}(X_i X_i^\top) = \Sigma_i$ and $\|X_i\|_{\psi_2} \leq B < \infty$. Let $\{\eta_k\}_{k=0}^{K+1} \subset \{0, \ldots, n\}$ be a collection of change points, such that $\eta_0 = 0$ and $\eta_{K+1} = n$ and that

$$\Sigma_{\eta_k+1} = \Sigma_{\eta_k+2} = \ldots = \Sigma_{\eta_{k+1}}, \text{ for any } k = 0, \ldots, K.$$

Assume the jump size $\kappa = \kappa(n)$ and the spacing $\Delta = \Delta(n)$ satisfy that

$$\inf_{k=1,\ldots,K+1} \{\eta_k - \eta_{k-1}\} \geq \Delta > 0,$$

and

$$\|\Sigma_{\eta_k} - \Sigma_{\eta_{k-1}}\|_{\mathrm{op}} = \kappa_k \geq \kappa > 0, \text{ for any } k = 1, \ldots, K+1.$$

The difficulty of change point localisation problem

- increases in p and B,
- decreases in Δ and κ.

Remarks

- A random variable $X \in \mathbb{R}$ is a sub-Gaussian random variable if

$$\|X\|_{\psi_2} = \sup_{k \geq 1} k^{-1/2}\{\mathbb{E}(|X|^k)\}^{1/k} < \infty.$$

A random vector $X \in \mathbb{R}^p$ is a sub-Gaussian random variable if

$$\|X\|_{\psi_2} = \sup_{v \in \mathcal{S}^{p-1}} \|v^\top X\|_{\psi_2} < \infty,$$

where $\mathcal{S}^{p-1}$ denote the sphere of the unit ball in $\mathbb{R}^p$.

- Note that there exists a relationship between κ and B^2 that

$$\kappa \leq \min_{k=1,\ldots,K+1} \|\Sigma_{\eta_k} - \Sigma_{\eta_{k-1}}\|_{\mathrm{op}} \leq 2 \max_{i=1,\ldots,n} \|\Sigma_i\|_{\mathrm{op}}$$
$$\leq 2 \max_{i=1,\ldots,n} \sup_{v \in \mathcal{S}^{p-1}} \mathbb{E}\left[(v^\top X_i)^2\right] \leq 4\|X_i\|^2_{\psi_2} \leq 4B^2.$$

- It also shows that $\|\Sigma_i\|_{\mathrm{op}} \leq 2B^2$.

Remarks (cont'd)

- Note that there exists a relationship between κ and B^2 that

$$\kappa \leq \min_{k=1,\ldots,K+1} \|\Sigma_{\eta_k} - \Sigma_{\eta_{k-1}}\|_{\mathrm{op}} \leq 2 \max_{i=1,\ldots,n} \|\Sigma_i\|_{\mathrm{op}}$$
$$\leq 2 \max_{i=1,\ldots,n} \sup_{v \in \mathcal{S}^{p-1}} \mathbb{E}\left[(v^\top X_i)^2\right] \leq 4\|X_i\|_{\psi_2}^2 \leq 4B^2.$$

- Signal-to-noise ratio κ/B^2 - invariant with respect to any multiplicity rescaling of the data by an arbitrary non-zero constant.

- Difference between the conditions in the covariance change point detection problems and the mean ones.

Definition 1 (Covariance CUSUM) For $X_1, \ldots, X_n \in \mathbb{R}^p$, an interval (s, e) with $0 \leq s < e - 1 < n$, and $t \in \{s+1, \ldots, e-1\}$, define the covariance CUSUM statistic as follows

$$\widetilde{S}_t^{s,e} = \sqrt{\frac{e-t}{(e-s)(t-s)}} \sum_{i=s+1}^{t} X_i X_i^\top - \sqrt{\frac{t-s}{(e-s)(e-t)}} \sum_{i=t+1}^{e} X_i X_i^\top.$$

Its expected value is

$$\widetilde{\Sigma}_t^{s,e} = \sqrt{\frac{e-t}{(e-s)(t-s)}} \sum_{i=s+1}^{t} \Sigma_i - \sqrt{\frac{t-s}{(e-s)(e-t)}} \sum_{i=t+1}^{e} \Sigma_i.$$

Algorithm 1 Binary Segmentation through Operator Norm. BSOP$((s, e), \tau)$

INPUT: $\{X_i\}_{i=s+1}^{e} \subset \mathbb{R}^{p \otimes (e-s)}$, $\tau > 0$.

Initial FLAG $\leftarrow 0$,

while $e - s > 2p\log(n) + 1$ and FLAG $= 0$ **do**

$a \leftarrow \max_{\lceil s + p\log(n) \rceil \le t \le \lfloor e - p\log(n) \rfloor} \|\widetilde{S}_t^{s,e}\|_{\mathrm{op}}$

if $a \le \tau$ **then**

FLAG $\leftarrow 1$

else

$b \leftarrow \arg\max_{\lceil s + p\log(n) \rceil \le t \le \lfloor e - p\log(n) \rfloor} \|\widetilde{S}_t^{s,e}\|_{\mathrm{op}}$

add b to the collection of estimated change points

BSOP$((s, b-1), \tau)$

BSOP$((b, e), \tau)$

end if

end while

OUTPUT: The collection of estimated change points.

Assumption 2 For a sufficiently large constant $C_\alpha > 0$ and sufficient small constant $c_\alpha > 0$, assume that $\Delta\kappa B^{-2} \geq C_\alpha n^{\Theta}$, $p \leq c_\alpha n^{8\Theta-7}/\log(n)$, where $\Theta \in (7/8, 1]$.

Remark (Generalising Assumption 2) Assumption 2 may be generalised by allowing for different types of scaling in n of κB^{-2}, Δ and p.

Say, $\kappa B^{-2} \succeq n^{\Theta_1}$, $\Delta \succeq n^{\Theta_2}$, and $p\log(n) \preceq n^{\Theta_3}$, for a given triplet of parameters $(\Theta_1, \Theta_2, \Theta_3) \subset S \subset [0, 1]^3$.

Theorem 1 (Consistency of BSOP) Under Assumptions 1 and 2, let $\mathcal{B} = \{\hat{\eta}_k\}_{k=1}^{\widehat{K}}$ be the collection of the estimated change points from the $\mathrm{BSOP}((0, n), \tau)$ algorithm, where $\tau > 0$ satisfies

$$B^2\sqrt{p\log(n)} + 2\sqrt{\epsilon_n}B^2 < \tau < C_1\kappa\Delta(e-s)^{-1/2},$$

for some constant $C_1 \in (0, 1)$ and where

$$\epsilon_n = C_2 B^2 \kappa^{-1} n^{5/2} \Delta^{-2}\sqrt{p\log(n)}$$

for some constant $C_2 > 0$. Then,

$$\mathbb{P}\left\{\widehat{K} = K; \quad \max_{k=1,\ldots,K} |\hat{\eta}_k - \eta_k| \leq \epsilon_n\right\} \geq 1 - n^3 9^p 2n^{-cp},$$

for some constant $c > 0$.

Theorem 1 (Consistency of BSOP) Under Assumptions 1 and 2, let $\mathcal{B} = \{\hat{\eta}_k\}_{k=1}^{\widehat{K}}$ be the collection of the estimated change points from the $\mathrm{BSOP}((0, n), \tau)$ algorithm, where $\tau > 0$ satisfies

$$B^2\sqrt{p\log(n)} + 2\sqrt{\epsilon_n}B^2 < \tau < C_1\kappa\Delta(e-s)^{-1/2},$$

for some constant $C_1 \in (0, 1)$ and where

$$\epsilon_n = C_2 B^2 \kappa^{-1} n^{5/2} \Delta^{-2}\sqrt{p\log(n)}$$

for some constant $C_2 > 0$. Then,

$$\mathbb{P}\left\{\widehat{K} = K; \quad \max_{k=1,\ldots,K} |\hat{\eta}_k - \eta_k| \leq \epsilon_n\right\} \geq 1 - n^3 9^p 2n^{-cp},$$

for some constant $c > 0$.

Theorem 1 (Consistency of BSOP) Under Assumptions 1 and 2, let $\mathcal{B} = \{\hat{\eta}_k\}_{k=1}^{\widehat{K}}$ be the collection of the estimated change points from the $\mathrm{BSOP}((0, n), \tau)$ algorithm, where $\tau > 0$ satisfies

$$B^2\sqrt{p\log(n)} + 2\sqrt{\epsilon_n}B^2 < \tau < C_1\kappa\Delta(e-s)^{-1/2},$$

for some constant $C_1 \in (0, 1)$ and where

$$\epsilon_n = C_2 B^2 \kappa^{-1} n^{5/2} \Delta^{-2} \sqrt{p\log(n)}$$

for some constant $C_2 > 0$. Then,

$$\mathbb{P}\left\{\widehat{K} = K; \quad \max_{k=1,\ldots,K} |\hat{\eta}_k - \eta_k| \leq \epsilon_n\right\} \geq 1 - n^3 9^p 2n^{-cp},$$

for some constant $c > 0$.

Theorem 1 (Consistency of BSOP) Under Assumptions 1 and 2, let $\mathcal{B} = \{\hat{\eta}_k\}_{k=1}^{\widehat{K}}$ be the collection of the estimated change points from the BSOP$((0, n), \tau)$ algorithm, where $\tau > 0$ satisfies

$$B^2 \sqrt{p \log(n)} + 2\sqrt{\epsilon_n} B^2 < \tau < C_1 \kappa \Delta (e - s)^{-1/2},$$

for some constant $C_1 \in (0, 1)$ and where

$$\epsilon_n = C_2 B^2 \kappa^{-1} n^{5/2} \Delta^{-2} \sqrt{p \log(n)}$$

for some constant $C_2 > 0$. Then,

$$\mathbb{P}\left\{ \widehat{K} = K; \quad \max_{k=1,\ldots,K} |\hat{\eta}_k - \eta_k| \leq \epsilon_n \right\} \geq 1 - n^3 9^p 2n^{-cp},$$

for some constant $c > 0$.

Theorem 1 (Consistency of BSOP) Under Assumptions 1 and 2, let $\mathcal{B} = \{\hat{\eta}_k\}_{k=1}^{\widehat{K}}$ be the collection of the estimated change points from the $\mathrm{BSOP}((0, n), \tau)$ algorithm, where $\tau > 0$ satisfies

$$B^2\sqrt{p\log(n)} + 2\sqrt{\epsilon_n}B^2 < \tau < C_1\kappa\Delta(e-s)^{-1/2},$$

for some constant $C_1 \in (0, 1)$ and where

$$\epsilon_n = C_2 B^2 \kappa^{-1} n^{5/2} \Delta^{-2} \sqrt{p\log(n)}$$

for some constant $C_2 > 0$. Then,

$$\mathbb{P}\left\{\widehat{K} = K; \quad \max_{k=1,\ldots,K} |\hat{\eta}_k - \eta_k| \leq \epsilon_n\right\} \geq 1 - n^3 9^p 2n^{-cp},$$

for some constant $c > 0$.

Theorem 1 (Consistency of BSOP) Under Assumptions 1 and 2, let $\mathcal{B} = \{\hat{\eta}_k\}_{k=1}^{\widehat{K}}$ be the collection of the estimated change points from the BSOP$((0, n), \tau)$ algorithm, where $\tau > 0$ satisfies

$$B^2 \sqrt{p \log(n)} + 2\sqrt{\epsilon_n} B^2 < \tau < C_1 \kappa \Delta (e - s)^{-1/2},$$

for some constant $C_1 \in (0, 1)$ and where

$$\epsilon_n = C_2 B^2 \kappa^{-1} n^{5/2} \Delta^{-2} \sqrt{p \log(n)}$$

for some constant $C_2 > 0$. Then,

$$\mathbb{P}\left\{ \widehat{K} = K; \quad \max_{k=1,\ldots,K} |\hat{\eta}_k - \eta_k| \leq \epsilon_n \right\} \geq 1 - n^3 9^p 2n^{-cp},$$

for some constant $c > 0$.

Theorem 1 (Consistency of BSOP) Under Assumptions 1 and 2, let $\mathcal{B} = \{\hat{\eta}_k\}_{k=1}^{\widehat{K}}$ be the collection of the estimated change points from the $\mathrm{BSOP}((0, n), \tau)$ algorithm, where $\tau > 0$ satisfies

$$B^2\sqrt{p\log(n)} + 2\sqrt{\epsilon_n}B^2 < \tau < C_1\kappa\Delta(e-s)^{-1/2},$$

for some constant $C_1 \in (0, 1)$ and where

$$\epsilon_n = C_2 B^2 \kappa^{-1} n^{5/2} \Delta^{-2} \sqrt{p\log(n)}$$

for some constant $C_2 > 0$. Then,

$$\mathbb{P}\left\{\widehat{K} = K; \quad \max_{k=1,\ldots,K} |\hat{\eta}_k - \eta_k| \leq \epsilon_n\right\} \geq 1 - n^3 9^p 2n^{-cp},$$

for some constant $c > 0$.

The behaviours of the population CUSUM statistics $\|\widetilde{\Sigma}_t^{s,e}\|_{\mathrm{op}}$.

Shadow vectors.

Improve?

Improve!

- In Assumption 2, we require $\Delta \succeq n^{7/8}$.

WBS, Fryzlewicz (2014).

- The error rate in Theorem 1 is $\epsilon_n = C_2 B^2 \kappa^{-1} n^{5/2} \Delta^{-2} \sqrt{p \log(n)}$.

Sample splitting, Wang and Samworth (2016).

The behaviours of the population CUSUM statistics $\|\widetilde{\Sigma}_t^{s,e}\|_{\mathrm{op}}$.

Shadow vectors.

Improve?

Improve!

- In Assumption 2, we require $\Delta \succeq n^{7/8}$.

  WBS, Fryzlewicz (2014).

- The error rate in Theorem 1 is $\epsilon_n = C_2 B^2 \kappa^{-1} n^{5/2} \Delta^{-2} \sqrt{p \log(n)}$.

  Sample splitting, Wang and Samworth (2016).

The behaviours of the population CUSUM statistics $\|\widetilde{\Sigma}_t^{s,e}\|_{\mathrm{op}}$.

Shadow vectors.

Improve?

Improve!

- In Assumption 2, we require $\Delta \succeq n^{7/8}$.

WBS, Fryzlewicz (2014).

- The error rate in Theorem 1 is $\epsilon_n = C_2 B^2 \kappa^{-1} n^{5/2} \Delta^{-2} \sqrt{p \log(n)}$.

Sample splitting, Wang and Samworth (2016).

Algorithm 2 Principal Component Estimation $\mathrm{PC}(\{X_i\}_{i=1}^n, \{(\alpha_m, \beta_m)\}_{m=1}^M)$

INPUT: $\{X_i\}_{i=1}^n$, $\{(\alpha_m, \beta_m)\}_{m=1}^M$
  for $m = 1, \ldots, M$ **do**
    if $\beta_m - \alpha_m > 2p\log(n) + 1$ **then**
      $d_m \leftarrow \arg\max_{\lceil \alpha_m + p\log(n) \rceil \le t \le \lfloor \beta_m - p\log(n) \rfloor} \|\widetilde{S}_t^{\alpha_m, \beta_m}\|_{\mathrm{op}}$
      $u_m \leftarrow \arg\max_{\|v\|_2 = 1} |v^\top \widetilde{S}_{d_m}^{\alpha_m, \beta_m} v|$
    else
      $u_m \leftarrow 0$
    end if
  end for
OUTPUT: $\{u_m\}_{m=1}^M$.

Algorithm 3 Wild Binary Segmentation through Independent Projection. WBSIP$((s, e), \{(\alpha_m, \beta_m)\}_{m=1}^M, \tau, \delta)$

INPUT: Two independent samples $\{W_i\}_{i=1}^n$, $\{X_i\}_{i=1}^n$, τ, δ.

$\{u_m\}_{m=1}^M = PC(\{W_i\}_{i=1}^n, \{(\alpha_m, \beta_m)\}_{m=1}^M)$
for $i \in \{s, \ldots, e\}$ **do**
  for $m = 1, \ldots, M$ **do**
    $Y_i(u_m) \leftarrow (u_m^\top X_i)^2$
  end for
end for
for $m = 1, \ldots, M$ **do**
  $(s'_m, e'_m) \leftarrow [s, e] \cap [\alpha_m, \beta_m]$ and $(s_m, e_m) \leftarrow (\lceil s'_m + \delta \rceil, \lfloor e'_m - \delta \rfloor)$
  if $e_m - s_m \geq 2\log(n) + 1$ **then**
    $b_m \leftarrow \arg\max_{s_m + \log(n) \leq t \leq e_m - \log(n)} |\widetilde{Y}_t^{s_m, e_m}(u_m)|$
    $a_m \leftarrow |\widetilde{Y}_{b_m}^{s_m, e_m}(u_m)|$
  else
    $a_m \leftarrow -1$
  end if
end for
$m^* \leftarrow \arg\max_{m=1,\ldots,M} a_m$
if $a_{m^*} > \tau$ **then**
  add b_{m^*} to the set of estimated change points
  WBSRP$((s, b_{m*}), \{(\alpha_m, \beta_m)\}_{m=1}^M, \tau)$
  WBSRP $((b_{m*} + 1, e), \{(\alpha_m, \beta_m)\}_{m=1}^M, \tau)$
end if

OUTPUT: The set of estimated change points.

Assumption 3 There exists a sufficiently large absolute constant $C > 0$ such that $\Delta\kappa^2 \geq Cp\log(n)B^4$.

Remarks

- $p\log(n) \leq \Delta(C^{-1}(\kappa B^{-2})^2)$.

Theorem 2 (Consistency of WBSIP) Let Assumptions 1 and 3 hold, and let $\{(\alpha_m,\ \beta_m)\}_{m=1}^{M} \subset (0, n)^{\otimes M}$ be a collection of intervals whose end points are drawn independently and uniformly from $\{1, \ldots, n\}$ and such that $\max_{1 \leq m \leq M}(\beta_m - \alpha_m) \leq C\Delta$ for an absolute constant $C > 0$. Set

$$\epsilon_n = C_1 B^4 \log(n) \kappa^{-2}$$

for a $C_1 > 0$. Suppose there exist $c_2, c_3 > 0$, sufficiently small, such that the input parameters τ and δ satisfy

$$B^2 \sqrt{\log(n)} < \tau < c_2 \kappa \sqrt{\Delta},$$
$$\epsilon_n < \delta \leq c_3 \Delta.$$

Then the collection of the estimated change points $\mathcal{B} = \{\hat{\eta}_k\}_{k=1}^{\widehat{K}}$ returned by WBSIP$((0, n), \{(\alpha_m, \beta_m)\}_{m=1}^{M}, \tau, \delta)$ satisfies

$$\mathbb{P}\left\{\widehat{K} = K; \quad \max_{k=1,\ldots,K} |\hat{\eta}_k - \eta_k| \leq \epsilon_n\right\}$$
$$\geq 1 - 2n^2 M n^{-c} - n^3 9^p 2n^{-cp} - \exp\left(\log(n/\Delta) - M\Delta^2/(16n^2)\right)$$

for some absolute constants $c > 0$.

Recall Assumption 3: $\Delta \geq CB^4\kappa^{-2}p\log(n)$.

Assumption 4 Let $X_1, \ldots, X_n \in \mathbb{R}^p$ be independent Gaussian random vectors such that $X_i \sim \mathcal{N}_p(0, \Sigma_i)$ with $\|\Sigma_i\|_{\mathrm{op}} \leq 2\sigma$. Let $\{\eta_k\}_{k=0}^{K+1} \subset \{0, \ldots, n\}$ be a collection of change points, such that $\eta_0 = 0$ and $\eta_{K+1} = n$ and that

$$\Sigma_{\eta_k+1} = \Sigma_{\eta_k+2} = \ldots = \Sigma_{\eta_{k+1}}, \text{ for any } k = 1, \ldots, K+1.$$

Assume there exists parameters $\kappa = \kappa(n)$ and $\Delta = \Delta(n)$ such that

$$\inf_{k=1,\ldots,K+1} \{\eta_k - \eta_{k-1}\} \geq \Delta > 0,$$

$$\|\Sigma_{\eta_k} - \Sigma_{\eta_{k-1}}\|_{\mathrm{op}} = \kappa_k \geq \kappa > 0, \text{ for any } k = 1, \ldots, K+1,$$

Lemma 3 Let $\{X_i\}_{i=1}^n$ be a time series satisfies Assumption 4 with only one change point and let $P^n_{\kappa,\Delta,\sigma}$ denote the corresponding joint distribution. Consider the class of distribution

$$\mathcal{P} = \left\{ P^n_{\kappa,\Delta,\sigma} : \frac{\sigma^4 p}{33\kappa^2} \leq \Delta \leq n/3, \kappa \leq \sigma^2/4 \right\},$$

Then

$$\inf_{\hat{\eta}} \sup_{P \in \mathcal{P}} E_P(|\hat{\eta} - \eta|) \geq n/6.$$

- If $\Delta \geq CB^4 p \log(n)/\kappa^2$ for sufficiently large $C > 0$, then it is possible to consistently locate the change points.
- If $\Delta = cB^4 p/\kappa^2$ for sufficiently small c, then it is impossible to estimate the change point consistently.

Lemma 4 Consider the class of distribution

$$\mathcal{Q} = \left\{ P^n_{\kappa,\Delta,\sigma} : \Delta < n/2, \kappa \leq \sigma^2/4 \right\},$$

Then

$$\inf_{\hat{\eta}} \sup_{P \in \mathcal{Q}} E_P(|\hat{\eta} - \eta|) \geq c\sigma^4 \kappa^{-2}.$$

Daren Wang, Y. and Alessandro Rinaldo (2017). Optimal Covariance Change Point Detection in High Dimension. (2017) arXiv preprint, arXiv:1712.09912.

p & n.

δ in WBSIP.

Time dependency.