Sequential testing and high-dimensional online change point detection

Yudong Chen

ongoing work with Richard J. Samworth (University of Cambridge) and Tengyao Wang (University College London)

Statistical Laboratory, University of Cambridge
Cantab Capital Institute for the Mathematics of Information

yc382@cam.ac.uk

June 13, 2019
• Change point detection
  \[ X_1, \ldots, X_{z-1} \overset{\text{iid}}{\sim} N_p(0, \sigma^2 I_p), \quad X_z, X_{z+1}, \ldots \overset{\text{iid}}{\sim} N_p(\theta, \sigma^2 I_p) \]

• (One-sided) sequential testing \[ X_1, X_2, \ldots \overset{\text{iid}}{\sim} N_p(\theta, \sigma^2 I_p) \]
  \[ H_0 : \theta = 0 \quad \text{vs.} \quad H_1 : \theta \neq 0 \]
  One-sided: never accepts \( H_0 \)

• Assume \( \sigma^2 \) known, \( z, \theta \) unknown

• Minimal signal strength and sparsity assumption on \( \theta \)
**Offline**: fixed sample size $n$, one or multiple mean shifts occur within $X_1, \ldots, X_n$

- CUSUM statistics (e.g. Wang and Samworth, 2018)
 Offline: fixed sample size $n$, one or multiple mean shifts occur within $X_1, \ldots, X_n$
CUSUM statistics (e.g. Wang and Samworth, 2018)
 Sequential: Observations $X_1, X_2, \ldots$ come in one at a time
 Declare change as quickly as possible after it takes place
 Online: In addition, computational complexity at each step does NOT depend on the current sample size
Quantities of interest

- Stopping time $N$, possibly $\infty$
- ‘Patience’ of the procedure: measured by the average run length till a false alarm under the null $E_0(N)$
- ‘Alertness’ of the procedure: measured by the average run length from change to reaction if a change is present, given the most adversarial null sequence before the change

$$\bar{E}_\theta(N) = \sup_{z \in \mathbb{N}} \text{ess sup } E_{z, \theta}[N - z + 1|X_1, \ldots, X_{z-1}]$$

- Goal:

$$\min_{N: E_0(N) \geq \gamma} \bar{E}_\theta(N)$$
Quantities of interest

change at $z=150$

declare at $N=190$
Lorden (1971) suggested that we can always build a sequential change point detection procedure with theoretical guarantees from an appropriate one-sided sequential testing procedure;

It is NOT clear that in general this approach will give us online procedures (computational complexity consideration);

In univariate exponential family case, we can construct an online change point detection algorithm according to Page (1954) and Lorden (1971) for composite alternative.
(One-sided) sequential testing $X_1, X_2, \ldots \sim \text{iid } N_p(\theta, \sigma^2 I_p)$

$H_0 : \theta = 0$ vs. $H_1 : \theta \neq 0$ (never accepts $H_0$)

Assume $\sigma^2 = 1$. Denote $S_t := X_1 + \ldots + X_t$. Let

$$N := \inf \{ t \in \mathbb{N} : t \geq 5, \|S_t\|_2^2 / t \geq p + \sqrt{p} \xi_t \},$$

where $\xi_t = C \left( \sqrt{\log (1/\epsilon)} \log \log t + \frac{\log(1/\epsilon) \log \log t}{\sqrt{p}} \right)$ for some large enough constant $C > 0$. 

Yudong Chen (Cambridge)  Sequential testing & change point detection  June 13, 2019  7 / 12
One-sided sequential testing

- Probability of ever stopping under the null: \( P_0(N < \infty) \leq \epsilon \)
- Upper bound (up to a log-log factor) on the average run length under the alternative:
  \[
  E_\theta(N) \lesssim \frac{\sqrt{p}}{\|\theta\|^2_2} \left( \sqrt{\log(1/\epsilon)} + \frac{\log(1/\epsilon)}{\sqrt{p}} \right)
  \]
- Working on the corresponding lower bound
- \( \longrightarrow \) sequential change point procedure
One-sided sequential testing

\[ S \sim p + \sqrt{p} \xi_t \]

\[ X_i \sim N(0, I_{100}) \]

\[ X_i \sim N(\theta, I_{100}) \]

Figure: Sequential testing procedure with \( p = 100, \epsilon = 1/8 \) and \( \vartheta = 1/2 \)
Future work

- Include sparsity in high dimensions
- Use sequential testing or otherwise to construct a change point detection procedure
- Spatial dependence between co-ordinates of the data stream
Thank you for listening!
