Department of ECE, IISc Bangalore

Bandit Change Point Detection

M. Tech Final Review

Braghadeesh Lakshminarayanan
Advisor: Dr. Aditya Gopalan

July 16, 2020

Outline

Introduction

Classical Change Point Detection

- **Change Point:** Time (Discrete) at which the probability distribution of the stochastic process changes.
- Formally,

$$X_1, X_2, \ldots, X_{\tau-1} \overset{i.i.d}{\sim} \mathbb{P}_{\theta^{(0)}}$$
$$X_\tau, X_{\tau+1}, \ldots \overset{i.i.d}{\sim} \mathbb{P}_{\theta^{(1)}}$$

τ is the change point, $\theta^{(0)}$ and $\theta^{(1)}$ are called pre-change and post-change parameter respectively

Goal

Detect the change in the distribution as quickly as possible

Classical Change Point Detection

Existing Work

- Setting: Both pre-change and post-change parameter are known.
  - Stopping rule: Likelihood Ratio Test.
  - Sub-versions: Multiple change points, replacing i.i.d assumption with Markovian, etc.
- Setting: Both pre-change and post-change parameter are unknown [14].
  - Stopping rule: Generalized Likelihood Ratio
- In classical change point detection, no sampling rule is devised.

What if there is a notion of sampling rule?

Bandit Change Point Detection

- Let $X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathbb{P}_{\theta_t}$ be the observations, observed sequentially.

$$\theta_t = \begin{cases} \theta^{(0)}, & \text{if } t < \tau \\ \theta^{(1)}, & \text{otherwise} \end{cases}$$

  $\theta^{(0)} \in \mathbb{R}^d$ is known and $\theta^{(1)} \in \Theta^{(1)} \subseteq \mathbb{R}^d$ is unknown

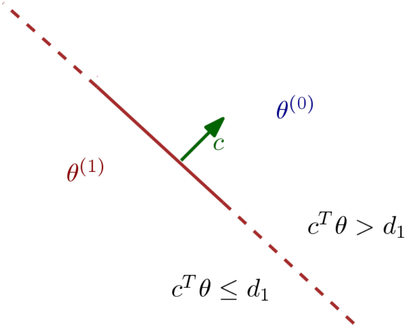
- **Actions:** $\mathcal{A} \subseteq \mathbb{R}^d \rightarrow$ given set of actions in $\mathbb{R}^d$
- **Observations(Linear):** $\mathbb{P}(X_t | A_t, \theta_t) = \mathcal{N}\left(\left\langle A_t, \theta_t \right\rangle, \sigma^2\right)$,

  $\sigma^2 \rightarrow$ Variance (Known)

Note

Action A_t controls the observation received at time t as $X_t = \langle A_t, \theta_t \rangle + \eta_t$, $\eta_t \overset{i.i.d}{\sim} \mathcal{N}(0, \sigma^2)$

- **Problem:** $\theta^{(0)}$ changes to $\theta^{(1)} \in \Theta^{(1)} \subseteq \mathbb{R}^d$ at an unknown time τ.
- **Assumption:** $\Theta^{(1)} = \{\theta \in \mathbb{R}^d \mid c^T\theta \leq d_1\}$



Hypotheses:

$$H_0 \quad : X_1, X_2, \ldots \overset{i.i.d}{\sim} \mathbb{P}_{\theta^{(0)}}$$

$$v.s$$

$$H_1^{(\tau)} \quad : X_1, X_2, \ldots, X_{\tau-1} \overset{i.i.d}{\sim} \mathbb{P}_{\theta^{(0)}}$$

$$X_\tau, X_{\tau+1}, \ldots \overset{i.i.d}{\sim} \mathbb{P}_{\theta^{(1)}}$$

False alarm control: $\underbrace{\mathbb{P}_{H_0}\left[\tau' < \infty\right]}_{\text{false alarm probability}} \leq \delta, \delta \in (0,1)$

Goal:

$$\min \quad \mathbb{E}_{H_1^{(\tau)}}[\tau' - \tau]$$

$$\text{s.t} \quad \mathbb{P}_{H_0}\left[\tau' < \infty\right] \leq \delta$$

where τ' = stopping time of an algorithm and τ = change point

Algorithms

- Algorithm consists of two parts:
  - **Stopping rule**: Specifies condition to stop
  - **Sensing rule**: Adaptively pick actions to quickly detect the change based on stopping rule
- Algorithms are classified based on stopping rule.
- Two novel algorithms
  - Confidence Set based approach
  - GLRT based approach

Confidence Set Approach

Stopping Rule (Based on Confidence Set)

Stopping rule 'S'

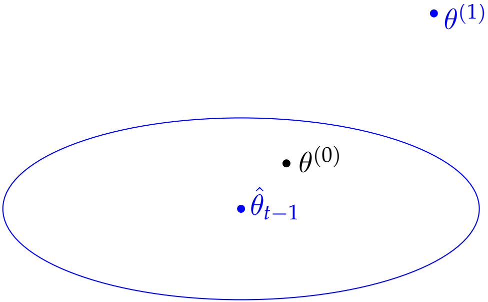
Build confidence set C_t with $\widehat{\theta}_t$ as the center, where $\widehat{\theta}_t$ is the point estimate of $\theta^{(0)}$, such that with high probability, $\theta^{(0)} \in C_t$. If $\theta^{(0)} \notin C_t$, then the change point has occurred before time t.

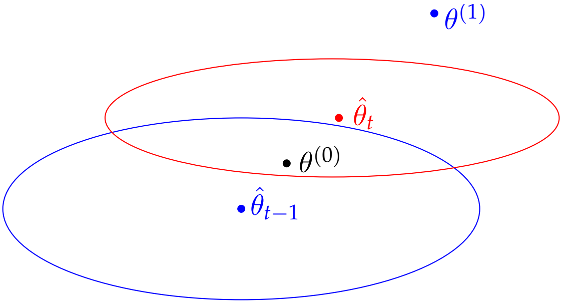
Lemma

'S' is a valid stopping rule

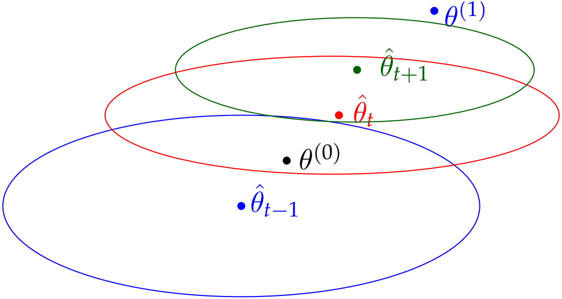
Proof.

Let $\theta^{(0)} \notin C_t$. Suppose the change has not occurred before time t. This implies that, with high probability $\theta^{(0)} \in C_t$, but this contradicts the fact that $\theta^{(0)} \notin C_t$. $\square$


$\theta^{(1)}$
$\theta^{(0)}$
$\hat{\theta}_{t-1}$



$\theta^{(1)}$
$\hat{\theta}_t$
$\theta^{(0)}$
$\hat{\theta}_{t-1}$


$\theta^{(1)}$
$\hat{\theta}_{t+1}$
$\hat{\theta}_t$
$\theta^{(0)}$
$\hat{\theta}_{t-1}$

- $C_t = \left\{ \theta \in \mathbb{R}^d : \left\| \theta - \hat{\theta}_{t-1} \right\|_{V_{t-1}} \leq \beta_t \right\}$
- $\beta_t^2 = 2 \log \left(\frac{1}{\delta} \right) + \log \left(\frac{|V_{t-1}(\lambda)|}{\lambda^d} \right)$
- $\widehat{\theta}_t = V_t^{-1} \sum_{s=1}^{t} A_s X_s, V_t = V_0 + \sum_{s=1}^{t} A_s A_s^T, V_0 = \lambda I_{d \times d}.$

Sensing Rules

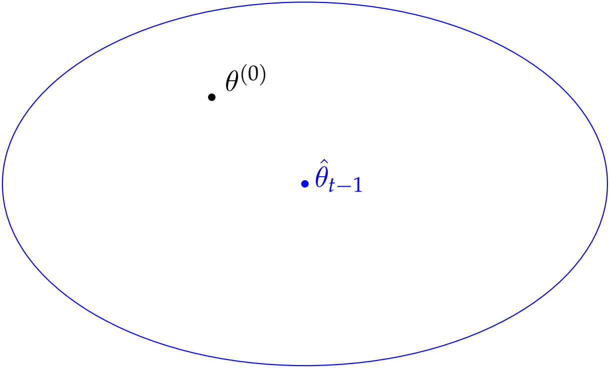
Intuition for a Sensing Rule

Exclude $\theta^{(0)}$ from the confidence set as early as possible, once the change $\theta^{(0)}$ to $\theta^{(1)}$ has occurred.

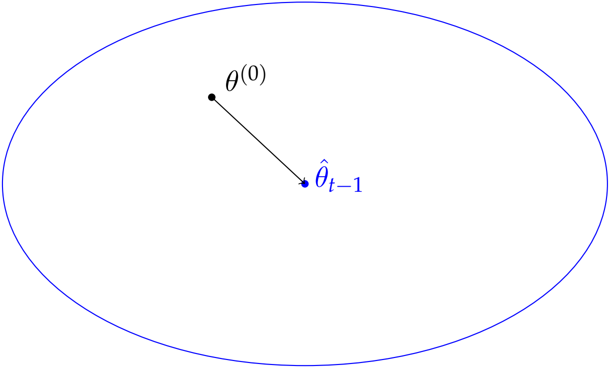
- Play successive actions, $A_t \in \mathcal{A}$, for $t = \tau, \tau + 1, \ldots, \tau^{'}$, such that $\theta^{(0)} \notin C_{\tau^{'}}$ and $\tau^{'} - \tau$ is as low as possible. τ = change point and $\tau^{'}$ = sensing rule's stopping time.

- Play $A_t \in \mathcal{A}$ such that angle between $\widehat{\theta}_t - \theta^{(0)}$ and A_t is minimum among all actions in $\mathcal{A}$

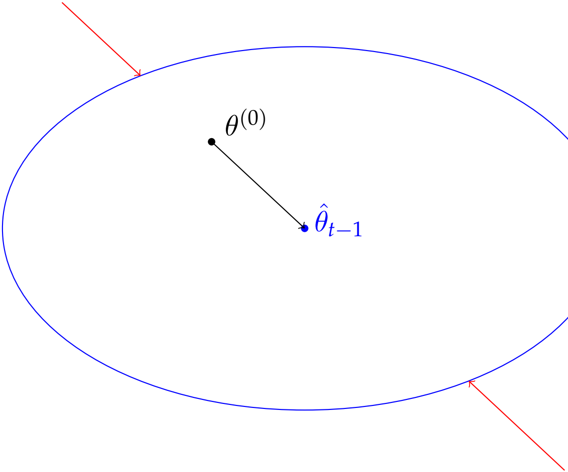
- Play $A_t \in \mathcal{A}$ such that angle between $\widehat{\theta}_t - \theta^{(0)}$ and A_t is minimum among all actions in $\mathcal{A}$



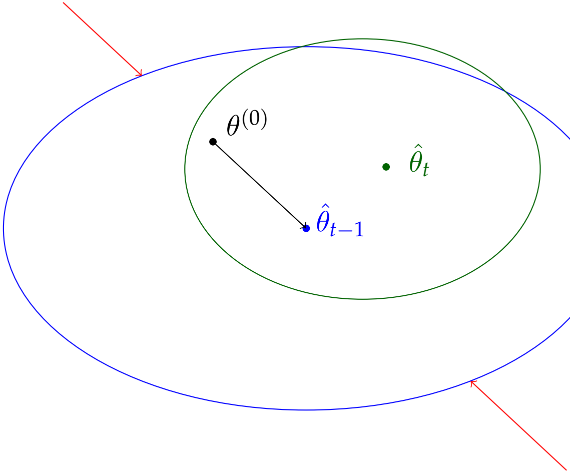
- Play $A_t \in \mathcal{A}$ such that angle between $\widehat{\theta}_t - \theta^{(0)}$ and A_t is minimum among all actions in $\mathcal{A}$



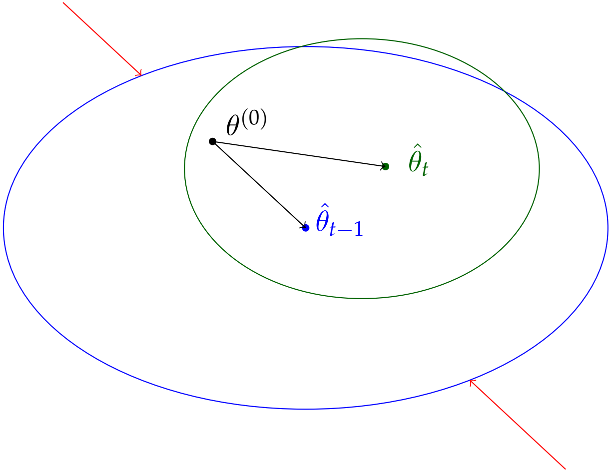
- Play $A_t \in \mathcal{A}$ such that angle between $\widehat{\theta}_t - \theta^{(0)}$ and A_t is minimum among all actions in $\mathcal{A}$



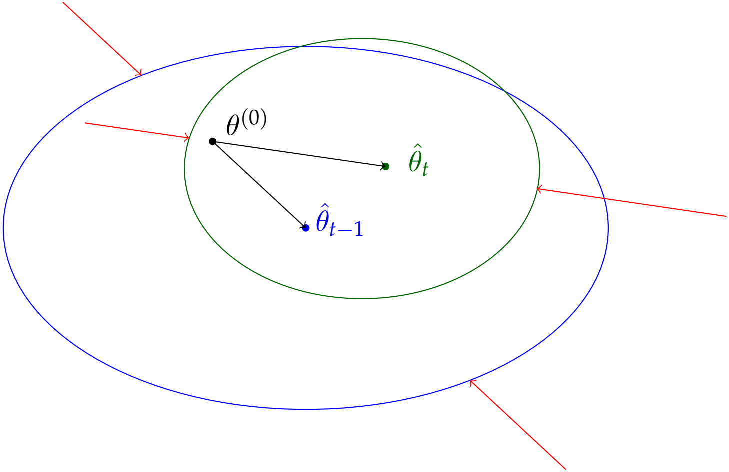
- Play $A_t \in \mathcal{A}$ such that angle between $\widehat{\theta}_t - \theta^{(0)}$ and A_t is minimum among all actions in $\mathcal{A}$



- Play $A_t \in \mathcal{A}$ such that angle between $\widehat{\theta}_t - \theta^{(0)}$ and A_t is minimum among all actions in $\mathcal{A}$

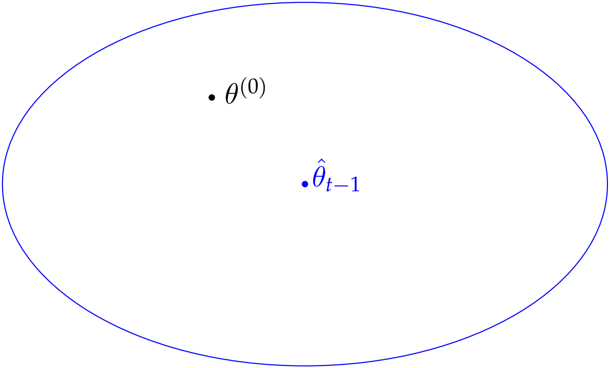


- Play $A_t \in \mathcal{A}$ such that angle between $\widehat{\theta}_t - \theta^{(0)}$ and A_t is minimum among all actions in $\mathcal{A}$

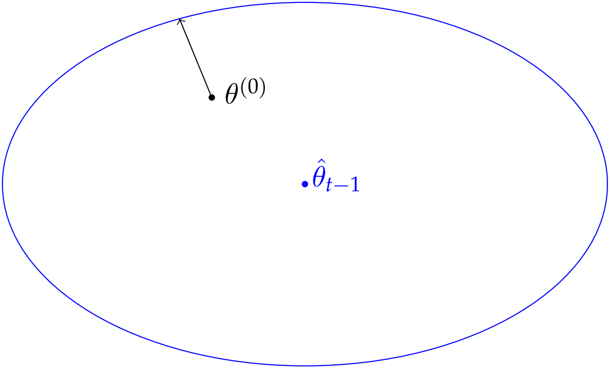


- Play $A_t \in \mathcal{A}$ such that angle between $\theta_{t_{min}} - \theta^{(0)}$ and A_t is minimum, where $\theta_{t_{min}}$ = closest boundary point of $\theta^{(0)}$

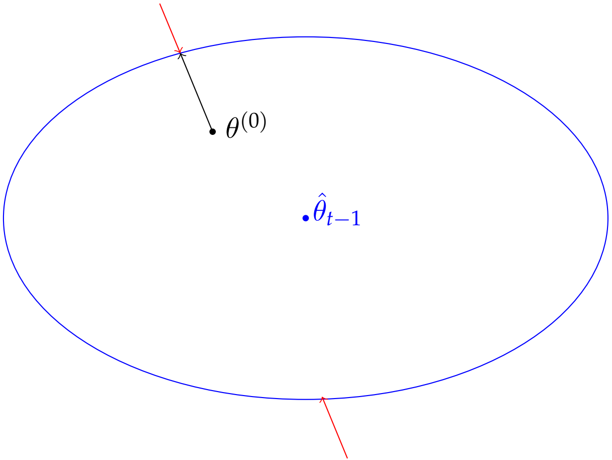
- Play $A_t \in \mathcal{A}$ such that angle between $\theta_{t_{min}} - \theta^{(0)}$ and A_t is minimum, where $\theta_{t_{min}}$ = closest boundary point of $\theta^{(0)}$



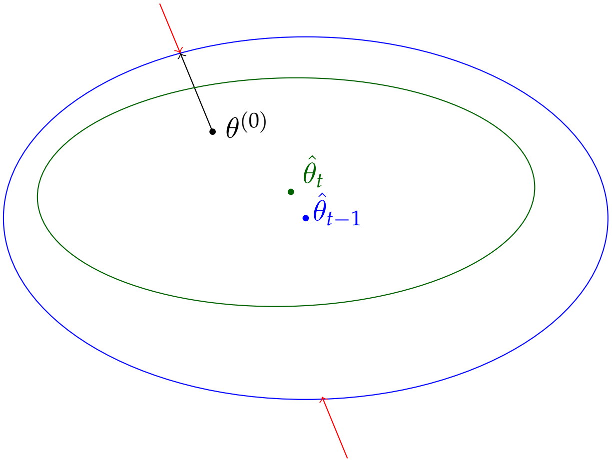
- Play $A_t \in \mathcal{A}$ such that angle between $\theta_{t_{min}} - \theta^{(0)}$ and A_t is minimum, where $\theta_{t_{min}}$ = closest boundary point of $\theta^{(0)}$



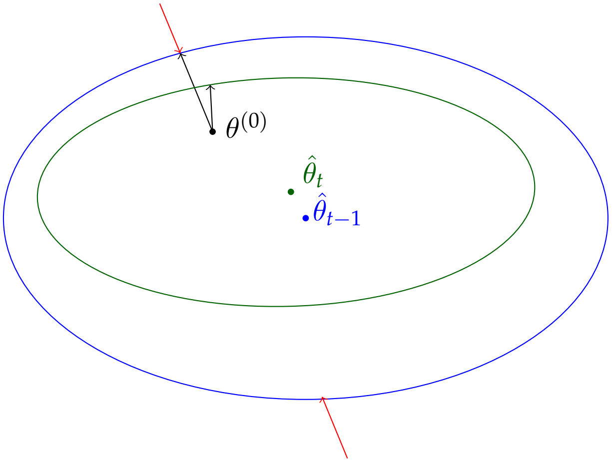
- Play $A_t \in \mathcal{A}$ such that angle between $\theta_{t_{min}} - \theta^{(0)}$ and A_t is minimum, where $\theta_{t_{min}}$ = closest boundary point of $\theta^{(0)}$



- Play $A_t \in \mathcal{A}$ such that angle between $\theta_{t_{min}} - \theta^{(0)}$ and A_t is minimum, where $\theta_{t_{min}}$ = closest boundary point of $\theta^{(0)}$

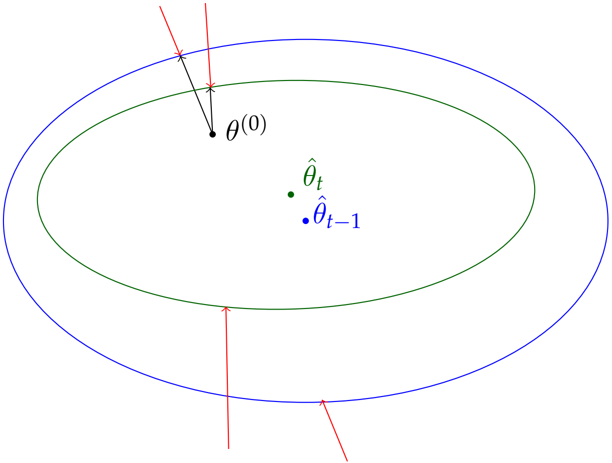


- Play $A_t \in \mathcal{A}$ such that angle between $\theta_{t_{min}} - \theta^{(0)}$ and A_t is minimum, where $\theta_{t_{min}}$ = closest boundary point of $\theta^{(0)}$



Sensing Rule 2

- Play $A_t \in \mathcal{A}$ such that angle between $\theta_{t_{min}} - \theta^{(0)}$ and A_t is minimum, where $\theta_{t_{min}}$ = closest boundary point of $\theta^{(0)}$



GLRT based approach

- GLR is given by

$$\max_{\tau' \leq t} \max_{\theta^{(1)} \in \Theta^{(1)}} \sum_{s=\tau'}^{t-1} \log \frac{\mathbb{P}_{\theta^{(1)}}(X_s)}{\mathbb{P}_{\theta^{(0)}}(X_s)}$$

- Assumption: $\Theta^{(1)} = \{\theta \in \mathbb{R}^d \mid c^T\theta \leq d_1\}$
- Solve inner optimization problem:

$$\max_{\theta^{(1)} \in \Theta^{(1)}} \sum_{s=\tau'}^{t-1} \log \frac{\mathbb{P}_{\theta^{(1)}}(X_s)}{\mathbb{P}_{\theta^{(0)}}(X_s)}$$

Inner Optimization Problem

- Defining

$$g_{t,\tau'}(\theta^{(1)}) = \sum_{s=\tau'}^{t-1} \log \frac{\mathbb{P}_{\theta^{(1)}}(X_s)}{\mathbb{P}_{\theta^{(0)}}(X_s)}$$

- Final expression for $g_{t,\tau'}(\theta^{(1)})$:

$$g_{t,\tau'}(\theta^{(1)}) = 2{\theta^{(1)}}^T b_{t,\tau'} - {\theta^{(1)}}^T V_{t,\tau'} \theta^{(1)} - 2{\theta^{(0)}}^T b_{t,\tau'} + {\theta^{(0)}}^T V_{t,\tau'} \theta^{(0)}$$

- $b_{t,\tau'}$ and $V_{t,\tau'}$ are defined as follows:

$$b_{t,\tau'} = \sum_{s=\tau'}^{t-1} X_s A_s \quad V_{t,\tau'} = \sum_{s=\tau'}^{t-1} A_s A_s^T$$

- The inner optimization problem is posed as:

$$\hat{\theta}^{(1)} = \underset{\theta^{(1)} \in \Theta^{(1)}}{\operatorname{argmax}} \, g_{t,\tau'}(\theta^{(1)})$$

- The above is same as

$$\hat{\theta}^{(1)} = \underset{\theta^{(1)} \in \Theta^{(1)}}{\operatorname{argmin}} \; {\theta^{(1)}}^T V_{t,\tau'} \theta^{(1)} - 2{\theta^{(1)}}^T b_{t,\tau'}$$

- Solution:

$$\hat{\theta}^{(1)} = V_{t,\tau'}^{-1} \left(b_{t,\tau'} - \left(\frac{b_{t,\tau'}^T V_{t,\tau'}^{-1} c - d_1}{c^T V_{t,\tau'}^{-1} c} \right) c \right)$$

Regularized ML Estimate

- **Note:** The inverse of $V_{t,\tau'}$ may not exist $\forall t, \tau'$
- Hence we solve:

$$\underset{\theta^{(1)} \in \Theta^{(1)}}{\operatorname{argmax}} \; g_{t,\tau'}(\theta^{(1)}) + \beta \left|\left|\theta^{(1)}\right|\right|_2^2$$

  where $\beta > 0$.
- Solution:

$$\hat{\theta}_{PE}^{(1)} = V_{t,\tau'}^{-1}(\beta) \left(b_{t,\tau'} - \left(\frac{b_{t,\tau'}^T V_{t,\tau'}^{-1}(\beta) c - d_1}{c^T V_{t,\tau'}^{-1}(\beta) c} \right) c \right)$$

  where,

$$V_{t,\tau'}(\beta) = V_{t,\tau'} + \beta I_{d \times d}$$

Stopping Rule

- Stopping rule is given by GLRT i.e,

$$\max_{\tau' \leq t} g_{t,\tau'}\left(\hat{\theta}_{PE}^{(1)}\right) \geq h_t$$

where h_t is the threshold

- The threshold is calculated subject to low probability false alarm. That is, under no change scenario ($\theta^{(0)} = \theta^{(1)}$),

$$\mathbb{P}\left\{\exists t \in \mathbb{N} : \max_{\tau' \leq t} g_{t,\tau'}\left(\hat{\theta}_{PE}^{(1)}\right) \geq h_t\right\} \leq \delta$$

where $\delta \in (0, 1)$.

- For fixed $\tau^{'}$

$$\mathbb{P}\left\{ g_{t,\tau^{'}}(\hat{\theta}_{PE}^{(1)}) \geq h_t \right\} \leq \delta$$

- Suppose there is no change, i.e, $\theta_t = \theta^{(0)}$, $\forall t$. Then the $g_{t,\tau^{'}}(\hat{\theta}_{PE}^{(1)})$ becomes:

$$g_{t,\tau^{'}}(\hat{\theta}_{PE}^{(1)}) = 2\sum_{s=\tau^{'}}^{t-1} \alpha_s \eta_s - \left|\left| \hat{\theta}_{PE}^{(1)} - \theta^{(0)} \right|\right|_{V_{t,\tau^{'}}}^{2},$$

where $\alpha_s = \left\langle A_s, \hat{\theta}_{PE}^{(1)} - \theta^{(0)} \right\rangle$

- Approximation: $\hat{\theta}_{PE}^{(1)} \approx V_{t,\tau'}^{-1} b_{t,\tau'}$ ($\beta \ll 1$)
- Then α_s is given by

$$\boxed{\alpha_s = \left\langle A_s, V_{t,\tau'}^{-1} \sum_{s=\tau'}^{t-1} A_s \eta_s \right\rangle}$$

- $\left|\left|\hat{\theta}_{PE}^{(1)} - \theta^{(0)}\right|\right|_{V_{t,\tau'}}^2$ is given by

$$\boxed{\left|\left|\hat{\theta}_{PE}^{(1)} - \theta^{(0)}\right|\right|_{V_{t,\tau'}}^2 = \left|\left|\sum_{s=\tau'}^{t-1} A_s \eta_s\right|\right|_{V_{t,\tau'}^{-1}}^2}$$

$$\mathbb{P}\left\{ ||S_{t-\tau'}||^2_{V^{-1}_{t,\tau'}} \geq h_t \right\} \leq \delta$$

where

$$S_{t-\tau'} = \sum_{l=\tau'}^{t-1} A_l \eta_l$$

and h_t is the threshold

- The above calculation is only for fixed $\tau^{'}$
- From [1, 12], the threshold h_t ($\tau^{'} = 1$) is
  $h_t = 2\log\left(\frac{1}{\delta}\right) + \log\left(\frac{|V_{1,t-1}(\beta)|}{\beta^d}\right)$
- Hence, intuitively, the approximate threshold h_t is
  $h_t = 2\log\left(\frac{1}{\delta}\right) + \log\left(\frac{\max_{\tau^{'}\leq t} |V_{t,\tau^{'}}(\beta)|}{\beta^d}\right)$, where $\beta \ll 1$

In the classical change point detection problem, the pre-change and post-change parameters are known and lower bound [13, 10] on the expected detection delay turns out to be

$$\mathbb{E}_{H_1}[\tau^{'} - \tau] \geq \frac{c_1 \log \dfrac{1}{\delta}}{D(\mathbb{P}_{\theta^{(1)}}, \mathbb{P}_{\theta^{(0)}})}$$

$D(.,.)$ denotes KL divergence, $c_1 > 0$

Given $\theta^{(1)}$, any possible sensing rule?

- First step: Analyse a randomized and non-adaptive sensing rule that picks action according to distribution π
- $\pi = (\pi_1, \ldots, \pi_K)$, π_i denotes prior on action $A_i \in \mathcal{A}$
- Sample X_t

$$X_t \sim \begin{cases} \mathcal{N}(A_1^T\theta_t, \sigma^2)\, w.p & \pi_1 \\ \vdots \\ \mathcal{N}(A_K^T\theta_t, \sigma^2)\, w.p & \pi_K \end{cases}$$

- KL divergence

$$D(\mathbb{P}_{\theta^{(1)}}, \mathbb{P}_{\theta^{(0)}}) = D\left(\sum_{i=1}^{K} \pi_i \mathbb{P}_{\theta^{(1)}}(i), \sum_{i=1}^{K} \pi_i \mathbb{P}_{\theta^{(0)}}(i)\right)$$

- $\mathbb{P}_{\theta^{(1)}}(i) = \mathcal{N}(A_i^T \theta^{(1)}, \sigma^2)$ and $\mathbb{P}_{\theta^{(0)}}(i) = \mathcal{N}(A_i^T \theta^{(0)}, \sigma^2)$, $\forall i \in [K]$
- Using the convexity of $D(.,.)$:

$$D\left(\sum_{i=1}^{K} \pi_i \mathbb{P}_{\theta^{(1)}}(i), \sum_{i=1}^{K} \pi_i \mathbb{P}_{\theta^{(0)}}(i)\right) \leq \sum_{i=1}^{K} \pi_i D(\mathbb{P}_{\theta^{(1)}}(i), \mathbb{P}_{\theta^{(0)}}(i))$$

- Finally,

$$\mathbb{E}_{H_1}[\tau^{'} - \tau] \geq \frac{c_1 \log \frac{1}{\delta}}{\sum_{i=1}^{K} \pi_i D(\mathbb{P}_{\theta^{(1)}}(i), \mathbb{P}_{\theta^{(0)}}(i))}$$

- So, the optimal action is to pick action A_i that has maximum $D(\mathbb{P}_{\theta^{(1)}}(i), \mathbb{P}_{\theta^{(0)}}(i))$.
- We use regularized plug-in estimate of $\theta^{(1)}$, i.e, $\hat{\theta}^{(1)}_{PE}$

Sensing Rule

Pick action A_t at time t that maximizes $D(\mathbb{P}_{\hat{\theta}^{(1)}_{PE}}(i), \mathbb{P}_{\theta^{(0)}}(i)), i \in \mathcal{A}$ and update $\hat{\theta}^{(1)}_{PE}, \forall t \in [T]$

Algorithm 1: Plug-in Estimate (Algorithm PE)

Initialization: $t = 1, \theta^{(0)}, \eta, T, c, d$;
for $t = 1, \dots T$ **do**
  for $\tau^{'} = 1, \dots t-1$ **do**
    b=0,V=0 **for** $s = \tau^{'}, \dots t-1$ **do**
      $b = b + X_s A_s$
      $V = V + A_s A_s^T$
    end
    $\hat{\theta}^{(1)} = \underset{\theta^{(1)} \in \Theta^{(1)}}{\operatorname{argmax}} g_{t,\tau'}(\theta^{(1)})$ **if** $(g_{t,\tau'}(\hat{\theta}^{(1)}) > \eta)$ **then**
      change = t **break**
    end
  end
  $A_t = \underset{A_i \in \mathcal{A}}{\operatorname{argmax}} D(\mathbb{P}_{\theta^{(1)}}(i), \mathbb{P}_{\theta^{(0)}}(i))$
  if $t < \tau$ **then**
    $X_t \sim \mathcal{N}(\langle A_t, \theta^{(0)} \rangle, \sigma^2)$
  else
    $X_t \sim \mathcal{N}(\langle A_t, \theta^{(1)} \rangle, \sigma^2)$
  end
end
Output: change

Numerical Results

Experiment Setup

- Number of dimension of the pre-change and post-change parameters $d = 2$ and number of arms/actions $K = 3 * d = 6$
- Pre-change parameter $\theta^{(0)} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$
- Post-change parameter $\theta^{(1)} = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$
- Hyper-plane parameters described by its normal c and intercept d_1 are taken to be $c = \begin{bmatrix} \frac{1}{\sqrt{2}} \\ \frac{1}{\sqrt{2}} \end{bmatrix}$ and $d_1 = 1$. Hyperplane: $c^T \theta^{(1)} \geq d_1$.

Experiment Setup

- Action set $\mathcal{A} = \begin{bmatrix} 0 & 1 & \frac{1}{\sqrt{2}} & -1 & 1 & \frac{-1}{\sqrt{2}} \\ 1 & 0 & \frac{1}{\sqrt{2}} & 1 & -1 & \frac{-1}{\sqrt{2}} \end{bmatrix}$
- Number of iterations = 100.
- Threshold, $h_t = 2\log\left(\frac{1}{\delta}\right) + \log\left(\frac{\max_{\tau' \leq t} |V_{t,\tau'}(\beta)|}{\beta^d}\right)$,

  $\delta = 0.01$, $\beta = 0.0001$.

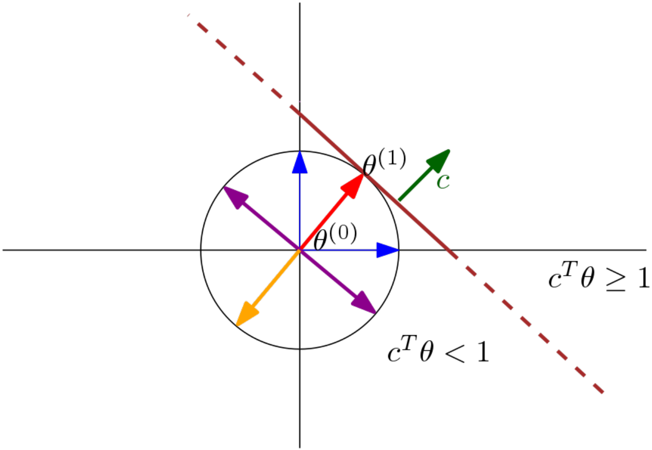


Figure: Description of action set and hyper-plane in the experiment setup 1.

Plots (Comparing CS and GLRT) I

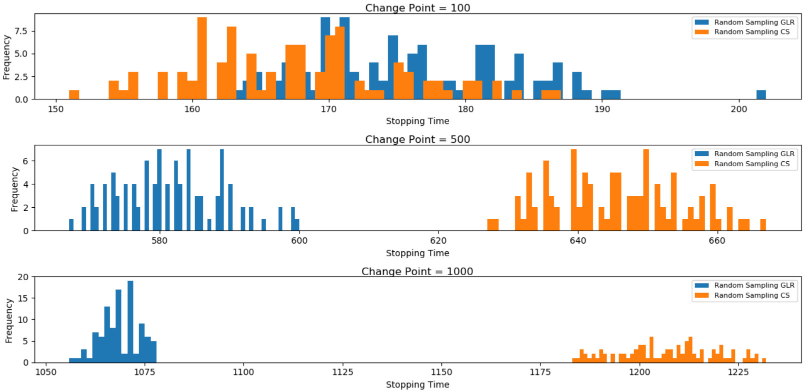


Figure: Histogram plot comparison

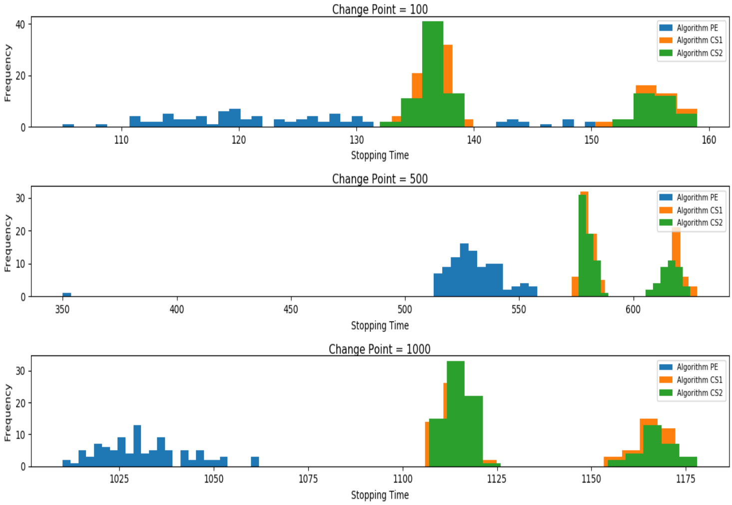


Figure: Histogram plot comparison

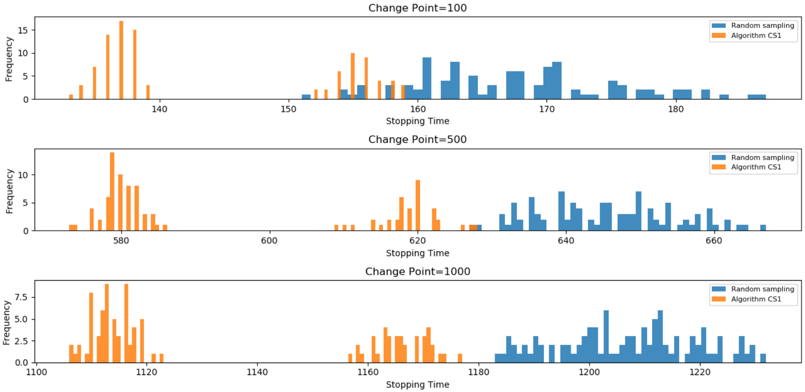


Figure: Histogram plot comparison

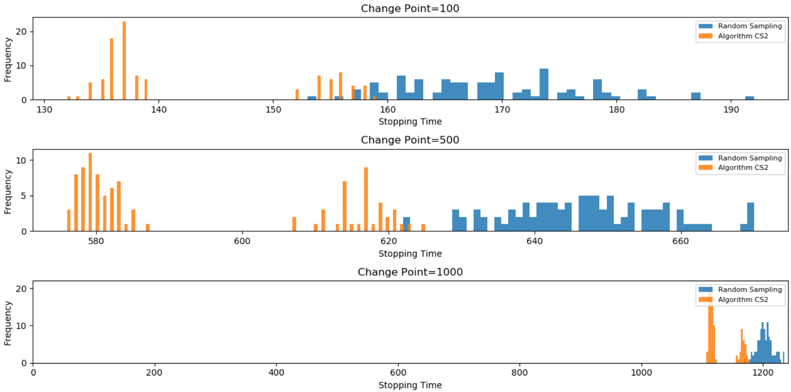


Figure: Histogram plot comparison

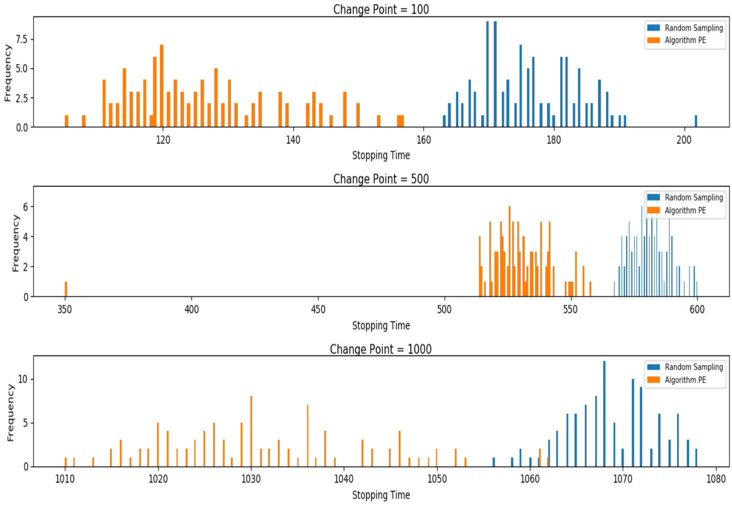


Figure: Histogram plot comparison

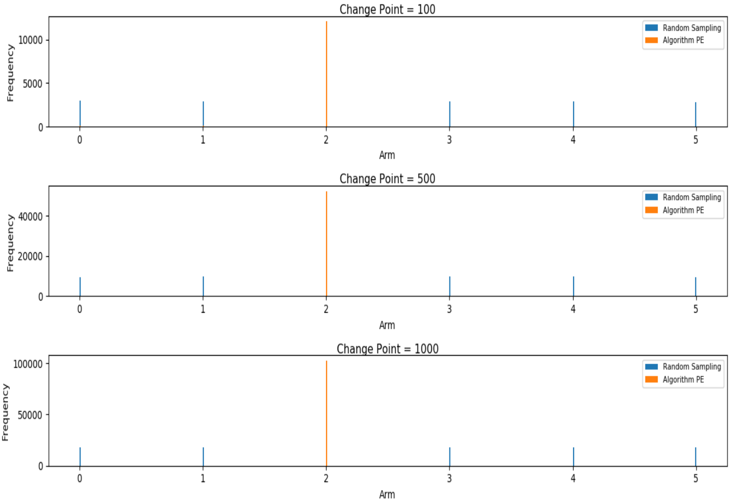


Figure: Arms pulled by Sensing Rule 1 and Random Sampling

Summary

- Figure 2 shows that GLRT based stopping rule is quickest in terms of detection delay (the difference between stopping time and actual change point) as compared to Confidence Set based stopping rule, for increasing change point, τ.
- Given GLRT based stopping rule, sensing rule *Algorithm PE* is quickest as compared to random sampling.
- Also, sensing rule *Algorithm PE* is quickest as compared to sensing rules *Algorithm CS1* and *Algorithm CS2*.
- Figure 7 shows that the sensing rule *Algorithm PE* plays the "best" action , action 3 in the action set $\mathcal{A}$ and detects the change point quickly as compared to Random Sampling.

References I

[1] Yasin Abbasi-yadkori, Dávid Pál, and Csaba Szepesvári. "Improved Algorithms for Linear Stochastic Bandits". In: *Advances in Neural Information Processing Systems 24* (2011), pp. 2312–2320.

[2] D.P. Bertsekas. *Convex Optimization Theory*. Athena Scientific optimization and computation series. Athena Scientific, 2009. ISBN: 9781886529311.

[3] Stephen Boyd and Lieven Vandenberghe. *Convex Optimization*. USA: Cambridge University Press, 2004. ISBN: 0521833787.

[4] Sébastien Bubeck. *Convex Optimization: Algorithms and Complexity*. 2014. arXiv: 1405.4980 [math.OC].

[5] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory*. Wiley-Interscience, New York, NY, USA, 1991.

References II

[6] Aurelien Garivier and Emilie Kaufmann. “Optimal Best Arm Identification with Fixed Confidence”. In: *Conference On Learning Theory*. 2016, pp. 998–1027.

[7] Aurélien Garivier, Pierre Ménard, and Gilles Stoltz. “Explore first, exploit next: The true shape of regret in bandit problems”. In: *Mathematics of Operations Research* 44.2 (2019), pp. 377–399.

[8] G.R. Grimmett and D.R. Stirzaker. *Probability and random processes*. Vol. 80. Oxford university press, 2001.

[9] P.G. Hoel, S.C. Port, and C.J. Stone. *Introduction to Probability Theory*. Houghton Mifflin series in statistics. Houghton Mifflin, 1971.

[10] T.L. Lai and Xing H. “Sequential change-point detection when the pre- and post-change parameters are unknown.”. In: *Sequential Analysis* 29 (2010).

[11] Tze Leung Lai. "Information bounds and quick detection of parameter changes in stochastic systems". In: *IEEE Transactions on Information Theory* 44.7 (1998), pp. 2917–2929.

[12] Tor Lattimore and Csaba Szepesvári. *Bandit Algorithms*. Cambridge University Press, 2018.

[13] G. Lorden. "Procedures for Reacting to a Change in Distribution". In: *The Annals of Mathematical Statistics* 42 (1971).

[14] Odalric-Ambrym Maillard. "Sequential change-point detection: Laplace concentration of scan statistics and non-asymptotic delay bounds". In: *Proceedings of the 30th International Conference on Algorithmic Learning Theory* 98 (2019).

[15] Yurii Nesterov. *Introductory Lectures on Convex Optimization: A Basic Course*. 1st ed. Springer Publishing Company, Incorporated, 2014.

[16] Pearson Egon Sharpe Neyman Jerzy and Pearson Karl. "On the problem of the most efficient tests of statistical hypotheses". In: *Philosophical Transactions of the Royal Society of London. Series A, Containing Papers of a Mathematical or Physical Character* 231 (1933), pp. 289–337.

[17] E.S. Page. "A test for a change in a parameter occurring at an unknown point". In: *Biometrika* 42 (1955), pp. 523–527.

[18] M. Pollak. "Average run lengths of an optimal method of detecting a change in distribution". In: *The Annals of Statistics* (1987), pp. 749–779.

[19] H. Vincent Poor. *An Introduction to Signal Detection and Estimation (2nd Ed.)* Berlin, Heidelberg: Springer-Verlag, 1994.

[20] Venugopal V Veeravalli. “Decentralized quickest change detection”. In: *IEEE Transactions on Information theory* 47 (2001), pp. 1657–1665.