Confirmatory Bayesian Online Change Point Detection in the Covariance Structure of Gaussian Processes

Jiyeon Han*, Kyowoon Lee*, Anh Tong and Jaesik Choi

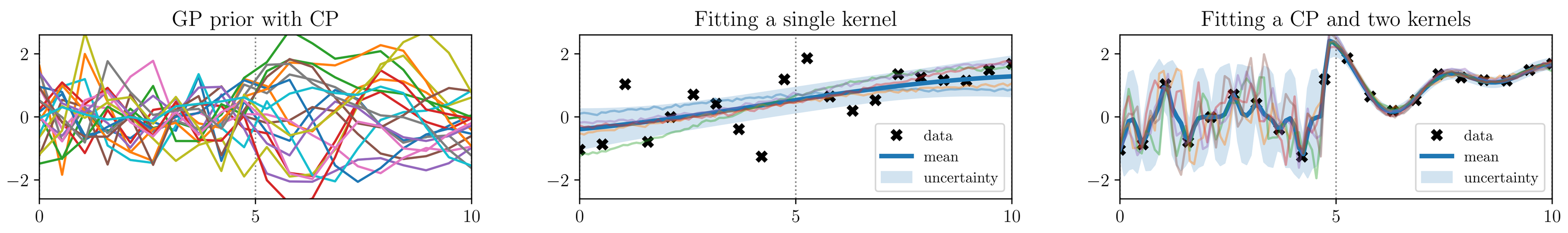
Ulsan National Institute of Science and Technology, Ulsan, Korea

Introduction

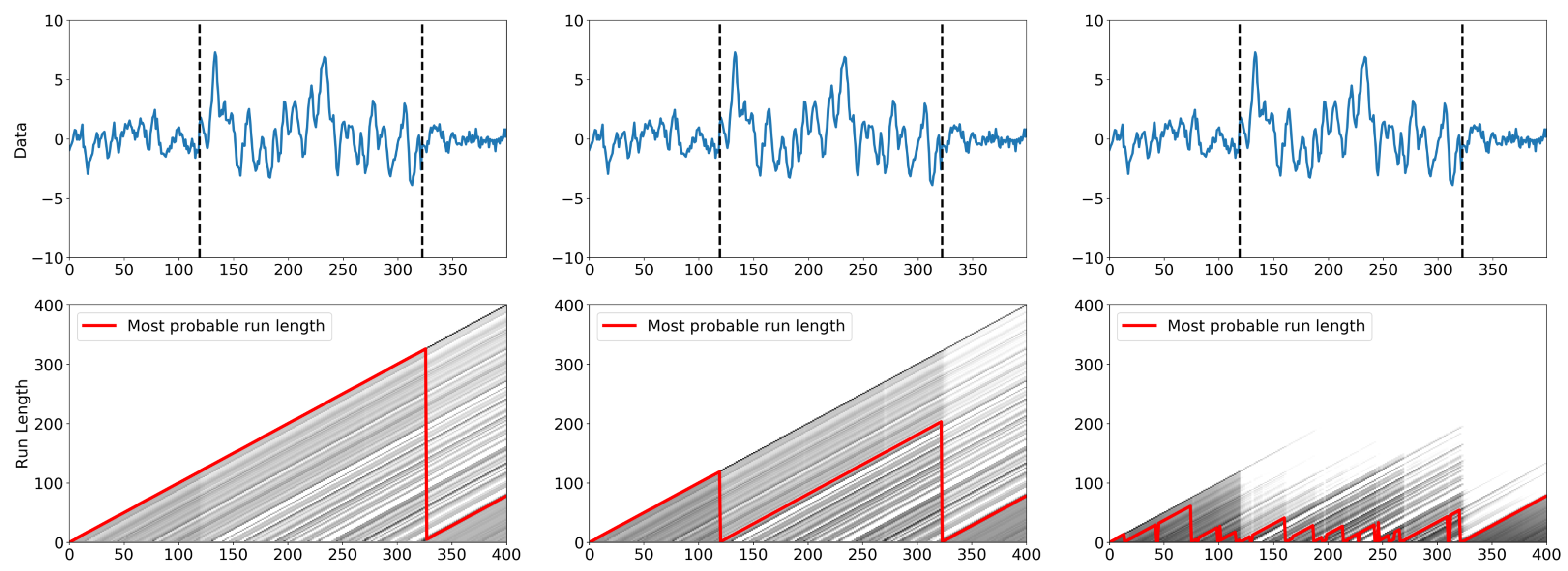
- **Goal**: Change point detection (CPD) in the covariance structure of Gaussian processes (GPs) and predict the next data.
- **Input**: Time series data.
- **Output**:
  - Binary detection of structural change in the underlying random process.
  - The distribution of the upcoming data.

Motivation

- It is difficult to define a change point objectively as it depends on the viewpoint
  - Define a **statistically correct** change point with a **hypothesis test**
- CPD of a covariance structure could affect the quality of a GP regression
  - Change the underlying predictive model after legitimate change point
  - Handle various types of changes with **covariance function** in GP



- Conventional Bayesian online change point detection algorithm (BOCPD) is highly **sensitive** to selected **hyperparameters**
  - Leverage a statistical test to control hyperparameter



w/o parameter tuned precisely | **w/** parameter tuned precisely | **w/o** parameter tuned precisely

Optimal Change Point Detection in Gaussian Processes (Keshavarz et al. 2018)

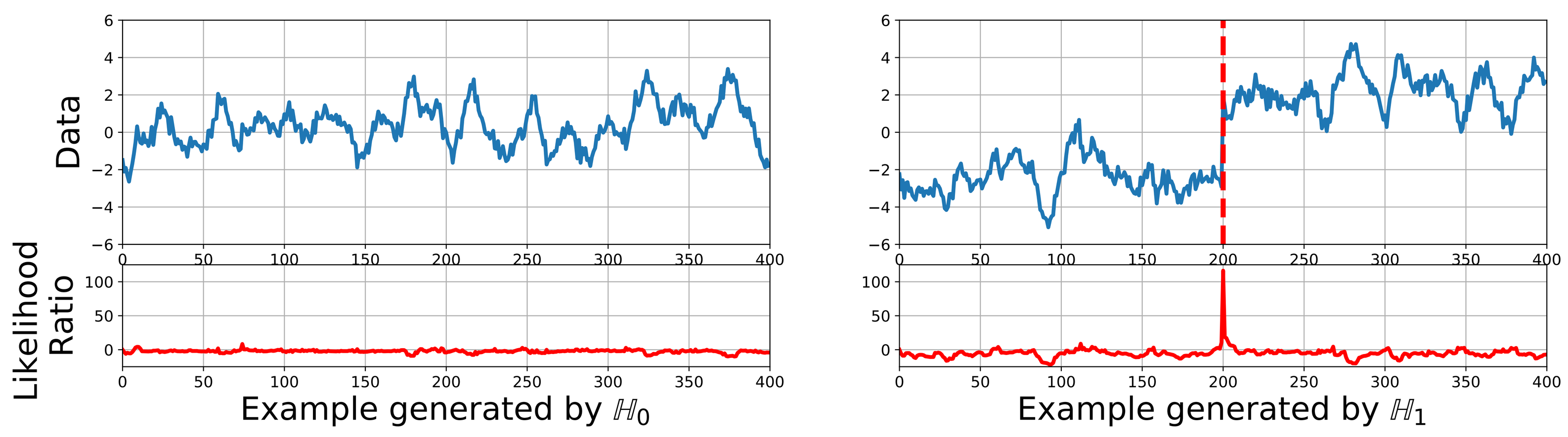
- For a time series $\boldsymbol{X} = \{X_k\}_{k=1}^n$, and $t \in \mathcal{C}_n \subseteq \{1, ..., n\}$

$$\text{(null hypothesis) } \mathbb{H}_0 : \mathbb{E}\boldsymbol{X} = \mathbf{0}_n, \text{ (alternative hypothesis) } \mathbb{H}_1 : \bigcup_{t \in \mathcal{C}_n} \mathbb{H}_{1,t}$$

for $\mathbb{H}_{1,t} : \exists\, b \neq 0,\ \mathbb{E}\boldsymbol{X} = \frac{b}{2}\zeta_t$ where $\zeta_t \in \mathbb{R}^n$ is given by $\zeta_t(k) := sign(k-t)$ for any $t \in \mathcal{C}_n$.

- Likelihood ratio is defined as

$$2\mathfrak{L} = 2 \cdot ln\left(\frac{\text{likelihood for } \mathbb{H}_1}{\text{likelihood for } \mathbb{H}_0}\right) = \max_{t \in \mathcal{C}_n} \left| \frac{(\zeta_t^T(\Sigma_n)^{-1}X)}{\sqrt{\zeta_t^T(\Sigma_n)^{-1}\zeta_t}} \right|^2.$$



Example generated by $\mathbb{H}_0$ | Example generated by $\mathbb{H}_1$

- Generalized likelihood ratio test (GLRT) is formulated as

$$\mathfrak{T}_{GLRT} = \mathbb{I}(2\mathfrak{L} \geq \mathfrak{R}_{n,\delta})$$

- Conditional detection error probability (CDEP) is bounded by δ

$$\varphi_n(\mathfrak{T}) = \mathbb{P}(\mathfrak{T} = 1 | \mathbb{H}_0) + \max_{t \in \mathcal{C}_n} \mathbb{P}(\mathfrak{T} = 0 | \mathbb{H}_{1,t}) \leq \delta$$

for $\mathfrak{R}_{n,\delta} = 1 + 2\left[log\left(\frac{4n}{\delta}\right) + \sqrt{log\left(\frac{4n}{\delta}\right)}\right]$ under the sufficient condition on b.

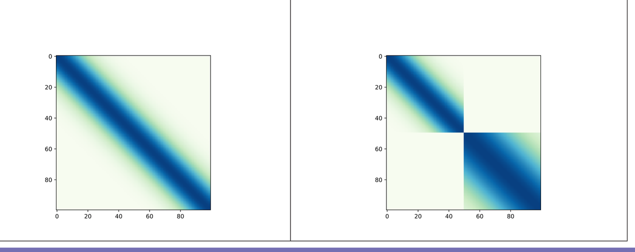
Covariance Detection in Gaussian Processes

- We conduct likelihood ratio test for covariance structural break.

$$\text{(null hypothesis) } \mathbb{H}_0 : Cov(X_i, X_j) = K(i,j), \text{ (alternative hypothesis) } \mathbb{H}_1 : \bigcup_{t \in \mathcal{C}_n} \mathbb{H}_{1,t}$$

with $\mathbb{H}_{1,t} : \mathrm{Cov}(X_i, X_j) = \begin{cases} K(i,j), & i,j < t \\ K'(i,j), & i,j \geq t \\ K''(i,j), & \text{otherwise} \end{cases}$.

$\Sigma = \begin{pmatrix} K_{aa} & K_{ab} \\ K_{ba} & K_{bb} \end{pmatrix}$	$\Sigma_t' = \begin{pmatrix} K_{aa} & 0 \\ 0 & K_{bb}' \end{pmatrix}$



Theorem

For $\mathfrak{R}_\delta$ such that $\mathfrak{R}_{n,\delta,\mathbb{H}_0} \leq \mathfrak{R}_\delta \leq \mathfrak{R}_{n,\delta,\mathbb{H}_1}$, the CDEP is bounded as

$$\varphi_n(\mathfrak{T}) = \mathbb{P}(\mathfrak{T} = 1 | \mathbb{H}_0) + \max_{t \in \mathcal{C}_n} \mathbb{P}(\mathfrak{T} = 0 | \mathbb{H}_{1,t}) \leq \delta$$

with properly set $\mathfrak{R}_{n,\delta,\mathbb{H}_0}$ and $R_{n,\delta,\mathbb{H}_1}$.

Proof. The concentration inequality on subgaussian random variable.

Confirmatory BOCPD

BOCPD (Adams et al. 2007)

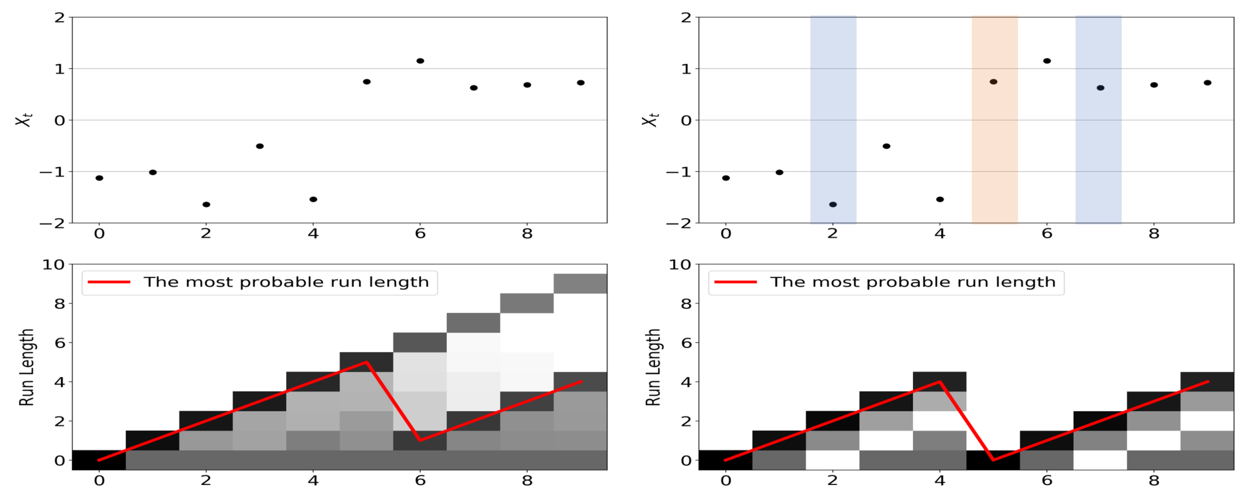
- BOCPD calculates the distribution of the next data by marginalizing over the possible change points.
- (run length) r_t: the number of time step up to time t after the most recent change point .
- Under conventional assumption, $\mathbb{P}(r_t = 0 | r_{t-1}, x_{t-1}^{(r)}) = 1/\lambda$ for some constant λ.

Confirmatory BOCPD

- Incorporate devised statistical test with BOCPD algorithm

$$\mathbb{P}(r_t = 0 | r_{t-1}, x_{t-1}^{(r)}) = \begin{cases} 1-\delta, & \tau^* = t \text{ and } \mathfrak{T}^* = 1 \\ \delta, & \mathfrak{T}^* = 0 \\ H_{const}, & \text{otherwise} \end{cases}$$

  - $\mathfrak{T}^* = 0$: there is a confirmed non-change point in the window around t.
  - $\mathfrak{T}^* = 1$: there is a confirmed change point in the window around t.
  - $\tau^* = t$: the likelihood is maximized at t in the window.
- CBOCPD corrects the BOCPD with an inappropriate hyperparameter



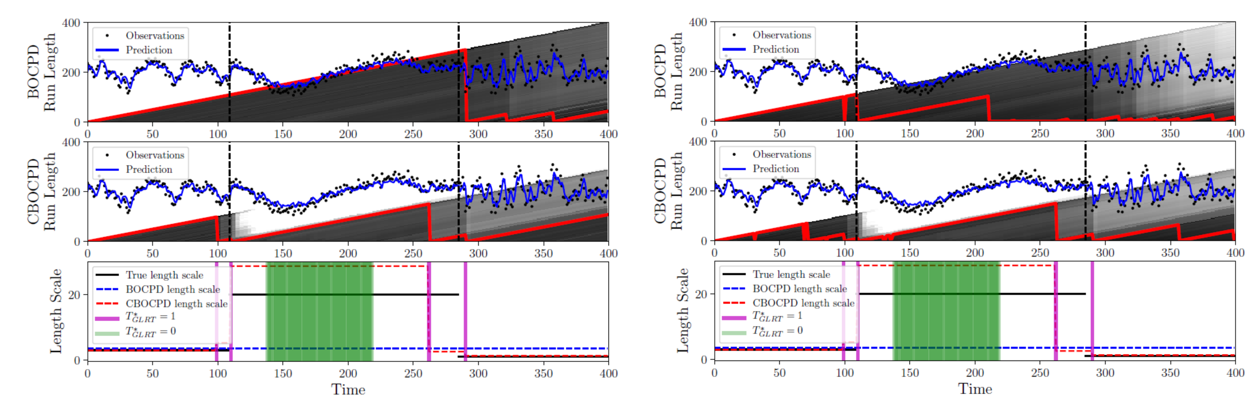
Theorem (Confirmatory BOCPD)

It can be shown that CBOCPD yields at most equal prediction error compared to BOCPD for both stationary case and non-stationary case under specified conditions.

Proof. We can derive this from the assumptions that more data is better for prediction in the stationary case whereas not a previous data is helpful for prediction at the change point.

Experimental Results

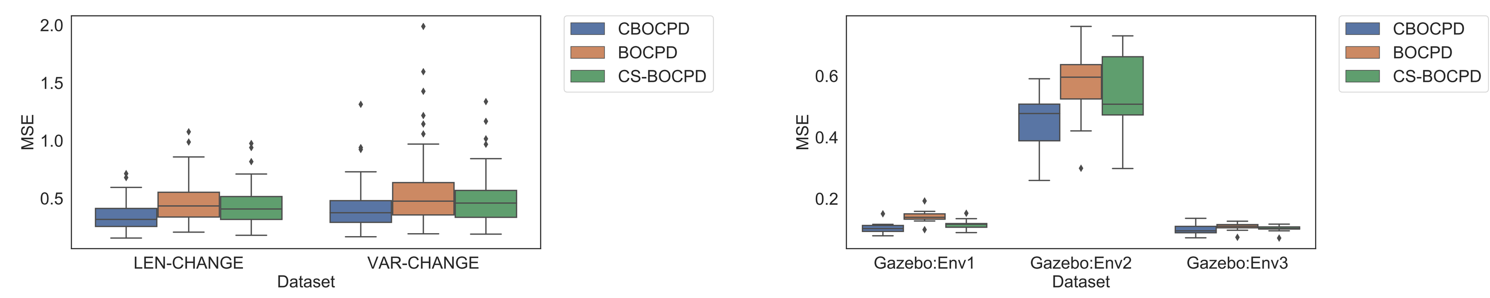
Qualitative Results



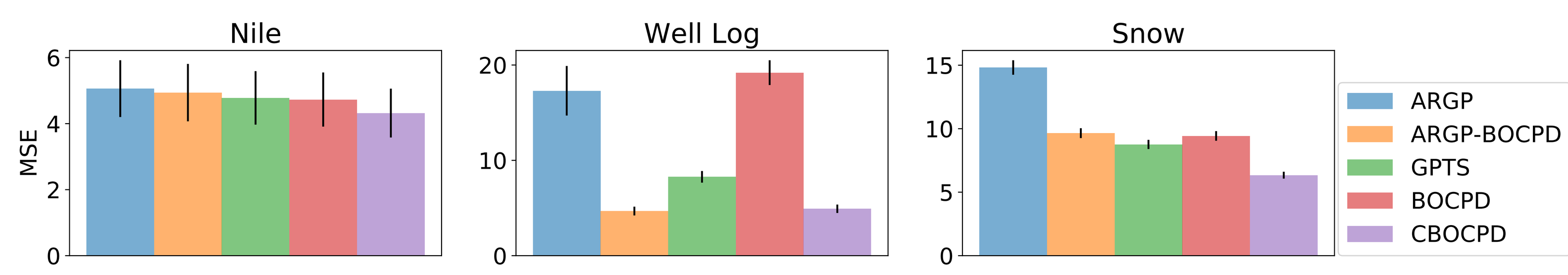
Hyperparameter which sets few change points | **Hyperparameter which sets many change points**

- CBOCPD identifies the covariance change with the help of a statistical test, when BOCPD captures the change too less or too many times.

Synthetic Data and Gazebo Robot Simulation Data



Real World Data



- Outperform other conventional BOCPDs on various datasets.

Conclusion

- We provide a new statistical test for detecting covariance change in GP.
- We propose a new algorithm, Confirmatory BOCPD, which is an improved version of BOCPD with embedded hypothesis tests.
- The proposed algorithm is applied to synthetic and real-world datasets and achieves the state-of-the-art performance.

Acknowledgements

- This work was supported by the Industrial Convergence Core Technology Development Program funded by the Ministry of Trade, Industry & Energy (MOTIE), Korea (No. 10063172) and Institute for Information & communications Technology Planning & Evaluation (IITP) grant funded by the Ministry of Science and ICT (MSIT), Korea (No. 2017-0-01779, XAI)

References

- Keshavarz, Hossein, Clayton Scott, and XuanLong Nguyen. "Optimal change point detection in Gaussian processes." Journal of Statistical Planning and Inference 193 (2018): 151-178.
- Saatçi, Y., and Turner, D R., and Rasmussen, E R. Gaussian process change point models. Proceedings of the International Conference on Machine Learning, ICML, pp. 927-934, 2010.
- Adams, Ryan Prescott and MacKay, David J.C. Bayesian online changepoint detection. arXiv preprint arXiv:0710.3742, 2007.

*Equal contribution

Email: {jiyeon,leekwoon,anhth,jaesik}@unist.ac.kr