Bayesian Online Changepoint Detection

Adams & MacKay, 2007

Elijah Whitham-Powell

Motivation

The Problem

Financial time series exhibit *regime changes*

We want to detect them **online, as they happen**

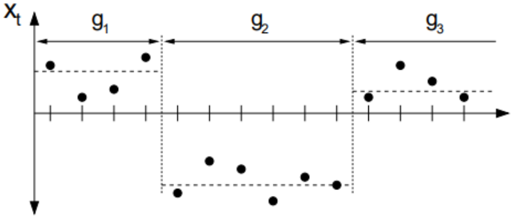
A Bayesian Approach

MMD: nonparametric, retrospective

BOCPD: parametric likelihood, maintains a **posterior over changepoint locations**

Problem Setup

- Data divided into non-overlapping segments
- Within each segment: i.i.d. with segment-specific parameters
- Boundaries between segments are **changepoints**
- New segments get fresh parameters from the prior



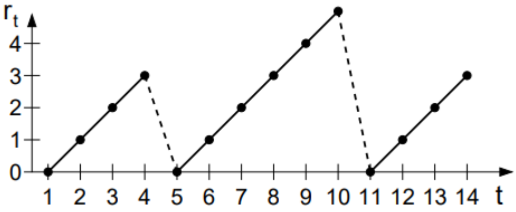
(a)

Run Length

The **run length** r_t counts time steps since the last changepoint.

At each step, only two outcomes:

1. Run **continues**: $r_t = r_{t-1} + 1$
2. **Changepoint**: $r_t = 0$



(b)

Resets to zero at each changepoint, then climbs linearly.

The Hazard Function

"Given that the current regime has lasted τ steps, what is the probability it ends now?"

$$H(\tau) = \frac{P_{\mathrm{gap}}(g = \tau)}{\sum_{t=\tau}^{\infty} P_{\mathrm{gap}}(g = t)}$$

This determines the **changepoint prior**:

$$P(r_t \mid r_{t-1}) = \begin{cases} H(r_{t-1} + 1) & \text{if } r_t = 0 \\ 1 - H(r_{t-1} + 1) & \text{if } r_t = r_{t-1} + 1 \end{cases}$$

Constant hazard (memoryless)

$$P_{\mathrm{gap}} \sim \mathrm{Geometric}(\lambda) \;\Rightarrow\; H(\tau) = \frac{1}{\lambda} \;\; \forall \, \tau$$

A modeling choice

Algorithm is fully general — any P_{gap} works. λ controls expected time between changepoints.

The Core Idea: Marginal Prediction

$$\boxed{P(x_{t+1} \mid x_{1:t}) = \sum_{r_t} P(x_{t+1} \mid r_t, \boldsymbol{x}_t^{(r)})\; P(r_t \mid x_{1:t})}$$

Predictive distribution
$P(x_{t+1} \mid r_t, \boldsymbol{x}_t^{(r)})$
"If the current regime has lasted r_t steps, what do I expect next?"

Run length posterior $P(r_t \mid x_{1:t})$
"How likely is each possible run length given everything observed?"

The Recursive Update

$$\boxed{P(r_t, x_{1:t}) = \sum_{r_{t-1}} \underbrace{P(r_t \mid r_{t-1})}_{\text{CP prior}} \; \underbrace{P(x_t \mid r_{t-1}, \boldsymbol{x}_t^{(r)})}_{\text{predictive likelihood}} \; \underbrace{P(r_{t-1}, x_{1:t-1})}_{\text{previous step}}}$$

Changepoint prior

Did the run continue or end?

Predictive likelihood

How well did this model predict?

Previous time step

Already computed — recursion!

Conjugate-Exponential Models

Why conjugate priors?

Need to update beliefs about segment parameters at every time step.

Conjugate priors allow this via **sufficient statistics** — a running tally rather than storing raw data.

Sufficient statistics update

$$\nu_t^{(r)} = \nu_{\text{prior}} + r_t$$

$$\chi_t^{(r)} = \chi_{\text{prior}} + \sum_{t' \in r_t} u(x_{t'})$$

ν counts observations; χ accumulates statistics.

Algorithm 1: The Full Procedure

1. **Initialize**

   $P(r_0 = 0) = 1$

   $\nu_1^{(0)} = \nu_{\text{prior}}, \quad \chi_1^{(0)} = \chi_{\text{prior}}$

2. **Observe datum** x_t

3. **Evaluate predictive probability**

   $\pi_t^{(r)} = P(x_t \mid \nu_t^{(r)}, \chi_t^{(r)})$

4. **Calculate growth probabilities**

   $P(r_t = r_{t-1} + 1, x_{1:t})$

   $= P(r_{t-1}, x_{1:t-1})\, \pi_t^{(r)}\, (1 - H(r_{t-1}))$

5. **Calculate changepoint probabilities**

   $P(r_t = 0, x_{1:t})$

   $= \sum_{r_{t-1}} P(r_{t-1}, x_{1:t-1})\, \pi_t^{(r)}\, H(r_{t-1})$

6. **Calculate evidence**

   $P(x_{1:t}) = \sum_{r_t} P(r_t, x_{1:t})$

7. **Determine run length distribution**

   $P(r_t \mid x_{1:t}) = P(r_t, x_{1:t}) \,/\, P(x_{1:t})$

8. **Update sufficient statistics**

   $\nu_{t+1}^{(0)} = \nu_{\text{prior}}, \quad \chi_{t+1}^{(0)} = \chi_{\text{prior}}$

   $\nu_{t+1}^{(r+1)} = \nu_t^{(r)} + 1$

   $\chi_{t+1}^{(r+1)} = \chi_t^{(r)} + u(x_t)$

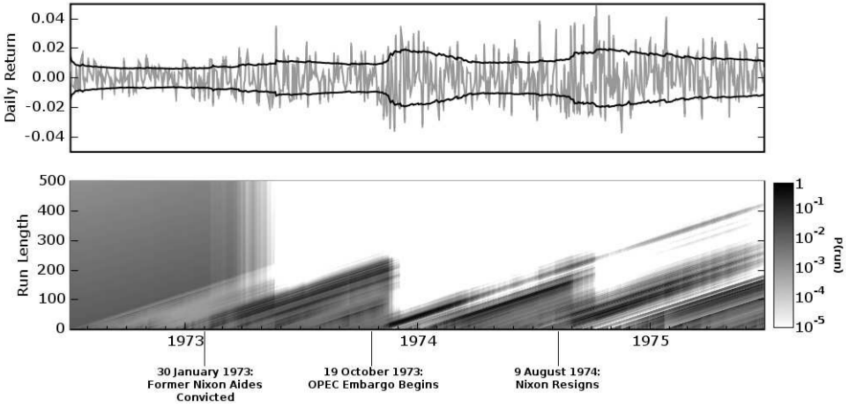
9. **Perform prediction**

   $P(x_{t+1} \mid x_{1:t})$

   $= \sum_{r_t} P(x_{t+1} \mid r_t, \boldsymbol{x}_t^{(r)})\, P(r_t \mid x_{1:t})$

10. **Return to step 2**

Application: Dow Jones 1972–75



Top: daily returns with predictive volatility. Bottom: posterior $P(r_t \mid x_{1:t})$ — the “staircase” shows growing runs; drops to zero indicate changepoints.

Next Steps: Regime Detection

I will implement BOCPD for market regime detection and compare with my existing MMD-based approach.

Online vs. Retrospective

BOCPD processes one point at a time; MMD compares windows after the fact

Parametric vs. Nonparametric

BOCPD assumes a distributional model; MMD detects any distributional shift

Posterior vs. Permutation Test

BOCPD gives a full distribution over changepoints; MMD uses p-values

Adams, R. P. & MacKay, D. J. C. (2007). Bayesian Online Changepoint Detection. arXiv:0710.3742

Questions?

References

Adams, R. P. & MacKay, D. J. C. (2007). Bayesian Online Changepoint Detection.
`https://arxiv.org/abs/0710.3742`

Whitham-Powell, E. (2025). MMD Regime Change Detection.
`https://whitham-powell.com/2025/12/23/mmd-regime-change.html`