Optimum Sequential Procedures

for

Detecting Changes in Processes

George V. Moustakides
INRIA-IRISA, Rennes, France

Outline

- The change detection (disorder) problem
- Overview of existing results
- Lorden’s criterion and the CUSUM test
- A modified Lorden criterion
- Optimality of CUSUM for Itô processes

The Change Detection (Disorder) Problem

We are observing sequentially a process ξ_t with the following statistics

$$\begin{aligned} \xi_t &\sim \mathbb{P}_\infty && \text{for } 0 \le t \le \tau \\ &\sim \mathbb{P}_0 && \text{for } \tau < t \end{aligned}$$

– Change time τ: deterministic (but unknown) or random.
– Probability measures $\mathbb{P}_\infty$, $\mathbb{P}_0$: known.

Detect the change "as soon as possible".

Applications include: systems monitoring; quality control; financial decision making; remote sensing (radar, sonar, seismology); occurrence of industrial accidents; speech/image/video segmentation; etc.

The observation process ξ_t is available sequentially; this can be expressed through the filtration:

$$\mathcal{F}_t = \sigma\{\xi_s : 0 \leq s \leq t\}.$$

For detecting the change we are interested in sequential schemes.

Any sequential detection scheme can be represented by a stopping time T (the time we stop and declare that the change took place).

The stopping time T is adapted to $\mathcal{F}_t$.
In other words, *at every time instant* t we perform a test (whether to stop and declare a change or continue sampling) using only the available information up to time t.

Overview of Existing Results

$\mathbb{P}_\tau$: the probability measure induced, when the change takes place at time τ.

$\mathbb{E}_\tau[\cdot]$: the corresponding expectation.

$\mathbb{P}_\infty$: all data under nominal régime.

$\mathbb{P}_0$: all data under alternative régime.

Optimality Criteria

They are basically comprised of two parts:

– The first measures the detection delay

– The second the frequency of false alarms

Possible approaches are Baysian and Min-max.

Bayesian Approach (Shiryayev):

τ is random and exponentially distributed

$$\inf_T \{c\,\mathbb{E}[(T-\tau)^+] + \mathbb{P}[T<\tau]\}$$

The Shiryayev test consists in computing the statistics $\pi_t = \mathbb{P}[\tau \le t | \mathcal{F}_t]$; and stop when

$$T_S = \inf_t \{t : \pi_t \ge \nu\}.$$

T_S is optimum (Shiryayev 1978):

– In discrete time: when ξ_n is i.i.d. before and after the change.

– In continuous time: when ξ_t is a Brownian Motion with constant drift before and after the change.

Min-Max Approach (Shiryayev-Roberts-Pollak):
τ is deterministic and unknown

$$\inf_T \sup_\tau \mathbb{E}_\tau[(T-\tau)^+ | T > \tau]; \text{ subject } \mathbb{E}_\infty[T] \geq \gamma.$$

Optimality results exists only for discrete time when ξ_n is i.i.d. before and after the change. Specifically if we define the statistics

$$S_n = (S_{n-1} + 1)\frac{f_0(\xi_n)}{f_\infty(\xi_n)},$$

where $f_\infty(\cdot), f_0(\cdot)$ the common pdf of the data before and after the change then (Yakir 1997) the stopping time

$$T_{SRP} = \inf_n\{n : S_n \geq \nu\}$$

is optimum.

Lorden's Criterion and the CUSUM Test

An alternative min-max approach consists in defining the following performance measure (Lorden 1971)

$$J(T) = \sup_{\tau} \operatorname{essup} \mathbb{E}_{\tau}[(T-\tau)^{+}|\mathcal{F}_{\tau}]$$

and solve the min-max problem

$$\inf_{T} J(T); \ \text{ subject to } \ \mathbb{E}_{\infty}[T] \geq \gamma.$$

The test closely related to Lorden's criterion and being to most popular one used in practice is the **Cumulative Sum** (CUSUM) test.

Define the CUSUM statistics y_t as follows:

$$u_t = \log\left(\frac{d\mathbb{P}_0}{d\mathbb{P}_\infty}(\mathcal{F}_t)\right); \;\; m_t = \inf_{0\le s\le t} u_s$$

$$y_t = u_t - m_t.$$

The CUSUM stopping time (Page 1954):

$$T_C = \inf_t\{t : y_t \ge \nu\}.$$

Optimality results:

– Discrete time: when ξ_n is i.i.d. before and after the change (Moustakides 1986, Ritov 1990).

– Continuous time: when ξ_t is a Brownian Motion with constant drift before and after the change (Shiryayev 1996, Beibel 1996).

A modified Lorden criterion

Our goal is to extend the optimality of CUSUM to Itô processes. For this it will be necessary to modify Lorden's criterion using the Kullback-Leibler Divergence (KLD).

Similar extension was proposed for the SPRT by Liptser and Shiryayev (1978).

Consider the process ξ_t

$$d\xi_t = \begin{cases} dw_t, & 0 \leq t \leq \tau \\ \alpha_t dt + dw_t, & \tau < t \end{cases}$$

where w_t is a standard Brownian motion with respect to $\mathcal{F}_t = \sigma(\xi_s; 0 \leq s \leq t)$; α_t is adapted to $\mathcal{F}_t$ and τ denotes the time of change.

To ξ_t we correspond the process u_t defined by

$$du_t = \alpha_t d\xi_t - 0.5\alpha_t^2 dt$$

which we like to play the role of the log-likelihood ratio $u_t = \log(d\mathbb{P}_0/d\mathbb{P}_\infty(\mathcal{F}_t))$. We therefore need to impose the following conditions:

1. $\mathbb{P}_0\left[\int_0^t \alpha_s^2 ds < \infty\right] = \mathbb{P}_\infty\left[\int_0^t \alpha_s^2 ds < \infty\right] = 1$

2. A "Novikov" condition, i.e. $\mathbb{E}_\infty[\exp(\int_{t_{n-1}}^{t_n} \alpha_s^2 ds)] < \infty$ where t_n strictly increasing with $t_n \to \infty$.

3. $\mathbb{P}_0\left[\int_0^\infty \alpha_s^2 ds = \infty\right] = \mathbb{P}_\infty\left[\int_0^\infty \alpha_s^2 ds = \infty\right] = 1$

From conditions 1 & 2 we have validity of Girsanov's theorem, therefore

$$\frac{d\mathbb{P}_0}{d\mathbb{P}_\infty}(\mathcal{F}_t) = e^{u_t}; \quad \frac{d\mathbb{P}_\tau}{d\mathbb{P}_\infty}(\mathcal{F}_t) = e^{u_t - u_\tau}.$$

Furthermore for the KLD we can write

$$\begin{aligned}
&\mathbb{E}_\tau \left[\log \left(\frac{d\mathbb{P}_\tau}{d\mathbb{P}_\infty}(\mathcal{F}_t) \right) \,\middle|\, \mathcal{F}_\tau \right] \\
&\quad = \mathbb{E}_\tau \left[\int_\tau^t \alpha_s dw_s + \int_\tau^t \frac{1}{2}\alpha_s^2 ds \,\middle|\, \mathcal{F}_\tau \right] \\
&\quad = \mathbb{E}_\tau \left[\int_\tau^t \frac{1}{2}\alpha_s^2 ds \,\middle|\, \mathcal{F}_\tau \right], \quad \text{for } 0 \le \tau \le t < \infty,
\end{aligned}$$

This suggests the following modification in Lorden's criterion

$$J(T) = \sup_{\tau\in[0,\infty)} \operatorname{essup} \mathbb{E}_\tau \left[\mathbb{1}_{\{T>\tau\}} \int_\tau^T \frac{1}{2}\alpha_t^2 dt \,\middle|\, \mathcal{F}_\tau \right],$$

and the corresponding min-max optimization

$$\inf_T J(T); \text{ subject } \mathbb{E}_\infty \left[\int_0^T \frac{1}{2}\alpha_t^2 dt \right] \geq \gamma.$$

The original and the modified criterion coincide when ξ_t is a Brownian motion with constant drift.

Let us form the CUSUM statistics y_t for the Itô process

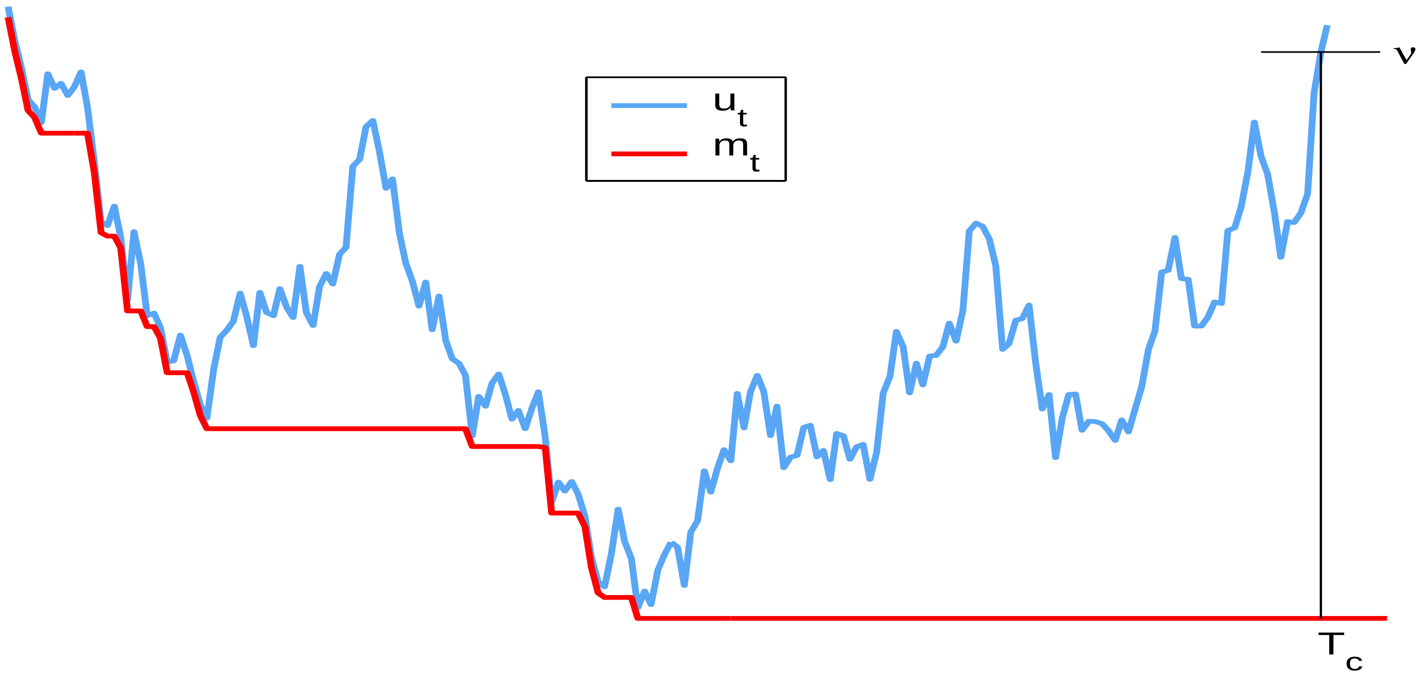
$$\begin{aligned} du_t &= \alpha_t d\xi_t - 0.5\alpha_t^2 dt \\ m_t &= \inf_{0 \le s \le t} u_s \\ y_t &= u_t - m_t \end{aligned}$$

and the optimum CUSUM test is

$$T_C = \inf_t \{t : y_t \ge \nu\}; \text{ where } \mathbb{E}_\infty \left[\int_0^{T_C} \frac{1}{2} \alpha_t^2 dt \right] = \gamma.$$

Since y_t has continuous paths we conclude that when the CUSUM test stops we will have: $y_{T_C} = \nu$.

Optimality of CUSUM for Itô processes



$u_t \geq m_t$ therefore $y_t = u_t - m_t \geq 0$.

m_t is nonincreasing and $dm_t \neq 0$ only when $u_t = m_t$ or $y_t = 0$.

If $f(y)$ continuous; $f(0) = 0$, then $\int_0^\infty f(y_t) dm_t = 0$.

If $f(y)$ is a twice continuously differentiable function with $f'(0)=0$, using standard Itô calculus we have

$$
\begin{aligned}
df(y_t) &= f'(y_t)(du_t - dm_t) + 0.5\alpha_t^2 f''(y_t)dt \\
&= f'(y_t)du_t + 0.5\alpha_t^2 f''(y_t)dt
\end{aligned}
$$

Theorem 1: T_C is a.s. finite and

$$
\mathbb{E}_\tau\left[\mathbb{1}_{\{T_C>\tau\}}\int_\tau^{T_C}\tfrac{1}{2}\alpha_t^2 dt\,\middle|\,\mathcal{F}_\tau\right] = [g(\nu)-g(y_\tau)]\mathbb{1}_{\{T_C>\tau\}}
$$

$$
\mathbb{E}_\infty\left[\mathbb{1}_{\{T_C>\tau\}}\int_\tau^{T_C}\tfrac{1}{2}\alpha_t^2 dt\,\middle|\,\mathcal{F}_\tau\right] = [h(\nu)-h(y_\tau)]\mathbb{1}_{\{T_C>\tau\}}.
$$

where

$$
g(y) = y + e^{-y} - 1; \qquad h(y) = e^{y} - y - 1.
$$

Since $g(y), h(y)$ are increasing and strictly convex with $g(0) = h(0) = 0$, we now conclude

$$\begin{aligned} J(T_C) &= \sup_\tau \text{essup} \mathbb{E}_\tau \left[\int_\tau^{T_C} \alpha_s^2 ds | \mathcal{F}_\tau \right] \\ &= \sup_\tau \text{essup} [g(\nu) - g(y_\tau)] \mathbb{1}_{\{T_C > \tau\}} \\ &= g(\nu) - g(0) = g(\nu) \end{aligned}$$

Similarly

$$\mathbb{E}_\infty \left[\int_0^{T_C} \alpha_s^2 ds \right] = h(\nu) - h(0) = h(\nu) = \gamma.$$

The threshold can thus be computed: $e^\nu - \nu - 1 = \gamma$.

Using again standard Itô calculus we have the following generalization of Theorem 1.

Corollary:

$$\mathbb{E}_\tau\left[\int_\tau^T \tfrac{1}{2}\alpha_t^2 dt \,\middle|\, \mathcal{F}_\tau\right] = \mathbb{E}_\tau\left[g(y_T) - g(y_\tau)|\mathcal{F}_\tau\right] \mathbb{1}_{\{T>\tau\}}$$

$$\mathbb{E}_\infty\left[\int_\tau^T \tfrac{1}{2}\alpha_t^2 dt \,\middle|\, \mathcal{F}_\tau\right] = \mathbb{E}_\infty\left[h(y_T) - h(y_\tau)|\mathcal{F}_\tau\right] \mathbb{1}_{\{T>\tau\}}$$

where T stopping time.

Remark 1: The false alarm constraint can be written as

$$\mathbb{E}_\infty\left[\int_0^T \tfrac{1}{2}\alpha_t^2 dt\right] = \mathbb{E}_\infty[h(y_T)] \geq \gamma$$

Remark 2: We can limit ourselves to stopping times that satisfy the false alarm constraint with equality, that is,

$$\mathbb{E}_\infty\left[\int_0^T \tfrac{1}{2}\alpha_t^2 dt\right] = \mathbb{E}_\infty[h(y_T)] = \gamma = h(\nu).$$

Remark 3: The modified performance measure $J(T)$ can be suitably lower bounded as follows

$$\begin{aligned} J(T) &= \sup_\tau \operatorname{essup} \mathbb{E}_\tau\left[\mathbb{1}_{\{T>\tau\}} \int_\tau^T \frac{1}{2}\alpha_t^2 dt \,\middle|\, \mathcal{F}_\tau\right] \\ &\geq \frac{\mathbb{E}_\infty\left[e^{y_T} g(y_T)\right]}{\mathbb{E}_\infty\left[e^{y_T}\right]}. \end{aligned}$$

Theorem 2: Any stopping time T that satisfies the false alarm constraint with equality has a performance measure $J(T)$ that is no less than $J(T_C) = g(\nu)$.

Proof: To show $J(T) \geq g(\nu)$, since

$$J(T) \geq \frac{\mathbb{E}_\infty\left[e^{y_T} g(y_T)\right]}{\mathbb{E}_\infty\left[e^{y_T}\right]},$$

it is sufficient to show that

$$\frac{\mathbb{E}_\infty\left[e^{y_T} g(y_T)\right]}{\mathbb{E}_\infty\left[e^{y_T}\right]} \geq g(\nu)$$

or equivalently: $\mathbb{E}_\infty\left[e^{y_T}\{g(y_T) - g(\nu)\}\right] \geq 0$

We recall that we consider stopping times with

$$\mathbb{E}_\infty\left[\int_0^T \frac{1}{2}\alpha_t^2 dt\right] = \mathbb{E}_\infty[h(y_T)] = \gamma = h(\nu),$$

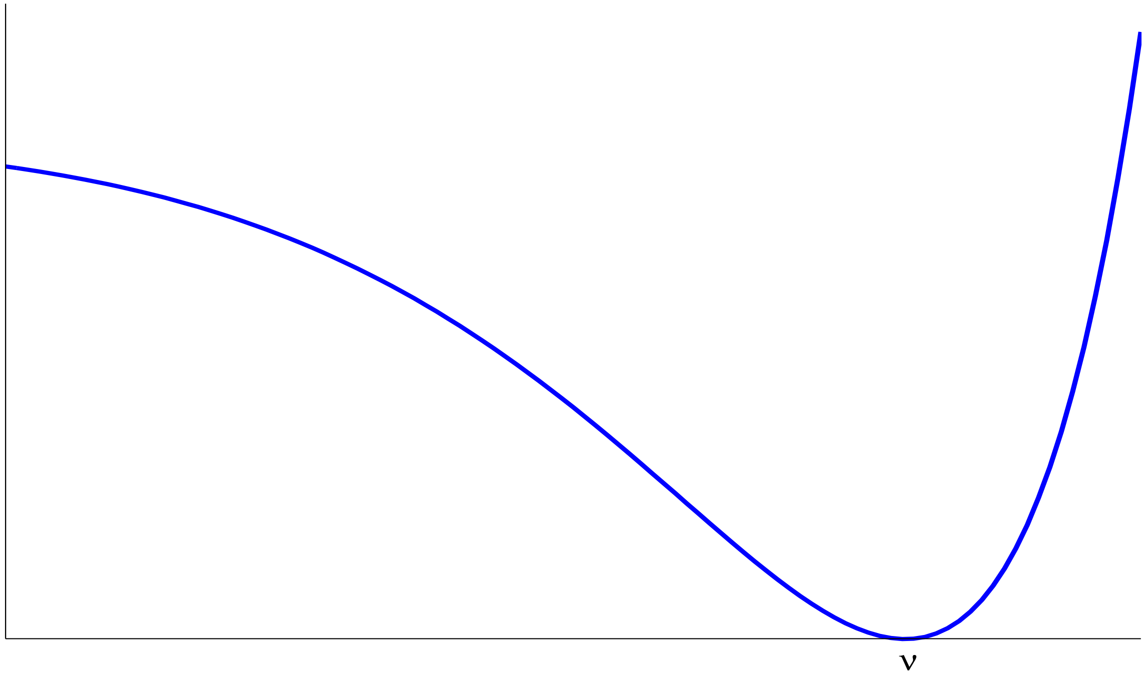
therefore the inequality we like to prove is equivalent to

$$\mathbb{E}_\infty\left[e^{y_T}\{g(y_T) - g(\nu)\} + h(\nu) - h(y_T)\right] \geq 0.$$

The function

$$p(y) = e^y\{g(y) - g(\nu)\} + h(\nu) - h(y)$$

for $y \geq 0$, can be shown to exhibit a global minimum at $y = \nu$



Because $p(\nu) = 0$, we conclude that $p(y) \geq 0$, thus

$$\mathbb{E}_\infty[p(y_T)] \geq 0$$

with equality iff $y_T = \nu$ (i.e. the CUSUM stopping time).

Conclusion

- We introduced a modification of Lorden's criterion based on the Kullback-Leibler Divergence for the problem of detecting changes in Itô processes.
- With the help of the new criterion we introduced a constrained min-max optimization problem that defines the optimum sequential scheme for the change detection problem.
- We demonstrated that the CUSUM test is the solution to the above optimization problem.