Tracking Threshold Crossing Times of a Gaussian Random Walk Through Correlated Observations

Aslan Tchamkerten

Telecom ParisTech

Joint work with Marat Burnashev, Russian Academy of Sciences (Moscow)

IWAP 2010

Outline

The Tracking Stopping Time problem

- $X = \{X_t\}_{t\geq 0}$ and stopping time τ on X
- Statistician has access to X through $Y = \{Y_t\}_{t\geq 0}$.
- Statistician wishes to find a stopping time η so that

$$\mathbb{E}|\eta - \tau|$$

is minimized.

Example: monitoring

- X_t: distance of an object from a barrier at time t
- τ: first time when $X_t = 0$
- Y_t: noisy measurements of X_t
- η: alarm time based on Y should be close to τ

Example: forecasting

- X_t: fatigue up to day t of a big manufacturing machine
- τ: first day t when X_t crosses critical fatigue threshold
- Machine replacement period: 10 days
- η: first day when new machine is operational

Wanted η close to τ because

- $\{\eta > \tau\}$: interrupted manufacturing process
- $\{\eta < \tau\}$: storage costs

$\Rightarrow$ TST problem with $Y_t = X_{t-10}$ if $t > 10$ and $X_t = 0$ else.

Example: change-point problem

- θ positive integers valued random variable
- $\{Y_t\}$ with $Y_t \sim P_0$ if $t < \theta$ and $Y_t \sim P_1$ if $t \geq \theta$
- Goal: find η such that $\mathbb{E}(\eta - \theta)_+$ is minimized while $\mathbb{P}(\eta < \theta) \leq \alpha$

Equivalent to TST problem:

- $\{X_t\}$ with $X_t = 0$ if $t < \theta$ and $X_t = 1$ if $t \geq \theta$.
- $\tau = \inf\{t \geq 1 : X_t = 1\}$ (i.e., $\tau = \theta$)
- Goal: find η such that $\mathbb{E}(\eta - \tau)_+$ minimized while $\mathbb{P}(\eta < \tau) \leq \alpha$

Change-point or tracking a stopping time?

C-P and TST are not equivalent:

for $k > t$

C-P: $\mathbb{P}(\theta = k | Y^t = y^t, \theta > t) = \mathbb{P}(\theta = k | \theta > t)$

TST: $\mathbb{P}(\tau = k | Y^t = y^t, \tau > t) \neq \mathbb{P}(\tau = k | \tau > t)$ in general

i.e., for the C-P problem,

if $\{\theta > t\}$, the first t samples Y^t are useless for predicting θ.

Bottom line

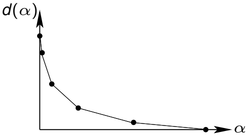
The TST problem formulation

- A useful generalization of the change-point problem.
- The change-point problem is known to be hard.
- Any hope?

A first step: algorithmic solution (Niesen and T, 2009)

Given finite alphabet process (X, Y) and $\tau \leq b$, the algorithm outputs

- $d(\alpha) = \min_{\tau:\mathbb{P}(\eta<\tau)\leq\alpha} \mathbb{E}(\eta - \tau)_+ \ \alpha \in [0, 1]$



- the corresponding optimal stopping times.

Under conditions on (X, Y) and τ, the algorithm is poly(b).

Problem formulation

Given

- $$X: \quad X_0 = 0 \qquad X_t = \sum_{i=1}^{t} V_i + st \quad t \geq 1$$

  $$Y: \quad Y_0 = 0 \qquad Y_t = X_t + \varepsilon \sum_{i=1}^{t} W_i \quad t \geq 1$$

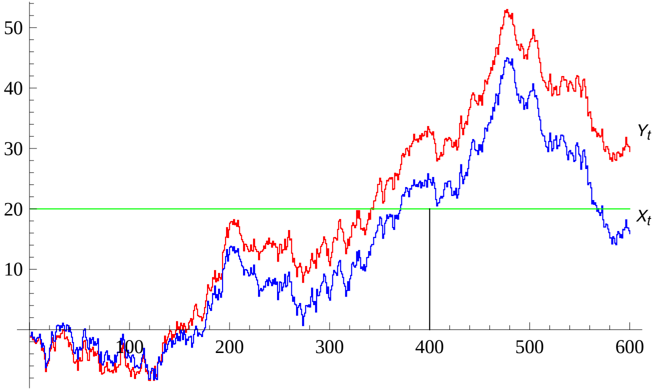
  with $s \geq 0$, $\varepsilon > 0$, and V_i's and W_i's i.i.d. $\mathcal{N}(0,1)$

- $$\tau = \inf\{t \geq 0 : X_t \geq T\} \quad T > 0$$

Find

$$\inf_{\eta} \mathbb{E}|\eta - \tau|$$

Problem formulation (cont.)



Theorem: asymptotics

For fixed $s > 0$ and $\varepsilon > 0$

$$\inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - \tau| = \inf_{\eta} \mathbb{E}|\eta - \tau| = \sqrt{\frac{2T\varepsilon^2}{\pi s^3(1+\varepsilon^2)}}\,[1 + o(1)]$$

as $T \to \infty$.

$\Rightarrow$ Causality doesn't come at the expense of delay (asym.).

Theorem: asymptotics (general)

Let $1/2 < q < 1$. In the asymptotic regime where $T/s \geq 2$,

$$s \left(\frac{T}{s}\right)^{q-1/2} \longrightarrow \infty \,,$$

and

$$\left(\frac{T}{s}\right)^{1-q} \frac{\varepsilon^2}{1+\varepsilon^2} \longrightarrow \infty \,,$$

we have

$$\inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - \tau| = \inf_{\eta} \mathbb{E}|\eta - \tau| = \sqrt{\frac{2T\varepsilon^2}{\pi s^3 (1+\varepsilon^2)}} \left[1 + o(1)\right] \,.$$

Theorem: upper bound

Fix $\varepsilon > 0$, $s > 0$, $T > 0$, and let

$$\hat{X}_0 = 0 \qquad \hat{X}_t = st + \frac{1}{1+\varepsilon^2}(Y_t - st) \qquad t \geq 1.$$

Then, $\eta = \inf\{t \geq 0 : \hat{X}_t \geq T\}$ satisfies

$$\mathbb{E}\,|\eta - \tau| \leq \sqrt{\frac{2T\varepsilon^2}{\pi(1+\varepsilon^2)s^3}} + \frac{6}{s}\left(\frac{T}{(2\pi s)^3}\right)^{1/4} + \sqrt{\frac{8(s+2)}{\pi s^3}} + 10 + \frac{20}{s}$$

Theorem: lower bound

Let $\varepsilon > 0$, $s > 0$, and $T/s \geq 2$. Then, for any integer n such that $1 \leq n < T/s$

$$\inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - \tau| \geq \sqrt{\frac{2n\varepsilon^2}{\pi s^2(1+\varepsilon^2)}} \left(1 - Q\left(\frac{T - sn}{\sqrt{n(1+\varepsilon)}}\right)\right) - \sqrt{\frac{2}{\pi s^3}} \left(T - sn + \sqrt{\frac{n}{2\pi}}\right)^{1/2} - 2 - \frac{4}{s}.$$

Theorem: case $s = 0$

If $s = 0$, $\varepsilon > 0$, and $T > 0$, then

$$\mathbb{E}|\eta - \tau|^p = \infty$$

for all $\eta = \eta(Y_0^\infty)$ and $p \geq 1/2$.

Extension to Brownian motion processes

Theorem

$$\begin{aligned} X: \quad & X_0 = 0 \qquad X_t = B_t + st \qquad t > 0 \\ Y: \quad & Y_0 = 0 \qquad Y_t = X_t + \varepsilon W_t \qquad t > 0 \end{aligned}$$

with $s \geq 0$, $\varepsilon > 0$, and $\{B_t\}$ and $\{W_t\}$ independent standard Brownian motions. Then,

$$\inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - \tau| = \inf_{\eta} \mathbb{E}|\eta - \tau| = \sqrt{\frac{2T\varepsilon^2}{\pi s^3(1+\varepsilon^2)}}\,[1 + o(1)] \ .$$

Bounds

Theorem

- for $\varepsilon > 0$, $s > 0$, and $T > 0$

$$\inf_{\eta} \mathbb{E}\,|\eta - \tau| \leq \sqrt{\frac{2T\varepsilon^2}{\pi(1+\varepsilon^2)s^3}} + \frac{6}{s}\left(\frac{T}{(2\pi s)^3}\right)^{1/4}$$

- for $\varepsilon > 0$, $s > 0$, $T/s \geq 2$ and any integer n such that $1 \leq n < T/s$

$$\inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - \tau| \geq \sqrt{\frac{2n\varepsilon^2}{\pi s^2(1+\varepsilon^2)}}\left(1 - Q\left(\frac{T - sn}{\sqrt{n(1+\varepsilon)}}\right)\right) - \sqrt{\frac{2}{\pi s^3}}\left(T - sn + \sqrt{\frac{n}{2\pi}}\right)^{1/2}.$$

TELECOM ParisTech

Upper bound: sketch

$$X: \quad X_0 = 0 \qquad X_t = \sum_{i=1}^{t} V_i + st \qquad t \geq 1$$

$$Y: \quad Y_0 = 0 \qquad Y_t = X_t + \varepsilon \sum_{i=1}^{t} W_i \qquad t \geq 1$$

- (X_t, Y_t) jointly Gaussian
- hence, linear estimator $\hat{X}_t = aY_t + b$ minimizes $\mathbb{E}|\hat{X}_t - X_t|$ (actually minimizes $\mathbb{E}|\hat{X}_t - X_t|^p$ for all $p \geq 1$),
- natural tracker: $\eta = \inf\{t \geq 1 : \hat{X}_t \geq T\}$

Upper bound: sketch (cont.)

- to evaluate $\mathbb{E}|\eta - \tau|$, first evaluate $\mathbb{E}|\hat{X}_\eta - X_\tau|$
- to evaluate $\mathbb{E}\hat{X}_\eta$ and $\mathbb{E}X_\tau$ use bounds on overshoot
- connect $\mathbb{E}|\hat{X}_\eta - X_\tau|$ to $\mathbb{E}|\eta - \tau|$ using Wald's equation ($\mathbb{E}\sum_{t=1}^{\mu} Z_t = \mathbb{E}\mu\mathbb{E}Z$)

Lower bound: sketch

Idea: reduce the problem to the estimation of X_n for $n \approx T/s$.

- $$\inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - \tau| \geq \frac{1}{s} \inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - X_n| + \text{ small}$$
- (X_n, Y_n) are jointly Gaussian, thus $X_n \overset{d}{=} aY_n + bV + c$ with $V \sim \mathcal{N}(0,1)$ independent of Y_n, hence

$$\frac{1}{s} \inf_{\eta(Y_0^\infty)} \mathbb{E}|\eta - X_n| \approx \sqrt{\frac{2n\varepsilon^2}{\pi s^2(1+\varepsilon^2)}}$$

TELECOM ParisTech

Summary

- the TST problem formulation naturally appears in
  - detection
  - prediction
  - communication
  - quality control
  - information econometrics
- Contribution: correlated gaussian random walks
  - non-asymptotic upper and lower bounds on minimum value of $\mathbb{E}|\eta - \tau|$
  - asymptotic characterization on minimum value of $\mathbb{E}|\eta - \tau|$ (extends to $\mathbb{E}|\eta - \tau|^p$, $p \geq 1$)
- A lot remains to be done:
  - more general delay penalty functions
  - Gaussianity is key for our results, what if non-gaussian?