Detection of a Stopping Time Through Noisy Observations

Aslan Tchamkerten

Telecom ParisTech

Joint work with U. Niesen (MIT)

IWSM 2009

Outline

The Tracking Stopping Time problem

Process $\{(X_i, Y_i)\}_{i \geq 1}$ whose law is known to Alice and Bob.

Alice

- chooses a stopping time S with respect to $\{X_1, X_2, \ldots\}$

Bob

- observes only $Y_1, Y_2, \ldots$
- knows the rule of S

Bob's goal

Find T so that $\mathbb{E}(T - S)^+$ is minimized while $\mathbb{P}(T < S) \leq \alpha$.

Example: monitoring

- X_n: distance at time n of an object from a barrier
- S: first time n when $X_n = 0$
- Y_n: noisy measurements of X_n

What time T should an alarm be raised?

$\mathbb{E}(T - S)^+ =$ "reaction time"

$\mathbb{P}(T < S) =$ false-alarm probability

Example: forecasting

- X_n: fatigue up to day n of a big manufacturing machine
- S: first day n for which X_n crosses critical fatigue threshold
- Machine replacement period: 10 days
- T: first day new machine is operational

Wanted T close to S because

- $T > S$: interrupted manufacturing process
- $T < S$: storage costs

$\Rightarrow$ TST problem with $Y_n = X_{n-10}$ if $n > 10$ and $X_n = 0$ else.

Example: change-point problem

- θ positive integers valued random variable
- $\{Y_i\}_{i\geq 1}$ i.i.d. with $Y_i \sim P_0$ if $i < \theta$ and $Y_i \sim P_1$ if $i \geq \theta$
- Goal: find T such that $\mathbb{E}(T-\theta)^+$ minimized while $\mathbb{P}(T < \theta) \leq \alpha$

Equivalent to:

- $\{X_i\}_{i\geq 1}$ with $X_i = 0$ if $i < \theta$ and $X_i = 1$ if $i \geq \theta$.
- $S = \inf\{i \geq 1 : X_i = 1\}$ (i.e., $S = \theta$)
- Goal: find T such that $\mathbb{E}(T-S)^+$ minimized while $\mathbb{P}(T < S) \leq \alpha$

Change-point or tracking a stopping time?

C-P and TST are not equivalent:

for $k > n$

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

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

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

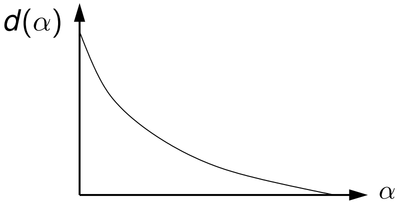
if $\{\theta > n\}$, the first n samples Y^n are useless for predicting θ.

TELECOM ParisTech

Goal

Given the law of $\{(X_i, Y_i)\}_{i\geq 1}$ and a s.t. S on $\{X_i\}_{i\geq 1}$ find

- the curve $d(\alpha) = \min_{T:\mathbb{P}(T<S)\leq\alpha} \mathbb{E}(T-S)^+$



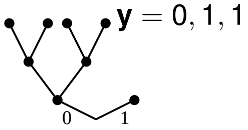
- the corresponding optimal stopping times.

Here

- discrete time, finite alphabet processes
- $S \leq \kappa$

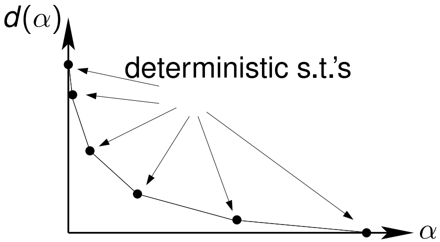
Stopping time $\equiv$ full tree

Ex: $X_i \in \{0, 1\}$, $S = \begin{cases} 1 & \text{if } X_1 = 1 \\ 3 & \text{else} \end{cases}$



What we can easily deduce

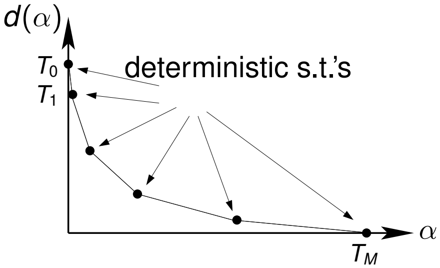
Epigraph of $d(\alpha) = \min_{T:\mathbb{P}(T<S)\leq\alpha} \mathbb{E}(T-S)^+$



- Decreasing
- Convex
- Finitely many extreme points: deterministic stopping times

$\Rightarrow$ $d(\alpha)$ is piecewise linear and convex.

Main result

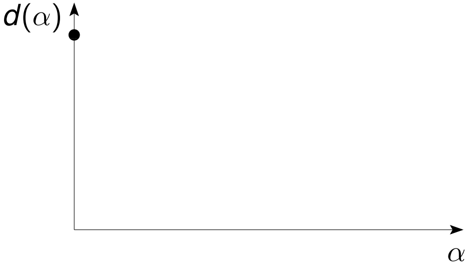


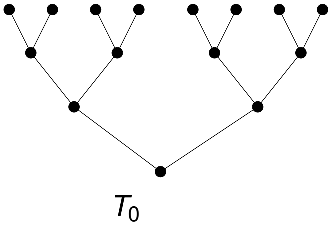
Algorithm for constructing $\{T_m\}_{m=0}^{M}$ using key property:

$$T_0 \geq T_1 \geq T_2 \geq \ldots \geq T_M$$

Construction of $\{T_m\}_{m=0}^{M}$

Tree pruning algorithm: $T_m \rightarrow T_{m+1}$

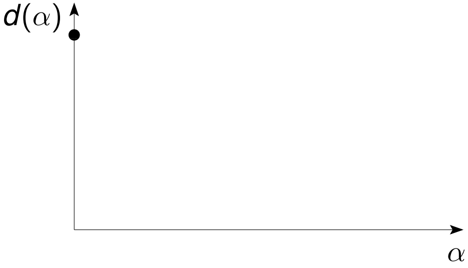


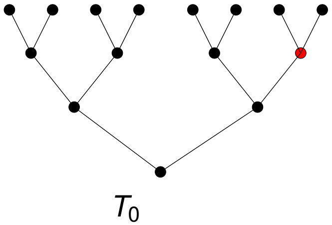


- cost associated to each intermediate nodes of T_m
- find nodes in T_m with maximum cost
- get T_{m+1} by cutting the branches above those nodes

Construction of $\{T_m\}_{m=0}^{M}$

Tree pruning algorithm: $T_m \rightarrow T_{m+1}$

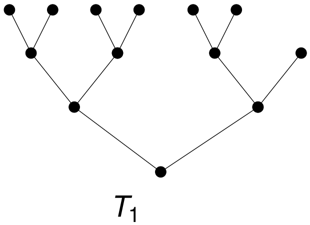




- cost associated to each intermediate nodes of T_m
- find nodes in T_m with maximum cost
- get T_{m+1} by cutting the branches above those nodes

Construction of $\{T_m\}_{m=0}^{M}$

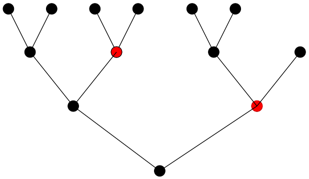
Tree pruning algorithm: $T_m \rightarrow T_{m+1}$



- cost associated to each intermediate nodes of T_m
- find nodes in T_m with maximum cost
- get T_{m+1} by cutting the branches above those nodes

Construction of $\{T_m\}_{m=0}^M$

Tree pruning algorithm: $T_m \to T_{m+1}$

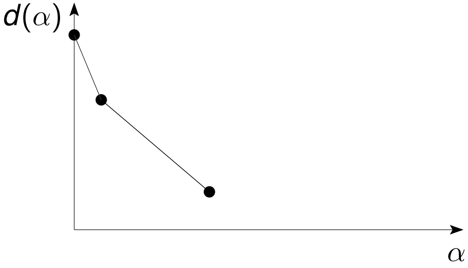


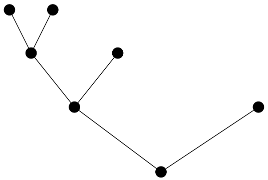
T_1

- cost associated to each intermediate nodes of T_m
- find nodes in T_m with maximum cost
- get T_{m+1} by cutting the branches above those nodes

Construction of $\{T_m\}_{m=0}^{M}$

Tree pruning algorithm: $T_m \rightarrow T_{m+1}$



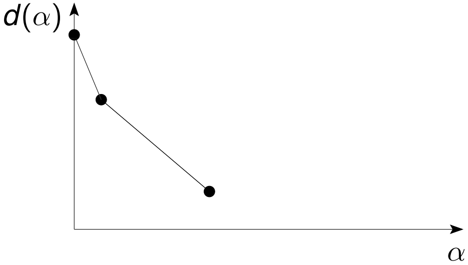


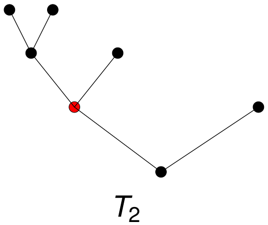
T_2

- cost associated to each intermediate nodes of T_m
- find nodes in T_m with maximum cost
- get T_{m+1} by cutting the branches above those nodes

Construction of $\{T_m\}_{m=0}^{M}$

Tree pruning algorithm: $T_m \rightarrow T_{m+1}$

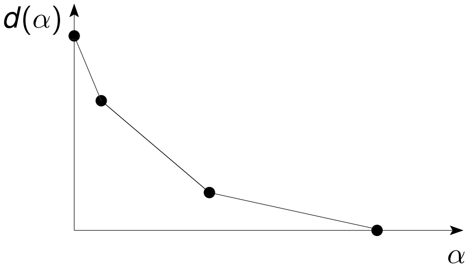


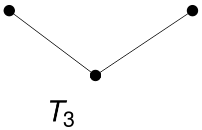


- cost associated to each intermediate nodes of T_m
- find nodes in T_m with maximum cost
- get T_{m+1} by cutting the branches above those nodes

Construction of $\{T_m\}_{m=0}^{M}$

Tree pruning algorithm: $T_m \to T_{m+1}$





- cost associated to each intermediate nodes of T_m
- find nodes in T_m with maximum cost
- get T_{m+1} by cutting the branches above those nodes

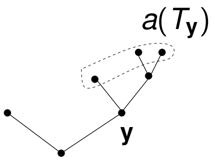
Cost function

Given a tree T_m define for all of its nodes $\mathbf{y}$

$$a(\mathbf{y}) = \mathbb{E}\left((|\mathbf{y}| - S)^+ \,\middle|\, \mathbf{Y} = \mathbf{y}\right) \mathbb{P}(\mathbf{Y} = \mathbf{y})$$
$$b(\mathbf{y}) = \mathbb{P}\left(|\mathbf{y}| < S \,\middle|\, \mathbf{Y} = \mathbf{y}\right) \mathbb{P}(\mathbf{Y} = \mathbf{y}) \, .$$

For each intermediate node $\mathbf{y}$ in T_m

$$r(\mathbf{y}) \triangleq \frac{a(T_\mathbf{y}) - a(\mathbf{y})}{b(\mathbf{y}) - b(T_\mathbf{y})} \, .$$



To get T_{m+1} cut branches of T_m above maximal cost intermediate nodes.

- Algorithm complexity: $\exp(O(\kappa))$
- Exhaustive search: $\exp(\exp(\Omega(\kappa)))$

Easy TST instances

Easy TST instances

S is permutation invariant if

$$\mathbb{P}(S \leq n | X^n = x^n) = \mathbb{P}(S \leq n | X^n = \pi(x^n))$$

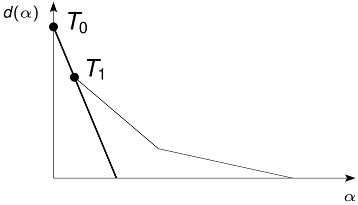
for any π and $n \geq 1$.

Theorem 1

If

- the (X_i, Y_i)'s are i.i.d.
- S is permutation invariant

then T_1 is obtained in poly(κ).



Easy TST instances

S is permutation invariant if

$$\mathbb{P}(S \leq n | X^n = x^n) = \mathbb{P}(S \leq n | X^n = \pi(x^n))$$

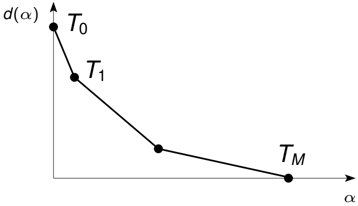
for any π and $n \geq 1$.

Theorem 2

If

- the (X_i, Y_i)'s are i.i.d.
- S and $\{T_m\}_{m=0}^{M}$ permutation invariant

then the algorithm is poly(κ).



Easy TST instances (cont.)

Example 1:

Easy TST instances (cont.)

Example 1:

- $\{(X_i, Y_i)\}_{i\geq 1}$ i.i.d. with $X_i \in \{0,1\}$
- $S \triangleq \inf\{i \geq 1 : X_i = 1 \quad \text{or} \quad i = \kappa\}$.

Note that (for $n < \kappa$)

- $\mathbb{P}(S = n) = p(1-p)^{n-1}$, where $p \triangleq \mathbb{P}(X = 1)$
- $$\mathbb{P}(Y_i = y_i | S = n) = \begin{cases} \mathbb{P}(Y_i = y_i | X_i = 0) & \text{if } i < n, \\ \mathbb{P}(Y_i = y_i | X_i = 1) & \text{if } i = n, \\ \mathbb{P}(Y_i = y_i) & \text{if } i > n. \end{cases}$$

TELECOM ParisTech

Easy TST instances (cont.)

Example 1:

- $\{(X_i, Y_i)\}_{i\geq 1}$ i.i.d. with $X_i \in \{0,1\}$
- $S \triangleq \inf\{i \geq 1 : X_i = 1 \quad \text{or} \quad i = \kappa\}$.

Note that (for $n < \kappa$)

- $\mathbb{P}(S = n) = p(1-p)^{n-1}$, where $p \triangleq \mathbb{P}(X = 1)$
- $$\mathbb{P}(Y_i = y_i | S = n) = \begin{cases} \mathbb{P}(Y_i = y_i | X_i = 0) & \text{if } i < n, \\ \mathbb{P}(Y_i = y_i | X_i = 1) & \text{if } i = n, \\ \mathbb{P}(Y_i = y_i) & \text{if } i > n. \end{cases}$$

$\Rightarrow$ change-point problem.

Easy TST problems (cont.)

Example 2: a pure TST problem

- $\{(X_i, Y_i)\}_{i\geq 1}$ be i.i.d. with $X_i \in \{0, 1\}$
- $S \triangleq \inf\{i \geq 1 : \sum_{j=1}^{i} X_j = 2\}$.

TELECOM ParisTech

Summary

- A change-point problem generalization appearing in:
  - detection
  - prediction
  - communication (e.g., feedback communication)
  - quality control
  - information econometrics
- Contribution:
  - Algorithmic solution (discrete time, finite alphabet)
  - Criterion for low complexity solution

A lot remains to be done:

- Algorithmically: explicit conditions for polynomial algorithm
- Analytically: exact results may be difficult to get. Given (X, Y) and S good lower bound on $\mathbb{E}(T-S)^2$ as a function of the correlation between X and Y and 'variability' of S?

Ref.: *U. Niesen & A. Tchamkerten, 'Tracking stopping times through noisy observations,' IEEE Transactions on Information Theory, January 2009*