The complexity of optimizing noisy functions on graphs

Mascotnum Workshop on "stochastic simulators"

Aurélien Garivier, joint project with Tomáš Kocák, Emilie Kaufmann, Pierre Ménard, Antoine Barrier, Aymen Al Marjani, etc.

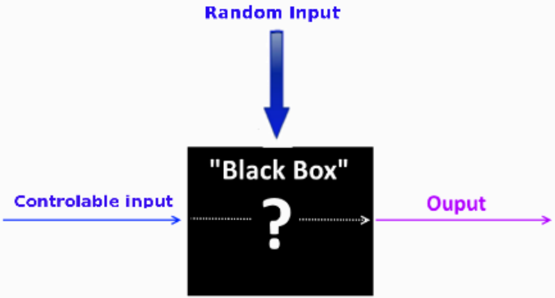
March 11^{th}, 2021

UMPA ENS DE LYON
LIP

Table of contents

PAC Optimization

Random Experiment



Random black box

$$F : \mathcal{X} \times \Omega \to \mathbb{R}$$

where $\mathcal{X}$ = space $\mathcal{X}$ of controlled variables and Ω = random space

Target function $f : \mathcal{X} \to \mathbb{R}$ $f(x) = \mathbb{E}\left[F(x, \cdot)\right]$ or some other functionnal of $F(x, \cdot)$

Typically (in the sequel): $F(x, \omega) = f(x) + \epsilon(\omega)$ where $\epsilon \sim \mathcal{N}(0, \sigma^2)$.

Black Box Optimization

Black-box interaction model:

- choose $X_1 = \phi_1(U_1)$, observe $Y_1 = F(X_1, \omega_1)$
- choose $X_2 = \phi_2(X_1, Y_1, U_2)$ observe $Y_2 = F(X_2, \omega_2)$,
- etc...

Target = optimize f: $f^* = \max_{\mathcal{X}} f$ (or min, or find level set, etc.)

Strategy: sampling rule $(\phi_t)_{t\geq 1}$, stopping time τ

PAC setting: for a risk δ and a tolerance ϵ, return $X_{\tau+1}$ such that

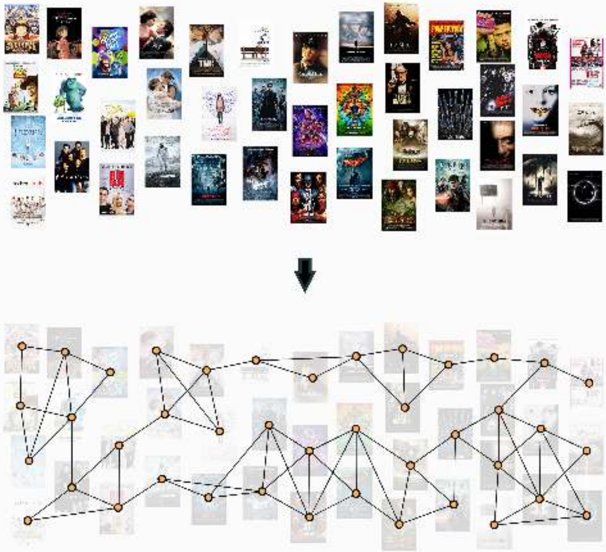
$$\mathbb{P}_f\Big(f(X_{\tau+1}) < f^* - \epsilon\Big) \leq \delta \ .$$

The set of correct answers is $\mathcal{X}_\epsilon(f) = \big\{x \in X : f(x) \geq f^* - \epsilon\big\}$.

Goal: find a strategy minimizing $\mathbb{E}_f[\tau]$ for all possible functions f
(no sub-optimal multiplicative constant!)

Optimizing on a Graph

Movie similarity graph:



Warm-up: 1.5 arms

A Simplistic and yet Interesting Example

$\mathcal{X} = \{0, 1\}$, $f = (0, \mu)$ ie. $F(0, \cdot) = 0$, $F(1, \cdot) \sim \mathcal{N}(\mu, \sigma^2)$

Here, $X_t = 1$ for $t \leq \tau$ and $Y_t \overset{iid}{\sim} \mathcal{N}(\mu, \sigma^2)$.

Equivalent to testing the overlapping hypotheses: $\mu \leq \epsilon$ vs $\mu \geq -\epsilon$.

Theorem

For every (ϵ, δ)-PAC strategy,

$$\liminf_{\delta \to 0} \frac{\mathbb{E}_\mu[\tau_\delta]}{\log(1/\delta)} \geq \frac{2\sigma^2}{\left(|\mu| + \epsilon\right)^2} .$$

Besides, the choice

$$\tau_\delta = \inf \left\{ t \in \mathbb{N} : \frac{t(|\hat{\mu}_t| + \epsilon)^2}{2\sigma^2} > 3 \ln(\ln(t) + 1) + \mathcal{T}(\ln(1/\delta)) \right\} ,$$

where $\mathcal{T}(x) \simeq x + 4 \ln\left(1 + x + \sqrt{2x}\right)$ and $X_{\tau+1} = \mathbb{1}\left\{\hat{\mu}_{\tau_\delta} > 0\right\}$ is such that $\forall \mu \in \mathbb{R}$

$$\limsup_{\delta \to 0} \frac{\mathbb{E}_\mu[\tau_\delta]}{\ln(1/\delta)} \leq \frac{2\sigma^2}{(|\mu| + \epsilon)^2} .$$

Lower Bounds: Information Inequalities

Idea: Cannot stop as long as an alternative hypothesis is likely to produce similar data.

Alternative hypotheses: $\text{Alt}(f) = \left\{g : \mathcal{X} \to \mathbb{R} \text{ such that } \mathcal{X}_\epsilon(f) \not\subset \mathcal{X}_\epsilon(g)\right\}$

and not $\mathcal{X}_\epsilon(f) \cap \mathcal{X}_\epsilon(g) = \emptyset$: just "one of the correct answers for f is incorrect for g".

Here,

- if $\mu > \epsilon$, $\text{Alt}(\mu) = (-\infty, -\epsilon)$
- if $\mu < -\epsilon$, $\text{Alt}(\mu) = (\epsilon, +\infty)$
- if $-\epsilon \le \mu \le \epsilon$, $\text{Alt}(\mu) = (-\infty, -\epsilon) \cup (\epsilon, +\infty)$

Change-of-Measure Lemma

For a given strategy, likelihood of the observation under function f: $\ell(Y_1, \dots, Y_t; f)$.

Likelihood ratio $L_t(f,g) := \log \dfrac{\ell(Y_1, \dots, Y_t; f)}{\ell(Y_1, \dots, Y_t; g)}$.

Lemma

Let f and g be two functions.

1. **Low-level form:** for all $x \in \mathbb{R}$, $n \in \mathbb{N}^*$, for every event $C \in \mathcal{F}_n$,

$$\mathbb{P}_g(C) \geq e^{-x} \left[\mathbb{P}_f(C) - \mathbb{P}_f \left(L_n(f,g) \geq x \right) \right] .$$

2. **High-level form:** for any stopping time τ and any event $C \in \mathcal{F}_\tau$,

$$\mathbb{E}_f \left[L_\tau(f,g) \right] = \mathrm{KL} \left(\mathbb{P}_f^{X_1,\dots,X_\tau}, \mathbb{P}_g^{X_1,\dots,X_\tau} \right) \geq \mathrm{kl}\big(\mathbb{P}_f(C), \mathbb{P}_g(C) \big) ,$$

where $\mathrm{kl}(x,y) = x \ln(x/y) + (1-x)\ln((1-x)/(1-y))$ is the binary relative entropy.

Proofs: Change-of-Measure Lemma

- Low-level form:

$$
\begin{aligned}
\mathbb{P}_g(C) &= \mathbb{E}_f\Big[\mathbb{1}_C \exp\big(-L_t(f,g)\big)\Big] \geq \mathbb{E}_f\Big[\mathbb{1}_C \mathbb{1}_{(L_t(f,g)<x)} e^{-L_t(f,g)}\Big] \\
&\geq e^{-x}\,\mathbb{P}_f\Big(C \cap \big(L_t(f,g) < x\big)\Big) \\
&\geq e^{-x}\left[\mathbb{P}_f(C) - \mathbb{P}_f\left(L_n(f,g) \geq x\right)\right].
\end{aligned}
$$

- High-level form: Information Theory (data-processing inequality)

$$
\begin{aligned}
\mathbb{E}_\mu\big[L_\tau(f,g)\big] &= \mathrm{KL}\left(\mathbb{P}_f^{X_1,\ldots,X_\tau}, \mathbb{P}_g^{X_1,\ldots,X_\tau}\right) \\
&\geq \mathrm{KL}\left(\mathbb{P}_f^{\mathbb{1}_C}, \mathbb{P}_g^{\mathbb{1}_C}\right) = \mathrm{kl}\big(\mathbb{P}_f(C), \mathbb{P}_g(C)\big).
\end{aligned}
$$

1.5 arm: Lower Bound for $|\mu| > \epsilon$

$\mathbb{E}_\mu[\tau] < \infty$ and Wald's lemma yield:

$$\mathbb{E}_\mu\left[L_\tau(\mu,\lambda)\right] = \mathbb{E}_\mu[\tau]\,\mathrm{KL}\big(\mathcal{N}(\mu,\sigma^2),\mathcal{N}(\lambda,\sigma^2)\big) = \mathbb{E}_\mu[\tau]\frac{(\mu-\lambda)^2}{2\sigma^2}\ .$$

Hence, by the high-level form of the lemma:

$$\mathbb{E}_\mu[\tau]\frac{(\mu-\lambda)^2}{2\sigma^2} \geq \mathrm{kl}\big(\mathbb{P}_\mu(C)\,,\mathbb{P}_\lambda(C)\big)\ .$$

For $\mu < -\epsilon$, choosing $\lambda = \epsilon$ and $C = \{X_{\tau+1} = 1\}$, which is such that $\mathbb{P}_\mu(C) \leq \delta$ and $\mathbb{P}_\lambda(C) \geq 1-\delta$, directly yields

$$\mathbb{E}_\mu[\tau]\frac{\big(|\mu|+\epsilon\big)^2}{2\sigma^2} \geq \mathrm{kl}(\delta, 1-\delta) \approx \log\frac{1}{\delta}\ .$$

Similarly, for $\mu > \epsilon$, we use $\lambda = -\epsilon$ and $C = (X_{\tau+1} = 0)$ so as to obtain the same inequality.

$\Longrightarrow$ non-asymptotic lower bound

1.5 arm: lower bound for $-\epsilon \leq \mu \leq \epsilon$

For a fixed $\eta > 0$. Introducing

$$n_\delta := \left\lfloor \frac{2\sigma^2(1-\eta)}{\left(|\mu|+\epsilon\right)^2} \ln\frac{1}{\delta} \right\rfloor \ ,$$

and the event $C_\delta = \left\{\tau_\delta \leq n_\delta\right\}$, we prove that

$$\mathbb{P}_{\boldsymbol{\mu}}\left(\tau_\delta \leq n_\delta\right) = \mathbb{P}_{\boldsymbol{\mu}}\left(C_\delta, X_{\tau+1} = 0\right) + \mathbb{P}_{\boldsymbol{\mu}}\left(C_\delta, X_{\tau+1} = 1\right) \to 0 \text{ when } \delta \to 0 \ .$$

Choosing $\lambda = \epsilon$ yields that $\mathbb{P}_\lambda(C_\delta, X_{\tau+1} = 0) \leq \mathbb{P}_\lambda(X_{\tau+1} = 0) \leq \delta$. The event $C_\delta \cap \left\{X_{\tau+1} = 0\right\}$ belongs to F_{n_δ}. Hence, for all $x \in \mathbb{R}$,

$$\delta \geq e^{-x}\left[\mathbb{P}_\mu\left(C_\delta, X_{\tau+1} = 0\right) - \mathbb{P}_\mu\left(L_{n_\delta}(\mu,\epsilon) \geq x\right)\right] \ ,$$

which can be rewritten as

$$\mathbb{P}_\mu\left(C_\delta, X_{\tau+1} = 0\right) \leq \delta e^x + \mathbb{P}_\mu\left(L_{n_\delta}(\mu,\epsilon) \geq x\right) \ .$$

The choice $x = (1-\eta/2)\ln(1/\delta)$ yields

$$\mathbb{P}_\mu\left(C_\delta, X_{\tau+1} = 0\right) \leq \delta^{\frac{\eta}{2}} + \mathbb{P}_\mu\left(\frac{L_{n_\delta}(\mu,\epsilon)}{n_\delta} \geq \frac{1-\eta/2}{1-\eta}\frac{(|\mu|+\epsilon)^2}{2\sigma^2}\right) \ .$$

By the law of large numbers and the fact that $n_\delta \to +\infty$ when $\delta \to 0$,

$$\frac{L_{n_\delta}(\mu,\epsilon)}{n_\delta} \xrightarrow{\delta\to 0} \mathbb{E}_\mu\left[\ln\frac{\ell(X_1;\mu)}{\ell(X_1;\epsilon)}\right] = \mathrm{KL}\left(\mathbb{P}_\mu^{X_1}, \mathbb{P}_\epsilon^{X_1}\right) = \frac{(\mu-\epsilon)^2}{2\sigma^2} < \frac{1-\eta/2}{1-\eta}\frac{\left(|\mu|+\epsilon\right)^2}{2\sigma^2} \ .$$

Strategy: GLRT

Idea: cannot decide as long as incompatible hypotheses remain δ-likely.

Generalized Likelihood Ratio Test:

$$\ln \frac{\max_{\mu \in \mathcal{R}} \ell(X_1, \ldots, X_t; \mu)}{\max_{\lambda \in \mathrm{Alt}(\hat{f}_t)} \ell(X_1, \ldots, X_t; \lambda)} = \ln \inf_{\lambda \in \mathrm{Alt}(\hat{\mu}_t)} \frac{\ell(X_1, \ldots, X_t; \hat{\mu}(t))}{\ell(X_1, \ldots, X_t; \lambda)}$$

$$= \inf_{\lambda \in \mathrm{Alt}(\hat{\mu}_t)} \frac{t\ (\hat{\mu}_t - \lambda)^2}{2\sigma^2}$$

where $\hat{\mu}(t) = \frac{1}{s} \sum_{i=1}^{t} X_i$. Here:

$$\tau_\delta = \inf \left\{ t \in \mathbb{N} : \lambda \in \max_{\mathrm{Alt}(\hat{\mu}_t)} \ln \frac{\ell(X_1, \ldots, X_t; \hat{\mu}(t))}{\ell(X_1, \ldots, X_t; \lambda)} > \beta(t, \delta) \right\}$$

$$= \inf \left\{ t \in \mathbb{N} : \frac{t(|\hat{\mu}_t| + \epsilon)^2}{2\sigma^2} > \beta(t, \delta) \right\}$$

and

$$X_{\tau+1} = 1\left\{ \hat{\mu}_{\tau_\delta} > 0 \right\}$$

Efficiency of the Sequential GLRT Procedure

Fix $\mu \in \mathbb{R}$ and let $\alpha \in [0, \epsilon)$.

$$\begin{aligned}
\mathbb{E}[\tau_\delta] &\leq \sum_{t=1}^{\infty} \mathbb{P}\left(t(|\hat{\mu}_t| + \epsilon)^2 \leq 2\sigma^2 \beta(t,\delta)\right) \\
&\leq \sum_{t=1}^{\infty} \mathbb{P}\left(|\hat{\mu}_t - \mu| > \alpha\right) + \sum_{t=1}^{\infty} \mathbb{P}\left(t(|\hat{\mu}_t| + \epsilon)^2 \leq 2\sigma^2 \beta(t,\delta), |\hat{\mu}_t - \mu| \leq \alpha\right) \\
&\leq \sum_{t=1}^{\infty} \mathbb{P}\left(|\hat{\mu}_t - \mu| > \alpha\right) + \sum_{t=1}^{\infty} \mathbb{P}\left(t(|\mu| - \alpha + \epsilon)^2 \leq 2\sigma^2 \beta(t,\delta), |\hat{\mu}_t - \mu| \leq \alpha\right) .
\end{aligned}$$

The first term is upper bounded by a constant (independent of δ), while the second is upper bounded by

$$T_0(\delta) = \inf\left\{T \in \mathbb{N}^* : \forall t \geq T, t(|\mu| - \alpha + \epsilon)^2 \leq 2\sigma^2 \beta(t,\delta)\right\} .$$

For $\beta(t,\delta) = 3\ln(\ln(t) + 1) + \mathcal{T}(\ln(1/\delta))$, since for every $\alpha \geq 0$, $\gamma \geq 1 + \alpha$ and $t > 0$,

$$t \geq \gamma + 2\alpha \ln(\gamma) \quad \Rightarrow \quad t \geq \gamma + \alpha \ln(t) ,$$

we have

$$T_0(\delta) = \frac{2\sigma^2}{(|\mu| - \alpha + \epsilon)^2} \ln\frac{1}{\delta} + o_{\delta \to 0}\left(\ln\frac{1}{\delta}\right) .$$

Letting α go to zero, one obtains, for all $\mu \in \mathbb{R}$,

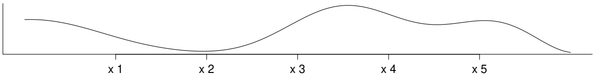
$$\limsup_{\delta \to 0} \frac{\mathbb{E}_\mu[\tau_\delta]}{\ln(1/\delta)} \leq \frac{2\sigma^2}{(|\mu| + \epsilon)^2} .$$

Discrete, Non-smooth Functions – The Vanilla Bandit Model

Best-Arm Identification with Fixed Confidence

K options = probability distributions $\boldsymbol{\nu} = (\nu_a)_{1 \leq a \leq K}$

$\nu_a \in \mathcal{F}$ exponential family parameterized by its expectation μ_a



At round t, you may:

- choose an option $A_t = \phi_t(A_1, X_1, \dots, A_{t-1}, X_{t-1}) \in \{1, \dots, K\}$
- observe a new independent sample $X_t \sim \nu_{A_t}$

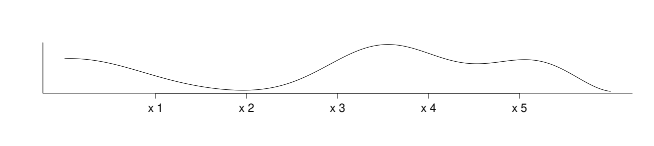
so as to identify the best option $a^* = \text{argmax}_a \, \mu_a$ and $\mu^* = \max_a \mu_a$
as fast as possible: stopping time τ .

Fixed-budget setting	Fixed-confidence setting
given $\tau = T$	minimize $\mathbb{E}[\tau]$
minimize $\mathbb{P}(\hat{a}_\tau \neq a^*)$	under constraint $\mathbb{P}(\hat{a}_\tau \neq a^*) \leq \delta$

Best-Arm Identification with Fixed Confidence

K options = probability distributions $\boldsymbol{\nu} = (\nu_a)_{1 \leq a \leq K}$

$\nu_a \in \mathcal{F}$ exponential family parameterized by its expectation μ_a



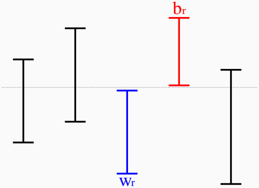
At round t, you may:

- choose an option $A_t = \phi_t\left(A_1, X_1, \ldots, A_{t-1}, X_{t-1}\right) \in \{1, \ldots, K\}$
- observe a new independent sample $X_t \sim \nu_{A_t}$

so as to identify the best option $a^* = \operatorname{argmax}_a \mu_a$ and $\mu^* = \max_a \mu_a$
as fast as possible: stopping time τ_δ .

Fixed-budget setting	Fixed-confidence setting
given $\tau = T$ minimize $\mathbb{P}(\hat{a}_\tau \neq a^*)$	minimize $\mathbb{E}[\tau_\delta]$ under constraint $\mathbb{P}(\hat{a}_\tau \neq a^*) \leq \delta$

Racing Strategy for $\epsilon = 0$, see [Kaufmann & Kalyanakrishnan '13]



$\mathcal{R} := \{1, \ldots, K\}$ set of remaining arms.

$r := 0$ current round

while $|\mathcal{R}| > 1$

- $r := r + 1$
- draw each $x \in \mathcal{R}$, compute $\hat{f}_r(x)$, the empirical mean of the r samples observed so far
- compute the empirical best and empirical worst arms:

$$b_r = \operatorname*{argmax}_{x \in \mathcal{R}} \hat{f}_r(x) \qquad w_r = \operatorname*{argmin}_{x \in \mathcal{R}} \hat{f}_r(x)$$

  as well as an ucb for w_r and an lcb for b_r.

- Elimination step: if

$$\ell_{b_r}(r) > u_{w_r}(r),$$

  then eliminate w_r : $\mathcal{R} := \mathcal{R} \backslash \{w_r\}$

end

Outpout: $\hat{x}$ the single element remaining in $\mathcal{R}$.

The case $\epsilon = 0$

Theorem [G. and Kaufmann, 2016]

For any δ-correct algorithm,

$$\mathbb{E}_{\boldsymbol{\mu}}[\tau] \geq T^*(\boldsymbol{\mu}) \,\mathrm{kl}(\delta, 1-\delta) \,,$$

where

$$\begin{aligned} T^*(\boldsymbol{\mu})^{-1} &= \sup_{w \in \Sigma_K} \inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} w(x) \,\mathrm{kl}\big(f(x), g(x)\big) \\ &= \max_{w \in \Sigma_K} \min_{y \neq x^*(f)} \inf_{f(y) \leq \lambda \leq f^*} w\big(x^*(f)\big) \,\mathrm{kl}\big(f^*, \lambda\big) + w(y) \,\mathrm{kl}\big(f(y), \lambda\big) \,. \end{aligned}$$

- A kind of game : you choose the proportions of draws $w \in \Sigma_K$, the opponent chooses the alternative confusing function.

➜ the optimal proportions of arm draws are

$$w^{*,f} = \operatorname*{argmax}_{w \in \Sigma_K} \inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} w(x) \,\mathrm{kl}\big(f(x), g(x)\big) \,.$$

Entropic Lower Bound (high-level form)

Assume that f and g have a different maximum $x^*(f) \neq x^*(g)$. Then

$$\begin{aligned}
\sum_{x \in \mathcal{X}} \mathbb{E}_f\left[N_x(\tau)\right] \mathrm{kl}\big(f(x), g(x)\big) &= \mathrm{KL}\Big(\mathbb{P}_f^{(X_1,\dots,X_\tau)}, \mathbb{P}_g^{(X_1,\dots,X_\tau)}\Big) \\
&\geq \mathrm{KL}\big(\mathbb{P}_f^{\mathbb{1}\{\hat{x}_\tau = x^*(f)\}}, \mathbb{P}_g^{\mathbb{1}\{\hat{x}_\tau = x^*(f)\}}\big) \\
&\geq \mathrm{kl}\Big(\mathbb{P}_f\left(\hat{x}_\tau = x^*(f)\right), \mathbb{P}_g\left(\hat{x}_\tau = x^*(f)\right)\Big) \\
&\geq \mathrm{kl}(1-\delta, \delta)\ .
\end{aligned}$$

Entropic Lower Bound

[Kaufmann, Cappé, G.'15],[G., Ménard, Stoltz '16]

For every δ-correct procedure, if $x^*(f) \neq x^*(g)$ then

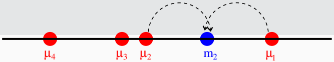
$$\sum_{x \in \mathcal{X}} \mathbb{E}_f\left[N_x(\tau)\right] \mathrm{kl}\big(f(x), g(x)\big) \geq \mathrm{kl}\big(1-\delta, \delta\big)\ .$$

Combining the Entropic Lower Bounds

Entropic Lower Bound

If $x^*(f) \neq x^*(g)$, any δ-correct algorithm satisfies

$$\sum_{x\mathcal{X}} \mathbb{E}_f \left[N_x(\tau) \right] \mathrm{kl}\big(f(x), g(x)\big) \geq \mathrm{kl}(\delta, 1-\delta).$$



Let $\mathrm{Alt}(f) = \{g : x^*(g) \neq x^*(f)\}$.

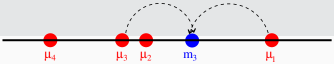
$$\begin{aligned}
\inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} \mathbb{E}_{\boldsymbol{\mu}}[N_x(\tau)] \, \mathrm{kl}\big(f(x), g(x)\big) &\geq \mathrm{kl}(\delta, 1-\delta) \\
\mathbb{E}_f[\tau] \times \inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} \frac{\mathbb{E}_f[N_x(\tau)]}{\mathbb{E}_f[\tau]} \, \mathrm{kl}\big(f(x), g(x)\big) &\geq \mathrm{kl}(\delta, 1-\delta) \\
\mathbb{E}_f[\tau] \times \left(\sup_{w \in \Sigma_K} \inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} w(x) \, \mathrm{kl}\big(f(x), g(x)\big) \right) &\geq \mathrm{kl}(\delta, 1-\delta)
\end{aligned}$$

Combining the Entropic Lower Bounds

Entropic Lower Bound

If $x^*(f) \neq x^*(g)$, any δ-correct algorithm satisfies

$$\sum_{x\mathcal{X}} \mathbb{E}_f\left[N_x(\tau)\right] \mathrm{kl}\big(f(x), g(x)\big) \geq \mathrm{kl}(\delta, 1-\delta).$$



Let $\mathrm{Alt}(f) = \{g : x^*(g) \neq x^*(f)\}$.

$$\inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} \mathbb{E}_{\boldsymbol{\mu}}[N_x(\tau)]\, \mathrm{kl}\big(f(x), g(x)\big) \quad \geq \quad \mathrm{kl}(\delta, 1-\delta)$$

$$\mathbb{E}_f[\tau] \times \inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} \frac{\mathbb{E}_f[N_x(\tau)]}{\mathbb{E}_f[\tau]}\, \mathrm{kl}\big(f(x), g(x)\big) \quad \geq \quad \mathrm{kl}(\delta, 1-\delta)$$

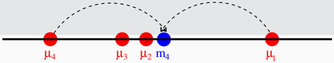
$$\mathbb{E}_f[\tau] \times \left(\sup_{w \in \boldsymbol{\Sigma}_K} \inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} w(x)\, \mathrm{kl}\big(f(x), g(x)\big) \right) \quad \geq \quad \mathrm{kl}(\delta, 1-\delta)$$

Combining the Entropic Lower Bounds

Entropic Lower Bound

If $x^*(f) \neq x^*(g)$, any δ-correct algorithm satisfies

$$\sum_{x\mathcal{X}} \mathbb{E}_f\left[N_x(\tau)\right] \mathrm{kl}\big(f(x), g(x)\big) \geq \mathrm{kl}(\delta, 1-\delta).$$



Let $\mathrm{Alt}(f) = \big\{g : x^*(g) \neq x^*(f)\big\}$.

$$\inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} \mathbb{E}_{\boldsymbol{\mu}}[N_x(\tau)] \,\mathrm{kl}\big(f(x), g(x)\big) \geq \mathrm{kl}(\delta, 1-\delta)$$

$$\mathbb{E}_f[\tau] \times \inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} \frac{\mathbb{E}_f[N_x(\tau)]}{\mathbb{E}_f[\tau]} \,\mathrm{kl}\big(f(x), g(x)\big) \geq \mathrm{kl}(\delta, 1-\delta)$$

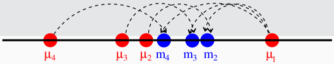
$$\mathbb{E}_f[\tau] \times \left(\sup_{w \in \Sigma_K} \inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} w(x) \,\mathrm{kl}\big(f(x), g(x)\big) \right) \geq \mathrm{kl}(\delta, 1-\delta)$$

Combining the Entropic Lower Bounds

Entropic Lower Bound

If $x^*(f) \neq x^*(g)$, any δ-correct algorithm satisfies

$$\sum_{x\mathcal{X}} \mathbb{E}_f\left[N_x(\tau)\right] \mathrm{kl}\big(f(x), g(x)\big) \geq \mathrm{kl}(\delta, 1-\delta).$$



Let $\mathrm{Alt}(f) = \big\{g : x^*(g) \neq x^*(f)\big\}$.

$$\inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} \mathbb{E}_{\boldsymbol{\mu}}[N_x(\tau)]\, \mathrm{kl}\big(f(x), g(x)\big) \geq \mathrm{kl}(\delta, 1-\delta)$$

$$\mathbb{E}_f[\tau] \times \inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} \frac{\mathbb{E}_f[N_x(\tau)]}{\mathbb{E}_f[\tau]}\, \mathrm{kl}\big(f(x), g(x)\big) \geq \mathrm{kl}(\delta, 1-\delta)$$

$$\mathbb{E}_f[\tau] \times \left(\sup_{w \in \Sigma_K} \inf_{g \in \mathrm{Alt}(f)} \sum_{x \in \mathcal{X}} w(x)\, \mathrm{kl}\big(f(x), g(x)\big)\right) \geq \mathrm{kl}(\delta, 1-\delta)$$

Information Complexity of PAC Optimization

Theorem

For any ϵ-PAC family of converging strategies,

$$\liminf_{\delta \to 0} \frac{\mathbb{E}_f[\tau_\delta]}{\ln(1/\delta)} \geq T_\epsilon^*(f) = \sup_{w \in \Sigma_K} \max_{x \in \mathcal{X}_\epsilon(f)} \min_{y \neq x} \inf_{\substack{(\lambda_x, \lambda_y): \\ \lambda_x \leq \lambda_y - \epsilon}} w(x) \frac{f(x) - g(x)}{2} + w(y) \frac{(f(y) - g(y)}{2},$$

where $\Sigma_K = \left\{ w \in [0,1]^{\mathcal{X}} : \sum_x w(x) = 1 \right\}$.

- Two armed-case: $T_\epsilon^*(f) = \frac{8\sigma^2}{(\operatorname{osc}(f)+\epsilon)^2}$.
- Approximation:

$$\sum_{y \neq x^*(f)}^{K} \frac{2\sigma^2}{\left(f^* + \epsilon - f(y)\right)^2} \leq T_\epsilon^*(f) \leq 2 \times \sum_{y \neq x^*(f)}^{K} \frac{2\sigma^2}{\left(f^* + \epsilon - f(y)\right)^2} .$$

- Computation: Denoting $f^{x,\epsilon}$ the bandit instance such that $f^{x,\epsilon}(y) = f(y)$ for all $y \neq x$ and $f^{x,\epsilon}(x) = f(x) + \epsilon$,

$$T_\epsilon^*(f) = \min_{x \in \mathcal{X}_\epsilon(f)} T_0\left(f^{x,\epsilon}\right) = T_0\left(f^{x^*(f),\epsilon}\right),$$

where $x^*(f) \in \mathcal{X}^*(f)$ is an optimal arm $\implies$ simple and efficient algorithm.

GLRT stopping rule

$$\begin{aligned} \tau_\delta &= \inf\left\{ t \in \mathbb{N} : \max_{x \in \mathcal{X}} \inf_{g \in \mathcal{R} \setminus \mathcal{R}_x} \sum_{y \in \mathcal{X}} N_y(t) d\left(\hat{f}_t(y), g(y)\right) > \beta(t, \delta) \right\} , \\ X_{\tau+1} &= \operatorname*{argmax}_{x=1,\dots,M} \inf_{g \in \mathcal{R} \setminus \mathcal{R}_x} \sum_{y \in \mathcal{X}} N_y(\tau_\delta) d\left(\hat{f}_{\tau_\delta}(y), \lambda_j\right)) = \operatorname*{argmax}_{x \in \mathcal{X}} \hat{f}_\tau(x) . \end{aligned}$$

Lemma

For any sampling rule, the parallel GLRT test $(\tau_\delta, X_{\tau+1})$ using the threshold function

$$\beta(t, \delta) = 3K \ln(1 + \ln t) + K\mathcal{T}\left(\frac{\ln(1/\delta)}{K}\right)$$

is δ-correct: for all $f \in \mathcal{R}$, $\mathbb{P}_f\left(\tau_\delta < \infty, \hat{X}_{\tau+1} \notin \mathcal{X}^*(f)\right) \leq \delta$.

Tracking the Optimal Proportions

Sampling rule:

$$X_{t+1} \in \begin{cases} \operatorname*{argmin}_{x \in U_t} N_x(t) \quad \text{if } U_t \neq \emptyset & (\textit{forced exploration}), \text{ or otherwise} \\ \operatorname*{argmax}_{x \in \mathcal{X}} t \times w_\epsilon^{*,\hat{f}_t(x)}(x) - N_x(t) & (\textit{tracking the plug-in estimate}). \end{cases}$$

An instance (f, ϵ) is regular if the set

$$\mathcal{W}_\epsilon^*(f) = \operatorname*{argmax}_{w \in \Sigma_K} \max_{x \in \mathcal{X}_\epsilon} \min_{y \neq y} \inf_{\substack{(\lambda_x, \lambda_y): \\ \lambda_x \leq \lambda_y - \epsilon}} w(x) \frac{\big(f(x) - g(x)\big)^2}{2} + w(y) \frac{\big(f(y) - g(y)\big)^2}{2}$$

is of cardinality one. For a regular instance, $\mathcal{W}_\epsilon^*(f) = \{w_\epsilon^*(f)\}$.

Lemma

Let (f, ϵ) be a regular instance of almost optimal best arm identification. Then, under the ϵ-Tracking sampling rule,

$$\mathbb{P}_f \left(\forall x \in \mathcal{X}, \quad \lim_{t \to \infty} \frac{N_x(t)}{t} = w_\epsilon^{*,f}(x) \right) = 1 \, .$$

Optimality of the Track-and-Stop Algorithm

Theorem

For every $\delta \in (0, 1]$, the ϵ-TaS(δ) algorithm is (ϵ, δ)-PAC. Moreover, for every instance (f, ϵ), if τ_δ denotes the stopping rule of ϵ-TaS(δ),

$$\limsup_{\delta \to 0} \frac{\mathbb{E}_f[\tau_\delta]}{\ln(1/\delta)} \leq T^*_\epsilon(f)\ .$$

Experiments

Two regular instances

$$\begin{array}{lclcl} f_1 & = & [0.2 \;\; 0.4 \;\; 0.5 \;\; 0.55 \;\; 0.7] & \epsilon_1 & = & 0.1 \\ f_2 & = & [0.4 \;\; 0.5 \;\; 0.6 \;\; 0.7 \;\; 0.75 \;\; 0.8] & \epsilon_2 & = & 0.15 \,, \end{array}$$

and one non-regular instance
$f_3 = [0.2 \;\; 0.3 \;\; 0.45 \;\; 0.55 \;\; 0.6 \;\; 0.6] \quad \epsilon_3 = 0.1$.

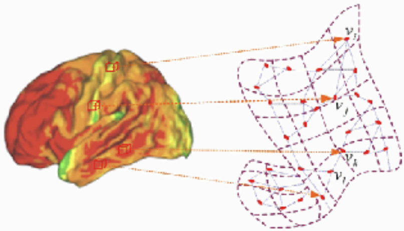
	$T^*(f)\ln(1/\delta)$	ϵ-TaS	KL-LUCB	UGapE	KL-Racing
f_1, $\epsilon_1 = 0.1$	97	171 (104)	322 (137)	324 (143)	372 (159)
f_2, $\epsilon_2 = 0.15$	108	162 (83)	345 (135)	344 (141)	402 (146)
f_3, $\epsilon_3 = 0.1$	531	501 (261)	1236 (403)	1199 (414)	1348 (436)

Table 1: Estimated values of $\mathbb{E}_{\boldsymbol{\mu}_i}[\tau_\delta]$ based on $N = 1000$ repetitions for different instances and algorithms (standard deviation indicated in parenthesis).

Graph Structure and Smooth Functions

Graph Smoothness

See [*Spectral bandits for smooth graph functions*, by M. Valko, R. Munos, B. Kveton, T. Kocák]



Src: A Spectral Graph Regression Model for Learning Brain Connectivity of Alzheimer's Disease ,by Hu, Cheng, Sepulcre and Li

Weighted graph structure with adjacency matrix $\boldsymbol{W} = (w_{x,y})_{x,y\in\mathcal{X}}$, and

$$S_G(f) \triangleq \sum_{x,y\in\mathcal{X}} w_{x,y} \frac{\big(f(x) - f(y)\big)^2}{2} = f^T \mathcal{L} f = \|f\|_{\mathcal{L}}^2 \leq R$$

for some (known) smoothness parameter R, where $\mathcal{L}$ is the **graph Laplacian**: $\mathcal{L}_{x,y} = -w_{x,y}$ for $x \neq y$ and $\mathcal{L}_{x,x} = \sum_{y\neq x} w_{x,y}$.

$\implies$ The values of f at two points $x, y \in \mathcal{X}$ are close if $w_{x,y}$ is large.

Complexity of Graph BAI

The set of considered signals is

$$\mathcal{M}_R = \left\{ f \in \mathbb{R}^K : f^\intercal \mathcal{L} f \leq R \right\}$$

Proposition

For any δ-correct strategy and any $R-$smooth function f,

$$\mathbb{E}_f[T_\delta] \geq T_R^*(f)\, \mathrm{kl}(\delta, 1-\delta)$$

where

$$T_R^*(f)^{-1} \triangleq \sup_{\boldsymbol{\omega} \in \Delta_K} \inf_{g \in \mathcal{A}_R(f)} \sum_{x \in \mathcal{X}} \omega_x \frac{\left(f(x) - g(x)\right)^2}{2} \tag{1}$$

for the set

$$\mathcal{A}_R(f) \triangleq \left\{ g \in \mathcal{M}_R : \exists x \in \mathcal{X} \setminus \{\mathcal{X}_*(f)\}, g(x) \geq g\left(\mathcal{X}_*(\boldsymbol{\mu})\right) \right\}$$

Computing the Complexity

Lemma

Let $\mathcal{D} \subseteq \mathbb{R}^K$ be a compact set. Then function $f : \Delta_K \to \mathbb{R}$ defined as $f(\boldsymbol{\omega}) = \inf_{d \in \mathcal{D}} \boldsymbol{\omega}^\top d$ Is a concave function and $d^*(\boldsymbol{\omega}) = \arg\min_{d \in \mathcal{D}} \boldsymbol{\omega}^\top d$ is a supergradient of f at $\boldsymbol{\omega}$.

Lemma

Let $h : \Delta_K \to \mathbb{R}$ be a function such that

$$h(w) = \inf_{g \in \mathrm{Alt}_R(f)} \sum_{x \in \mathcal{X}} w(x) \, \frac{\left(f(x) - g(x)\right)^2}{2}$$

Then function h is L-Lipschitz with respect to $\|\cdot\|_1$ for any

$$L \geq \max_{x, y \in \mathcal{X}} \frac{\left(f(x) - f(y)\right)^2}{2}.$$

Mirror Ascent Algorithm

One can compute the best response g_w^* and therefore, the supergradient of concave function $h(w) = \sum_{x \in \mathcal{X}} \omega_x k(f(x), g_w^*(x))$ at w.

Proposition

For $(\Delta_K, \|\cdot\|_1)$, $w_1 = (\frac{1}{K}, \ldots, \frac{1}{K})^\intercal$, $Q = \sqrt{\log K}$, for any $L \geq \max_{x,u \in \mathcal{X}} \frac{\left(f(x) - f(y)\right)^2}{2}$, mirror ascent with generalized negative entropy as the mirror map

$$\Phi(w) = \sum_{x \in \mathcal{X}} w(x) \log\left(w(x)\right)) - w(x)$$

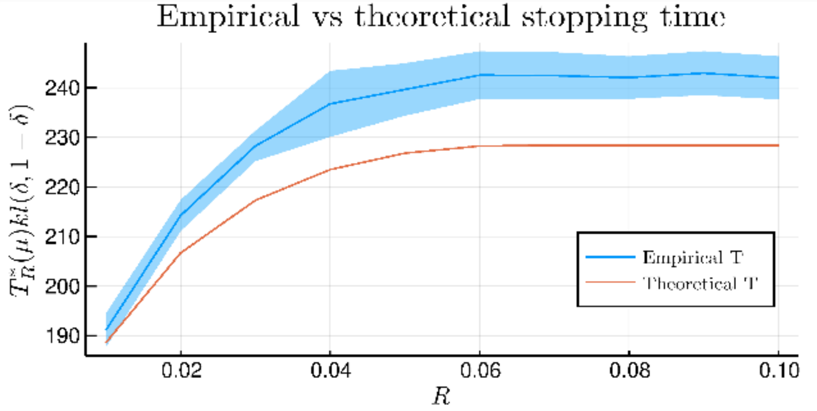
and $\eta = \frac{1}{L}\sqrt{\frac{2\log K}{t}}$ satisfies:

$$h(w^*) - h\left(\frac{1}{t}\sum_{s=1}^{t} w_s\right) \leq L\sqrt{\frac{2\log K}{t}}.$$

SpectralTaS

```
Input and initialization:
    𝓛 : graph Laplacian
    δ : confidence parameter
    R : upper bound on the smoothness of f
    Play each x ∈ 𝒳 once and observe rewards F(x, ω)
    f̂_t = empirical estimate of f
while Stopping Rule not satisfied do
    Compute ω*(f̂_t) by mirror ascent
    Choose X_t according to Tracking Sampling Rule
    Obtain reward F(X_t, ω_t)
    Update f̂_t accordingly
end while
    Output arm x̂ = argmax_{x∈𝒳} f̂(x)
```

Experiment


Empirical vs theoretical stopping time
$T_R^*(\mu) kl(\delta, 1-\delta)$
240
230
220
210
200
190
0.02
0.04
0.06
0.08
0.10
R
Empirical T
Theoretical T

Optimizing a Continuous Function

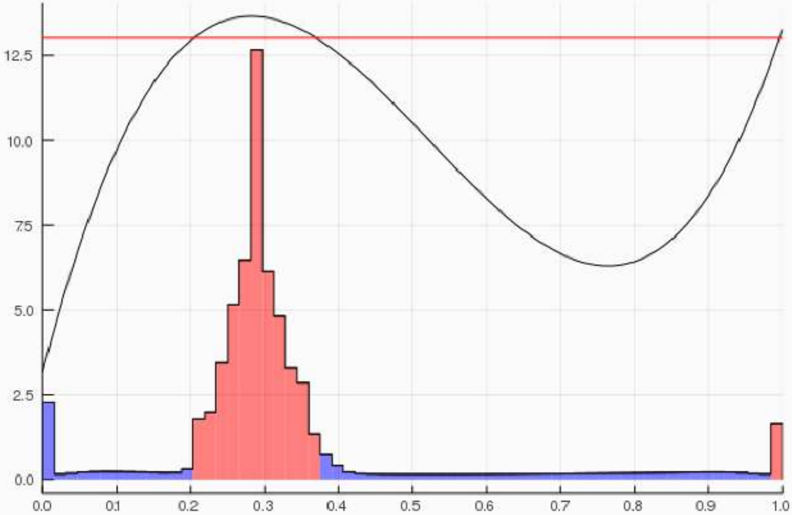
- Need a way to encode the *regularity* of the function.
- Discretization: gaph = grid

- Idea 1: Lipschitz $\implies$ neighbors means differ by at most L
- Idea 2: Graph Laplacian regularity

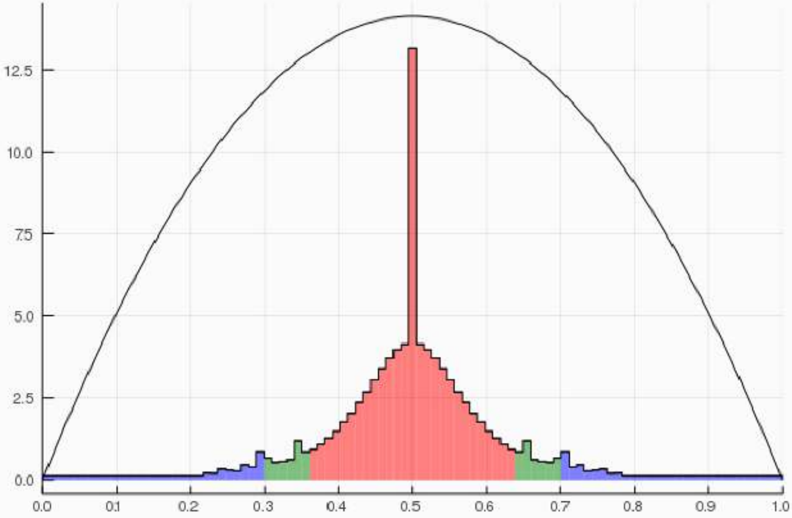
$$\implies \int_{R^d} \|\nabla f(x)\|^2 \, dx \leq C \, .$$

- Question: what happens when the grid refines? Is there a *limit*?
  - Limiting Complexity of ϵ, δ-PAC optimization?
  - Limiting density of optimal arm draws?
  - How do known methods compare?

Example



Example (large ϵ)



Example (small ϵ)