VROOM:
A very robust online optimisation method

Victor Gabillon, Rasul Tutunov, Michal Valko,
Haitham Bou Ammar

AISTATS 2020
Problem setting: black-box optimisation

In budgeted online optimisation, a learner optimises \( f : \mathcal{X} \to \mathbb{R} \). We consider a general case where \( f \) is decomposable as,

\[
 f = \frac{1}{n} \sum_{t=1}^{n} f_t.
\]

At each round \( t \in \{1, \ldots, n\} \), the learner chooses an element \( x_t \in \mathcal{X} \) and observes a real number \( y_t \), where \( y_t = f_t(x_t) \). no gradient, zero-order optimisation

Objective: Study the optimisation problem under different assumption on the \( f_1, \ldots, f_n \)
In budgeted online optimisation, a learner optimises \( f : \mathcal{X} \rightarrow \mathbb{R} \). We consider a general case where \( f \) is decomposable as,

\[
f = \frac{1}{n} \sum_{t=1}^{n} f_t.
\]

At each round \( t \in \{1, \ldots, n\} \),
the learner chooses an element \( x_t \in \mathcal{X} \) and observes a real number \( y_t \), where \( y_t = f_t(x_t) \). 

**Objective:** Study the optimisation problem under different assumption on the \( f_1, \ldots, f_n \)
Problem setting: black-box optimisation

In budgeted online optimisation, a learner optimises $f : \mathcal{X} \rightarrow \mathbb{R}$. We consider a general case where $f$ is decomposable as,

$$f = \frac{1}{n} \sum_{t=1}^{n} f_t.$$

At each round $t \in \{1, \ldots, n\}$, the learner chooses an element $x_t \in \mathcal{X}$ and observes a real number $y_t$, where $y_t = f_t(x_t)$. **no gradient, zero-order optimisation**

**Objective:** Study the optimisation problem under different assumption on the $f_1, \ldots, f_n$
**Assumptions: Two regimes**

**Stochastic feedback**: At any round, we have $f_t = \bar{f} + \epsilon_t$ with $\epsilon_t$ distributed (i.i.d.) over rounds.

\[ \mathbb{E}[\epsilon_t] = 0 \quad \text{and} \quad |\epsilon_t| \leq b. \quad (1) \]

**Non-stochastic feedback** we minimally assume:

\[ |f_{t'}(x) - f_t(x)| \leq b \quad \text{for all} \quad t, t' \quad \text{and} \quad x \in \mathcal{X}. \quad (2) \]

Actually we will sometimes rephrase this condition as the equivalent condition $|f_t(x)| \leq f_{\text{max}}$ for all $x \in \mathcal{X}$ and $t \in [n]$. 


The regret

The learner recommends after round \( n \), the element \( x(n) \) and minimises the **simple regret** \( r_n \).

**Stochastic case:** Expected regret

\[
\mathbb{E}_f[r_n] \triangleq \mathbb{E}_{f_1,\ldots,f_n} \left[ \sup_{x \in \mathcal{X}} f(x) - \mathbb{E}_{x(n)}[f(x(n))] \right] \\
= \sup_{x \in \mathcal{X}} \bar{f}(x) - \mathbb{E}_{x(n)}[\bar{f}(x(n))].
\]

**Non-stochastic setting:** A regret for any sequence \( f_1, \ldots, f_n \)

\[
 r_n \triangleq \sup_{x \in \mathcal{X}} f(x) - \mathbb{E}_{x(n)}[f(x(n))],
\]
Introducing the tools and the minimal assumptions
Partitioning

- For any depth $h$, $\mathcal{X}$ is partitioned in $K^h$ cells $(\mathcal{P}_{h,i})_{0 \leq K^h - 1}$.
- $K$-ary tree $\mathcal{T}$ where depth $h = 0$ is the whole $\mathcal{X}$.

An example of partitioning in one dimension with $K = 3$. 
Tree-based learners: use the partitioning to explore $f$ (uniformly)
Tree-based learners: use the partitioning to explore $f$ (uniformly)
Tree-based learners: use the partitioning to explore \( f \) (uniformly)
Tree-based learners: use the partitioning to explore $f$ (uniformly)
The assumption and the smoothness

Assumption (on the local smoothness around $x^*$)

For any global optimum $x^*$, there exists $\nu > 0$ and $\rho \in (0, 1)$, ($\nu$, $\rho$ depend on $x^*$), such that $\forall h \in \mathbb{N}, \forall x \in \mathcal{P}_{h, i^*_h}$,

$$f(x) \geq f(x^*) - \nu \rho^h.$$

- The smoothness is local, around a $x^*$.

- This guaranties that the algorithm will not under-estimate by more than $\nu \rho^h$ the value of optimal cell $\mathcal{P}_{h, i^*_h}$ if it observes $f(x)$ with $x \in \mathcal{P}_{h, i^*_h}$.

- Now for the opposite question: How much none optimal cells have values $\nu \rho^h$-close to optimal and therefore indiscernible from it? Let us count them!


The smoothness and the near-optimal dimension

**Definition**

\[ \mathcal{N}_h(3\nu \rho^h) \leq \]

where \( \mathcal{N}_h(\varepsilon) \) is the number of cells \( \mathcal{P}_{h,i} \) of depth \( h \) such that

\[ \sup_{x \in \mathcal{P}_{h,i}} f(x) \geq f(x^*) - \varepsilon. \]
The smoothness and the near-optimal dimension

Definition

For any $\nu > 0$, $C > 1$, and $\rho \in (0, 1)$, the near-optimality dimension $d(\nu, C, \rho)$ of $f$ with respect to the partitioning $P$, is $d(\nu, C, \rho) \equiv \inf \{ d' \in \mathbb{R}_+ : \forall h \geq 0, N_h(3\nu \rho^h) \leq C \rho^{-d'} \}$, where $N_h(\varepsilon)$ is the number of cells $P_{h,i}$ of depth $h$ such that $\sup_{x \in P_{h,i}} f(x) \geq f(x^*) - \varepsilon$.
The smoothness and the near-optimal dimension

Let us bound \( N_h(3\nu\rho^h) \) as a function of the depth \( h \).

**Definition**

For any \( \nu > 0 \), \( C > 1 \), and \( \rho \in (0, 1) \), the near-optimality dimension \( d(\nu, C, \rho) \) of \( f \) with respect to the partitioning \( \mathcal{P} \), is

\[
\inf \left\{ d' \in \mathbb{R}^+ : \forall h \geq 0, \ N_h(3\nu\rho^h) \leq C \rho^{-d'h} \right\},
\]

where \( N_h(\varepsilon) \) is the number of cells \( \mathcal{P}_{h,i} \) of depth \( h \) such that \( \sup_{x \in \mathcal{P}_{h,i}} f(x) \geq f(x^*) - \varepsilon \).
The smoothness and the near-optimal dimension

Let us bound $N_h(3\nu \rho^h)$ as a function of the depth $h$.

- $\rho^{-d'h}$ controls how $N_h(3\nu \rho^h)$ explodes with $h$ if $d' > 0$.
- $N_h(3\nu \rho^h)$ is simply bounded, $\forall h$, by a constant $C$ if $d' = 0$.

**Definition**

$$N_h(3\nu \rho^h) \leq C \rho^{-d'h}$$

where $N_h(\varepsilon)$ is the number of cells $P_{h,i}$ of depth $h$ such that $\sup_{x \in P_{h,i}} f(x) \geq f(x^*) - \varepsilon$. 

Let's us bound $N_h(3\nu \rho^h)$ as a function of the depth $h$.

- $\rho^{-d'h}$ controls how $N_h(3\nu \rho^h)$ explodes with $h$ if $d' > 0$.
- $N_h(3\nu \rho^h)$ is simply bounded, $\forall h$, by a constant $C$ if $d' = 0$.

**Definition**

For any $\nu > 0$, $C > 1$, and $\rho \in (0, 1)$, the near-optimality dimension $d(\nu, C, \rho)$ of $f$ with respect to the partitioning $\mathcal{P}$, is

$$d(\nu, C, \rho) \triangleq \inf \left\{ d' \in \mathbb{R}^+ : \forall h \geq 0, \ N_h(3\nu \rho^h) \leq C \rho^{-d'h} \right\},$$

where $N_h(\varepsilon)$ is the number of cells $\mathcal{P}_{h,i}$ of depth $h$ such that $\sup_{x \in \mathcal{P}_{h,i}} f(x) \geq f(x^*) - \varepsilon$. 
Previous work

Previous approaches under similar assumptions with unknown smoothness \((\nu, \rho)\):

<table>
<thead>
<tr>
<th>Method</th>
<th>(b = 0)</th>
<th>Stochastic ((b &gt; 0))</th>
</tr>
</thead>
<tbody>
<tr>
<td>StroquOOL</td>
<td>((\frac{1}{n})\frac{1}{d})</td>
<td>((\frac{1}{n})\frac{1}{d+2})</td>
</tr>
<tr>
<td>SequOOL</td>
<td>((\frac{1}{n})\frac{1}{d})</td>
<td>(\times)</td>
</tr>
<tr>
<td>Uniform(s)</td>
<td>(\frac{1}{n}\frac{\log \frac{1}{\rho}}{\log K})</td>
<td>(\frac{1}{n}\frac{1}{\log K + \log \frac{1}{\rho} + 2})</td>
</tr>
</tbody>
</table>

- We characterise the rates of the uniform strategy under non-stochastic setting.
- We will introduce VROOM.
Previous work

Previous approaches under similar assumptions with unknown smoothness \((\nu, \rho)\):

<table>
<thead>
<tr>
<th></th>
<th>(b = 0)</th>
<th>stochastic ((b &gt; 0))</th>
<th>non-stochastic</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>StroquOOL</strong></td>
<td>((\frac{1}{n})^{\frac{1}{d}})</td>
<td>((\frac{1}{n})^{\frac{1}{d+2}})</td>
<td>(\times)</td>
</tr>
<tr>
<td><strong>SequOOL</strong></td>
<td>((\frac{1}{n})^{\frac{1}{d}})</td>
<td>(\times)</td>
<td>(\times)</td>
</tr>
<tr>
<td><strong>Uniform(s)</strong></td>
<td>(\frac{\log \frac{1}{\rho}}{\log K})</td>
<td>(1/n) (\frac{1}{\log K} + 2)</td>
<td>?</td>
</tr>
<tr>
<td><strong>VROOM</strong></td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

- We characterise the rates of the uniform strategy under non-stochastic setting.
- We will introduce VROOM.
Challenges

• In non-stochastic setting, a learner has to employ internal randomisation, \( P(x_t \in \mathcal{P}_{h,i}) \). Candidates estimators are:
  
  — \( \hat{f}_{h,i}(t) \triangleq \frac{1}{T_{h,i}(t)} \sum_{s=1}^{T_{h,i}(t)} y_s \) is easily biased by an adversary.
  
  — \( \tilde{f}_{h,i}(t) \triangleq \frac{y_t 1_{x_t \in \mathcal{P}_{h,i}}}{P(x_t \in \mathcal{P}_{h,i})} \). unbiased / high variance if \( P(x_t \in \mathcal{P}_{h,i}) \approx 0 \). Ex: a uniform exploration can lead to a \( K^h \).

**Challenge I:** How to control potentially large estimator variances (especially in the stochastic setting)?

• The confidence interval of estimate \( \sum_{t=1}^{n} \tilde{f}_{h,i}(t) \), varies with \( h \) (number of pulls & variance).

  Cross validation techniques as in StroquOOL, are biased against an adversary.

**Challenge II:** How to recommend an optimum \( x(n) \) capable of operating successfully in both feedback settings?
Now: The Algorithms

- Robust Uniform strategies

- VROOM, best of both worlds?
Robust uniform strategies

Parameters: $\mathcal{P} = \{\mathcal{P}_{h,i}\}, b, n, f_{\text{max}}$. Set $\delta = \frac{4b}{f_{\text{max}} \sqrt{n}}$.

For $t = 1, \ldots, n$

Evaluate a point $x_t$ sampled from $U_{\mathcal{P}}(\mathcal{P}_{0,1})$.

Output $x(n) \sim U(\mathcal{P}_{h(n), i(n)})$

where $(h(n), i(n)) = \arg \max \tilde{F}_{h,i}(n) - B_{h,i}^{\text{adv}}(n)$

Figure: The Robuni algorithm

- The algorithm uses a lower confidence bound estimator: $\tilde{F}_{h,i}(n) - B_{h,i}^{\text{adv}}(n)$ where
- $\tilde{F}_{h,i}(n)$ is an unbiased estimates
- $B_{h,i}^{\text{adv}}(n)$ is the width of the confidence interval of that estimate
Robust uniform strategies

**Theorem** (Upper bounds for Robuni)

Any $f_1, \ldots, f_n$ such that $|f_t(x)| \leq f_{\text{max}}$ for all $x \in \mathcal{X}$ and $t \in [n]$. Let $f = \frac{1}{n} \sum_{t=1}^{n} f_t$, with associated $(\nu, \rho)$.

$$
\mathbb{E}[r_n] = O\left(\log(n/\delta) \left(\frac{K}{n \rho^2}\right) \frac{1}{\log K} \frac{1}{\log 1/\rho} + 2\right)
$$
<table>
<thead>
<tr>
<th>Method</th>
<th>$b = 0$</th>
<th>Stochastic ($b &gt; 0$)</th>
<th>Non-stochastic</th>
</tr>
</thead>
<tbody>
<tr>
<td>StroquOOL</td>
<td>$(\frac{1}{n})^{\frac{1}{d}}$</td>
<td>$(\frac{1}{n})^{\frac{1}{d+2}}$</td>
<td>$\times$</td>
</tr>
<tr>
<td>SequOOL</td>
<td>$(\frac{1}{n})^{\frac{1}{d}}$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>Uniform(s)</td>
<td>$\frac{1}{n} \frac{\log 1}{\log K}$</td>
<td>$\frac{1}{n} \frac{1}{\log 1/\rho + 2}$</td>
<td>?</td>
</tr>
<tr>
<td>VRoom</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
<tr>
<td></td>
<td>$b = 0$</td>
<td>stochastic ($b &gt; 0$)</td>
<td>non-stochastic</td>
</tr>
<tr>
<td>------------------</td>
<td>--------------------------</td>
<td>----------------------</td>
<td>----------------</td>
</tr>
<tr>
<td>StroquOOL</td>
<td>$(\frac{1}{n})^{\frac{1}{d}}$</td>
<td>$(\frac{1}{n})^{\frac{1}{d+2}}$</td>
<td>$\times$</td>
</tr>
<tr>
<td>SequOOL</td>
<td>$(\frac{1}{n})^{\frac{1}{d}}$</td>
<td>$\times$</td>
<td>$\times$</td>
</tr>
<tr>
<td>Uniform(s)</td>
<td>$\frac{1}{n} \frac{\log \frac{1}{\rho}}{\log K}$</td>
<td>$\frac{1}{n} \frac{\log K}{\log 1/\rho} + 2$</td>
<td>?</td>
</tr>
<tr>
<td>VR00M</td>
<td>?</td>
<td>?</td>
<td>?</td>
</tr>
</tbody>
</table>

- The rates of Uniform extends to the non-stochastic case!

- Best of both worlds?: Can we obtain the rates in the stochastic setting. and in the non-stochastic setting.
Zipf exploration: Open best $\frac{n}{h}$ cells at depth $h$
• need to pull more each $x$ to limit uncertainty
• **tradeoff:** the more you pull each $x$ the shallower you can explore
Noisy case: Stroquil (Bartlett et al. 2019)

At depth $h$:

- order the cells by decreasing value and
- open the $i$-th best cell with $m = \frac{n}{h_i}$ estimations
Parameters: $\mathcal{P} = \{\mathcal{P}_{h,i}\}, \ b, n, f_{\text{max}}$. Set $\delta = \frac{4b}{f_{\text{max}} \sqrt{n}}$.

For $t = 1, \ldots, n$

For each depth $h \in [\lfloor \log_K(n) \rfloor]$, rank the cells by decreasing order of $\hat{f}_{h,i}^-(t-1)$: Rank cell $\mathcal{P}_{h,i}$ as $\langle i \rangle_{h,t}$.

$x_t \sim \mathcal{U}_\mathcal{P}(\mathcal{P}_{h_t,i_t})$ where

$$p_{h,i,t} \triangleq \mathbb{P}(\mathcal{P}_{h_t,i_t} = \mathcal{P}_{h,i}) \triangleq \frac{1}{h\langle i \rangle_{h,t} \log_K(n)}$$

Output $x(n) \sim \mathcal{U}_\mathcal{P}(\mathcal{P}_{h(n),i(n)})$

where $(h(n), i(n)) \leftarrow \arg\max_{(h,i)} \tilde{F}_{h,i}(n) - B_{h,i}(n)$
**Theorem**  Upper bounds for VR0OM

In the **non-stochastic** setting,:

\[ \mathbb{E}[r_n] = \tilde{O} \left( \frac{1}{n \log \frac{1}{\rho}} + 2 \right) \]

Moreover in the **stochastic** setting, we have,

\[ \mathbb{E}[r_n] = \tilde{O} \left( \frac{1}{n} \right)^\max \left( \frac{1}{d + 3}, \frac{1}{\log K \log \frac{1}{\rho} + 2} \right) \]
Discussion

• Is the rate $\frac{1}{d + 3}$ optimal? Lowerbound?

• Contrary to StroquOOL, VROOM requires the knowledge of $b$. Can we get rid of this assumption.

• Can we obtain results for the deterministic setting ($b=0$)? (without knowledge $b=0$)
Thank you!