Luckiness in Multi-Scale Online Learning

Wouter M. Koolen

2nd AI & Mathematics workshop
University of Twente
Friday 2nd June, 2023
JOIN OUR EVENTS!

February 14-15
BOOT CAMP
A whirlwind overview of machine learning theory aimed at PhD students

February 14
LAUNCH LECTURE
By distinguished international machine learning theoreticians

March 8, 22 & April 5
SEMINAR++ MEETING
Exposition, open problem, brainstorm

April 12
MID-SEMESTER LECTURE
By distinguished international machine learning theoreticians

April 19 & May 10, 24 & June 7, 21
SEMINAR++ MEETING
Exposition, open problem, brainstorm

For the complete programme see:
cwi.nl/semesterprogramme
Team effort

Muriel Felipe Pérez-Ortiz
PhD student at CWI
1. Motivation

2. Theory

3. Application
Motivating Example

Every week I face the choice

- Total travel time \( \leq \) best fixed carrier + small learning overhead
- By choosing my carrier adaptively (possibly randomised)
- With full information of past service
- Without relying on i.i.d. assumption

2h ± 10 min
2h ± 30 min
Motivating Example

Every week I face the choice

2h ± 10 min

2h ± 30 min

I want:

• Total travel time ≤ best fixed carrier + small learning overhead
• By choosing my carrier adaptively (possibly randomised)
• With full information of past service
• Without relying on i.i.d. assumption
Why is this important/interesting

- Fundamental problem with strong connections to
  - martingale deviation inequalities
  - convex optimisation and duality
  - (stochastic) gradient descent
  - **uncertainty quantification**
  - bandit problems (partial information)
  - reinforcement learning
  - **game theory** (saddle point computation)
  - differential privacy
  - Boosting
  - ...
Why is this important/interesting

- Fundamental problem with strong connections to
  - martingale deviation inequalities
  - convex optimisation and duality
  - (stochastic) gradient descent
  - **uncertainty quantification**
  - bandit problems (partial information)
  - reinforcement learning
  - **game theory** (saddle point computation)
  - differential privacy
  - Boosting
  - ...

- Theory well-developed for **single loss scale**. (Freund and Schapire, 1997; De Rooij et al., 2014; Koolen, Grünwald, and Van Erven, 2016)
Why is this important/interesting

- Fundamental problem with strong connections to
  - martingale deviation inequalities
  - convex optimisation and duality
  - (stochastic) gradient descent
  - \textit{uncertainty quantification}
  - bandit problems (partial information)
  - reinforcement learning
  - \textit{game theory} (saddle point computation)
  - differential privacy
  - Boosting
  - ...

- Theory well-developed for \textit{single loss scale}. (Freund and Schapire, 1997; De Rooij et al., 2014; Koolen, Grünwald, and Van Erven, 2016)

- Similar treatment for \textit{multi-scale} was lacking.
  - Existing algorithm templates too rigid
  - No multi-scale Bernstein Inequality
Supervised Learning Theory

What is **Statistical Learning**

- Receive batch of i.i.d. labelled examples
- Output predictor for new data (e.g. by ERM)
- Prove risk bound using **concentration** (PAC, VC dim)
Supervised Learning Theory

What is **Statistical Learning**

- Receive batch of i.i.d. labelled examples
- Output predictor for new data (e.g. by ERM)
- Prove risk bound using **concentration** (PAC, VC dim)

What is **Online Learning**

- Cancel the i.i.d. assumption (⇒ adversary)
- Algorithm predicts/updates sequentially (e.g. by online gradient descent)
- Prove regret bound using **game theory** (minimax)
Supervised Learning Theory

What is **Statistical Learning**

- Receive batch of i.i.d. labelled examples
- Output predictor for new data (e.g. by ERM)
- Prove risk bound using **concentration** (PAC, VC dim)

What is **Online Learning**

- Cancel the i.i.d. assumption (⇒ adversary)
- Algorithm predicts/updates sequentially (e.g. by online gradient descent)
- Prove regret bound using **game theory** (minimax)

We say a **single algorithm** exploits **Luckiness** if

- Regret bounded (by minimax rate)
- Risk bounded by **fast rate** if i.i.d. with margin
Supervised Learning Theory

What is Statistical Learning

• Receive batch of i.i.d. labelled examples
• Output predictor for new data (e.g. by ERM)
• Prove risk bound using **concentration** (PAC, VC dim)

What is Online Learning

• Cancel the i.i.d. assumption (⇒ adversary)
• Algorithm predicts/updates sequentially (e.g. by online gradient descent)
• Prove regret bound using **game theory** (minimax)

We say a **single algorithm** exploits **Luckiness** if

• Regret bounded (by minimax rate)
• Risk bounded by **fast rate** if i.i.d. with margin

A learning problem is **Multi-Scale** if

• range of losses varies wildly between predictions
The learner is **uncertain** about the overall **best predictor**.

Need to **maintain uncertainty**. Vague: many implementations.

Online learning provides a crisp framework with a scalar objective.

Hence it informs us about **optimal/good/appropriate** ways to maintain uncertainty.

The answer is **far from** Bayesian (or perhaps **profound generalisation**)
1. Motivation

2. Theory

3. Application
Fix number $K$ of actions with loss ranges $\sigma \in [0, \infty)^K$.

**Protocol**

for $t = 1, 2, \ldots$

- Learner picks probability distribution $w_t \in \Delta_K$ on actions
- Adversary sets action losses $\ell_t \in \mathbb{R}^K$ with $|\ell^k_t| \leq \sigma_k$
- Learner incurs expected loss $w_t^T \ell_t$
Fix number $K$ of actions with loss ranges $\sigma \in [0, \infty)^K$

**Protocol**

for $t = 1, 2, \ldots$

1. Learner picks probability distribution $w_t \in \Delta_K$ on actions
2. Adversary sets action losses $\ell_t \in \mathbb{R}^K$ with $|\ell^k_t| \leq \sigma_k$
3. Learner incurs expected loss $w_t^T \ell_t$

**Definition (Regret)**

The regret after $T$ rounds with respect to action $k$ is

$$R^k_T = \sum_{t=1}^{T} w_t^T \ell_t - \sum_{t=1}^{T} \ell^k_t$$
Fix number $K$ of actions with loss ranges $\sigma \in [0, \infty)^K$

### Protocol

for $t = 1, 2, \ldots$

- Learner picks probability distribution $w_t \in \Delta_K$ on actions
- Adversary sets action losses $\ell_t \in \mathbb{R}^K$ with $|\ell^k_t| \leq \sigma_k$
- Learner incurs expected loss $w_t^T \ell_t$

### Definition (Regret)

The regret after $T$ rounds with respect to action $k$ is

$$R^k_T = \sum_{t=1}^{T} w_t^T \ell_t - \sum_{t=1}^{T} \ell^k_t$$

### Question

Can Learner keep $R^k_T \leq \sigma_k \sqrt{T}$?
Consider Follow-the-Regularised-Leader (FTRL) template

\[ w_t = \arg\min_{w \in \Delta_K} \langle w, L_{t-1} \rangle + D_\eta(w, u) \]

with cumulative losses \( L_t = \sum_{s=1}^{t} \ell_s \) and multi-scale entropy

\[ D_\eta(w, u) = \sum_k w_k \ln \frac{w_k}{u_k} - w_k + u_k \]

Theorem (Bubeck et al., 2019) FTRL with learning rate \( \eta_k = \frac{1}{\sigma_k q^2 \ln K T} \) has regret bounded by

\[ R_k T \leq \sigma_k \sqrt{T \ln K} \]

Matching worst-case regret lower bound.
Consider Follow-the-Regularised-Leader (FTRL) template

\[ w_t = \arg\min_{w \in \Delta_K} \langle w, L_{t-1} \rangle + D_\eta(w, u) \]

with cumulative losses \( L_t = \sum_{s=1}^{t} \ell_s \) and multi-scale entropy

\[ D_\eta(w, u) = \sum_k \frac{w_k \ln \frac{w_k}{u_k} - w_k + u_k}{\eta_k}. \]

**Theorem (Bubeck et al., 2019)**

FTRL with learning rate \( \eta_k = \frac{1}{\sigma_k} \sqrt{\frac{2 \ln K}{T}} \) has regret bounded by

\[ R_T^{k} \leq \sigma_k \sqrt{T \ln K}. \]

Matching worst-case regret lower bound.
Are we there yet?

**Luckiness?**

What if losses $\ell_1, \ell_2, \ldots$ turn out to be i.i.d. after all?
Are we there yet?

<table>
<thead>
<tr>
<th>Luckiness?</th>
</tr>
</thead>
<tbody>
<tr>
<td>What if losses $\ell_1, \ell_2, \ldots$ turn out to be i.i.d. after all?</td>
</tr>
</tbody>
</table>

Actually, the worst-case lower bound example is of that kind.
Are we there yet?

**Luckiness?**

What if losses $\ell_1, \ell_2, \ldots$ turn out to be i.i.d. after all?

Actually, the worst-case lower bound example is of that kind.

**Luckiness with Margin?**

Losses sampled i.i.d. $\ell_t \sim \mathbb{P}$, where mean loss vector $\mu = \mathbb{E}[\ell]$ exhibits positive gap $\Delta = \mu(2) - \mu(1) > 0$
Are we there yet?

### Luckiness?
What if losses $\ell_1, \ell_2, \ldots$ turn out to be i.i.d. after all?

Actually, the worst-case lower bound example is of that kind.

### Luckiness with Margin?

Losses sampled i.i.d. $\ell_t \sim P$, where mean loss vector $\mu = \mathbb{E}[\ell]$ exhibits positive gap $\Delta = \mu(2) - \mu(1) > 0$

Still no cigar. $\Omega(\sqrt{T})$ regret.
Are we there yet?

**Luckiness?**
What if losses $\ell_1, \ell_2, \ldots$ turn out to be i.i.d. after all?

Actually, the worst-case lower bound example is of that kind.

**Luckiness with Margin?**
Losses sampled i.i.d. $\ell_t \sim P$, where mean loss vector $\mu = \mathbb{E}[\ell]$ exhibits positive gap $\Delta = \mu(2) - \mu(1) > 0$

Still no cigar. $\Omega(\sqrt{T})$ regret.

**Strong Contrast**
For same-scale case there is a single algorithm with

- Worst-case regret $\sqrt{T \ln K}$ (matching lower bound)
- Stochastic+gap regret $O(1/\Delta)$, a constant(!)
- Interpolates spectrum by data-dependent $\sqrt{V_T \ln K}$ bound.
Main Result

Muscada Algorithm

\[ w_t := \arg\min_{w \in \Delta_K} \langle w, L_{t-1} + \mu_{t-1} \rangle + D_{\eta_{t-1}}(w, u) \]

where

\[ \mu_t^k = \sigma_k \sqrt{v_t \ln K} \]

\[ v_t = 4 \sum_{s=1}^{t} \frac{\text{var} \tilde{w}_s (\ell_s)}{\langle \tilde{w}_s, \sigma^2 \rangle} \]

with \( \tilde{w}_t^k \propto w_t^k \eta_{t-1}^k \)

\[ \eta_{t}^k = \frac{1}{\sigma_k \sqrt{2 \ln K / v_t}} \]
Main Result

Muscada Algorithm

\[ w_t := \arg\min_{w \in \Delta_K} \langle w, L_{t-1} + \mu_{t-1} \rangle + D_{\eta_{t-1}} (w, u) \]

where

\[ \mu_t^k = \sigma_k \sqrt{v_t \ln K} \]

\[ v_t = 4 \sum_{s=1}^{t} \frac{\text{var} \tilde{w}_s (\ell_s)}{\langle \tilde{w}_s, \sigma^2 \rangle} \]

with \( \tilde{w}_t^k \propto w_t^k \eta_{t-1}^k \)

\[ \eta_t^k = \frac{1}{\sigma_k \sqrt{\frac{2 \ln K}{v_t}}} \]

Theorem (Main Result)

Muscada guarantees \( R_T^k \leq \mu_T^k \).

- Sharpens worst-case regret bound of Bubeck et al. 2019 as \( v_t \leq t \).
- In i.i.d. setting with gap \( \Delta \), expected regret is constant

\[ \mathbb{E}[\mu_T^{k*}] \leq \frac{K \sigma_{\max}^2}{\Delta} \]
Anatomy of the Muscada weights

\[ w_{t+1}^k = \frac{1}{K} \exp \left( -\frac{L_t^k + \lambda_t^*}{\sigma_k \sqrt{v_t}} \right) \]

where \( \lambda_t^* \) ensures normalisation
Anatomy of the Muscada weights

\[ w_{t+1}^k = \frac{1}{K} \exp \left( -\frac{L^k_t + \lambda^*_t}{\sigma_k \sqrt{\nu_t}} \right) \]

where \( \lambda^*_t \) ensures normalisation

- More loss, less weight
- Evidence in loss decays with “time” \( \nu_t \)
- Loss and normalisation \( \lambda^*_t \) affects large scales \( \sigma_k \) less.
1. Motivation

2. Theory

3. Application
Now suppose I am paranoid about travel time.

<table>
<thead>
<tr>
<th></th>
<th>10</th>
<th>90</th>
</tr>
</thead>
<tbody>
<tr>
<td>0</td>
<td></td>
<td>−10</td>
</tr>
</tbody>
</table>
Now suppose I am paranoid about travel time.

What is the saddle point? (Hint: it is pure)
Now suppose I am paranoid about travel time.

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>90</td>
</tr>
<tr>
<td></td>
<td></td>
</tr>
<tr>
<td>0</td>
<td>-10</td>
</tr>
</tbody>
</table>

What is the saddle point? (Hint: it is pure)

Multi-scale: 🚄 and 🦦 of scale 90 while 🚀 and 🍁 of scale 10.
In General

Problem

Given (large) payoff matrix $M$. Compute an $\epsilon$-equilibrium $(p, q)$:

$$
\max_j p^T Me_j - \min_i e_i^T M q \leq \epsilon.
$$

Popular approach (Freund and Schapire, 1999)

Run online learners $p_t$ and $q_t$ on loss vectors $Mq_t$ and $-M^T p_t$. 

Does multi-scale knowledge help?

Yes, sub-optimality gap improves from $\sigma_{\text{max}}/\sqrt{T}$ to $\sigma_{\text{saddle-point}}/\sqrt{T}$.

With optimism (Rakhlin and Sridharan, 2013), empirically $\sigma_{\text{saddle-point}}/T$. 


### Problem

*Given (large) payoff matrix $M$. Compute an $\epsilon$-equilibrium $(p, q)$:

$$
\max_j p^T M e_j - \min_i e_i^T M q \leq \epsilon.
$$

### Popular approach (Freund and Schapire, 1999)

Run online learners $p_t$ and $q_t$ on loss vectors $Mq_t$ and $-M^T p_t$.

### Question

Does multi-scale knowledge help?
In General

<table>
<thead>
<tr>
<th>Problem</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>Given (large) payoff matrix</em> $M$. Compute an $\epsilon$-equilibrium $(p, q)$:</td>
</tr>
<tr>
<td>$$\max_j p^T M e_j - \min_i e_i^T M q \leq \epsilon.$$</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Popular approach (Freund and Schapire, 1999)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Run online learners $p_t$ and $q_t$ on loss vectors $Mq_t$ and $-M^Tp_t$.</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Question</th>
</tr>
</thead>
<tbody>
<tr>
<td>Does multi-scale knowledge help?</td>
</tr>
</tbody>
</table>

Yes, sub-optimality gap improves from $\sigma_{\text{max}}/\sqrt{T}$ to $\sigma_{\text{saddle-point}}/\sqrt{T}$.

With optimism (Rakhlin and Sridharan, 2013), empirically $\sigma_{\text{saddle-point}}/T$. 
Thanks!


