Gains and Losses are Fundamentally Different in Regret Minimization — The Sparse Case

Joon Kwon
Université Pierre-et-Marie-Curie – Paris 6
Paris, France

Vianney Perchet
INRIA & Université Paris-Diderot – Paris 7
Paris, France

Workshop on Learning
National University of Singapore

November 20th, 2015

Framework

- $d \geqslant 1$ integer.
- Set of actions/decisions for the player: $[d] := \{1, \dots, d\}$.
- At stage $t = 1, \dots, T$,
  - Player chooses action $i_t \in \{1, \dots, d\}$.
  - Nature reveals gain vector $g_t \in [0, 1]^d$.
  - Player gets $g_t^{(i_t)}$.
- player chooses $x_t \in \Delta([d])$, draws $i_t \sim x_t$. Expected gain: $\langle g_t | x_t \rangle$.
- A strategy/algorithm $\sigma = (\sigma_t)_{1 \leqslant t \leqslant T}$

$$x_t = \sigma_t(x_1, i_1, g_1, \dots, x_{t-1}, i_{t-1}, g_{t-1}).$$

The Regret

$$\limsup_{T\to+\infty} \frac{1}{T} \Bigg(\underbrace{\max_{i\in[d]} \sum_{t=1}^{T} g_t^{(i)} - \sum_{t=1}^{T} \langle g_t | x_t \rangle}_{:=R_T} \Bigg) \leqslant 0$$

- **Introduced**: Hannan (1957)
- **Surveys**: Cesa-Bianchi–Lugosi (2006), Rakhlin–Tewari (2008), Shalev-Shwartz (2011), Hazan (2012), Bubeck–Cesa-Bianchi (2012),...

The Minimax Regret

- T: number of stages
- d: number of actions

$$\min_{\sigma} \max_{(g_t)_t} R_T \quad \text{is of order} \quad \sqrt{T \log d}$$

- **Upper bound**: Cesa-Bianchi (1997)
- **Lower bound**: Cesa-Bianchi, Freund, Haussler, Helmbold, Schapire, Warmuth (1997)

Gains and Losses are Equivalent

- Nature chooses loss vectors $\ell_t \in [0,1]^d$

$$R_T = \sum_{t=1}^{T} \ell_t^{(i_t)} - \min_{i \in [d]} \sum_{t=1}^{T} \ell_t^{(i)}$$

- $g_t^{(i)} := 1 - \ell_t^{(i)}$
- $\ell_t \in [0,1]^d \implies g_t \in [0,1]^d$.

$$\max_{i \in [d]} \sum_{t=1}^{T} g_t^{(i)} - \sum_{t=1}^{T} g_t^{(i_t)} = \sum_{t=1}^{T} \ell_t^{(i_t)} - \min_{i \in [d]} \sum_{t=1}^{T} \ell_t^{(i)}$$

A Sparsity Assumption

Let $s \geqslant 1$ be an integer.

Assumption

All gain (resp. loss) vectors are s-sparse, i.e. have at most s nonzero components.

Example

$d = 3$ and $s = 1$.

$$g_1 = \begin{pmatrix} 0 \\ 0 \\ 1 \end{pmatrix} \qquad g_2 = \begin{pmatrix} \frac{1}{2} \\ 0 \\ 0 \end{pmatrix} \qquad g_3 = \begin{pmatrix} 0 \\ \frac{1}{3} \\ 0 \end{pmatrix}$$

$$\ell_1 := \begin{pmatrix} 1 \\ 1 \\ 1 \end{pmatrix} - g_1 = \begin{pmatrix} 1 \\ 1 \\ 0 \end{pmatrix} \quad \rightsquigarrow \quad \text{not 1-sparse}$$

Minimax Regrets

$$\begin{pmatrix} s \text{ actions} \\ (\text{sparsity } s) \end{pmatrix} \underset{\text{easier}}{\leqslant} \begin{pmatrix} d \text{ actions} \\ \text{sparsity } s \end{pmatrix} \underset{\text{easier}}{\leqslant} \begin{pmatrix} d \text{ actions} \\ \text{no sparsity} \end{pmatrix}$$

$$\sqrt{T \log s} \quad \leqslant \quad \text{minimax regret} \quad \leqslant \quad \sqrt{T \log d}.$$

Gains: $\sqrt{T \log s}$

Losses: $\sqrt{Ts \frac{\log d}{d}}$

Algorithms used to achieve minimax regrets

Gains

$$\sqrt{T \log s}$$

Online Mirror Descent with

$$h_p(x) = \begin{cases} \frac{1}{2} \|x\|_p^2 & \text{if } x \in \Delta([d]) \\ +\infty & \text{otherwise} \end{cases}$$

$$p = 1 + \frac{1}{2 \log s - 1}$$

Losses

$$\sqrt{Ts \frac{\log d}{d}}$$

Exponential Weights Algorithm with

$$\eta = \log \left(1 + \sqrt{\frac{2d \log d}{sT}} \right).$$

The Bandit Setting

For stages $t = 1, \ldots, T$,

- Player chooses action $i_t \in [d]$.
- Nature only reveals $g_t^{(i_t)}$.
- Player gets gain $g_t^{(i_t)}$.

Theorem

Minimax Regret is of order $\sqrt{Td}$

- **Upper bound**: Audibert and Bubeck (2009)
- **Lower bound**: Auer, Cesa-Bianchi, Freund and Schapire (2002)

Upper and Lower Bounds

Without sparsity: $\sqrt{Td}$

	Gains	Losses
Upper bound	$\sqrt{Td}$	$\sqrt{Ts \log \frac{d}{s}}$
Lower bound	$\sqrt{Ts}$	$\sqrt{Ts}$

- If the Player knows gain vectors are s-sparse, he can choose to right strategy to achieve $\sqrt{T \log s}$.
- What if is s is unknown ? Can he still take advantage of sparsity?
- The Player knows vectors are 1000-sparse. But if they actually turn out to be 10-sparse, ... ?

YES

Theorem (K. & Perchet (2015))

There exists a strategy which guarantees a $\sqrt{T \log s^*}$ regret bound, where $s^* = \max_{1 \leqslant t \leqslant T} \|g_t\|_0$.

- You don't know the sparsity level of the gain vectors.
- Just play the aforementionned strategy.
- If the gain vectors turn out to be s-sparse, then you will achieve:

$$R_T \lesssim \sqrt{T \log s}.$$

Analog result for losses

Recap

<table>
<tr><td></td><td colspan="2">Full information</td><td colspan="2">Bandit</td></tr>
<tr><td></td><td>Gains</td><td>Losses</td><td>Gains</td><td>Losses</td></tr>
<tr><td>Upper bound</td><td rowspan="2">$\sqrt{T \log s}$</td><td rowspan="2">$\sqrt{Ts\frac{\log d}{d}}$</td><td>$\sqrt{Td}$</td><td>$\sqrt{Ts \log \frac{d}{s}}$</td></tr>
<tr><td>Lower bound</td><td>$\sqrt{Ts}$</td><td>$\sqrt{Ts}$</td></tr>
<tr><td></td><td colspan="2">can be achieved without knowledge of s</td><td>big gap
open problem</td><td>minor gap
$\nearrow$ in d ?</td></tr>
<tr><td></td><td colspan="2"></td><td colspan="2">without knowledge of s... ?</td></tr>
</table>