New Potential-Based Bounds for the Geometric-Stopping Version of Prediction with Expert Advice

Vladimir A. Kobzar[1]
joint work with Robert V. Kohn[2] and Zhilei Wang[2]

[1]NYU Center for Data Science

[2]Courant Institute, NYU

MSML 2020

Outline

- Background on prediction with expert advice
- Our contribution
  - Upper bounds framework
  - Lower bounds framework

Background

Prediction with expert advice

- Focus on the *geometric stopping* (GS) version: the *stopping time* T is sampled from a geometric distribution G with mean $\frac{1}{\delta}$
- At each round, the predictor (*player*) uses guidance from a collection of N *experts* with the goal of minimizing the difference (*regret*) between the player's loss and that of the best performing expert in hindsight
- The expert losses are determined by the *adversary*

Strategies and bounds

- Player strategies (leading to upper bounds) are known and often based on potentials (e.g., exp. weights)
- No previously known lower bounds for GS for general N

Our contribution

Previously, we developed a PDE-based viewpoint on player and adversary potentials used for the *fixed horizon* (FH) version of the expert problem (where T is fixed) [Kobzar et al., 2020]

This paper extends this viewpoint to the GS version

- Specifically, if an adversary for FH does not depend on time (stationary), it can be used for GS
- *Technically*: Given a FH potential, its Laplace transform gives a GS potential
- *Intuition*: This transform is the expectation w/r/t the Exp distribution (limit of G when $\delta \to 0$)
- *Key result:* We obtain the first lower bounds for general N for GS

Definitions

Prediction with expert advice: In each period $t \in [T]$,

- the *player* determines the mix of N experts to follow - distribution $p_t \in \Delta_N$;
- the *adversary* allocates losses to them - distribution a_t over $[-1, 1]^N$; and
- expert losses $q_t \in [-1, 1]^N$ are sampled from a_t, player's choice of expert $I_t \in [N]$ is sampled from p_t, and both samples are revealed to both parties.

Instantaneous regret: $r_\tau = q_{I_\tau,\tau}\mathbb{1} - q_\tau$

Accumulated regret: $x_t = \sum_{\tau<t} r_\tau$

Final regret

FH: $R_T(p, a) = \mathbb{E}_{p,a} \max_i x_{i,T}$
GS: $R(p, a) \ = \mathbb{E}_G R_T(p, a)$

Our viewpoint on potentials uses the value function

- Focus first on player strategies/upper bounds
- Assume player p is Markovian: depends only on the cumulative regret x
- *Value function* v_p: expected final-time regret achieved by p if the game starts with realized regret x (and the adversary behaves optimally)
  - Characterized by

$$v_p(x) = \delta \max_i x_i + (1 - \delta) \max_a \mathbb{E}_{a,p} \; v_p(x + r)$$

GS upper bound potentials/players

- Our *upper bound potential* is a function $\hat{w} : \mathbb{R}^N \to \mathbb{R}$, nondecr. in x_i, which solves

$$\begin{cases} \hat{w}(x) \geq \max_i x_i + \frac{1-\delta}{2\delta} \max_{q \in [-1,1]^N} \langle D^2 \hat{w}(x) \cdot q, q \rangle & \text{(1a)} \\ \hat{w}(x + c\mathbb{1}) = \hat{w}(x) + c & \text{(1b)} \end{cases}$$

- The associated player $p = \nabla \hat{w}$
- Leads to an upper bound on v_p if $\hat{w}(x) - \max_i x_i$ is uniformly bounded below (we'll later sketch of the proof assuming $\hat{w}(x) - \max_i x_i \geq 0$ for simplicity)
- This upper bounds the regret since $v_p(0) = max_a R(a, p)$

Constructing a GS upper bound potential from a FH one

Illustrate by the exponential weights

- The FH potential $w^e(x,t) = \Phi(x) + kt$ where $\Phi(x) = \frac{1}{\eta}\log(\sum_{i=1}^N e^{\eta x_i})$
- Associated with the exponential weights player $p^e = \nabla w^e$
- The standard FH upper bound: $\max_a R_T(a, p^e) \leq \Phi(0) + \frac{1}{2}\eta T$.
- Thus, taking $k = \frac{1}{2}\eta$ ensures that $\max_a R_T(a, p^e) \leq w^e(0, T)$

The Laplace transform gives the GS potential

$$\hat{w}^e(x) = \int_0^\infty e^{-t} w^e(x,t)dt = \Phi(x) + k$$

- $\Phi(x) \geq \max_i x_i$ and $\langle D^2 \hat{w}^e(x) \cdot q, q\rangle = \langle D^2\Phi \cdot q, q\rangle \leq \eta$
- Taking $k = \frac{1-\delta}{2\delta}\eta$ ensures $\hat{w}^e(x) \geq \max_i x_i + \frac{1-\delta}{2\delta}\max_{q\in[-1,1]^N}\langle D^2\hat{w}^e(x)\cdot q, q\rangle$
- Also $\Phi(x + c\mathbb{1}) = \Phi(x) + c$; thus $\hat{w}^e$ satisfies our definition of a GS upper bound potential
- Since Φ is convex, $0 \leq \langle D^2\Phi \cdot q, q\rangle$. Therefore, $\hat{w}^e(x) - \max_i x_i \geq 0$.

Proof of $v_p \leq \hat{w}$

- *Issue*: want to use induction backwards ("verification" argument), but don't know the final time T
- *Solution*: introduce a new problem, which is the same as the original problem except that it ends at t_0 (if it doesn't end earlier in accordance with the GS condition)
- The difference in regret relative to the original problem$\to 0$ as $t_0 \to \infty$.
- Thus, it suffices to bound the value function g of the new problem.
- It is given by a dynamic program

$$g(x, t_0) = \max_i x_i$$

$$g(x, t) = \delta \max_i x_i + (1-\delta) \min_p \mathbb{E}_{a,p}\, g(x + r, t+1) \text{ if } t \leq t_0 - 1$$

Proof of $v_p \leq \hat{w}$: step 1 - controlling increase of $\hat{w}$

- As a reminder $r = q_I \mathbb{1} - q$
- By the linearity along $\mathbb{1}$, Taylor's thm, and the PDE-based definition of $\hat{w}$

$$\begin{aligned} &\mathbb{E}_{p,a}\left[\hat{w}(x+r)\right] - \hat{w}(x) = \mathbb{E}_a\left[p \cdot q + \hat{w}(x-q)\right] - \hat{w}(x) \\ &\leq \frac{1}{2} \max_{q \in [-1,1]^N} \langle D^2 \hat{w} \cdot q, q \rangle \leq \frac{\delta}{1-\delta}(\hat{w}(x) - \max_i x_i) \end{aligned}$$

  where the choice of p eliminated the 1st-order term for all q: $p \cdot q - \nabla \hat{w} \cdot q = 0$
- Rearranging the foregoing,

$$-\hat{w}(x) + \delta \max_i x_i + (1-\delta) \max_a \mathbb{E}_{a,p}\ \hat{w}(x+r) \leq 0$$

Proof of $v_p \leq \hat{w}$: step 2 ("verification" argument)

Show $g \leq \hat{w}$ by induction (and therefore $v_p \leq \hat{w}$)

- Initialization: $g(x, t_0) \leq \hat{w}(x)$ **[since $g(x, t_0) = \max_i x$ and $\hat{w}(x) - \max_i x_i \geq 0$]**
- Inductive hypothesis: $g(x + r, t + 1) \leq \hat{w}(x + r)$

$$\begin{aligned}
\hat{w}(x) &\geq \hat{w}(x) + (-\hat{w}(x) + \delta \max_i x_i + (1 - \delta) \max_a \mathbb{E}_{a,p}\ \hat{w}(x + r)) && \textbf{[by step 1]} \\
&\geq \delta \max_i x_i + (1 - \delta) \max_a \mathbb{E}_{a,p}\ [g(x + r, t + 1)] && \textbf{[by the hypothesis]} \\
&= g(x, t) && \textbf{[by the DP]}
\end{aligned}$$

- Exp. weights: $\hat{w}^e(0) = \frac{\log N}{\eta} + \frac{1-\delta}{2\delta}\eta = \sqrt{\frac{2(1-\delta)\log N}{\delta}}$ for $\eta = \sqrt{\frac{2\delta \log N}{1-\delta}}$
- This improves on the best known regret bound $\sqrt{\frac{2 \log N}{\delta}}$ [Gravin et al., 2017]

Our framework also works for lower bounds

- Adversary a is Markovian & "balanced" ($\mathbb{E}_a q_i = \mathbb{E}_a q_j$)
- It's value function v_a is similar to the player value function v_p
- *Lower bound potential* is a function $\hat{u} : \mathbb{R}^N \to \mathbb{R}$ which solves

$$\begin{cases} \hat{u} \leq \max_i x_i + \dfrac{1-\delta}{2\delta} \mathbb{E}_a \langle D^2 \hat{u}(x) \cdot q, q \rangle & \text{(2a)} \\ \hat{u}(x + c\mathbb{1}) = \hat{u}(x) + c & \text{(2b)} \end{cases}$$

- $\hat{u} \leq v_a$ (modulo discretization error E)
- Regret bound $\hat{u}(0) - E \leq v_a(0) = \min_p R(a, p)$
- In estimating the expected value of $u(x + r) - u(x)$, the dependence on p is in the 1st-order Taylor term, which gets eliminated since a is balanced
- The dependence on a remains at the 2nd order

Heat-based adversary a^h

- a^h is a uniform distribution over the following set S $\left\{q \in \{\pm 1\}^N \mid \sum_{i=1}^N q_i = \pm 1\right\}$ for N odd or $\left\{q \in \{\pm 1\}^N \mid \sum_{i=1}^N q_i = 0\right\}$ for N even
- Potential $\hat{u}$ is the Laplace transform of the sol'n of the linear heat equation

$$\begin{cases} u_t + \kappa \Delta u = 0 \\ u(x,0) = \max_i x \end{cases} \qquad u(x,t) = \alpha \int e^{-\frac{\|y\|^2}{2\sigma^2}} \max_k (x_k - y_k) dy$$

  where $\alpha = (2\pi\sigma^2)^{-\frac{N}{2}}$ and $\sigma^2 = -2\kappa t$.
- Satisfies our def'n of a lower bound potential for a well-chosen κ
- The leading order asymptotics of our lower bound $\hat{u}(0) = \Omega\left(\sqrt{\frac{\log N}{\delta}}\right)$ matches that of the exponential weights upper bound
- Optimal leading order term for $N = 2$
- Also give a nonasymptotic guarantee $\hat{u}(0) - E \leq v_{a^h}(0)$
- The discretization error E is computed explicitly and is $O\left(N\sqrt{N} \wedge \sqrt{N}\left(1 + \log \frac{1}{\delta}\right)\right)$

Conclusion

- We provide easily-checked conditions for a function to be useful as a lower-bound or an upper bound potential
- Using the Laplace transform, we construct potentials for the GS problem from potentials used for the FH version
- Lower bound potentials correspond to strategies for adversary
- We obtain the first known lower bound in the GS setting for general N associated with a simple randomized strategy
- Our framework also leads in some cases to improved upper bounds

Questions? vladimir.kobzar@nyu.edu

Acknowledgements

- V.A.K and R.V.K. are supported, in part, by NSF grant DMS-1311833.
- V.A.K. is also supported by the Moore-Sloan Data Science Environment at NYU.

References

Nick Gravin, Yuval Peres, and Balasubramanian Sivan. Tight Lower Bounds for Multiplicative Weights Algorithmic Families. In *44th International Colloquium on Automata, Languages, and Programming (ICALP 2017)*, volume 80, pages 48:1–48:14, Warshaw, Poland, 2017.

Vladimir A. Kobzar, Robert V. Kohn, and Zhilei Wang. New Potential-Based Bounds for Prediction with Expert Advice. In *Proceedings of the 33rd Annual Conference on Learning Theory (COLT)*, forthcoming, 2020.