New Potential-Based Bounds for 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

COLT 2020

Outline

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

Background on prediction with expert advice

- Fixed horizon version
- At each round, the predictor (*player*) uses guidance from a collection of *experts* with the goal of minimizing the difference (*regret*) between the player's loss and that of the best performing expert in hindsight
- Adversarial model (no distributional assumptions)
- Player strategies often use potentials (e.g., exp. weights)
- Adversary strategies are randomized (e.g., i.i.d. coin flips)

Our contribution

- Rakhlin et al. [2012]: Potentials arise as relaxations of the minimax value
- Rokhlin [2017]: A related PDE-based viewpoint on player potentials
- Our conceptual advances
  1. This viewpoint extends to adversaries, leading to lower bounds
  2. More technically: Finding better regret bounds $\equiv$ finding "better" solutions of certain PDEs
  3. Practical impact: Understanding how potentials work gives guidance on choosing good potentials
- Applying these advances, we obtain not only a fresh perspective, but also improved bounds

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 (vector-valued) regret: $r_\tau = q_{I_\tau,\tau}\mathbb{1} - q_\tau$
Accumulated (vector-valued) regret: $x_t = \sum_{\tau<t} r_\tau$
Final-time regret: $R_T(p,a) = \mathbb{E}_{p,a}\left[\max_i x_{i,T}\right]$

Potential-based player strategies are familiar

- Exponential weights potential: $\Phi(x) = \frac{1}{\eta} \log(\sum_i e^{\eta x_i})$
- Bounds the regret: $\max_i x_i \leq \Phi(x)$ for all x.
- The player p^e controls the increase in Φ as the game proceeds
- $p^e = \nabla \Phi$ eliminates the 1st-order Taylor term in the expansion of Φ, as the regret evolves, for all adversary's choices of q.

Our viewpoint on potentials uses the value function

- Convention: $T < 0$ is the starting and 0 is the final time
- Assume p is Markovian: at time t can depend only on t and x
- Value function v_p: expected final-time regret achieved by p if the game starts at time t with realized regret x (assuming the adversary behaves optimally)
- Characterized by a dynamic program

$$v_p(x, 0) = \max_i x_i$$
$$v_p(x, t) = \max_a \mathbb{E}_{a,p_t} \; v_p(x + r, t + 1) \text{ for } t \leq -1$$

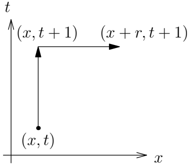
Upper bound potential and associated player

- General def'n: an *upper bound potential* is a function $w : \mathbb{R}^N \times \mathbb{R}_{\leq 0} \to \mathbb{R}$, nondecr. in x_i, which solves

$$
\begin{cases}
w_t + \frac{1}{2} \max_{q \in [-1,1]^N} \langle D^2 w \cdot q, q \rangle \leq 0 & \text{(1a)} \\
w(x,0) \geq \max_i x_i & \text{(1b)} \\
w(x + c\mathbb{1}, t) = w(x,t) + c & \text{(1c)}
\end{cases}
$$

- The associated player $p = \nabla w$
- Leads to an upper bound $v_p \leq w$ (will sketch the proof next)
- Bounds regret above since $v_p(0,T) = max_a R_T(a,p)$.
- *Exponential weights*: $w^e(x,t) = \Phi(x) - \frac{1}{2}\eta t$ where $\Phi(x) = \frac{1}{\eta} \log(\sum_i e^{\eta x_i})$.

Proof of $v_p \le w$: step 1 - controlling increase of w



- Due to $\Delta x = r$: By the linearity along $\mathbb{1}$ and Taylor's thm,

$$\mathbb{E}_{p_t,a}\left[w(x+r,t+1)\right] - w(x,t+1)$$
$$\le \frac{1}{2}\max_{q\in[-1,1]^N}\langle D^2 w\cdot q, q\rangle \qquad [\le\ \eta/2 \text{ for } w^e]$$

  where p eliminated 1st-order term for all q: $p_t\cdot q - \nabla w\cdot q = 0$

- Due to Δt is:

$$w_t\ [= -\eta/2 \text{ for } w^e]$$

- By the PDE, $\max_a\ \mathbb{E}_{p_t,a}\left[w(x+r,t+1)\right] - w(x,t) \le 0$

Proof of $v_p \leq w$: step 2

Show $v_p \leq w$ by induction

- Initialization: $v_p(x,0) \leq w(x,0)$ by the final value of w
- Inductive hypothesis: $v_p(x+r,t+1) \leq w(x+r,t+1)$

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

- Exp. weights: $w^e(0,T) = \frac{1}{\eta}\log N + \frac{1}{2}\eta|T| = \sqrt{2|T|\log N}$ for $\eta = \sqrt{\frac{2\log N}{|T|}}$ (best known regret bound for $[-1,1]^N$ losses)

Our framework also works for lower bounds

- Adversary a is Markovian & "balanced" ($\mathbb{E}_{a_t} q_i = \mathbb{E}_{a_t} q_j$)
- Value function v_a for the adversary a has a similar DP characterization
- *Lower bound potential* is also a very similar object—$u : \mathbb{R}^N \times \mathbb{R}_{\leq 0} \to \mathbb{R}$ which solves

$$\begin{cases} u_t + \frac{1}{2}\mathbb{E}_a \langle D^2 u \cdot q, q \rangle \geq 0 & \text{(2a)} \\ u(x, 0) \geq \max_i x_i & \text{(2b)} \\ u(x + c\mathbb{1}, t) = u(x, t) + c & \text{(2c)} \end{cases}$$

- Since a is balanced, 1st-order Taylor term is eliminated:

$$\mathbb{E}_{p,a_t}[q_I - \nabla u \cdot q] = \langle p - \nabla u, \mathbb{E}_{a_t} q \rangle = 0$$

- We used $\nabla u \cdot \mathbb{1} = 1$ by (2c) and $p \cdot \mathbb{1} = 1$ since $p \in \Delta_N$
- $u \leq v_a$ (modulo error E from higher order terms)
- Regret bound $u(0, T) - E(T) \leq v_a(0, T) = \min_p R_T(a, p)$

Classic randomized adversary a^r

- a^r: each $q_i \sim$ i.i.d. Rademacher r.v.
- Potential is 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 $\sigma^2 = -2\kappa t$.

- $\mathbb{E}_{a^r} \langle D^2 u \cdot q, q \rangle = \mathbb{E}_{a^r} [\sum_{i,j} \partial_{ij} u \ q_i q_j] = \Delta u$
- Satisfies the def'n of potential for $\kappa = \frac{1}{2}$.
- Same leading order regret as in the classic CLT-based proof:

$$u(0,T) = \sqrt{-2\kappa T} \mathbb{E}_G \max G_i = \sqrt{|T|} \mathbb{E}_G \max G_i$$

- But we give a new nonasymptotic guarantee $u(0,T) - E(T) \leq v_{a^r}(0,T)$ where $E(T) = O(N\sqrt{N} \wedge \sqrt{N \log N} + \sqrt{N} \log |T|)$

Comb adversary a^c

- Ranks the experts by their losses and advances even-ranked experts with prob. $1/2$ and odd-ranked experts with prob. $1/2$
  - Gravin et al. [2016] conjectured a^c is optimal for general N
  - Abbasi-Yadkori et al. [2017] and Bayraktar et al. [2020] confirmed that for $N = 3$ and 4, respectively.
- We do not resolve this conjecture for general N
- But we show that for the same heat potential with $\kappa = 1/2$, $\mathbb{E}_{a^c}\langle D^2 u \cdot q, q\rangle \geq \Delta u$
- Thus, the a^c is at least as powerful as a^r, in particular a^c is doubly asymptotically optimal in T and N
- Previously, this was only known for a^r

Improved heat-based adversary a^h

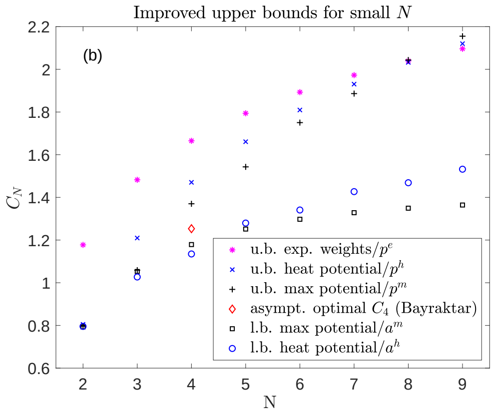
- a^h is a uniform distribution over the following set S
  $\left\{q \in \{\pm 1\}^N \mid \sum_{i=1}^N q_i = \pm 1\right\}$ for N odd or
  $\left\{q \in \{\pm 1\}^N \mid \sum_{i=1}^N q_i = 0\right\}$ for N even
- Best known leading order (in T) constant: improved from $\frac{1}{2}$ to

$$\kappa_h = \begin{cases} 1 & \text{if } N = 2 \\ \frac{1}{2} + \frac{1}{2N} & \text{if } N \text{ is odd} \\ \frac{1}{2} + \frac{1}{2N-2} & \text{otherwise.} \end{cases}$$

- Asymptotically optimal for $N = 2$

Another class of potentials: max potential

- Explicit sol'n of $\begin{cases} u_t + \kappa \max_i \partial_i^2 u = 0 \\ u(x,0) = \max_i x \end{cases}$
- Asymptotically optimal for $N = 2$ and 3
- For $N = 3$, improve the best known nonasymptotic bound
- Obtain lower and upper bound potentials for different κ's
- For small N and large T, the max-based player outperforms p^e



Extension

- The discussion here focused on a fixed horizon problem
- In a separate paper Kobzar et al. [2020], we extended this framework to the *geometric stopping* problem where the final time T is sampled from a geometric distribution

Conclusion

- Offer a fresh PDE-based perspective on how to find potentials and why they work
- It applies to lower as well as upper bounds
- Solutions of certain PDEs are good potentials

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 I

Yasin Abbasi-Yadkori, Peter L. Bartlett, and Victor Gabillon. Near Minimax Optimal Players for the Finite-Time 3-Expert Prediction Problem. In *Advances in Neural Information Processing Systems 30*, pages 3033–3042, 2017.

Erhan Bayraktar, Ibrahim Ekren, and Xin Zhang. Finite-time 4-expert prediction problem. *Communications in Partial Differential Equations*, 2020. doi: 10.1080/03605302.2020.1712418.

Nick Gravin, Yuval Peres, and Balasubramanian Sivan. Towards optimal algorithms for prediction with expert advice. In *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, pages 528–547, Arlington, Virginia, 2016.

References II

Vladimir A. Kobzar, Robert V. Kohn, and Zhilei Wang. New Potential-Based Bounds for the Geometric-Stopping Version of Prediction with Expert Advice. In *Proceedings of the 1st Annual Conference on Mathematical and Scientific Machine Learning*, forthcoming, 2020.

Alexander Rakhlin, Ohad Shamir, and Karthik Sridharan. Relax and Randomize: From Value to Algorithms. In *Advances in Neural Information Processing Systems 25*, pages 2141–2149. 2012.

Dmitry B. Rokhlin. PDE approach to the problem of online prediction with expert advice: a construction of potential-based strategies. *International Journal of Pure and Applied Mathematics*, 114(4):907–915, 2017.