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

Vladimir A. Kobzar\textsuperscript{1}, Robert V. Kohn\textsuperscript{2} and Zhilei Wang\textsuperscript{2}

\textsuperscript{1}NYU Center for Data Science, \textsuperscript{2}Courant Institute of Mathematical Sciences

\textbf{Prediction with expert advice}

In each \( t \in [T] \),
- 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 \( q_t \sim a_t \), player’s choice of expert \( I_t \sim p_t \); these samples revealed to both parties.

\textbf{Our contribution}

Previously we developed a PDE viewpoint for the fixed horizon (FH) version of the problem where the \textit{stopping time} \( T \) is fixed (COLT 2020)

This paper (MSML 2020) extends this viewpoint to the geometric stopping (GS) version where the \textit{stopping time} \( T \sim G \) and \( G = \text{Geom} (\mu) \)

- Specifically, if an FH adversary does not depend on time (stationary), it can be used for GS
- \textbf{Technically}: Given a FH potential, its Laplace transform gives a GS potential
- \textbf{Intuition}: This transform is the expectation \( w/\mu \) the Exp distribution (limit of \( G \) when \( \delta \to 0 \))
- \textbf{Key result}: Obtain the first lower bounds for general \( N \) associated with a simple randomized strategy

\textbf{Definitions}

- \textbf{Instantaneous regret}: \( r_t = q_t r_t - q_t \)
- \textbf{Accumulated regret}: \( x_t = \sum_{\tau \leq t} r_t \)
- \textbf{Final regret}: \( \text{FH} - \text{FH} = e^{\frac{\pi}{2}} \min_{x_t, x_{t+1}} T \)

GS - \( R(p, a) = e^{\frac{\pi}{2}} \min_{x_t, x_{t+1}} T \)

\textbf{Lower bound potentials/adversaries}

- \textbf{Adversary} a Markovian & “balanced”: \( v_q = \sum_{q'} v_q (q') \)
- Use the value function \( v_q \) for this adversary
- \textbf{Lower bound potential} is a function \( \hat{u} : \mathbb{R} \to \mathbb{R} \) which solves

\[ \hat{u}(x) \geq \max_{x_t + \frac{1}{2}, \max_{p \in S_{t+1}} (D^2 \hat{u}(x) \cdot q, q)} \]

\[ \hat{u}(x + c) = \hat{u}(x) + \epsilon \]

- The associated player \( p = \nabla \hat{u} \)
- Leads to an upper bound on \( v_q \) if \( \hat{u}(x) - \max x_t \)

is uniformly bounded below
- \textbf{Regret} upper bound since \( v_q (0) = \max_a R(a, p) \)

- \textbf{The associated player} \( p = \nabla \hat{u} \)
- \textbf{Regret} upper bound since \( v_q (0) = \max_a R(a, p) \)

\textbf{Results}

\textbf{Upper bound potentials/players}

A function \( \hat{u} : \mathbb{R}^N \to \mathbb{R} \), nondecr. in \( x_t \), which solves

\[ \hat{u}(x) \geq \max_{x_t + \frac{1}{2}, \max_{p \in S_{t+1}} (D^2 \hat{u}(x) \cdot q, q)} \]

\[ \hat{u}(x + c) = \hat{u}(x) + \epsilon \]

- The associated player \( p = \nabla \hat{u} \)
- Leads to an upper bound on \( v_q \) if \( \hat{u}(x) - \max x_t \)

is uniformly bounded below
- \textbf{Regret} upper bound since \( v_q (0) = \max_a R(a, p) \)

\textbf{Proof of} \( v_p \leq \hat{u} \): Idea

- \textbf{Issue}: want to use induction backwards (“verification” argument), but don’t know \( T \)
- \textbf{Sol’n}: introduce a new problem, which is the same except that it ends at \( t_0 \) (if it doesn’t end earlier in accordance with the GS condition)
- \textbf{The difference} in regret relative to the original problem \( \to 0 \text{ as } t_0 \to \infty \)
- \textbf{Suffices} to bound the value \( g \) of the new problem.
- \textbf{It is given by} a dynamic program:

\[ g(x, t) = \max_{x_t \text{ and } d} g(x_t, d, t = t_0 - 1, \]

\[ g(p, a) = \sum_{p \in S_{t+1}} \max_{x_t \text{ and } d} g(x_t, d, t = t_0 - 1) \]

\textbf{Laplace tr.: FH\textrightarrow GS potential}

Illustrated by the exponential weights example: \( \hat{w}(x) = \frac{1}{2} \log (\sum_{x_t} e^{0\mu}) \)

- \( \Phi(x) \geq \max x_t \text{ and } (D^2 \Phi \cdot q, q) \leq \eta \)
- Also \( \Phi(x + c) = \Phi(x) + c \)
- Thus, taking \( k = \frac{1}{2} \frac{1}{2} \) ensures \( \hat{w}(x) \)

satisfies our def’n of a GS upper bound potential
- \textbf{Since} \( \Phi \) is convex, \( 0 \leq (D^2 \Phi \cdot q, q) \).
- Thus \( \hat{w}(x) - \max x_t \geq 0 \).

- \textbf{Control increase of} \( \hat{w} \) as the game evolves: the choice \( p = \nabla \hat{u} \)

eliminates the 1st-order Taylor term in this evolution for all \( q \)
- \textbf{Show} \( g \leq \hat{w} \) by induction (and thus \( v_q \leq \hat{u} \))

\textbf{Heat-based adversary}

- \( \Phi \) is a uniform distribution over the following set

\[ \text{FH} \{ q \in \{ \pm 1 \}^N \mid \text{Sum}_q q = \pm 1 \} \]

- For \( N \) odd
- \[ \text{FH} \{ q \in \{ \pm 1 \}^N \mid \text{Sum}_q q = 0 \} \]

- For \( N \) even
- \textbf{Potential} \( \hat{u} \) is the Laplace transform of the sol’n of the linear heat equation

\[ \dot{u} + \alpha u(x, t) = \alpha \int e^{\frac{\pi}{2}} \max_{x_k \text{ and } d} g(x_t, d, t = t_0 - 1) \]

\[ u(x, t) = \max_{x_t \text{ and } d} g(x_t, d, t = t_0 - 1) \]

\textbf{Proof of} \( v_p \leq \hat{u} \): “verification” arg.

- Satisfies our def’n of a lower bound potential for a well-chosen \( \kappa \)
- \textbf{The leading order asymptotics} of our lower bound \( \hat{u}(0) = \Omega (\sqrt{N} \max_{x_k \text{ and } d} g(x_t, d, t = t_0 - 1) \)

\textbf{Potential} \( \hat{u} \text{ of the exponential weights upper bound}
- \textbf{Optimal leading order term} for \( N = 2 \)
- Also give a nonasymptotic guarantee

\[ \hat{u}(0) - E \leq v_p(0) \]

\textbf{The discretization error} \( E \) is computed explicitly and is \( O(\sqrt{N} \wedge \sqrt{N}) (1 + \log \frac{1}{N}) \)

\textbf{Acknowledgements}

NSF grant DMS-1311833; Moore-Sloan Data Science Environment at NYU