A PDE APPROACH TO MIXED STRATEGIES PREDICTION WITH EXPERT ADVICE

EXTENDED ABSTRACT

NADEJDA DRENSKA
COURANT INSTITUTE, NYU

1. Overview

This work investigates a discrete model problem from online machine learning using methods from PDEs and optimal control. The overall area is 'prediction with expert advice' - a framework in which an agent tries to use 'expert advice' to invest optimally (for the worst case scenario) against an adversarial market. A discrete time iterative process involves decision making at every step; the goal for mathematical analysis is to understand the optimal decision and its consequences over a long period of time. Our general approach is 'numerical analysis in reverse' - interpreting the discrete formulation as a numerical scheme for an appropriate PDE. We prove that the solution to the discrete problem is asymptotically close to the unique solution of the PDE and thus use knowledge of the PDE solution to inform the optimal strategy of the original setup.

This work focuses on the following classical problem in online machine learning. There are two players - an 'investor' and a 'market' - and a fixed number n of 'experts'. The market chooses which experts win or lose at every time step. The investor chooses which expert to listen to at each time step. The measure of effectiveness of the investor's strategy is regret minimization: performance under the metric of 'regret', or distance between a investor's performance and that of an 'expert'. Thus, the investor's goal is to accumulate overall winnings as close as possible to those of the best performing expert at the 'end' of the game (assuming that the market works against him). There are two variants: one with a fixed stopping time ('the finite horizon problem') and one where the stopping time is random with a constant probability of stopping δ at every time step ('the geometric stopping problem'). The goal is to identify the optimal strategies of the investor and the market, as well as the associated value function. A recent paper [2] by Gravin, Peres, and Sivan analyzes this problem in considerable detail. In particular it studies the character of the optimal strategies and it obtains an exact solution for the the geometric stopping problem with 3 experts and payoff function $\varphi(x) = \max_k\{x_k\}$. That work uses discrete methods and connections to random walks.

Our work provides a continuous time - continuous space analog of [2]. One perspective is that it shows that PDE and continuous methods can be used to provide an understanding parallel to the one with discrete methods and random walks. A particular advantage is that my treatment is not limited to the classical payoff function in online machine learning literature, namely regret with respect to the best expert $\varphi(x) = \max_k\{x_k\}$. In fact, it works for a more general class of payoff functions, namely functions φ that are uniformly continuous, symmetric in their dependent variables x_k, and with $\varphi(x_1+c, ..., x_n+c) = \varphi(\vec{x})+c$ to reflect a more general linear regret. For this general class of functions, the same analysis as in the specific case applies.

To put this work in context, there are other instances in the literature where scaling limits of multistep decision processes lead to parabolic or elliptic PDEs. Some examples include the work of Kohn and Serfaty on two-person game interpretations of motion by curvature [3] and many other PDE problems [4]. Peres, Sheffield, Schramm, and Wilson show connections between the 'tug-of-war' game and the infinity-Laplacian [5], and the p-Laplacian [6]. These works have seen many extensions, see e.g. [7], [8], [9], [10]. Some features the current work shares with these articles are:

- an optimal control perspective, based on a suitable dynamic programming principle; and
- use of viscosity solutions - a natural tool for highly nonlinear PDEs that emerge from optimal control problems.

This work - just like the papers just cited - focuses on a particular class of two-person games.

2. The Mixed Strategies Problem

Rules of the game. The mixed strategies problem is a two-player zero-sum game, with mixed strategies. Players are called 'market' and 'player'. At each step the market makes a choice (according to a probability distribution $\tilde{p}$ on the space of possible outcomes $\{0,1\}^n$) which experts make a gain - indicated by 1 - and which experts make a loss - indicated by 0. The outcome is recorded in a vector $v \in \{0,1\}^n$. Meanwhile, the player chooses one of the n experts, say k, according to a probability vector $\vec{\alpha} = (\alpha_1, \alpha_2, ..., \alpha_n)$, where $\sum \alpha_i = 1$ and $\alpha_i \geq 0$, and records the same gain or loss as expert k, i.e. v_k. The state variables x_j for this game are investor's regret with respect to the j expert, so the state variables change as $\Delta x_j = v_j - v_k$ if the investor chooses to follow expert k.

The finite horizon problem is to determine the expected regret (value function) of the investor and the associated optimal strategies for the players, provided that the game ends at an a priori fixed time T and starts at time t, $t \leq T$ with initial regret vector x. One can write the value function through a dynamic programming principle (DPP). Since one wants to capture the behavior over many time steps, we consider an appropriate scaling limit: a scaling of ε for space steps (so each gain is 0 or ε, instead of initial gains of 0 or 1) and time steps of size ε^2. Through dynamic programming principle, the discrete finite horizon formulation becomes:

$$w_\varepsilon(t,x) = \min_{player} \max_{market} \mathbf{E}[w(t+\varepsilon^2, x+\varepsilon\Delta x)].$$

$$w_\varepsilon(T,x) = \varphi(x).$$

The scaled DPP above turns out to be a semi-discrete numerical scheme for an associated parabolic PDE:

$$w_t(t,x) + \frac{1}{2}\max_{v\in\{0,1\}^n}(v\cdot\nabla)^2 w = 0.$$

Theorem 1. *The final-time nonlinear degenerate parabolic PDE above has a unique at most linear growth viscosity solution w, by* [11]*. Moreover, the discrete value functions of the game w_ε converge uniformly in ε to the viscosity solution w:*

$$\lim_{\varepsilon\to 0}\sup_x |w_\varepsilon(t,x) - w(t,x)| = 0.$$

If w is twice differentiable with bounded third derivative, then the convergence above is linear in ε.

The geometric-stopping problem is to determine the expected regret (value function) of the investor provided the game starts at regret vector x. The game either stops with a constant probability $\delta, 0<\delta<1$, and is evaluated with a payoff of $\varphi(x)$, or it continues for at least one more round, with investor and market playing against each other optimally. One can write the value function through a dynamic programming principle (DPP). In order to consider an appropriate scaling limit, we introduce a scaling of ε for space steps (so each gain is 0 or ε, instead of initial gains of 0 or 1). Then, the discrete geometric-stopping formulation becomes:

$$u_\varepsilon(x) = \delta\varphi(x) + (1-\delta)\min_{player}\max_{market}\mathbf{E}[u_\varepsilon(x+\varepsilon\Delta x)].$$

I scale δ as $\delta = \frac{\varepsilon^2}{2}$ in order to obtain an associated elliptic PDE:

$$u(x) = \varphi(x) + \max_{v\in\{0,1\}^n}(v\cdot\nabla)^2 u. \tag{1}$$

Theorem 2. *The nonlinear degenerate elliptic PDE above has a unique at most linear growth viscosity solution u by* [11]*. Moreover, the discrete value functions of the game u_ε converge uniformly in ε to the viscosity solution u:*

$$\lim_{\varepsilon\to 0}\sup_x |u_\varepsilon(x) - u(x)| = 0.$$

If in addition u is twice differentiable with bounded third derivatives, then the convergence above is linear in ε.

Theorem 3. *In the case of $n=3$ experts the unique viscosity solution of the PDE (3) with $\varphi(x) = \max_k\{x_k\}$) has an explicit formula. In the region where $x_1 \geq x_2 \geq x_3$ it is*

$$u(x) = x_1 + \frac{1}{2}e^{x_2-x_1} + \frac{1}{6}e^{2x_3-x_2-x_1}$$

and the values at other x are determined by symmetry. Since u is twice differentiable with bounded third derivatives, the convergence of u_ε to u is linear in ε.

Methods

- *Define a discrete approximation scheme*: $\mathcal{F}_\varepsilon(u_\varepsilon, x)$ is associated with the DPP of the discrete game:

$$\mathcal{F}_\varepsilon(u_\varepsilon, x) := u_\varepsilon(x) - \varphi(x) - \frac{1-2\varepsilon^2}{2\varepsilon^2} \min_{player} \max_{market} \mathbf{E}[u_\varepsilon(x+\varepsilon\Delta x) - u_\varepsilon(x)].$$

  The condition $\mathcal{F}_\varepsilon(u_\varepsilon, x) = 0$ is just a simple rearrangement of the DPP.
- *Existence of a solution to the scheme* is nontrivial. Constructing it relies on two ingredients: a time dependent problem which is run to equilibrium and a contraction mapping argument.
- *Convergence of the scheme* is obtained through standard viscosity technology: the scheme above is stable, monotone, and consistent, hence it converges to the PDE (by a variant of [12] to accommodate at most linear growth solutions).
- *Optimal strategies* are obtained in the process of obtaining the PDE. The player's choice is $\alpha_i = \partial_i u$ (or $\alpha_i = \partial_i w$). The market finds a v so maximize $\max_{v\in\{0,1\}^n}(v\cdot\nabla)^2 u$ and chooses this v with probability 0.5 and its complement $(1-v_1, 1-v_2, ..., 1-v_n)$ with probability 0.5.
- *Probability matching.* In the geometric stopping version of the game with 3 experts, our explicit solution (Theorem 5) permits direct verification of the following fact: the investor should follow expert k with probability equal to the probability that this expert winds up performing best (assuming the market plays optimally). A similar property in the discrete setting was established in [2]. The generalization of this statement to more experts and the finite horizon setting remains open.

Implications for Machine Learning. The machine learning literature offers explicit optimal algorithms for relatively few problems of the type considered here; instead it often uses the performance of standard algorithms (such as the weighted majority algorithm, or a follow the perturbed leader algorithm) as benchmarks to provide upper or lower bounds. The strategy that the PDE approach provides is different: more problem specific, but optimal (at least in the limit $\varepsilon \to 0$). Moreover, the PDE approach handles a more general class of performance measures $\varphi(x)$, not just the standard comparison to the best performing expert.

References

[1] N. Cesa-Bianchi and G. Lugosi, *Prediction, Learning, and Games.* New York, NY, USA: Cambridge University Press, 2006.

[2] N. Gravin, Y. Peres, and B. Sivan, "Towards optimal algorithms for prediction with expert advice," in *Proceedings of the Twenty-Seventh Annual ACM-SIAM Symposium on Discrete Algorithms*, SODA '16, (Philadelphia, PA, USA), pp. 528–547, Society for Industrial and Applied Mathematics, 2016.

[3] R. Kohn and S. Serfaty, "A deterministic-control-based approach motion by curvature," *Communications on Pure and Applied Mathematics*, vol. 59, no. 3, pp. 344–407, 2006.

[4] R. V. Kohn and S. Serfaty, "A deterministic-control-based approach to fully nonlinear parabolic and elliptic equations," *Communications on Pure and Applied Mathematics*, vol. 63, no. 10, pp. 1298–1350, 2010.

[5] Y. Peres, O. Schramm, S. Sheffield, and D. B. Wilson, "Tug-of-war and the infinity Laplacian," *J. Amer. Math. Soc.*, vol. 22, no. 1, pp. 167–210, 2009.

[6] Y. Peres and S. Sheffield, "Tug-of-war with noise: A game-theoretic view of the p -laplacian," *Duke Math. J.*, vol. 145, pp. 91–120, 10 2008.

[7] S. N. Armstrong and C. K. Smart, "A finite difference approach to the infinity Laplace equation and tug-of-war games," *Trans. Amer. Math. Soc.*, vol. 364, no. 2, pp. 595–636, 2012.

[8] A. Naor and S. Sheffield, "Absolutely minimal lipschitz extension of tree-valued mappings," *Mathematische Annalen*, vol. 354, no. 3, pp. 1049–1078, 2012.

[9] T. Antunovic, Y. Peres, S. Sheffield, and S. Somersille, "Tug-of-war and infinity laplace equation with vanishing neumann boundary condition," *Communications in Partial Differential Equations*, vol. 37, no. 10, pp. 1839–1869, 2012.

[10] M. Lewicka and J. J. Manfredi, "The obstacle problem for the p-laplacian via optimal stopping of tug-of-war games," *Probability Theory and Related Fields*, pp. 1–30, 2015.

[11] M. G. Crandall, H. Ishii, and P.-L. Lions, "User's guide to viscosity solutions of second order partial differential equations," *Bull. Amer. Math. Soc. (N.S.)*, vol. 27, no. 1, pp. 1–67, 1992.

[12] G. Barles and P. E. Souganidis, "Convergence of approximation schemes for fully nonlinear second order equations," *Asymptotic analysis*, vol. 4, no. 3, pp. 271–283, 1991.