Prediction without probability: a PDE approach to a two-player game from machine learning

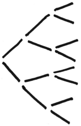
Robert V. Kohn
Courant Institute, NYU

Joint work with Nadejda Drenska

Indiana University
October, 2017

Prediction without probability

A thread from machine learning: prediction with expert advice.



VERSION 1: (Not today's focus, but still a natural starting point)

- a time series – eg a binomial stock price tree;
- some notion of gain/loss due to good/bad predictions (eg buy or stell stock);
- *N* experts (eg public or private algorithms based on recent history);
- investor's goal: do as well as the (retrospectively) best-performing expert – or at least, don't fall too far behind;
- focus on worst-case scenario (malevolent market), so probabilities are irrelevant.

Not today's focus – that would be a different talk (eg thesis of Kangping Zhu).

Today's focus

VERSION 2: Investor has no mind of his own – he just integrates the advice of many experts. So let's ignore any underlying time series.

- N experts
- investor's action: at each time step, "choose an expert to follow"

  to allow mixtures: investor chooses a prob distrn on $\{1, \ldots, N\}$ (follow expert j with prob p_j)
- market's action: at each time step, "choose which experts receive gains" (eg for 3 experts, vector of gains can be $(1, 0, 0)$ or $(1, 1, 0)$ or $\ldots$)

  to allow mixtures: market chooses a prob distrn on $\{0, 1\}^N$

One interpretation: experts $\Leftrightarrow$ market sectors,
investor's probabilities $\Leftrightarrow$ portfolio allocations.

Prediction as a 2-player game

Recall: investor chooses a prob distrn (follow expert j with prob p_j); mkt chooses a prob distrn on the 2^N expert gain scenarios.

This is a 2-player, zero-sum game. The state variables are

$$x_j = j\text{th expert's gain} - \text{investor's gain} = \text{regret wrt } j\text{th expert.}$$

The investor's value function is:

$$u(x,t) = \text{expected final time regret, under worst-case scenario.}$$

The dynamic programming principle says (if game ends at time T):

$$\begin{aligned} u(x,t) &= \min_{\substack{\text{investor's} \\ \text{choices}}} \max_{\substack{\text{market's} \\ \text{choices}}} \mathbb{E}[u(x+\Delta x, t+1)] \quad \text{for } t < T \\ u(x,T) &= \phi(x) = \max\{x_1, \cdots, x_N\} \end{aligned}$$

Note 1: Other choices of ϕ are possible. Mainly, we'll use that ϕ is increasing in each x_i with linear growth at ∞, and $\phi(x_1+c,\ldots,x_N+c) = \phi(x)+c$.

Note 2: If stopping is random (Poisson) rather than deterministic then value function depends on space alone, and dyn prog prin becomes

$$w(x) = \delta\phi(x) + (1-\delta) \min_{\substack{\text{investor's} \\ \text{choices}}} \max_{\substack{\text{market's} \\ \text{choices}}} \mathbb{E}[w(x+\Delta x)]$$

where δ = stopping probability.

Prediction as a 2-player game

Recall: investor chooses a prob distrn (follow expert j with prob p_j); mkt chooses a prob distrn on the 2^N expert gain scenarios.

This is a 2-player, zero-sum game. The state variables are

$$x_j = j\text{th expert's gain} - \text{investor's gain} = \text{regret wrt } j\text{th expert.}$$

The investor's value function is:

$$u(x,t) = \text{expected final time regret, under worst-case scenario.}$$

The dynamic programming principle says (if game ends at time T):

$$u(x,t) = \min_{\substack{\text{investor's}\\ \text{choices}}} \max_{\substack{\text{market's}\\ \text{choices}}} \mathbb{E}[u(x+\Delta x, t+1)] \quad \text{for } t < T$$

$$u(x,T) = \phi(x) = \max\{x_1, \cdots, x_N\}$$

Note 1: Other choices of ϕ are possible. Mainly, we'll use that ϕ is increasing in each x_i with linear growth at ∞, and $\phi(x_1 + c, \ldots, x_N + c) = \phi(x) + c$.

Note 2: If stopping is random (Poisson) rather than deterministic then value function depends on space alone, and dyn prog prin becomes

$$w(x) = \delta\phi(x) + (1-\delta) \min_{\substack{\text{investor's}\\ \text{choices}}} \max_{\substack{\text{market's}\\ \text{choices}}} \mathbb{E}[w(x+\Delta x)]$$

where δ = stopping probability.

Prediction as a 2-player game

Recall: investor chooses a prob distrn (follow expert j with prob p_j); mkt chooses a prob distrn on the 2^N expert gain scenarios.

This is a 2-player, zero-sum game. The state variables are

$$x_j = j\text{th expert's gain} - \text{investor's gain} = \text{regret wrt } j\text{th expert.}$$

The investor's value function is:

$$u(x,t) = \text{expected final time regret, under worst-case scenario.}$$

The dynamic programming principle says (if game ends at time T):

$$\begin{aligned} u(x,t) &= \min_{\substack{\text{investor's}\\ \text{choices}}} \max_{\substack{\text{market's}\\ \text{choices}}} \mathbb{E}[u(x+\Delta x, t+1)] \quad \text{for } t < T \\ u(x,T) &= \phi(x) = \max\{x_1, \cdots, x_N\} \end{aligned}$$

Note 1: Other choices of ϕ are possible. Mainly, we'll use that ϕ is increasing in each x_i with linear growth at ∞, and $\phi(x_1+c,\ldots,x_N+c) = \phi(x)+c$.

Note 2: If stopping is random (Poisson) rather than deterministic then value function depends on space alone, and dyn prog prin becomes

$$w(x) = \delta\phi(x) + (1-\delta) \min_{\substack{\text{investor's}\\ \text{choices}}} \max_{\substack{\text{market's}\\ \text{choices}}} \mathbb{E}[w(x+\Delta x)]$$

where δ = stopping probability.

Prediction as a PDE problem

We are interested in long-time behavior. In the ML lit, a typical question is: estimate $u(0,t)$ when $T - t$ is large, and give an easily-implemented strategy that does almost as well.

Continuum limits were designed for this. For example: the behavior of a random walk after many time steps is captured by considering the assoc diffusion process. Same idea is useful here; so we introduce a small parameter ε:

- gains are ε or 0 (rather than 1 or 0); time step is ε^2
- scaled version is equiv to unscaled version, if $\phi(\lambda x) = \lambda\phi(x)$
- for random stopping variant, the stopping prob δ should be $\sim \varepsilon^2$

Claim:

- There is a meaningful PDE limit.
- In finding it, we learn about both players' optimal strategies.
- In some cases (eg time-dependent version with 2 experts, and random stopping version with 3 experts) we know the PDE soln explicitly. (So we know the optimal strategies explicitly.)

Relation to the ML literature

- This game is a well-studied model problem. But our PDE viewpoint is new.
- ML lit gives upper and lower bounds, by considering particular strategies – eg, for unscaled problem, $u(0,t) \sim C_N\sqrt{T-t}$. PDE gives optimal prefactor.
- Our attn was drawn by a recent paper *Towards optimal algorithms for prediction with expert advice* (N Gravin, Y Peres, B Sivan, Proc SODA '16). Their treatment is discrete, and this talk is roughly its PDE analogue.
- For more ML perspective on prediction with expert advice, see *Prediction, Learning, & Games*, N Cesa-Bianchi and G Lugosi, Cambridge Univ Press, 2006.

Relation to the ML literature

- This game is a well-studied model problem. But our PDE viewpoint is new.
- ML lit gives upper and lower bounds, by considering particular strategies – eg, for unscaled problem, $u(0,t) \sim C_N \sqrt{T-t}$. PDE gives optimal prefactor.
- Our attn was drawn by a recent paper *Towards optimal algorithms for prediction with expert advice* (N Gravin, Y Peres, B Sivan, Proc SODA '16). Their treatment is discrete, and this talk is roughly its PDE analogue.
- For more ML perspective on prediction with expert advice, see *Prediction, Learning, & Games*, N Cesa-Bianchi and G Lugosi, Cambridge Univ Press, 2006.

Mathematical context

Key features:

- a multiperiod decision-making process;
- two players (the investor and the market);
- decisions via worst-case analysis (hence the min-max);
- both players see the same value function (a zero-sum game).

About 10 years ago, two problems sharing these features were considered at length, involving

(1) a two-person game interpretation of motion by curvature (Kohn-Serfaty, CPAM 2006)

(2) a two-person game interpretation of the infinity Laplacian (Peres-Schramm-Sheffield-Wilson, JAMS 2009).

Techniques used: mainly from optimal control (dynamic programming, Hamilton-Jacobi-Bellman equation, viscosity solutions).

PDE's are 2nd order (although we're not doing stochastic control). You'll see why in a moment . . .

Mathematical context

Key features:

- a multiperiod decision-making process;
- two players (the investor and the market);
- decisions via worst-case analysis (hence the min-max);
- both players see the same value function (a zero-sum game).

About 10 years ago, two problems sharing these features were considered at length, involving

(1) a two-person game interpretation of motion by curvature (Kohn-Serfaty, CPAM 2006)

(2) a two-person game interpretation of the infinity Laplacian (Peres-Schramm-Sheffield-Wilson, JAMS 2009).

Techniques used: mainly from optimal control (dynamic programming, Hamilton-Jacobi-Bellman equation, viscosity solutions).

PDE's are 2nd order (although we're not doing stochastic control). You'll see why in a moment . . .

Mathematical context

Key features:

- a multiperiod decision-making process;
- two players (the investor and the market);
- decisions via worst-case analysis (hence the min-max);
- both players see the same value function (a zero-sum game).

About 10 years ago, two problems sharing these features were considered at length, involving

(1) a two-person game interpretation of motion by curvature (Kohn-Serfaty, CPAM 2006)

(2) a two-person game interpretation of the infinity Laplacian (Peres-Schramm-Sheffield-Wilson, JAMS 2009).

Techniques used: mainly from optimal control (dynamic programming, Hamilton-Jacobi-Bellman equation, viscosity solutions).

PDE's are 2nd order (although we're not doing stochastic control).

You'll see why in a moment . . .

Finding the PDE

Returning to our problem, let's find the associated PDE. Overall strategy is familiar:

SIMPLE VERSION: Scaled DPP defines value function u_ε. We expect $u_\varepsilon \to u$. Find the PDE by replacing u_ε by u in DPP and using Taylor expansion.

FANCIER VERSION: Scaled DPP is a semi-discrete numerical scheme for the desired PDE. The simple version finds the PDE for which it is a *consistent* numerical scheme.

Some notation for the players' choices at a given time step:

investor's choice : follow expert k with prob p_k

market's choice : prob distr of experts' gains $\varepsilon(g_1, \ldots, g_N)$

where $g = (g_1, \ldots, g_N)$ is a random variable taking values in $\{0,1\}^N$.

If the investor follows expert k, then the regret increment is

$$\Delta x = \varepsilon(g_1 - g_k, \ldots, g_N - g_k) = \varepsilon(g - g_k\vec{1})$$

Scaled dyn prog prin:

$$u_\varepsilon(x,t) = \min_{\substack{p_k \geq 0 \\ \sum p_k = 1}} \max_{\substack{\text{prob distr on} \\ g \in \{0,1\}^N}} \sum_{k=1}^{N} p_k \, \mathbb{E}_g[u_\varepsilon(x + \varepsilon(g - g_k\vec{1}), t + \varepsilon^2)]$$

Finding the PDE

Returning to our problem, let's find the associated PDE. Overall strategy is familiar:

SIMPLE VERSION: Scaled DPP defines value function u_ε. We expect $u_\varepsilon \to u$. Find the PDE by replacing u_ε by u in DPP and using Taylor expansion.

FANCIER VERSION: Scaled DPP is a semi-discrete numerical scheme for the desired PDE. The simple version finds the PDE for which it is a *consistent* numerical scheme.

Some notation for the players' choices at a given time step:

investor's choice :	follow expert k with prob p_k
market's choice :	prob distr of experts' gains $\varepsilon(g_1, \ldots, g_N)$

where $g = (g_1, \ldots, g_N)$ is a random variable taking values in $\{0, 1\}^N$.

If the investor follows expert k, then the regret increment is

$$\Delta x = \varepsilon(g_1 - g_k, \ldots, g_N - g_k) = \varepsilon(g - g_k \vec{1})$$

Scaled dyn prog prin:

$$u_\varepsilon(x, t) = \min_{\substack{p_k \geq 0 \\ \sum p_k = 1}} \max_{\substack{\text{prob distr on} \\ g \in \{0,1\}^N}} \sum_{k=1}^{N} p_k \, \mathbb{E}_g[u_\varepsilon(x + \varepsilon(g - g_k \vec{1}), t + \varepsilon^2)]$$

Finding the PDE

Returning to our problem, let's find the associated PDE. Overall strategy is familiar:

SIMPLE VERSION: Scaled DPP defines value function u_ε. We expect $u_\varepsilon \to u$. Find the PDE by replacing u_ε by u in DPP and using Taylor expansion.

FANCIER VERSION: Scaled DPP is a semi-discrete numerical scheme for the desired PDE. The simple version finds the PDE for which it is a *consistent* numerical scheme.

Some notation for the players' choices at a given time step:

investor's choice :	follow expert k with prob p_k
market's choice :	prob distr of experts' gains $\varepsilon(g_1, \ldots, g_N)$

where $g = (g_1, \ldots, g_N)$ is a random variable taking values in $\{0, 1\}^N$.

If the investor follows expert k, then the regret increment is

$$\Delta x = \varepsilon(g_1 - g_k, \ldots, g_N - g_k) = \varepsilon(g - g_k \vec{1})$$

Scaled dyn prog prin:

$$u_\varepsilon(x, t) = \min_{\substack{p_k \geq 0 \\ \sum p_k = 1}} \; \max_{\substack{\text{prob distr on} \\ g \in \{0,1\}^N}} \sum_{k=1}^{N} p_k \, \mathbb{E}_g[u_\varepsilon(x + \varepsilon(g - g_k \vec{1}), t + \varepsilon^2)]$$

Finding the PDE, cont'd

Substitute u_ε by u (soln of anticipated PDE) in DPP:

$$u(x,t) \approx \min_{\substack{p_k \geq 0 \\ \sum p_k = 1}} \ \max_{\substack{\text{prob distr on} \\ g \in \{0,1\}^N}} \sum_{k=1}^{N} p_k \, \mathbb{E}_g[u(x + \varepsilon(g - g_k \vec{1}), t + \varepsilon^2)].$$

RHS = $u(x,t) + \varepsilon$[terms involving $\partial_k u$] $+ \varepsilon^2$[terms involving $\partial^2_{ij} u$ and u_t] $+ \ldots$

Zeroth order term $u(x,t)$ cancels LHS.

First order term seems to dominate. But min-max of first-order term alone is a linear programming problem. Its value is 0, achieved (only) when

investor's choice is $p_k = \partial_k u / (\partial_1 u + \cdots \partial_N u)$;
market's choices are balanced: $\mathbb{E}[g_1] = \ldots = \mathbb{E}[g_N]$.

Consistency check: we expect $\partial_k u \geq 0$, since $u(x,T) = \phi(x)$ is monotone increasing in each x_k.

The investor's strategies are fully determined but the market's strategies are not, so we must continue . . .

Finding the PDE, cont'd

Second order term gives

$$u_t + \max_{\mathbb{E}[g_j] \text{ indep of j}} \frac{1}{2} \sum_{k=1}^{N} p_k \mathbb{E}\left[\langle D^2 u, (g - g_k \vec{1}) \otimes (g - g_k \vec{1}) \rangle \right] = 0$$

in which $p_k = \partial_k u / (\partial_1 u + \cdots \partial_N u)$.

This can be greatly simplified, using that

(a) each g_j takes only the values 0 or 1;

(b) $u(x + c\vec{1}, t) = u(x, t) + c$ (proved by induction, since the final-time function has this property);

(c) we are maximizing a linear function over a convex set.

$N = 2$ is misleadingly simple: (b) implies $(\partial_1 + \partial_2)u = 1$ and $\partial_{11} u = \partial_{22} u$; PDE simplifies to $u_t + \frac{1}{2} p_1 \partial_{22} u + \frac{1}{2} p_2 \partial_{11} u = 0$, or equivalently

$$u_t + \frac{1}{4} \Delta u = 0 \text{ for } t < T, \text{ with } u = \max\{x_1, x_2\} \text{ at } t = T.$$

Market advances each expert with prob $\frac{1}{2}$. (A discrete version of this was understood by T. Cover in the 1960's.)

Finding the PDE, cont'd

Second order term gives

$$u_t + \max_{\mathbb{E}[g_j] \text{ indep of j}} \frac{1}{2} \sum_{k=1}^{N} p_k \mathbb{E}\left[\langle D^2 u, (g - g_k \vec{1}) \otimes (g - g_k \vec{1})\rangle\right] = 0$$

in which $p_k = \partial_k u / (\partial_1 u + \cdots \partial_N u)$.

This can be greatly simplified, using that

(a) each g_j takes only the values 0 or 1;

(b) $u(x + c\vec{1}, t) = u(x, t) + c$ (proved by induction, since the final-time function has this property);

(c) we are maximizing a linear function over a convex set.

$N = 2$ is misleadingly simple: (b) implies $(\partial_1 + \partial_2)u = 1$ and $\partial_{11} u = \partial_{22} u$; PDE simplifies to $u_t + \frac{1}{2} p_1 \partial_{22} u + \frac{1}{2} p_2 \partial_{11} u = 0$, or equivalently

$$u_t + \frac{1}{4} \Delta u = 0 \text{ for } t < T, \text{ with } u = \max\{x_1, x_2\} \text{ at } t = T.$$

Market advances each expert with prob $\frac{1}{2}$. (A discrete version of this was understood by T. Cover in the 1960's.)

Finding the PDE, cont'd

The PDE is nonlinear for $N \geq 3$.

When $N = 3$: we have $(\partial_1 + \partial_2 + \partial_3)u = 1$, which implies $\partial_{11}u = (\partial_2 + \partial_3)^2 u$, etc; PDE reduces to

$$u_t + \tfrac{1}{2}\max\{\partial_{11}u, \partial_{22}u, \partial_{33}u\} = 0.$$

- If max is achieved at $\partial_{11}u$ then market's strategy is: "advance expert 1 with prob $\frac{1}{2}$, advance all but 1 with prob $\frac{1}{2}$".

General N: PDE is

$$u_t + \tfrac{1}{2}\max_k \max_{i_1,\ldots,i_k}\{(\partial_{i_1} + \cdots + \partial_{i_k})^2 u\} = 0.$$

- If max is at $(\partial_{i_1} + \cdots + \partial_{i_k})^2 u$, then market's strategy is: "advance experts $i_1, \ldots, i_k$ with prob $\frac{1}{2}$, advance all the others $\frac{1}{2}$".

Finding the PDE, cont'd

The PDE is nonlinear for $N \geq 3$.

When $N = 3$: we have $(\partial_1 + \partial_2 + \partial_3)u = 1$, which implies $\partial_{11}u = (\partial_2 + \partial_3)^2 u$, etc; PDE reduces to

$$u_t + \frac{1}{2}\max\{\partial_{11}u, \partial_{22}u, \partial_{33}u\} = 0.$$

- If max is achieved at $\partial_{11}u$ then market's strategy is: "advance expert 1 with prob $\frac{1}{2}$, advance all but 1 with prob $\frac{1}{2}$".

General N: PDE is

$$u_t + \frac{1}{2}\max_k \max_{i_1,\ldots,i_k}\{(\partial_{i_1} + \cdots + \partial_{i_k})^2 u\} = 0.$$

- If max is at $(\partial_{i_1} + \cdots + \partial_{i_k})^2 u$, then market's strategy is: "advance experts $i_1, \ldots, i_k$ with prob $\frac{1}{2}$, advance all the others $\frac{1}{2}$".

How the PDE emerges

Focus on $N = 3$ as illustrative example: $u_t + Lu = 0$ with

$$Lu = \max_{\mathbb{E}[g_j] \text{ indep of j}} \frac{1}{2} \sum_{k=1^N} p_k \mathbb{E}\left[\langle D^2 u, (g - g_k \vec{1}) \otimes (g - g_k \vec{1})\rangle\right]$$

in which $p_k = \partial_k u$.

STEP 1: Let

$$a_0 = \text{Prob}\{(0,0,0) \text{ or } (1,1,1)\}, \quad a_1 = \text{Prob}\{(1,0,0) \text{ or } (0,1,1)\}$$
$$a_2 = \text{Prob}\{(0,1,0) \text{ or } (1,0,1)\}, \quad a_3 = \text{Prob}\{(0,0,1) \text{ or } (1,1,0)\}$$

and ignore the constraint of balance ($E[g_j]$ indep of j). Then RHS becomes

$$\frac{1}{2} \max_{\substack{a_j \geq 0 \\ \sum a_j = 1}} \left\{ \begin{array}{l} a_1[(1-p_1)\partial_{11} u + p_1(\partial_2 + \partial_3)^2 u] \\ +a_2[(1-p_2)\partial_{22} u + p_2(\partial_1 + \partial_3)^2 u] \\ +a_3[(1-p_3)\partial_{33} u + p_3(\partial_1 + \partial_2)^2 u]. \end{array} \right\}$$

STEP 2: If coefft of a_1 is largest, then optimal choice is $a_1 = 1$. Consistent with balance, by taking $\text{Prob}\{(1,0,0)\} = \frac{1}{2}$ and $\text{Prob}\{(0,1,1)\} = \frac{1}{2}$. Thus: if coefft of a_1 is largest,

$$Lu = \tfrac{1}{2}[(1-p_1)\partial_{11} u + p_1(\partial_2 + \partial_3)^2 u].$$

How the PDE emerges

Focus on $N = 3$ as illustrative example: $u_t + Lu = 0$ with

$$Lu = \max_{\mathbb{E}[g_j] \text{ indep of j}} \frac{1}{2} \sum_{k=1^N} p_k \mathbb{E}\left[\langle D^2 u, (g - g_k \vec{1}) \otimes (g - g_k \vec{1})\rangle\right]$$

in which $p_k = \partial_k u$.

STEP 1: Let

$$a_0 = \text{Prob}\{(0,0,0) \text{ or } (1,1,1)\}, \quad a_1 = \text{Prob}\{(1,0,0) \text{ or } (0,1,1)\}$$
$$a_2 = \text{Prob}\{(0,1,0) \text{ or } (1,0,1)\}, \quad a_3 = \text{Prob}\{(0,0,1) \text{ or } (1,1,0)\}$$

and ignore the constraint of balance ($E[g_j]$ indep of j). Then RHS becomes

$$\frac{1}{2} \max_{\substack{a_j \geq 0 \\ \sum a_j = 1}} \left\{ \begin{array}{l} a_1[(1-p_1)\partial_{11} u + p_1(\partial_2 + \partial_3)^2 u] \\ +a_2[(1-p_2)\partial_{22} u + p_2(\partial_1 + \partial_3)^2 u] \\ +a_3[(1-p_3)\partial_{33} u + p_3(\partial_1 + \partial_2)^2 u]. \end{array} \right\}$$

STEP 2: If coefft of a_1 is largest, then optimal choice is $a_1 = 1$. Consistent with balance, by taking $\text{Prob}\{(1,0,0)\} = \frac{1}{2}$ and $\text{Prob}\{(0,1,1)\} = \frac{1}{2}$. Thus: if coefft of a_1 is largest,

$$Lu = \tfrac{1}{2}[(1-p_1)\partial_{11} u + p_1(\partial_2 + \partial_3)^2 u].$$

How the PDE emerges, cont'd

Thus far: if coefft of a_1 is largest, $Lu = \frac{1}{2}[(1-p_1)\partial_{11}u + p_1(\partial_2+\partial_3)^2 u]$

STEP 3: From special structure of ϕ (and induction in time) we have $u(x + c\vec{1}, t) = u(x,t) + c$, so $(\partial_1 + \partial_2 + \partial_3)u = 1$. Thus

$$\partial_{11}u + \partial_1(\partial_2+\partial_3)u = 0 \quad \text{and} (\partial_2+\partial_3)\partial_1 u + (\partial_2+\partial_3)^2 u = 0,$$

whence $(\partial_2+\partial_3)^2 u = \partial_{11}u$ and

$$\text{coefft of } a_1 = (1-p_1)\partial_{11}u + p_1\partial_{11}u = \partial_{11}u.$$

CONCLUSION: Arguing similarly for coeffts of a_2 and a_3, we get

$$Lu = \tfrac{1}{2}\max\{\partial_{11}u, \partial_{22}u, \partial_{33}u, 0\}.$$

But 0 should never be optimal: worst-case regret should increase with time.

$N = 4$ is similar, but analogue of step 1 involves

$a_0 = \text{Prob}\{(0,0,0,0) \text{ or } (1,1,1,1)\}$, $a_1 = \text{Prob}\{(1,0,0,0) \text{ or } (0,1,1,1)\}$
$a_2 = \text{Prob}\{(0,1,0,0) \text{ or } (1,0,1,1)\}$, $a_3 = \text{Prob}\{(0,0,1,0) \text{ or } (1,1,0,1)\}$
$a_4 = \text{Prob}\{(0,0,0,1) \text{ or } (1,1,1,0)\}$ $b_{12} = \text{Prob}\{(1,1,0,0) \text{ or } (0,0,1,1)\}$
$b_{13} = \text{Prob}\{(1,0,1,0) \text{ or } (0,1,0,1)\}$ $b_{14} = \text{Prob}\{(1,0,0,1) \text{ or } (0,1,1,0)\}$

How the PDE emerges, cont'd

Thus far: if coefft of a_1 is largest, $Lu = \frac{1}{2}[(1-p_1)\partial_{11}u + p_1(\partial_2+\partial_3)^2 u]$

STEP 3: From special structure of ϕ (and induction in time) we have $u(x+c\vec{1},t) = u(x,t)+c$, so $(\partial_1+\partial_2+\partial_3)u = 1$. Thus

$$\partial_{11}u + \partial_1(\partial_2+\partial_3)u = 0 \quad \text{and} (\partial_2+\partial_3)\partial_1 u + (\partial_2+\partial_3)^2 u = 0,$$

whence $(\partial_2+\partial_3)^2 u = \partial_{11}u$ and

$$\text{coefft of } a_1 = (1-p_1)\partial_{11}u + p_1\partial_{11}u = \partial_{11}u.$$

CONCLUSION: Arguing similarly for coeffts of a_2 and a_3, we get

$$Lu = \frac{1}{2}\max\{\partial_{11}u, \partial_{22}u, \partial_{33}u, 0\}.$$

But 0 should never be optimal: worst-case regret should increase with time.

$N=4$ is similar, but analogue of step 1 involves

$$\begin{aligned} a_0 &= \text{Prob}\{(0,0,0,0) \text{ or } (1,1,1,1)\}, & a_1 &= \text{Prob}\{(1,0,0,0) \text{ or } (0,1,1,1)\} \\ a_2 &= \text{Prob}\{(0,1,0,0) \text{ or } (1,0,1,1)\}, & a_3 &= \text{Prob}\{(0,0,1,0) \text{ or } (1,1,0,1)\} \\ a_4 &= \text{Prob}\{(0,0,0,1) \text{ or } (1,1,1,0)\} & b_{12} &= \text{Prob}\{(1,1,0,0) \text{ or } (0,0,1,1)\} \\ b_{13} &= \text{Prob}\{(1,0,1,0) \text{ or } (0,1,0,1)\} & b_{14} &= \text{Prob}\{(1,0,0,1) \text{ or } (0,1,1,0)\} \end{aligned}$$

How the PDE emerges, cont'd

Thus far: if coefft of a_1 is largest, $Lu = \frac{1}{2}[(1-p_1)\partial_{11}u + p_1(\partial_2+\partial_3)^2 u]$

STEP 3: From special structure of ϕ (and induction in time) we have $u(x + c\vec{1}, t) = u(x,t) + c$, so $(\partial_1 + \partial_2 + \partial_3)u = 1$. Thus

$$\partial_{11}u + \partial_1(\partial_2+\partial_3)u = 0 \quad \text{and} (\partial_2+\partial_3)\partial_1 u + (\partial_2+\partial_3)^2 u = 0,$$

whence $(\partial_2+\partial_3)^2 u = \partial_{11}u$ and

$$\text{coefft of } a_1 = (1-p_1)\partial_{11}u + p_1\partial_{11}u = \partial_{11}u.$$

CONCLUSION: Arguing similarly for coeffts of a_2 and a_3, we get

$$Lu = \tfrac{1}{2}\max\{\partial_{11}u, \partial_{22}u, \partial_{33}u, 0\}.$$

But 0 should never be optimal: worst-case regret should increase with time.

$N = 4$ is similar, but analogue of step 1 involves

$$\begin{array}{ll}
a_0 = \text{Prob}\{(0,0,0,0) \text{ or } (1,1,1,1)\}, & a_1 = \text{Prob}\{(1,0,0,0) \text{ or } (0,1,1,1)\} \\
a_2 = \text{Prob}\{(0,1,0,0) \text{ or } (1,0,1,1)\}, & a_3 = \text{Prob}\{(0,0,1,0) \text{ or } (1,1,0,1)\} \\
a_4 = \text{Prob}\{(0,0,0,1) \text{ or } (1,1,1,0)\} & b_{12} = \text{Prob}\{(1,1,0,0) \text{ or } (0,0,1,1)\} \\
b_{13} = \text{Prob}\{(1,0,1,0) \text{ or } (0,1,0,1)\} & b_{14} = \text{Prob}\{(1,0,0,1) \text{ or } (0,1,1,0)\}
\end{array}$$

The random-stopping version

Recall that if stopping is random (Poisson, rate δ), the value function satisfies:

$$w_\varepsilon(x) = \delta\phi(x) + (1-\delta) \min_{\substack{\text{investor's}\\ \text{choices}}} \max_{\substack{\text{market's}\\ \text{choices}}} \mathbb{E}[w_\varepsilon(x+\Delta x)].$$

Stopping rate should be of order ε^2, so

- game lasts $O(\varepsilon^{-2})$ steps $\Rightarrow$ typical regret at stopping is $O(1)$;
- δ doesn't affect the order ε min-max calculation;
- δ interacts with $O(\varepsilon^2)$ Taylor expansion terms.

PDE is elliptic, with source term $\phi(x) = \max_k\{x_k\}$, and the same 2nd order operator as before.

If $N = 3$ and $\delta = \frac{1}{\lambda}\varepsilon^2$, then PDE is $w - \frac{1}{2}\lambda \max_k\{\partial_{kk} w\} = \phi$.

Surpisingly, the solution is explicit: when $x_1 > x_2 > x_3$,

$$w(x) = x_1 + \frac{1}{c}\left(\frac{1}{2}e^{c(x_2-x_1)} + \frac{1}{6}e^{c(2x_3-x_2-x_1)}\right) \quad \text{with } c = \sqrt{2/\lambda}.$$

Another surprise: market has (at least) two optimal strategies. In fact: from the formula, when $x_1 > x_2 > x_3$ we have $\partial_{11} w = \partial_{22} w$.

Analogous soln for discrete pbm is in Gravin, Peres, and Sivan, SODA '16.

The random-stopping version

Recall that if stopping is random (Poisson, rate δ), the value function satisfies:

$$w_\varepsilon(x) = \delta\phi(x) + (1-\delta) \min_{\substack{\text{investor's}\\\text{choices}}} \max_{\substack{\text{market's}\\\text{choices}}} \mathbb{E}[w_\varepsilon(x+\Delta x)].$$

Stopping rate should be of order ε^2, so

- game lasts $O(\varepsilon^{-2})$ steps $\Rightarrow$ typical regret at stopping is $O(1)$;
- δ doesn't affect the order ε min-max calculation;
- δ interacts with $O(\varepsilon^2)$ Taylor expansion terms.

PDE is elliptic, with source term $\phi(x) = \max_k\{x_k\}$, and the same 2nd order operator as before.

If $N = 3$ and $\delta = \frac{1}{\lambda}\varepsilon^2$, then PDE is $w - \frac{1}{2}\lambda \max_k\{\partial_{kk} w\} = \phi$.

Surpisingly, the solution is explicit: when $x_1 > x_2 > x_3$,

$$w(x) = x_1 + \frac{1}{c}\left(\frac{1}{2}e^{c(x_2-x_1)} + \frac{1}{6}e^{c(2x_3-x_2-x_1)}\right) \quad \text{with } c = \sqrt{2/\lambda}.$$

Another surprise: market has (at least) two optimal strategies. In fact: from the formula, when $x_1 > x_2 > x_3$ we have $\partial_{11} w = \partial_{22} w$.

Analogous soln for discrete pbm is in Gravin, Peres, and Sivan, SODA '16.

The random-stopping version, cont'd

In general: if time-dependent PDE is $u_t + Lu = 0$ with $u = \phi$ at $t = T$, then random-stopping PDE (with $\delta = \varepsilon^2/\lambda$) is $u - \lambda L w = \phi$.

- $N = 3$ soln is explicit because market's optimal strategy is very simple: when expert j is ahead, advance him with prob $\frac{1}{2}$, and advance all the others with prob $\frac{1}{2}$.

  Thus: pde soln has reflection symmetry across planes where experts' order changes. Moreover, in sector $x_1 > x_2 > x_3$ it solves linear eqn $w - \frac{1}{2}\lambda\partial_{11} w = x_1$.

- In ML lit, a common test-strategy for the investor uses probabilities p_j depending exponentially on the experts' performance. For $N = 3$, the investor's optimal strategy $p_j = \partial_j w$ has this character (since w is a sum of exponentials). It also has a simple interpretation: p_j = probability that expert j is ahead when the game stops.

- $N = 4$ is different. In fact, if market's optimal strategy were directly analogous to $N = 3$ (depending only on identity of leading expert) then w would solve $w - \frac{1}{2}\lambda \max_j\{\partial_{jj} w\} = \phi$. Soln of this PDE is explicit (generalizing $N = 3$ calcn), and it does not solve $w - \lambda L w = \phi$.

The random-stopping version, cont'd

In general: if time-dependent PDE is $u_t + Lu = 0$ with $u = \phi$ at $t = T$, then random-stopping PDE (with $\delta = \varepsilon^2/\lambda$) is $u - \lambda L w = \phi$.

- $N = 3$ soln is explicit because market's optimal strategy is very simple: when expert j is ahead, advance him with prob $\frac{1}{2}$, and advance all the others with prob $\frac{1}{2}$.

  Thus: pde soln has reflection symmetry across planes where experts' order changes. Moreover, in sector $x_1 > x_2 > x_3$ it solves linear eqn $w - \frac{1}{2}\lambda\partial_{11}w = x_1$.

- In ML lit, a common test-strategy for the investor uses probabilities p_j depending exponentially on the experts' performance. For $N = 3$, the investor's optimal strategy $p_j = \partial_j w$ has this character (since w is a sum of exponentials). It also has a simple interpretation: p_j = probability that expert j is ahead when the game stops.

- $N = 4$ is different. In fact, if market's optimal strategy were directly analogous to $N = 3$ (depending only on identity of leading expert) then w would solve $w - \frac{1}{2}\lambda \max_j\{\partial_{jj} w\} = \phi$. Soln of this PDE is explicit (generalizing $N = 3$ calcn), and it does not solve $w - \lambda L w = \phi$.

The random-stopping version, cont'd

In general: if time-dependent PDE is $u_t + Lu = 0$ with $u = \phi$ at $t = T$, then random-stopping PDE (with $\delta = \varepsilon^2/\lambda$) is $u - \lambda Lw = \phi$.

- $N = 3$ soln is explicit because market's optimal strategy is very simple: when expert j is ahead, advance him with prob $\frac{1}{2}$, and advance all the others with prob $\frac{1}{2}$.

  Thus: pde soln has reflection symmetry across planes where experts' order changes. Moreover, in sector $x_1 > x_2 > x_3$ it solves linear eqn $w - \frac{1}{2}\lambda\partial_{11}w = x_1$.

- In ML lit, a common test-strategy for the investor uses probabilities p_j depending exponentially on the experts' performance. For $N = 3$, the investor's optimal strategy $p_j = \partial_j w$ has this character (since w is a sum of exponentials). It also has a simple interpretation: p_j = probability that expert j is ahead when the game stops.

- $N = 4$ is different. In fact, if market's optimal strategy were directly analogous to $N = 3$ (depending only on identity of leading expert) then w would solve $w - \frac{1}{2}\lambda \max_j\{\partial_{jj}w\} = \phi$. Soln of this PDE is explicit (generalizing $N = 3$ calcn), and it does <u>not</u> solve $w - \lambda Lw = \phi$.

Rigorous results

(1) Our PDE's have at most one viscosity soln (w lin growth at ∞).

- from standard viscosity-solution theory

(2) For final-time pbm, $\lim_{\varepsilon\to 0} u_\varepsilon$ exists and solves the PDE (in the viscosity sense).

- from Barles-Souganidis thm on conv of numerical schemes

(3) For the random-stopping problem, w_ε exists (ie the scaled DPP has a solution); also, $\lim_{\varepsilon\to 0} w_\varepsilon$ exists and solves PDE.

- get w_ε as steady state of an assoc time-dependent pbm; convergence via Barles-Souganidis.

(4) Uniform estimates on $u_\varepsilon(x,t)$ and $w_\varepsilon(x)$ give stability, and also qualitative results about PDE solutions; for example, $\sup_x |w(x) - \phi(x)| \leq C$.

Rigorous results

(1) Our PDE's have at most one viscosity soln (w lin growth at ∞).

- from standard viscosity-solution theory

(2) For final-time pbm, $\lim_{\varepsilon \to 0} u_\varepsilon$ exists and solves the PDE (in the viscosity sense).

- from Barles-Souganidis thm on conv of numerical schemes

(3) For the random-stopping problem, w_ε exists (ie the scaled DPP has a solution); also, $\lim_{\varepsilon \to 0} w_\varepsilon$ exists and solves PDE.

- get w_ε as steady state of an assoc time-dependent pbm; convergence via Barles-Souganidis.

(4) Uniform estimates on $u_\varepsilon(x,t)$ and $w_\varepsilon(x)$ give stability, and also qualitative results about PDE solutions; for example, $\sup_x |w(x) - \phi(x)| \leq C$.

Rigorous results

(1) Our PDE's have at most one viscosity soln (w lin growth at ∞).

- from standard viscosity-solution theory

(2) For final-time pbm, $\lim_{\varepsilon \to 0} u_\varepsilon$ exists and solves the PDE (in the viscosity sense).

- from Barles-Souganidis thm on conv of numerical schemes

(3) For the random-stopping problem, w_ε exists (ie the scaled DPP has a solution); also, $\lim_{\varepsilon \to 0} w_\varepsilon$ exists and solves PDE.

- get w_ε as steady state of an assoc time-dependent pbm; convergence via Barles-Souganidis.

(4) Uniform estimates on $u_\varepsilon(x, t)$ and $w_\varepsilon(x)$ give stability, and also qualitative results about PDE solutions; for example, $\sup_x |w(x) - \phi(x)| \leq C$.

Rigorous results

(1) Our PDE's have at most one viscosity soln (w lin growth at ∞).

- from standard viscosity-solution theory

(2) For final-time pbm, $\lim_{\varepsilon\to 0} u_\varepsilon$ exists and solves the PDE (in the viscosity sense).

- from Barles-Souganidis thm on conv of numerical schemes

(3) For the random-stopping problem, w_ε exists (ie the scaled DPP has a solution); also, $\lim_{\varepsilon\to 0} w_\varepsilon$ exists and solves PDE.

- get w_ε as steady state of an assoc time-dependent pbm; convergence via Barles-Souganidis.

(4) Uniform estimates on $u_\varepsilon(x,t)$ and $w_\varepsilon(x)$ give stability, and also qualitative results about PDE solutions; for example, $\sup_x |w(x) - \phi(x)| \le C$.

Rigorous results – some hints about the methods

Focus for simplicity on the final-time version: $u_t + \mathcal{L}u = 0$ for $t < T$, with $w = \phi$ at $t = T$.

(1) STABILITY Let $\tilde{\phi}$ be a mollified version of ϕ (so $\tilde{\phi}$ shares the structural features of ϕ, but is C^3). Then time-stepping changes $\max_x |u(x,t) - \tilde{\phi}(x)|$ by at most $C\varepsilon^2$. So

$$\max_x |u(x,t) - \tilde{\phi}(x)| \leq C_1 + C_2(T - t).$$

(2) CONSISTENCY The formal calculation shows (when done more carefully) that if $u(t,x)$ is smooth and satisfies the structural conditions (monotone in each x_j, and $u(t, x + c\vec{1}) = u(t,x) + c$) then

$$\frac{\min_{\text{investor}} \max_{market} E[u(t+\varepsilon^2, x + \varepsilon\Delta x)] - u(t,x)}{\varepsilon^2} \approx u_t + \mathcal{L}u.$$

Main point: structural conds assure that 2nd order part doesn't depend on player's probabilities. Thus: in the min-max, first-order term can really be handled separately (even for $\varepsilon > 0$).

Rigorous results – some hints about the methods

Focus for simplicity on the final-time version: $u_t + \mathcal{L}u = 0$ for $t < T$, with $w = \phi$ at $t = T$.

(1) STABILITY Let $\tilde{\phi}$ be a mollified version of ϕ (so $\tilde{\phi}$ shares the structural features of ϕ, but is C^3). Then time-stepping changes $\max_x |u(x,t) - \tilde{\phi}(x)|$ by at most $C\varepsilon^2$. So

$$\max_x |u(x,t) - \tilde{\phi}(x)| \le C_1 + C_2(T - t).$$

(2) CONSISTENCY The formal calculation shows (when done more carefully) that if $u(t,x)$ is smooth and satisfies the structural conditions (monotone in each x_j, and $u(t, x + c\vec{1}) = u(t,x) + c$) then

$$\frac{\min_{\text{investor}} \max_{market} E[u(t+\varepsilon^2, x + \varepsilon \Delta x)] - u(t,x)}{\varepsilon^2} \approx u_t + \mathcal{L}u.$$

Main point: structural conds assure that 2nd order part doesn't depend on player's probabilities. Thus: in the min-max, first-order term can really be handled separately (even for $\varepsilon > 0$).

Stepping back

Mathematical messages

- This approach to prediction leads, in a suitable limit, to some interesting nonlinear PDE. Their solutions determine the optimal strategies (at least if the solutions are smooth enough).
- For classic goal of minimizing regret ($\phi(x) = \max_k\{x_k\}$), explicit solns are available in some cases (deterministic stopping – 2 experts; random stopping – 3 experts.) What about other cases, and other ϕ?

Machine learning messages

- ML literature is mainly discrete. This example suggests that PDE can help. But full impact is far from clear.
- In ML, focus is usually on easy-to-implement strategies (based eg on schemes for weighting experts, using past performance). PDE soln, if known, permits comparison to the optimal strategy.
- For ML, focus is often on asymptotics as $\#$ experts $\to \infty$. PDE, by contrast, are inconvenient in high dimensions (unless explicit solns are available).

Stepping back

Mathematical messages

- This approach to prediction leads, in a suitable limit, to some interesting nonlinear PDE. Their solutions determine the optimal strategies (at least if the solutions are smooth enough).
- For classic goal of minimizing regret ($\phi(x) = \max_k\{x_k\}$), explicit solns are available in some cases (deterministic stopping – 2 experts; random stopping – 3 experts.) What about other cases, and other ϕ?

Machine learning messages

- ML literature is mainly discrete. This example suggests that PDE can help. But full impact is far from clear.
- In ML, focus is usually on easy-to-implement strategies (based eg on schemes for weighting experts, using past performance). PDE soln, if known, permits comparison to the optimal strategy.
- For ML, focus is often on asymptotics as $\#$ experts $\to \infty$. PDE, by contrast, are inconvenient in high dimensions (unless explicit solns are available).