Optimal Online Prediction in Adversarial Environments

Peter Bartlett
EECS and Statistics
UC Berkeley

http://www.cs.berkeley.edu/∼bartlett

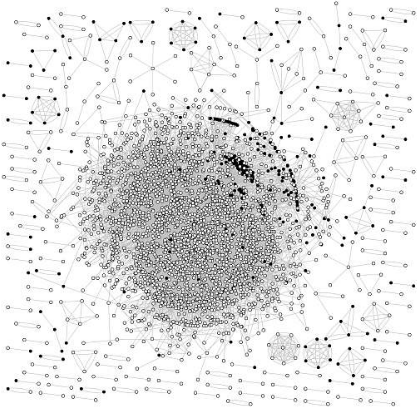
Online Prediction

- Probabilistic Model
  - **Batch**: independent **random** data.
  - Aim for small **expected** loss subsequently.
- Adversarial Model
  - **Online**: Sequence of interactions with an **adversary**.
  - Aim for small **cumulative** loss throughout.

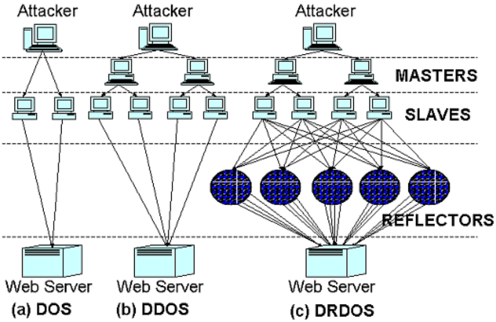
Online Learning: Motivations

1. Adversarial model is appropriate for
   - Computer security.
   - Computational finance.


SPAM
SPAM*** What IS
SPAM*** Viagra
SPAM*** work
SPAM
The



Web Spam Challenge (www.iw3c2.org)


Attacker
Attacker
Attacker
MASTERS
SLAVES
REFLECTORS
Web Server
Web Server
Web Server
(a) DOS
(b) DDOS
(c) DRDOS

ACM


1.7765
1.7900
1.7855
1.7810
1.7765
11.12
14.56

Online Learning: Motivations

2. Understanding statistical prediction methods.

- Many statistical methods, based on *probabilistic assumptions*, can be effective in an adversarial setting.
- Analyzing their performance in adversarial settings provides perspective on their robustness.
- We would like violations of the probabilistic assumptions to have a limited impact.

Online Learning: Motivations

3. Online algorithms are also effective in probabilistic settings.

- Easy to convert an online algorithm to a batch algorithm.
- Easy to show that good online performance implies good i.i.d. performance, for example.

Prediction in Probabilistic Settings

- i.i.d. $(X, Y), (X_1, Y_1), \ldots, (X_n, Y_n)$ from $\mathcal{X} \times \mathcal{Y}$.
- Use data $(X_1, Y_1), \ldots, (X_n, Y_n)$ to choose $f_n : \mathcal{X} \to \mathcal{A}$ with small risk,

$$R(f_n) = \mathbf{E}\ell(Y, f_n(X)).$$

Online Learning

- Repeated game:

$$\text{Player chooses } a_t$$
$$\text{Adversary reveals } \ell_t$$

- Example: $\ell_t(a_t) = \text{loss}(y_t, a_t(x_t))$.
- Aim: minimize $\sum_t \ell_t(a_t)$, compared to the best (in retrospect) from some class:

$$\text{regret} = \sum_t \ell_t(a_t) - \min_{a \in \mathcal{A}} \sum_t \ell_t(a).$$

- Data can be adversarially chosen.

Outline

1. An Example from Computational Finance: The Dark Pools Problem.
2. Bounds on Optimal Regret for General Online Prediction Problems.

The Dark Pools Problem

- Computational finance: adversarial setting is appropriate.
- Online algorithm improves on best known algorithm for probabilistic setting.

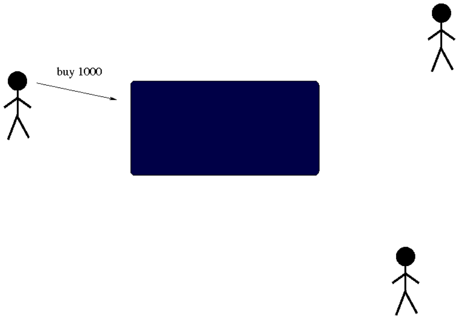
Joint work with Alekh Agarwal and Max Dama.

Dark Pools

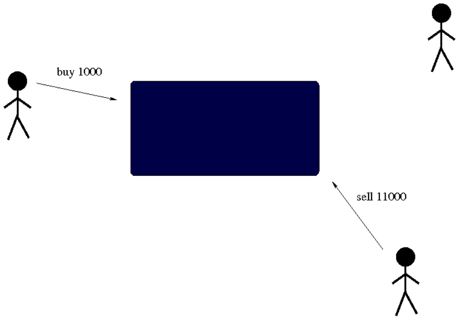
Instinet, Chi-X, Knight Match, ... International Securities Exchange, Investment Technology Group (POSIT),

- Crossing networks.
- Alternative to open exchanges.
- Avoid market impact by hiding transaction size and traders' identities.

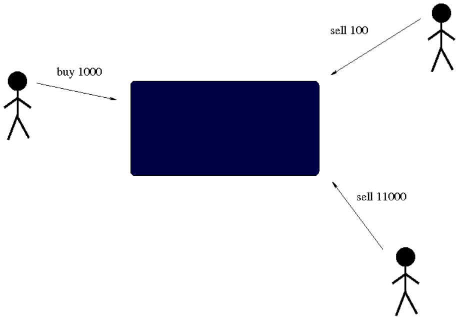
Dark Pools



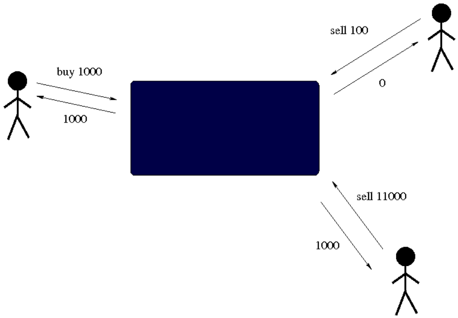
Dark Pools



Dark Pools



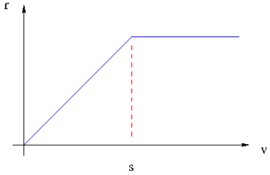
Dark Pools



Allocations for Dark Pools

The problem: Allocate orders to several dark pools so as to maximize the volume of transactions.

- Volume V^t must be allocated across K venues: $v_1^t, \ldots, v_K^t$, such that $\sum_{k=1}^{K} v_k^t = V^t$.
- Venue k can accommodate up to s_k^t, transacts $r_k^t = \min(v_k^t, s_k^t)$.



- The aim is to maximize $\sum_{t=1}^{T} \sum_{k=1}^{K} r_k^t$.

Allocations for Dark Pools: Probabilistic Assumptions

Previous work: (Ganchev, Kearns, Nevmyvaka and Wortman, 2008)

- Assume venue volumes are i.i.d.: $\{s_k^t,\ k = 1, \ldots, K,\ t = 1, \ldots, T\}$.
- In deciding how to allocate the first unit, choose the venue k where $Pr(s_k^t > 0)$ is largest.
- Allocate the second and subsequent units in decreasing order of venue tail probabilities.
- Algorithm: estimate the tail probabilities (Kaplan-Meier estimator—data is censored), and allocate as if the estimates are correct.

Allocations for Dark Pools: Adversarial Assumptions

Why i.i.d. is questionable:

- one party's gain is another's loss
- volume available now affects volume remaining in future
- volume available at one venue affects volume available at others

In the adversarial setting, we allow an arbitrary sequence of venue capacities (s_k^t), and of total volume to be allocated (V^t). The aim is to compete with any fixed allocation order.

Continuous Allocations

We wish to maximize a sum of (unknown) concave functions of the allocations:

$$J(v) = \sum_{t=1}^{T} \sum_{k=1}^{K} \min(v_k^t, s_k^t),$$

subject to the constraint $\sum_{k=1}^{K} v_k^t \leq V^t$.
The allocations are parameterized as distributions over the K venues:

$$x_t^1, x_t^2, \ldots \in \Delta_{K-1} = (K-1)\text{-simplex}.$$

Here, x_t^1 determines how the first unit is allocated, x_t^2 the second, ...

The algorithm allocates to the kth venue: $v_k^t = \sum_{v=1}^{V^t} x_{t,k}^v$.

Continuous Allocations

We wish to maximize a sum of (unknown) concave functions of the distributions:

$$J = \sum_{t=1}^{T} \sum_{k=1}^{K} \min(v_k^t(x_{t,k}^v), s_k^t).$$

Want small regret with respect to an arbitrary distribution x^v, and hence w.r.t. an arbitrary allocation.

$$\text{regret} = \sum_{t=1}^{T} \sum_{k=1}^{K} \min(v_k^t(x_k^v), s_k^t) - J.$$

Continuous Allocations

We use an exponentiated gradient algorithm:

Initialize $x^v_{1,i} = \frac{1}{K}$ for $v = \{1, \ldots, V\}$.
for $t = 1, \ldots, T$ **do**
  Set $v^t_k = \sum_{v=1}^{V^T} x^v_{t,k}$.
  Receive $r^t_k = \min\{v^t_k, s^t_k\}$.
  Set $g^v_{t,k} = \nabla_{x^v_{t,k}} J$.
  Update $x^v_{t+1,k} \propto x^v_{t,k} \exp(\eta g^v_{t,k})$.
end for

Continuous Allocations

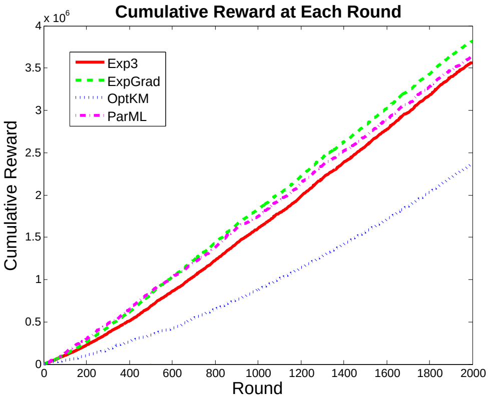
Theorem: For all choices of $V^t \leq V$ and of s_k^t, ExpGrad has regret no more than $3V\sqrt{T \ln K}$.

Continuous Allocations

Theorem: For all choices of $V^t \leq V$ and of s_k^t, ExpGrad has regret no more than $3V\sqrt{T \ln K}$.

Theorem: For every algorithm, there are sequences V^t and s_k^t such that regret is at least $V\sqrt{T \ln K}/16$.

Experimental results



Continuous Allocations: i.i.d. data

- Simple online-to-batch conversions show ExpGrad obtains per-trial utility within $O(T^{-1/2})$ of optimal.
- Ganchev et al bounds:
  per-trial utility within $O(T^{-1/4})$ of optimal.

Discrete allocations

- Trades occur in quantized parcels.
- Hence, we cannot allocate arbitrary values.
- This is analogous to a multi-arm bandit problem:
  - We cannot directly obtain the gradient at the current x.
  - But, we can estimate it using importance sampling ideas.

Theorem: There is an algorithm for discrete allocation with expected regret $\tilde{O}((VTK)^{2/3})$.
Any algorithm has regret $\tilde{\Omega}((VTK)^{1/2})$.

Dark Pools

- Allow adversarial choice of volumes and transactions.
- Per trial regret rate superior to previous best known bounds for probabilistic setting.
- In simulations, performance comparable to (correct) parametric model's, and superior to nonparametric estimate.

Outline

1. An Example from Computational Finance: The Dark Pools Problem.
2. Bounds on Optimal Regret for General Online Prediction Problems.

Optimal Regret for General Online Decision Problems

- Parallels between probabilistic and online frameworks.
- Tools for the analysis of probabilistic problems: Rademacher averages.
- Analogous results in the online setting:
  - Value of dual game.
  - Bounds in terms of Rademacher averages.
- Open problems.

Joint work with Jake Abernethy, Alekh Agarwal, Sasha Rakhlin, Karthik Sridharan and Ambuj Tewari.

Prediction in Probabilistic Settings

- i.i.d. $(X, Y), (X_1, Y_1), \ldots, (X_n, Y_n)$ from $\mathcal{X} \times \mathcal{Y}$.
- Use data $(X_1, Y_1), \ldots, (X_n, Y_n)$ to choose $f_n : \mathcal{X} \to \mathcal{A}$ with small risk,

$$R(f_n) = P\ell(Y, f_n(X)),$$

ideally not much larger than the minimum risk over some comparison class F:

$$\text{excess risk} = R(f_n) - \inf_{f \in F} R(f).$$

Parallels between Probabilistic and Online Settings

- Prediction with i.i.d. data:
  - Convex F, strictly convex loss, $\ell(y, f(x)) = (y - f(x))^2$:

$$\sup_P \left(\mathbb{P}R(\hat{f}) - \inf_{f \in F} R(f) \right) \approx \frac{C(F) \log n}{n}.$$

  - Nonconvex F, or (not strictly) convex loss, $\ell(y, f(x)) = |y - f(x)|$:

$$\sup_P \left(\mathbb{P}R(\hat{f}) - \inf_{f \in F} R(f) \right) \approx \frac{C(F)}{\sqrt{n}}.$$

- Online convex optimization:
  - Convex $\mathcal{A}$, strictly convex ℓ_t:

$$\text{per trial regret} \approx \frac{c \log n}{n}.$$

  - ℓ_t (not strictly) convex:

$$\text{per trial regret} \approx \frac{c}{\sqrt{n}}.$$

Tools for the analysis of probabilistic problems

For $f_n = \arg\min_{f \in F} \sum_{t=1}^{n} \ell(Y_t, f(X_t))$,

$$R(f_n) - \inf_{f \in F} P\ell(Y, f(X)) \leq 2 \sup_{f \in F} \left| \frac{1}{n} \sum_{t=1}^{n} \ell(Y_t, f(X_t)) - P\ell(Y, f(X)) \right|.$$

So supremum of empirical process, indexed by F, gives upper bound on excess risk.

Tools for the analysis of probabilistic problems

Typically, this supremum is concentrated about

$$\begin{aligned}
&\mathbb{P} \sup_{f \in F} \left| \frac{1}{n} \sum_{t=1}^{n} \left(\ell(Y_t, f(X_t)) - P\ell(Y, f(X)) \right) \right| \\
&= \mathbb{P} \sup_{f \in F} \left| \mathbb{P}' \frac{1}{n} \sum_{t=1}^{n} \left(\ell(Y_t, f(X_t)) - \ell(Y_t', f(X_t')) \right) \right| \\
&\leq \mathbf{E} \sup_{f \in F} \left| \frac{1}{n} \sum_{t=1}^{n} \epsilon_t \left(\ell(Y_t, f(X_t)) - \ell(Y_t', f(X_t')) \right) \right| \\
&\leq 2\mathbf{E} \sup_{f \in F} \left| \frac{1}{n} \sum_{t=1}^{n} \epsilon_t \ell(Y_t, f(X_t)) \right|,
\end{aligned}$$

where (X_t', Y_t') are independent, with same distribution as (X, Y), and ϵ_t are independent Rademacher (uniform ± 1) random variables.

Tools for the analysis of probabilistic problems

That is, for $f_n = \arg\min_{f \in F} \sum_{t=1}^{n} \ell(Y_t, f(X_t))$, with high probability,

$$R(f_n) - \inf_{f \in F} P\ell(Y, f(X)) \leq c\mathbf{E} \sup_{f \in F} \left| \frac{1}{n} \sum_{t=1}^{n} \epsilon_t \ell(Y_t, f(X_t)) \right|,$$

where ϵ_t are independent Rademacher (uniform ± 1) random variables.

- Rademacher averages capture complexity of $\{(x, y) \mapsto \ell(y, f(x)) : f \in F\}$: they measure how well functions align with a random $(\epsilon_1, \ldots, \epsilon_n)$.
- Rademacher averages are a key tool in analysis of many statistical methods: related to covering numbers (Dudley) and combinatorial dimensions (Vapnik-Chervonenkis, Pollard), for example.
- A related result applies in the online setting...

Online Decision Problems

We have:

- a set of actions $\mathcal{A}$,
- a set of loss functions $\mathcal{L}$.

At time t,

- Player chooses distribution P_t on decision set $\mathcal{A}$.
- Adversary chooses $\ell_t \in \mathcal{L}$ ($\ell_t : \mathcal{A} \to \mathbb{R}$).
- Player incurs loss $P_t \ell_t$.

Regret is value of game:

$$V_n(\mathcal{A}, \mathcal{L}) = \inf_{P_1} \sup_{\ell_1} \cdots \inf_{P_n} \sup_{\ell_n} \mathbf{E} \left(\sum_{t=1}^{n} \ell_t(a_t) - \inf_{a \in \mathcal{A}} \sum_{t=1}^{n} \ell_t(a) \right),$$

where $a_t \sim P_t$.

Optimal Regret in Online Decision Problems

Theorem

$$V_n = \sup_P P \left(\sum_{t=1}^{n} \inf_{a_t \in \mathcal{A}} \mathbf{E} \left[\ell_t(a_t) | \ell_1, \ldots, \ell_{t-1} \right] - \inf_{a \in \mathcal{A}} \sum_{t=1}^{n} \ell_t(a) \right),$$

where P is distribution over sequences $\ell_1, \ldots, \ell_n$.

- Follows from von Neumann's minimax theorem.
- Dual game: adversary plays first by choosing P.

Optimal Regret in Online Decision Problems

Theorem

$$V_n = \sup_P P\left(\sum_{t=1}^{n} \inf_{a_t \in \mathcal{A}} \mathbf{E}\left[\ell_t(a_t) | \ell_1, \ldots, \ell_{t-1}\right] - \inf_{a \in \mathcal{A}} \sum_{t=1}^{n} \ell_t(a)\right),$$

where P is distribution over sequences $\ell_1, \ldots, \ell_n$.

- Value is the difference between minimal (conditional) expected loss and minimal empirical loss.
- If P were i.i.d., the expression would be the difference between the minimal expected loss and minimal empirical loss.

Optimal Regret in Online Decision Problems

Theorem

$$V_n \leq 2 \sup_{\ell_1} \mathbf{E}_{\epsilon_1} \cdots \sup_{\ell_n} \mathbf{E}_{\epsilon_n} \sup_{a \in \mathcal{A}} \sum_{t=1}^{n} \epsilon_t \ell_t(a),$$

where $\epsilon_1, \ldots, \epsilon_n$ are independent Rademacher (uniform ± 1-valued) random variables.

- Compare to bound involving Rademacher averages in the probabilistic setting:

$$\text{excess risk} \leq c \mathbf{E} \sup_{f \in F} \left| \frac{1}{n} \sum_{t=1}^{n} \epsilon_t \ell(Y_t, f(X_t)) \right| .$$

- In the adversarial case, the choice of ℓ_t is deterministic, and can depend on $\epsilon_1, \ldots, \epsilon_{t-1}$.
- Proof idea similar to i.i.d. case, but using a *tangent sequence* (dependent on previous ℓ_ts).

Optimal Regret: Lower Bounds

- Rakhlin, Sridharan and Tewari recently considered the case of prediction with absolute loss:

$$\ell_t(a_t) = |y_t - a_t(x_t)|,$$

and showed (almost) corresponding lower bounds:

$$\frac{c_1 R_n(\mathcal{A})}{\log^{3/2} n} \leq V_n \leq c_2 R_n(\mathcal{A}),$$

where

$$R_n(\mathcal{A}) = \sup_{x_1} \mathbf{E}_{\epsilon_1} \cdots \sup_{x_n} \mathbf{E}_{\epsilon_n} \sup_{a \in \mathcal{A}} \sum_{t=1}^{n} \epsilon_t a(x_t).$$

Optimal Regret: Open Problems

- The bounds on regret of an optimal strategy in the online framework might be loose:
  In the probabilistic setting, the supremum of the empirical process can be a loose bound on the excess risk. If the variance of the excess loss can be bounded in terms of its expectation (for example, in regression with a strongly convex loss and a convex function class, or in classification with a margin condition on the conditional class probability), then we can get better (optimal) rates with *local Rademacher averages*.
  Is there an analogous result in the online setting?

Optimal Regret: Open Problems

- These results bound the regret of an optimal strategy, but they are not constructive.
  In what cases can we efficiently solve the optimal online prediction optimization problem?

Outline

1. An Example from Computational Finance: The Dark Pools Problem.
   - Adversarial model is appropriate.
   - Online strategy improves on the regret rate of previous best known method for probabilistic setting.
2. Bounds on Optimal Regret for General Online Prediction Problems.
   - Parallels between probabilistic and online frameworks.
   - Tools for the analysis of probabilistic problems: Rademacher averages.
   - Bounds on optimal online regret using Rademacher averages.