Dueling Convex Optimization

Aadirupa Saha[1], Tomer Koren[2], Yishay Mansour[2]
[1] Microsoft Research, New York City
[2] Blavatnik School of Computer Science, Tel Aviv University, & Google Research Tel Aviv.

38th International Conference on Machine Learning, 2021

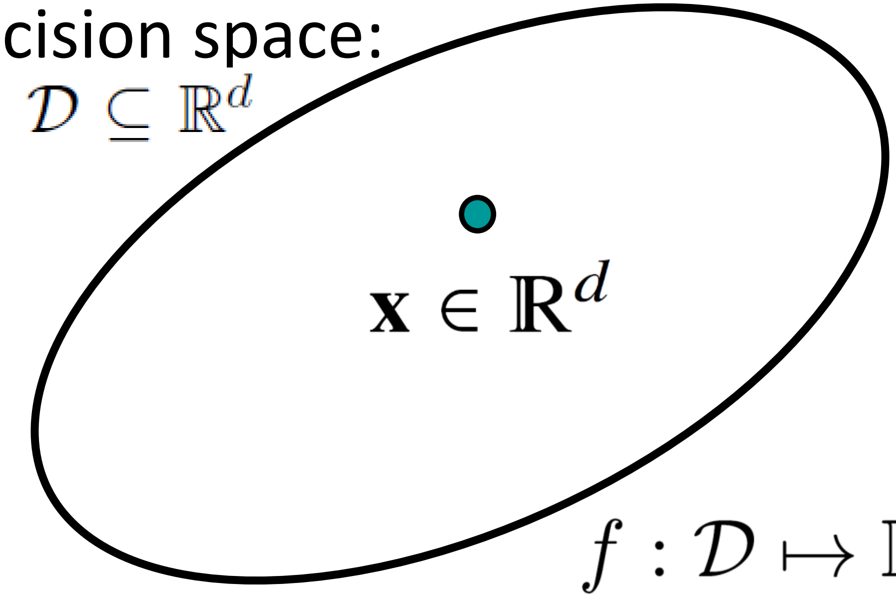
Problem Overview

Dueling Convex Optimization

Problem: Online Convex Optimization

Classical problem of Online Convex Optimization (OCO)

Decision space:
$\mathcal{D} \subseteq \mathbb{R}^d$



$f : \mathcal{D} \mapsto \mathbb{R}^d$
Convex

$\{f_1, f_2, \ldots, f_t\} : \mathcal{D} \mapsto \mathbb{R}^d$

At round t = 1,2,…,T

- Environment chooses f_t
- Play point $x_t \in D$
- Receive feedback at x_t

End

First order feedback:
Gradient information $\nabla f_t(\mathbf{x}_t)$

Zeroth order feedback:
Function value $f_t(\mathbf{x}_t)$

Objective: Fast Convergence to Optimal Point

Function Optimization (minimization): Find a point $\mathbf{x} \in \mathbb{R}^d$

Estimation error, given (ε,δ) ϵ [0,1]: $Pr(\bar{f}(\mathbf{x}) - \bar{f}(\mathbf{x}^*) < \epsilon) \geq 1 - \delta$

where, averaged function: $\bar{f}(x) := \sum_{t=1}^{T} \frac{\mathbb{E}_{f_t \sim \mathcal{P}}[f_t(x)]}{T}$

and, true minimizer: $\mathbf{x}^* := \arg\min_{\mathbf{z} \in \mathbb{R}^d} \bar{f}(z)$

with least possible #queries (T)

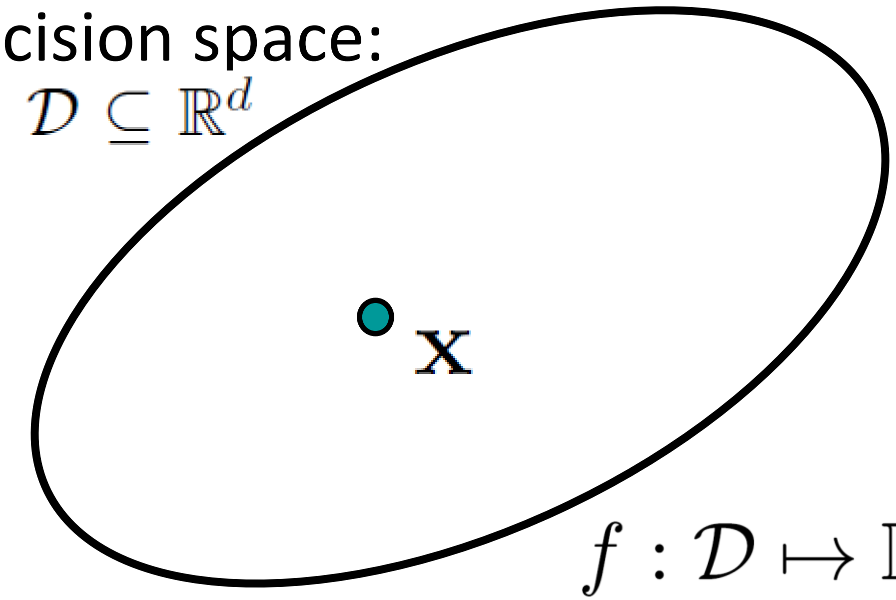
Same objective
with
Dueling/Comparison feedback?

Our Setup: Optimization from 0/1 feedback

Classical problem of Online Convex Optimization (OCO)

Decision space:

$\mathcal{D} \subseteq \mathbb{R}^d$



$f : \mathcal{D} \mapsto \mathbb{R}^d$

Convex

$\{f_1, f_2, \ldots, f_t\} : \mathcal{D} \mapsto \mathbb{R}^d$

At round t = 1,2,…,T

- Environment chooses f_t
- Play point $x_t \in D$
- Receive feedback at x_t

End

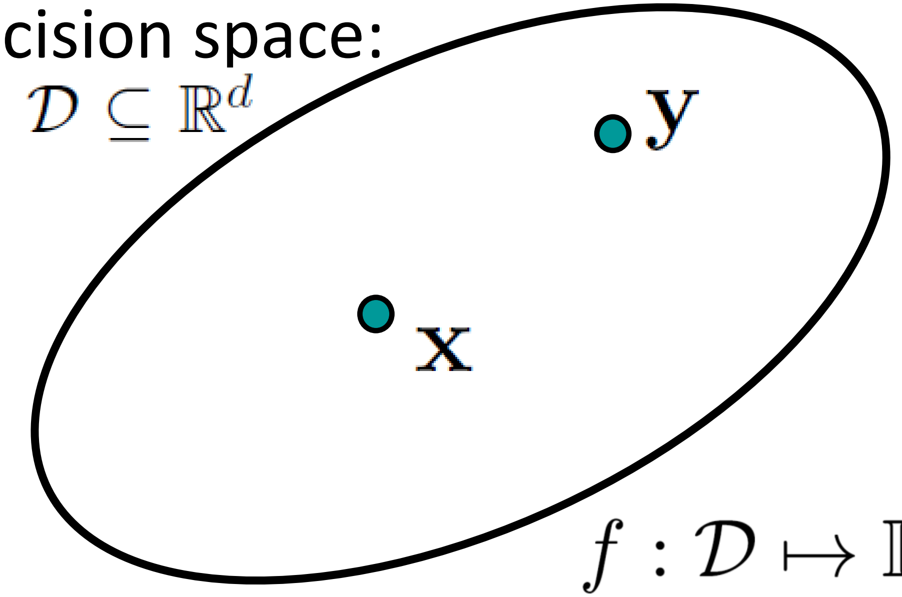
rst order feedback:
Gradient information $\nabla f_t(\mathbf{x}_t)$

eroth order feedback:
Function value $f_t(\mathbf{x}_t)$

Our Setup: Optimization from 0/1 feedback

Classical problem of Online Convex Optimization (OCO)

Decision space: $\mathcal{D} \subseteq \mathbb{R}^d$



$f : \mathcal{D} \mapsto \mathbb{R}^d$
Convex

$\{f_1, f_2, \ldots, f_t\} : \mathcal{D} \mapsto \mathbb{R}^d$

At round t = 1,2,…,T

- Environment chooses f_t
- Play *pair* $(x_t, y_t) \in D$
- See comparison: $o_t = 1(f(x_t) > f(y_t))$

End

0/1 Comparison feedback: Much "weaker"!!

Prior works: Almost none!

1. Jamieson et al. *Query complexity of derivative-free optimization.* NeurIPS 2012.

   - Only strongly-convex+ smooth functions

2. Kumagai W. *Regret analysis for continuous dueling bandit.* NeurIPS 2015.

   - Extremely restricted function class: Twice continuously differentiable, L-Lipschitz, strongly convex and smooth and a thrice differentiable, rotation symmetric preference map.

Main Challenge

- No gradient information!
- 1 bit feedback: Impossible to estimate gradient magnitude!

Impossibility result:
Non-stationary function sequence

Impossibility: Non-stationary or Stochastic Sequences

Simple two-point decision space: $\mathcal{D}$ ... $, \mathbf{x}_2\}$

$$\mathcal{F} = \{f_1, f_2\} \qquad f_t = \begin{cases} f_1, & \text{w}\ldots \\ \ldots \end{cases}$$

Instance $\mathcal{I}$ (x* ... = x$_2$)

$\mathbb{E}_{f_t \sim \mathcal{P}}[f_t(\mathbf{x}_1)] = 0.0099$

$\mathbb{E}_{f_t \sim \mathcal{P}}[f_t(\mathbf{x}_2)] = 0.001$

$f_1(\mathbf{x})$ ...

...me!

$$f_2(\mathbf{x}) = \begin{cases} \ldots \\ 1, & \ldots\text{wise} \end{cases}$$

$$f_2(\mathbf{x}) = \begin{cases} 0, & \text{if } \mathbf{x} = \mathbf{x}_1 \\ 0.1 & \text{otherwise} \end{cases}$$

$$Pr_{f_t \sim \mathcal{P}}(\mathbf{x}_2 \succ \mathbf{x}_1) = 0.99$$

Same!

No ways to distinguish!

Summary of Results

(Fixed function $f_t = f, \forall t \in [T]$)

Our results: Fixed function $f_t=f, \forall t \in [T]$

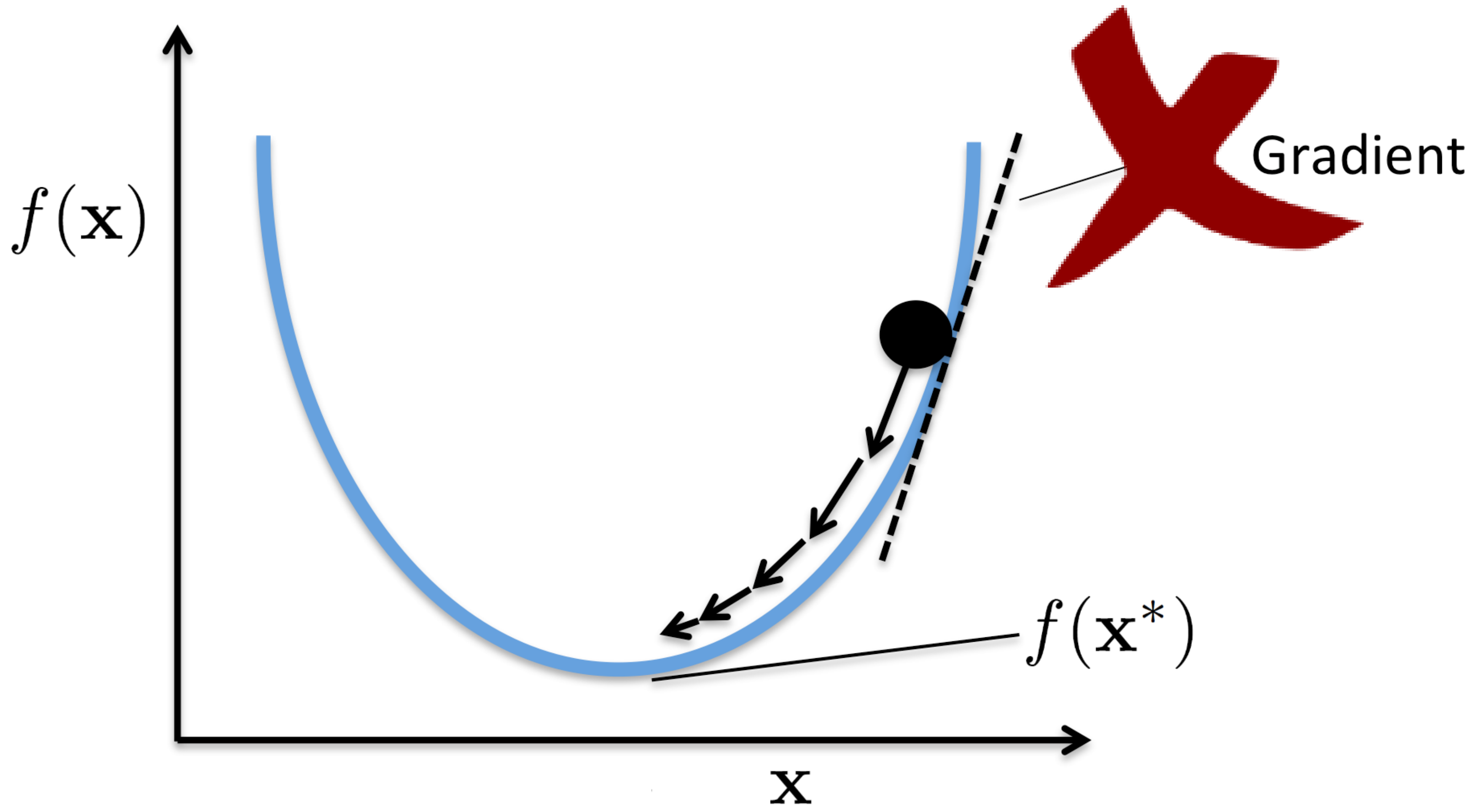
Rate of Convergence (Sample-Complexity of our algorithms)

Observe: $\boldsymbol{o_t} = \mathbf{1}(\boldsymbol{f}(\boldsymbol{x_t} > \boldsymbol{f}(\boldsymbol{y_t}))$	Noisy: $P(o_t = 1(f(x_t) > f(y_t)) > \frac{1}{2} + \nu$
- Only smooth functions $O\left(\frac{d\beta D}{\epsilon}\right),\ D = \|\mathbf{x}_1 - \mathbf{x}^*\|^2$ - Faster convergence with Strong convexity $O\left(\frac{d\beta}{\alpha}\left(\log_2\left(\frac{\alpha}{\epsilon}\right) + \|\mathbf{x}_1 - \mathbf{x}^*\|^2\right)\right)$	- Smooth functions $O\left(\frac{d\beta D}{\epsilon\nu^2}\log\frac{d\beta D}{\epsilon\nu^2\delta}\right)$ - Strong convex+ Smooth (Faster convergence) $O\left(\frac{d\beta}{\nu^2\alpha}\left(\log_2\left(\frac{\alpha}{\epsilon}\right)+D\right)\right)\log\frac{d\beta D\log(\alpha/\epsilon)}{\nu^2\delta}$

Solution approach: Main Idea

Can we hope for Gradient descent?

Impossible to estimate gradient magnitude
from 0/1 comparison feedback!



But at least Gradient direction?

Key finding:

For any vector $g \in R^d$: $$\mathbb{E}[\text{sign}(\mathbf{g} \cdot \mathbf{u})\mathbf{u}] = \frac{c}{\sqrt{d}} \frac{\mathbf{g}}{||\mathbf{g}||}$$

where, $u \sim$ sphere of unit radius in R^d

$c \in [1/20, 1]$: constant

Normalized Gradient Estimate

How to estimate Normalized Gradient ?

Given any point $x \in R^d$:

$x + \gamma u$ | $x - \gamma u$

$u \sim$ sphere of unit radius in R^d

↓

Query duel: $(x+\gamma u, x-\gamma u)$

↓

Receive: $\text{sign}(f(x+\gamma u)-f(x-\gamma u))$

↓

Estimate: $P(\text{sign}(f(\mathbf{x}+\gamma\mathbf{u}) - f(\mathbf{x}-\gamma\mathbf{u}))\mathbf{u} = \text{sign}(\nabla f(\mathbf{x}) \cdot \mathbf{u})\mathbf{u}) > 1 - \lambda$

where, $\lambda = O\left(\frac{3\beta\gamma\sqrt{d}}{\|\nabla f(\mathbf{x})\|}\right)$

Proposed Algorithms

- Smooth functions
- Faster convergence with Strong convexity

Any β-Smooth function

Algorithm 1 β-*NGD*$(\mathbf{x}_1, \eta, \gamma, T_\epsilon)$

Input: Initial point: $\mathbf{x}_1 \in \mathbb{R}^d$ such that $D := \|\mathbf{x}_1 - \mathbf{x}^*\|^2$ (assume known), Learning rate η, Perturbation parameter γ, Query budget T_ϵ
(Recall the desired error tolerance is $\epsilon > 0$)
Initialize Current minimum $\tilde{\mathbf{x}}_1 \in \mathbb{R}^d$
for $t = 1, 2, 3, \ldots, T_\epsilon$ **do**
  Sample $\mathbf{u}_t \sim \text{Unif}(\mathcal{S}_d(1))$
  $\mathbf{x}'_t := \mathbf{x}_t + \gamma \mathbf{u}_t$
  $\mathbf{y}'_t := \mathbf{x}_t - \gamma \mathbf{u}_t$
  Play the duel $(\mathbf{x}'_t, \mathbf{y}'_t)$, and receive binary preference feedback $o_t = \mathbf{1}(f(\mathbf{x}'_t) < f(\mathbf{y}'_t))$. Set $o'_t = 2o_t - 1$.
  Update $\mathbf{x}_{t+1} \leftarrow \mathbf{x}_t - \eta \mathbf{h}_t$, where $\mathbf{h}_t = o'_t \mathbf{u}_t$
  Query the pair $(\mathbf{x}_{t+1}, \tilde{\mathbf{x}}_t)$.
  Update $\tilde{\mathbf{x}}_{t+1} \leftarrow \begin{cases} \mathbf{x}_{t+1} & \text{if } o'_t = -1 \\ \tilde{\mathbf{x}}_t & \text{otherwise } (o'_t = +1) \end{cases}$
end for
Return $\tilde{\mathbf{x}}_{T+1}$

NGD Estimation

Update

Tracking the minimum

Convergence Guarantee:

$$\mathbb{E}[f(\tilde{\mathbf{x}}_{T+1})] - f(\mathbf{x}^*) \leq \epsilon$$

in $O\left(\frac{d\beta D}{\epsilon}\right)$ queries!

α-Strongly convex + β-Smooth function

Algorithm 2 (α, β)-*NGD*(ϵ)

1: **Input:** Error tolerance $\epsilon > 0$

2: **Initialize** Initial point: $\mathbf{x}_1 \in \mathbb{R}^d$ such that $D := \|\mathbf{x}_0 - \mathbf{x}^*\|^2$ (assume known).
Phase counts $k_\epsilon := \lceil \log_2 \left(\frac{\alpha}{\epsilon}\right) \rceil$, $t \leftarrow \frac{800 d\beta}{(\sqrt{2}-1)\alpha}$
$\eta_1 \leftarrow \frac{\sqrt{\epsilon_1}}{20\sqrt{d\beta}}$, $\epsilon_1 = \frac{400 d\beta D}{(\sqrt{2}-1)t_1} = 1$, $t_1 = t\|\mathbf{x}_1 - \mathbf{x}^*\|^2$
$\gamma_1 \leftarrow \frac{(\epsilon_1/\beta)^{3/2}}{240\sqrt{2}d(D+\eta_1 t_1)^2\sqrt{\log 480\sqrt{\beta d}(D+\eta_1 t_1)/\sqrt{2\epsilon_1}}}$.

3: Update $\mathbf{x}_2 \leftarrow \beta$-*NGD*$\left(x_1, \eta_1, \gamma_1, t_1\right)$

4: **for** $k = 2, 3, \ldots, k_\epsilon$ **do**

5: $\eta_k \leftarrow \frac{\sqrt{\epsilon_k}}{20\sqrt{d\beta}}$, $\epsilon_k = \frac{400 d\beta}{(\sqrt{2}-1)t_k}$, $t_k = 2t$
$\gamma_k \leftarrow \frac{(\epsilon_k/\beta)^{3/2}}{240\sqrt{2}d(1+\eta_k t_k)^2\sqrt{\log 480\sqrt{\beta d}(1+\eta_k t_k)/\sqrt{2\epsilon_k}}}$.

6: Update $\mathbf{x}_{k+1} \leftarrow \beta$-*NGD*$\left(x_k, \eta_k, \gamma_k, t_k\right)$

7: **end for**

8: Return $\tilde{\mathbf{x}} = \mathbf{x}_{k_\epsilon+1}$

Phasewise blackbox

Strong-convexity property:

$$\frac{\alpha}{2}\|\mathbf{x}^* - \mathbf{x}\|^2 \leq f(\mathbf{x}) - f(\mathbf{x}^*)$$

Convergence Guarantee:

$$\mathbb{E}[f(\tilde{\mathbf{x}}_{T+1})] - f(\mathbf{x}^*) \leq \epsilon$$

in $O\left(\frac{d\beta}{\alpha}\left(\log_2\left(\frac{\alpha}{\epsilon}\right) + \|\mathbf{x}_1 - \mathbf{x}^*\|^2\right)\right)$ queries!

Faster convergence!!

Robustness: Noisy Feedback

General Setup: Noisy 0/1 feedback

Observe correct comparison with probability $\frac{1}{2} + \nu$

$$Pr(o_t = \mathbf{1}(f(\mathbf{y}_t) > f(\mathbf{x}_t))) = 1/2 + \nu$$

Noise parameter

Estimate correct Sign: Resampling

Algorithm 3 sign-recovery$(\mathbf{x}, \mathbf{y}, \delta)$

Input: Dueling pair: $(\mathbf{x}, \mathbf{y})$. Desired confidence $\delta \in [0, 1]$. **Initialize** $w \leftarrow 0$
for t = 1,2, . . . **do**
  Play $(\mathbf{x}, \mathbf{y})$.
  Receive $o_t \leftarrow$ noisy-preference$(\mathbf{1}(f(\mathbf{x}) < f(\mathbf{y})))$
  Update $w \leftarrow w + o_t$, $p_t(\mathbf{x}, \mathbf{y}) \leftarrow \frac{w}{t}$.
  $\texttt{conf}_t := \sqrt{\frac{\log(8t^2/\delta)}{2t}}$
  $l_t(\mathbf{x}, \mathbf{y}) := p_t(\mathbf{x}, \mathbf{y}) - \texttt{conf}_t$
  $l_t(\mathbf{y}, \mathbf{x}) := 1 - p_t(\mathbf{x}, \mathbf{y}) - \texttt{conf}_t$
  if either $l_t(\mathbf{x}, \mathbf{y}) > 1/2$ **or** $l_t(\mathbf{y}, \mathbf{x}) > 1/2$: Break.
end for
Compute $o \leftarrow \begin{cases} 1 & \text{if } l_t(\mathbf{x}, \mathbf{y}) > 1/2 \\ 0 & \text{otherwise} \end{cases}$
Return o

Convergence rates:

- Smooth functions

$$O\left(\frac{d\beta D}{\epsilon \nu^2} \log \frac{d\beta D}{\epsilon \nu^2 \delta}\right)$$

- Better convergence with Strong convexity

$$O\left(\frac{d\beta}{\nu^2 \alpha}\left(\log_2\left(\frac{\alpha}{\epsilon}\right) + D\right)\right) \log \frac{d\beta D \log(\alpha/\epsilon)}{\nu^2 \delta}$$

$$Pr\left(o = \mathbf{1}(f(\mathbf{x}) - f(\mathbf{y})) \; and \; t = O\left(\frac{1}{\nu^2} \log \frac{1}{\nu^2 \delta}\right)\right) > 1 - \frac{\delta}{2}$$

Recovers correct sign w.h.p.

In a nutshell:

- Problem formulation: Dueling convex optimization
- Impossibility result for non-stationary setup!
- Proposed algorithms: α-Strongly convex + β-Smooth functions
- Robust to noisy preferences

Future Works:

- General preference feedback? Extension to subsetwise feedback?
- Regret guarantees?
- Understanding fundamental performance limits of optimization with dueling feedback?

Thanks!

Questions @ aasa@microsoft.com