No-Regret Learning in Convex Games

Geoff Gordon, Amy Greenwald, Casey Marks

Background

- No-regret algorithms can learn well, even in adversarial environments
- Seems useful for learning in a repeated game
- Well-known result: no-regret learners reach minimax equilibrium in zero-sum normal-form games
- More recent: no-*internal*-regret learners reach *correlated* equilibrium in general-sum normal-form games

Normal-form picture

Regret

no external	$\Leftarrow$	no internal
$\Updownarrow$		$\Updownarrow$
coarse correlated	$\Leftarrow$	correlated

Equilibrium

This talk

- What about games with more structure?
  - E.g., extensive-form games
  - E.g., classification or regression
  - In general, *convex games*

Convex games (known)

Regret

no external $\Leftarrow$		$\Leftarrow$ no swap
$\Updownarrow$		$\Updownarrow$
coarse cor. $\Leftarrow$		$\Leftarrow$ correlated

Equilibrium

Convex games (contrib)

Regret

$$
\begin{array}{ccccccccc}
\text{no external} & \Leftarrow & \text{no EF} & \Leftarrow & \text{no linear} & \Leftarrow & \text{no FE} & \Leftarrow & \text{no swap} \\
\Updownarrow & & \Updownarrow & & \Updownarrow & & \Updownarrow & & \Updownarrow \\
\text{coarse cor.} & \Leftarrow & \text{EFCE} & \Leftarrow & \text{“linear”} & \Leftarrow & \text{“FE”} & \Leftarrow & \text{correlated}
\end{array}
$$

Equilibrium

Convex games (contrib)

Regret

no external	$\Leftarrow$	no EF	$\Leftarrow$	no linear	$\Leftarrow$	no FE	$\Leftarrow$	no swap
$\Updownarrow$		$\Updownarrow$		$\Updownarrow$		$\Updownarrow$		$\Updownarrow$
coarse cor.	$\Leftarrow$	EFCE	$\Leftarrow$	“linear”	$\Leftarrow$	“FE”	$\Leftarrow$	correlated

Equilibrium

all equivalent in matrix games

Convex games (contrib)

Regret

no external	$\Leftarrow$	no EF	$\Leftarrow$	no linear	$\Leftarrow$	no FE	$\Leftarrow$	no swap
$\Updownarrow$		$\Updownarrow$		$\Updownarrow$		$\Updownarrow$		$\Updownarrow$
coarse cor.	$\Leftarrow$	EFCE	$\Leftarrow$	“linear”	$\Leftarrow$	“FE”	$\Leftarrow$	correlated

efficient algorithms

Equilibrium

Convex games (contrib)

Regret

FE $\Rightarrow$ swap if "careful"

no external	$\Leftarrow$	no EF	$\Leftarrow$	no linear	$\Leftarrow$	no FE	$\Leftarrow$	no swap
$\Updownarrow$		$\Updownarrow$		$\Updownarrow$		$\Updownarrow$		$\Updownarrow$
coarse cor.	$\Leftarrow$	EFCE	$\Leftarrow$	"linear"	$\Leftarrow$	"FE"	$\Leftarrow$	correlated

efficient algorithms

Equilibrium

Outline

- Convex games and OCPs
- Regret and Φ-regret
- Algorithm
- Making it fast
- Summary & related work

Convex games

- Generalization of normal-form games, extensive-form games, …
- N players, convex action sets A_i
- Loss(i) = $c(a_{-i}) \cdot a_i$
- E.g., A_i = simplex: normal-form game

Online convex programs

- Single agent's view of repeated convex game
- Convex feasible region A
- Alternately choose action, see cost vector
  - $a_1 \to c_1 \to a_2 \to c_2 \to \ldots$
- Total cost = $a_1 \cdot c_1 + a_2 \cdot c_2 + \ldots$
- If A = simplex: "expert advice" problem

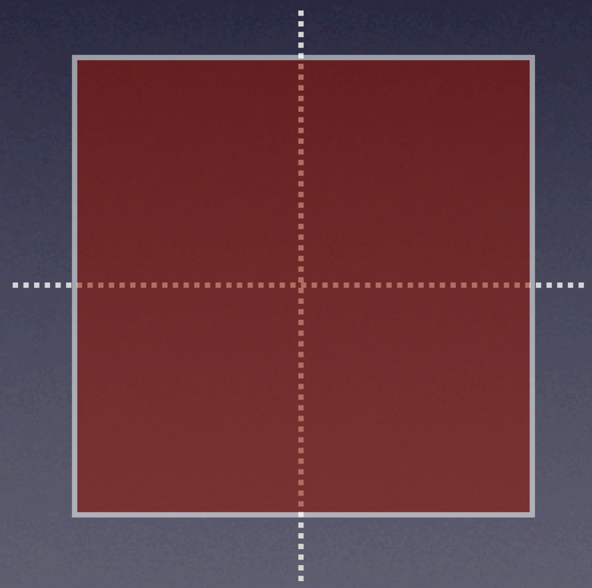
External regret

$$\rho_t = \sum c_t \cdot a_t - \min_{a \in A} \sum c_t \cdot a$$

= (observed cost) – (cost of best action post-hoc)

Action transformations

- Function $\varphi: A \mapsto A$
  - maps OCP's feasible region to itself
- E.g., linear fn mapping unit square into itself

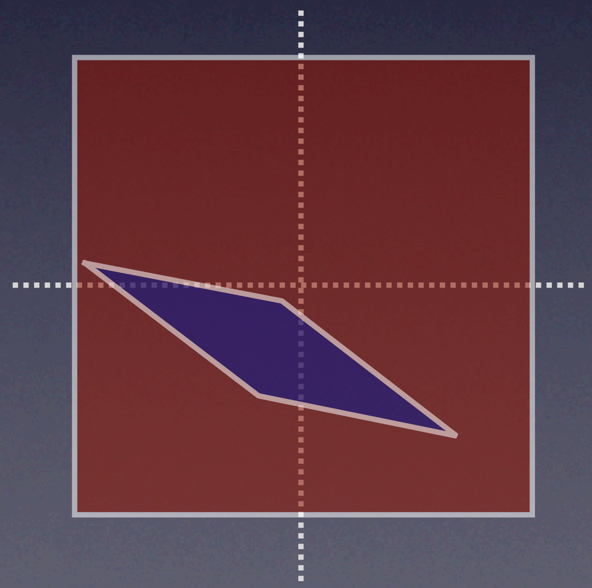


$$\begin{pmatrix} \pm p & \pm(1-p) \\ \pm q & \pm(1-q) \end{pmatrix}$$

$$p, q \in [0,1]$$

Action transformations

- Function φ: $A \mapsto A$
  - maps OCP's feasible region to itself
- E.g., linear fn mapping unit square into itself



$$\begin{pmatrix} \pm p & \pm(1-p) \\ \pm q & \pm(1-q) \end{pmatrix}$$

$p, q \in [0,1]$

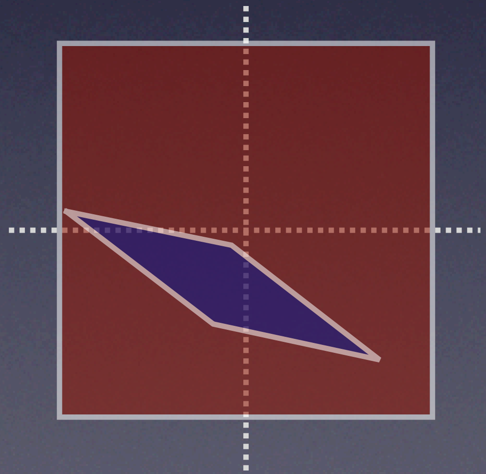
Φ-regret

$$\rho_t = \sum c_t \cdot a_t - \min_{\varphi \in \Phi} \sum c_t \cdot \varphi(a_t)$$

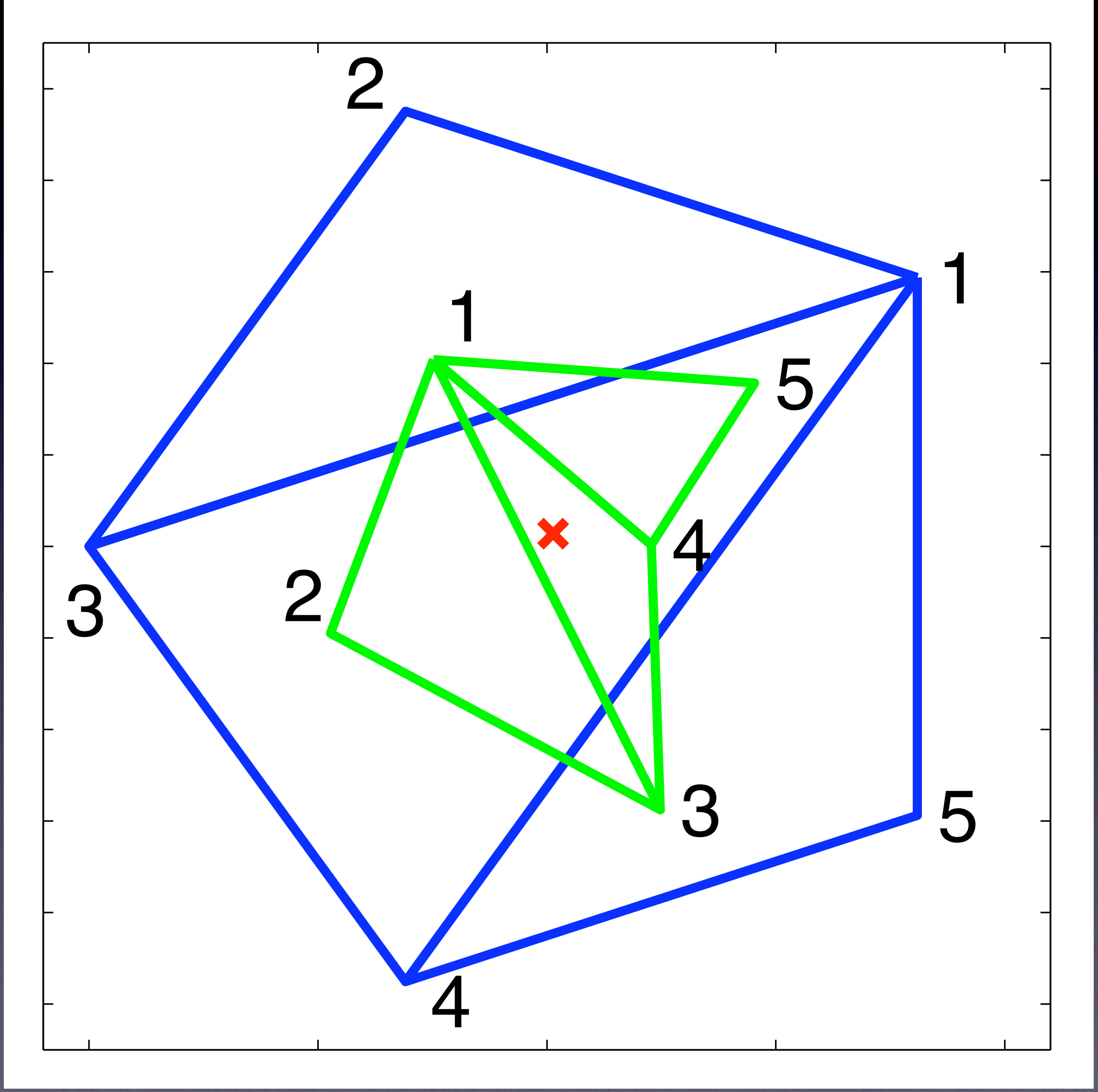
= (observed cost) – (cost of best transformation post-hoc)

Regret examples

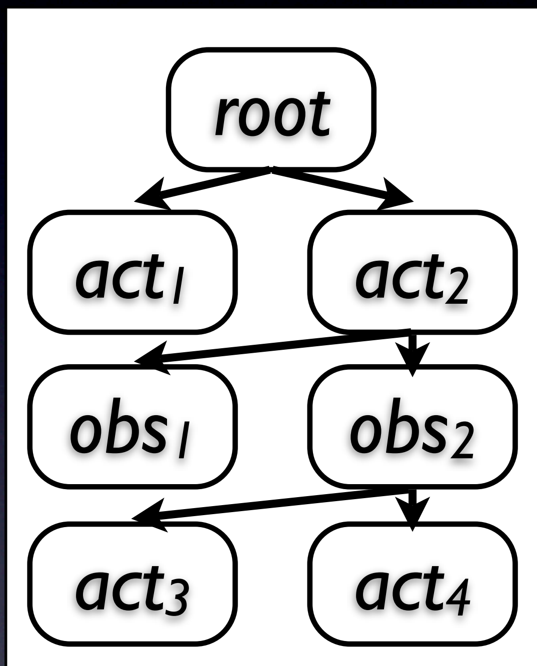
- Φ = constant transformations (= external)
- Φ = all measurable transformations (= swap)
- Φ = linear transformations



Φ = finite element



Φ = EF transformations



$$\begin{bmatrix} 1-a & d & 0 & 0 \\ a & 1-d & 0 & 0 \\ b & 0 & 1-d-e & f \\ c & 0 & e & 1-d-f \end{bmatrix}$$

- Extensive-form correlated equilibrium
- No time for these—come to the poster!

Algorithm idea

Algorithm idea

- Reduce the no-Φ-regret problem for *A* to a no-external-regret problem on a more-complicated feasible region

Algorithm idea

- Reduce the no-Φ-regret problem for A to a no-external-regret problem on a more-complicated feasible region
- Namely, Φ, considered as a subset of a vector space

Algorithm

$$a_1 \to c_1 \to a_2 \to c_2 \to a_3 \to c_3 \quad \ldots$$

- Get first play φ_1 from NER subroutine
- For t = 1, 2, …
  - Find fixed point a_t of φ_t, play a_t
  - Observe c_t
  - Construct $m_t(\varphi) = c_t \cdot \varphi(a_t)$
  - Give m_t, φ_t to NER subroutine, get φ_{t+1}

Algorithm

$$a_1 \rightarrow c_1 \rightarrow a_2 \rightarrow c_2 \rightarrow a_3 \rightarrow c_3 \quad \ldots$$

$$\varphi_1 \quad m_1 \quad \varphi_2 \quad m_2 \quad \varphi_3 \quad m_3 \quad \ldots$$

- Get first play φ_1 from NER subroutine
- For t = 1, 2, …
  - Find fixed point a_t of φ_t, play a_t
  - Observe c_t
  - Construct $m_t(\varphi) = c_t \cdot \varphi(a_t)$
  - Give m_t, φ_t to NER subroutine, get φ_{t+1}

Algorithm

$$\begin{array}{cccccccccc}
a_1 & \rightarrow & c_1 & a_2 & \rightarrow & c_2 & a_3 & \rightarrow & c_3 & \ldots \\
\uparrow & & \downarrow & \uparrow & & \downarrow & \uparrow & & \downarrow & \\
\varphi_1 & & m_1 & \rightarrow \varphi_2 & & m_2 & \rightarrow \varphi_3 & & m_3 & \ldots
\end{array}$$

- Get first play φ_1 from NER subroutine
- For t = 1, 2, …
  - Find fixed point a_t of φ_t, play a_t
  - Observe c_t
  - Construct $m_t(\varphi) = c_t \cdot \varphi(a_t)$
  - Give m_t, φ_t to NER subroutine, get φ_{t+1}

Algorithm

$$\begin{array}{cccccccccc}
a_1 & \to & c_1 & & a_2 & \to & c_2 & & a_3 \to c_3 & \ldots \\
\uparrow & & \downarrow & & \uparrow & & \downarrow & & \uparrow \quad \downarrow & \\
\varphi_1 & & m_1 & \to & \varphi_2 & & m_2 & \to & \varphi_3 \quad m_3 & \ldots
\end{array}$$

- Get first play φ_1 from NER subroutine
- For t = 1, 2, …
  - Find fixed point a_t of φ_t, play a_t
  - Observe c_t
  - Construct $m_t(\varphi) = c_t \cdot \varphi(a_t)$
  - Give m_t, φ_t to NER subroutine, get φ_{t+1}

Theorem

- The algorithm achieves no Φ-regret
- It runs in poly time if its subroutines do

Proof of no Φ-regret

$\forall\, \varphi \in \Phi,\ \sum m_t(\varphi_t) \ \leq\ \sum m_t(\varphi) + o(T)$

$\sum c_t \cdot \varphi_t(a_t) \ \leq\ \sum c_t \cdot \varphi(a_t) + o(T)$

$\sum c_t \cdot a_t \ \leq\ \sum c_t \cdot \varphi(a_t) + o(T)$

Making it fast

- Φ may be a complex set
- Expensive to achieve no external regret
- Solution: use (half of) the kernel trick to get efficient linear representation

Half of kernel trick

- Nonlinear function $\varphi(a)$
- Represent as $\varphi(a) = M_\varphi \circ K(a)$

adjustable linear fn

fixed nonlinearity

Kernelized $\Phi \Rightarrow$ fast

- For any $\varphi \in \Phi$
  - $m_t(\varphi) = c_t \cdot M_\varphi K(a_t) = \mathrm{tr}((K(a_t)\, c_t^T)\, M_\varphi)$
- I.e., $m_t(\varphi)$ is a linear function of M_φ
- And, $\mathscr{M}$ = { feasible M_φ } is convex
- Standard OCP $\Rightarrow$ standard OCP algorithms

Theorem

- If we play only vertices of mesh, and achieve no FE-regret, we also have no swap regret
- We can extend our algorithm by "warping" its plays to nearby vertices, while preserving regret guarantees

$\Rightarrow$ efficient algorithm to learn CE

Summary

Regret FE $\Rightarrow$ swap if "careful"

no external	$\Leftarrow$	no EF	$\Leftarrow$	no linear	$\Leftarrow$	no FE	$\Leftarrow$	no swap
$\Updownarrow$		$\Updownarrow$		$\Updownarrow$		$\Updownarrow$		$\Updownarrow$
coarse cor.	$\Leftarrow$	EFCE	$\Leftarrow$	"linear"	$\Leftarrow$	"FE"	$\Leftarrow$	correlated

efficient algorithms **Equilibrium**

Summary

Regret

FE ⇒ swap if "careful"

no external ⇐ no EF ⇐ no linear ⇐ no FE ⇐ no swap

⇕ ⇕ ⇕ ⇕ ⇕

coarse cor. ⇐ EFCE ⇐ "linear" ⇐ "FE" ⇐ correlated

efficient algorithm

Equilibrium

1st efficient no-EF-regret learner

Summary

Regret

FE $\Rightarrow$ swap if "careful"

no external $\Leftarrow$ no EF $\Leftarrow$ no linear $\Leftarrow$ no FE $\Leftarrow$ no swap

$\Updownarrow$ $\Updownarrow$ $\Updownarrow$ $\Updownarrow$ $\Updownarrow$

coarse cor. $\Leftarrow$ EFCE $\Leftarrow$ "linear" $\Leftarrow$ "FE" $\Leftarrow$ correlated

efficient algorithm

Equilibrium

1st efficient no-EF-regret learner

exponentially-more-efficient CE

Related work

- Blum & Mansour [2005]
  - fixed-point trick for normal-form learners
- Stoltz & Lugosi [2007]
  - nice analysis of Φ-regret in OCPs
  - first algorithm for Φ = swap

Related work

- Hazan & Kale [simultaneous w/ us]
  - proposed non-kernelized version of algorithm
  - nice reduction between finding fixed points and achieving no Phi-regret
  - lack of kernel trick slows implementation by arbitrarily large factor

Thanks!