No-Regret Learning in Convex Games

Geoff Gordon, Amy Greenwald, Casey Marks, and Martin Zinkevich

Introduction

- The connection between regret and equilibria is well understood in matrix games.
- Most research is focused on external and internal/swap regret.
- Corresponding learning algorithms learn coarse correlated and correlated equilibria, respectively.

Introduction

- We explore this connection in convex games.
- We find a much richer set of varieties of regret.
- In matrix games, elements of this richer set are all equivalent (insofar as we can apply them to matrix games).
- In convex games, we show they are distinct.

Introduction

- We present a general schema for algorithms that minimize regret for this richer set.
- We show how to implement it efficiently in two interesting cases.
- One of these cases leads to an efficient algorithm for learning correlated equilibria in repeated convex games.

Overview

- Games, Regret, and Equilibria
- Minimizing Finite-Element Regret

Games, Regret, and Equilibiria

One-Shot Game

A one-shot game $\Gamma = \left(N, \{A_i\}_{i=1}^{N}, \{\mathcal{R}_i\}_{i=1}^{N}, \{r_i\}_{i=1}^{N}\right)$, where

- $N \geq 1$ is the (finite) number of players,
- A_i is the set of actions available to player i,
- $\mathcal{R}_i$ is the set of rewards available to player i, and
- $r_i : (\bigotimes_j A_j) \to \mathcal{R}_i$ is the reward function for player i,

so that if each player j “plays” action a_j, player i gets reward $r_i(a_1, a_2, \ldots, a_N)$.

Kinds of Games

- Matrix game: each A_i is a finite set
- Experts game: each A_i a simplex (set of distributions over a finite set)
- Convex game: each A_i is a convex set and each r_i is linear in its ith argument
- Corner game: play only corners

Transformations

A transformation is a (measurable) mapping from A to itself ($\phi : A \rightarrow A$)

- Φ_{SWAP}: the set of all transformations
- Φ_{FE}: will be defined later in the talk
- Φ_{LIN}: the set of linear transformations
- Φ_{EXT}: the set of constant (“external”) transformations

In general convex games,

$$\Phi_{\text{EXT}} \subset \Phi_{\text{LIN}} \subset \Phi_{\text{F-E}} \subset \Phi_{\text{SWAP}}$$

In experts games,

$$\Phi_{\text{EXT}} \subset \Phi_{\text{LIN}} = \Phi_{\text{F-E}} \subset \Phi_{\text{SWAP}}$$

Φ-Equilibria

Definition 1 *Given a game and a collection of sets of transformations,* $\langle \Phi_i \rangle_{i \in N}$*, a probability distribution* q *over* $\mathcal{A}$ *is a* $\{\Phi_i\}$**-equilibrium** *if*

$$\mathbb{E}\left[r_i(\phi(a_i), a_{\neg i}) - r_i(a)\right] \leq 0 \quad \forall i \in N,\ \forall \phi \in \Phi_i$$

Φ-Equilibria

If each Φ_i uses the same set of transformations,

- Φ_{SWAP}-equilibria = correlated equilibria
- Φ_{EXT}-equilibria = coarse correlated equilibria

In convex games,

$$\Phi_{\text{EXT}}(\mathrm{CCE}) \subset \Phi_{\text{LIN}} \subset \Phi_{\text{F-E}} \subset \Phi_{\text{SWAP}}(\mathrm{CE})$$

In experts games,

$$\Phi_{\text{EXT}}(\mathrm{CCE}) \subset \Phi_{\text{LIN}} = \Phi_{\text{F-E}} \subset \Phi_{\text{SWAP}}(\mathrm{CE})$$

Repeated Games

Given a one-shot game Γ, we define a repeated game Γ^∞.

In each sequential round t,

1. each player i chooses action $a_i^{(t)}$
2. each player observes the actions of all other players $a_j^{(t)}$
3. each player receives payoff $r_i\left(a_1^{(t)}, a_2^{(t)}, \ldots, a_N^{(t)}\right)$

Regret

Given a player i and a transformation ϕ for that player, at each round t the **instantaneous regret** is calculated with respect to the joint action played at that round:

$$\rho_{i,\phi}^{(t)} = r\left(\phi\left(a_i^{(t)}\right), a_{-i}^{(t)}\right) - r\left(a^{(t)}\right) \quad (1)$$

If a player's algorithm guarantees that

$$\sup_{\phi \in \Phi} \frac{1}{T} \sum_{t=1}^{T} \rho_{i,\phi}^{(t)} \to (-\infty, 0]$$

with probability 1, then we say that it is **no-Φ-regret**

No Regret Properties

In convex games,

$$\text{(CCE)}\ \Phi_{\text{EXT}} \Leftarrow \Phi_{\text{LIN}} \Leftarrow \Phi_{\text{F-E}} \Leftarrow \Phi_{\text{SWAP}}\ \text{(CE)}$$

In experts games,

$$\text{(CCE)}\ \Phi_{\text{EXT}} \Leftarrow \Phi_{\text{LIN}} \Leftrightarrow \Phi_{\text{F-E}} \Leftarrow \Phi_{\text{SWAP}}\ \text{(CE)}$$

Convergence

Theorem 2 (Foster and Vohra) *In a repeated matrix game, if all players play no-swap-regret algorithms, then the empirical distribution of play converges to the set of correlated equilibria with probability 1.*

Stoltz and Lugosi prove the existence of an algorithm that minimizes swap regret and ensures convergence to correlated equilibria in repeated convex games. However, they do not explicitly construct such an algorithm. Constructing an algorithm according to their proof of existence would be prohibitively expensive (run time would grow unboundedly with t).

Corner Games

Definition 3 *A corner game is a convex game with each player's action set restricted to the corners of its feasible region.*

Proposition 4 *A CE of the corner game is a CE of the convex game.*

Proposition 5 *For all correlated equilibria in the convex game, there exists a payoff-equivalent correlated equilibrium in the corner game.*

CE of Convex Games

Theorem 6 (GGMZ) *If, in a repeated convex game, each agent*

- *plays only corners and*
- *and uses an algorithm that achieves no-swap-regret for the corner game,*

then the empirical distribution of play converges to the set of correlated equilibria of the convex game with probability 1.

No Regret Properties

In convex games (corners only),

$$(\mathrm{CCE})\ \Phi_{\mathsf{EXT}} \Leftarrow \Phi_{\mathsf{LIN}} \Leftarrow \Phi_{\mathsf{F\text{-}E}} \Leftrightarrow \Phi_{\mathsf{SWAP}}\ (\mathrm{CE})$$

In experts games (corners only),

$$(\mathrm{CCE})\ \Phi_{\mathsf{EXT}} \Leftarrow \Phi_{\mathsf{LIN}} \Leftrightarrow \Phi_{\mathsf{F\text{-}E}} \Leftrightarrow \Phi_{\mathsf{SWAP}}\ (\mathrm{CE})$$

Online Convex Programming

Online Convex Programming

- convex compact action space $A \in \mathbb{R}^d$
  (for convenience, we add an extra dimension whose value is always 1)
- bounded loss vector space $L \subseteq \mathbb{R}^d$
- The net loss for an action is given by a dot product.

Special Case: Experts Problem

- feasible region is probability simplex in d dimensions

Regret

Given a set of transformations Φ, an algorithm's Φ-regret is

$$\rho_t^{\Phi} = \sup_{\phi \in \Phi} \sum_{\tau=1}^{t} (l_\tau \cdot a_\tau - l_\tau \cdot \phi(a_t))$$

and is "no-Φ-regret" if

$$\sum_{\tau=1}^{t} l_\tau \cdot a_\tau \leq \sum_{\tau=1}^{t} l_\tau \cdot \phi(a_\tau) + g(t, A, L, \Phi) \qquad \forall \phi \in \Phi, \forall t \geq 1$$

where $g(t, A, L, \Phi)$ is $o(t)$ for any fixed A, L, and Φ.

Goal

- Known: Algorithms that minimize external regret in OCPs, e.g., Lagrangian Hedging (Gordon06), GIGA (Zinkevich03)
- Goal: Derive an algorithm that minimizes finite-element-regret in OCPs.

Key Idea #1

Key Idea #1

Key Idea #1: represent Φ as the composition of a **fixed** nonlinear continuous “feature” function with an **adjustable** linear function

$$\Phi = \{\phi_C \mid C \in \mathcal{C}\}$$

$$\phi_C(a) = CB(a)$$

Here B is our feature function, which maps the feasible region $A \subset \mathbb{R}^d$ to a p-dimensional feature space, while $\mathcal{C}$ is a set of $d \times p$ matrices which map the feature space back down to the d-dimensional feasible region. (Often, $p \gg d$.) We assume B is continuous.

Linear Transformations

Choose $B =$ identity, so $\phi_C = C$.

Example: any matrix that maps A into itself
e.g., if A is a simplex, the set of linear transformations can be represented by the set of stochastic matrices

Barycentric Coordinates

- Barycentric coordinate/feature mapping on polyhedral feasible region A.
- B is a fixed nonlinear function that encodes a triangulation/tessellation.
- $B(a)$ is a point in higher-dimensional space called the Barycentric coordinate space.

Barycentric Coordinates

Formally,

- choose a triangulation
- choose a numbering from corners of the polyhedron to dimensions in the Barycentric coordinate space $B(A)$

Intuitively, $B(a)$ tells you what triangle a is in, and where in that triangle:

- i.e., which corners and what their weights are
- i.e., $d+1$ coordinates in $\mathbb{R}^n$ that are nonzero, and $d+1$ weights summing to 1

Finite-element Transformations

Given a B, each transformation corresponds to a linear mapping from $B(A)$ back down to A.

Consider mapping the corners of the square $1 \mapsto 2 \mapsto 3 \mapsto 4 \mapsto 1$ as follows:

$$C = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 \end{pmatrix}$$

So, each column of a matrix lists the coordinates to which the corresponding corner of the feasible region is mapped.

Intuitively, each transformation corresponds to choosing a point inside the (polyhedral) feasible region for each corner to map to; everything else follows, according to B.

Key Idea #2

Algorithm

Given: a subroutine that minimizes external regret

Key Idea #2: Instead of minimizing Φ-regret on $A \subseteq \mathbb{R}^d$ directly, we minimize external regret on $\mathcal{C} \subseteq \mathbb{R}^{d \times p}$.

Algorithm

High-Level Recipe: for each time t,

1. play a real action $a_t \in A$ and receives a real loss vector $l_t \in A$
2. construct a fictitious loss vector in $\mathbb{R}^{d \times p}$
3. send the loss vector to an external-regret minimizing subroutine
4. subroutine constructs a fictitious action in $\mathcal{C}$
5. use that fictitious action to construct a real action $a_{t+1} \in A$

Algorithm

Questions

- Q1: How do we construct the fictitious loss vector?
- Q2: How do we construct the real action?
- Q3: How do we efficiently minimize external regret in a high dimensional space?

Q1

Q1: how do we construct the fictitious loss vector?
A1: based on the real action a_t and real loss vector l_t:

$$m_t = l_t B(a_t)^T$$

What's going on here?

- Dotting m_t with a transformation/action in the higher-dimensional space gives you the loss associated with that transformation.
- You can interpret it as the loss you would get by performing that transformation on the real action: i.e.,

$$m_t \cdot C = \mathrm{tr}(B(a_t){l_t}^{\mathrm{T}}C) = \mathrm{tr}({l_t}^{\mathrm{T}}CB(a_t)) = {l_t}^{\mathrm{T}}CB(a_t)$$

Q1, Example

Example: loss vector is $\langle 2\ 4 \rangle$

play a corner: e.g., play $\langle 0\ 1\ 0\ 0 \rangle$

$$\begin{pmatrix} 0 & 2 & 0 & 0 \\ 0 & 4 & 0 & 0 \end{pmatrix}$$

play a non-corner: e.g., $\langle 0\ \frac{1}{2}\ \frac{1}{2}\ 0 \rangle$

$$\begin{pmatrix} 0 & 1 & 1 & 0 \\ 0 & 2 & 2 & 0 \end{pmatrix}$$

Q2

Q2: how do we construct the real action $a_{t+1} \in A$?

A2: based on the fictitious action, C_t

let a_{t+1} be an arbitrary fixed point of ϕ_{C_t}, where
$\phi_{C_t}(a) = C_t B(a)$

Theorem

Theorem 7 *For any convex compact feasible region A and bounded loss vector space L, and for any set of transformations $\Phi : A \mapsto A$, each one represented as the composition of a fixed nonlinear continuous feature function and an adjustable linear function, $\mathcal{A}$ achieves no Φ-regret whenever its subroutine $\mathcal{A}'$ achieves no external regret.*

Proof

Note that:

$$\sum_{t=1}^{T} m_t \cdot C_t \leq \sum_{t=1}^{T} m_t \cdot C + f(T, \mathcal{C}, L\mathcal{B}^{\mathrm{T}}) \qquad \forall C \in \mathcal{C}$$

where $L\mathcal{B}^{\mathrm{T}} = \{lb^{\mathrm{T}} \mid l \in L, b \in \mathcal{B}\}$, and f is sublinear in T.
So

$$\sum_{t=1}^{T} {l_t}^{\mathrm{T}} C_t B(a_t) \leq \sum_{t=1}^{T} {l_t}^{\mathrm{T}} C B(a_t) + f(T, \mathcal{C}, L\mathcal{B}^{\mathrm{T}}) \qquad \forall C \in \mathcal{C}$$

But, since $C_t B(a_t) = \phi_{C_t}(a_t) = a_t$, and since each $\phi \in \Phi$ can be represented as $\phi(a) = CB(a)$ with $C \in \mathcal{C}$, this implies

$$\sum_{t=1}^{T} {l_t}^{\mathrm{T}} a_t \leq \sum_{t=1}^{T} {l_t}^{\mathrm{T}} \phi(a_t) + f(T, \mathcal{C}, L\mathcal{B}^{\mathrm{T}}) \qquad \forall \phi \in \Phi$$

which is exactly the required no-Φ-regret guarantee.

Q3

Q3: how do we efficiently minimize external regret in a high dimensional space?

A3: for finite-element we can factor $\mathcal{C}$ (each corner's destination is independent) so we can separately run n copies of any NER algorithm for A. Each one is typically, $O(d^3)$, so the whole thing is $O(nd^3)$.

(Our approach is related to Blum and Mansour, 2005.)

Why do I care?

- We just showed how to efficiently minimize finite-element regret.
- Now we will remind you why this is worthwhile.

Back to Convex Games

Remark 8 *In the corner game, each player faces an ODP: an OCP in which only corners can be played.*

Lemma 9 *Minimizing finite element regret in an OCP, while playing only corners, minimizes swap regret in the corresponding ODP.*

PROOF: Every swap transformation in the ODP can be expressed as a finite element transformation in the OCP. □

Theorem 7 (GGMZ, again) *If, in a repeated convex game, each agent plays only corners and uses a finite-element regret-minimizing algorithm, then the empirical distribution of play converges to the set of correlated equilibria of the convex game with probability 1.*

Take-away Message

We have developed what is to our knowledge the first efficient algorithm for learning correlated equilibria in convex games.

(Gordon, Greenwald, Marks, and Zinkevich. No-regret learning in convex games. Technical Report CS-07-10, Brown University, Department of Computer Science, October 2007.)

Final Notes

- Extensive-form games can be expressed efficiently as convex games.
- Open question: What set of transformations corresponds to extensive-form correlated equilibria?