

Deterministic Calibration and Nash Equilibrium

Sham Kakade

University of Pennsylvania

With Dean Foster (tomorrow at 10:30!)

The Central Question

Can we view a Nash Equilibrium as a convergent point of a “sensible” learning process?

Few general results, short of exhaustive search.

Some Convergence Results

- Extensive list of special cases
 - zero sum
 - 2 player, 2 action
 - potential games
- “Exhaustive search” results:
 - Hypothesis Testing [Foster & Young]
 - Distributed Search [Hart & MasColell]

Outline

- Background on learning in games
- Online Prediction:
 - the calibration task
 - Theorem 1: There exists a deterministic weakly calibrated algorithm.
- Game Theory:
 - Learning using the “public” algorithm
 - Game Theory: If players use the same weakly calibrated algorithm, the joint frequency of play converges into the set of convex combinations of NE.

A Learning Process

1. Players make predictions of other players using:
 - the joint history of play
 - private utility functions
 2. Players then take best responses.
- - Fictitious play is the most well studied example
 - What constitutes “good” predictions?

Calibration

"Suppose in a long (conceptually infinite) sequence of weather forecasts, we look at all those days for which the forecast probability of precipitation was, say, close to some given value p and ... determine the long run proportion f of such days on which the forecast event (rain) in fact occurred. ... if $f=p$ the forecaster may be termed well calibrated." Dawid [1982]

not stringent condition: 0 1 0 1 0 1 0 ...

Is calibration always possible?

- Yes, with private randomization [Foster & Vohra]
- But how does the forecaster announce his predictions?
-
- **Theorem:** [Foster & Vohra] If players make calibrated predictions in the learning process, then convergence is to correlated equilibria.

Outline

- Background on learning in games
- Online Prediction:
 - the calibration task
 - Theorem 1: There exists a deterministic weakly calibrated algorithm.
- Game Theory:
 - Learning using the “public” algorithm
 - Game Theory: If players use the same weakly calibrated algorithm, the joint frequency of play converges into the set of convex combinations of NE.

The Online Prediction Setting

- Finite Outcome Space $\{1, 2, \dots, m\}$
- Sequence of Outcomes X_1, X_2, \dots where X_t is of the form $(0, 0, 1, 0, 0)$
- The empirical frequency at time t is $\frac{1}{t} \sum_{i=1}^t X_i$
- A forecasting method F maps histories to probability forecasts. At time t ,

$$f_t = F(X_1, X_2, \dots, X_{t-1})$$

Calibration Error

- $I\{f_t \approx p\}$ is the indicator function which is 1 iff f_t is ε -close to p , and 0 otherwise.
- For a sequence X , the calibration error is
- $$R_T(p, X) = \frac{1}{T} \sum_{t=1}^T I\{f_t \approx p\} (X_t - f_t)$$
- This is an "internal" regret (directional).
- F is calibrated if $\forall X, \forall p, R_T(p, X) \rightarrow 0$
- Deterministic F are not calibrated [Oakes; Dawid]
- Randomized F can be [Foster & Vohra]

Calibration Error

- $I\{f_t \approx p\}$ is the indicator function which is 1 iff f_t is ε -close to p , and 0 otherwise.
- For a sequence X , the calibration error is
- $$R_T(p, X) = \frac{1}{T} \sum_{t=1}^T I\{f_t \approx p\} (X_t - f_t)$$
- This is an "internal" regret (directional).
- F is calibrated if $\forall X, \forall p, R_T(p, X) \rightarrow 0$
- Deterministic F are not calibrated [Oakes; Dawid]
- Randomized F can be [Foster & Vohra]

"Weak" Calibration Error

- Replace $I\{f_t \approx p\}$ with a continuous (bounded) "test function" w

- For X , the weak calibration error w.r.t. w is

- $$R_T(w, X) = \frac{1}{T} \sum_{t=1}^T w(f_t) (X_t - f_t)$$

- Again, a notion of internal regret

- F is **weakly** calibrated if $\forall X, \forall w, R_T(w, X) \rightarrow 0$

- Definition using weak convergence of measures

A Deterministic Algorithm Exists

- **Theorem 1:** There exists a deterministic, weakly calibrated F .
- The proof is constructive.
- How do we use this algorithm to calibrated in the standard sense?

Corollary: Randomized Rounding of F

Say deterministic F makes predictions:

0.8606, 0.2387, 0.5751, 0.40051 ...

- Randomly round the forecast onto ε -spaced grid $\{p\}$
- **Corollary 1:** Almost surely, in the limit, $\forall X \forall p$
- $|R_T(p, X)| < 2\varepsilon$
- F makes “public” predictions

The Algorithm: Forecast the Fixed Point

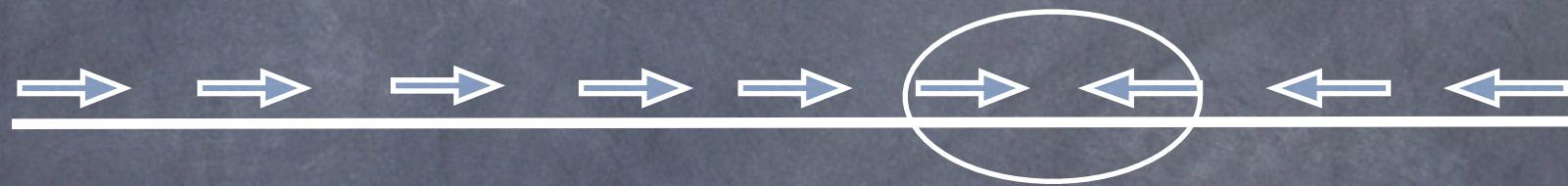
- For a grid of points $\{p\}$, each point has associated 'vector' error $R_T(p)$



- This defines a mapping: $p \mapsto p + R_T(p)$

The Algorithm: Forecast the Fixed Point

- For a grid of points $\{p\}$, each point has associated 'vector' error $R_T(p)$



- This defines a mapping: $p \mapsto p + R_T(p)$
- Forecast any fixed point of this interpolated function

The Proof

- Use Blackwell's Approachability Theorem
- "Simpler" proof than most
 - geometric property satisfies approachability condition
- Take grid size to 0 (convergence rate exponential in dim)
- Some comments:
 - this generalizes to an internal regret algorithm for the online convex programming problem

Outline

- Background on learning in games
- Online Prediction:
 - the calibration task
 - Theorem 1: There exists a deterministic weakly calibrated algorithm.
- Game Theory:
 - Learning using the “public” algorithm
 - Game Theory: If players use the same weakly calibrated algorithm, the joint frequency of play converges into the set of convex combinations of NE.

Game Theory: The Setting

- Players: $1, 2, \dots, n$
- Action set for player i : A_i
- Action spaces: $A = \prod A_j$ and $A_{-i} = \prod_{j \neq i} A_j$
- Utility function for player i : $u_i : A \rightarrow [0, 1]$
- If p is a distribution over A , then p_{-i} is a distribution over A_{-i}

Definitions of Equilibrium

p is a **Nash equilibrium** if

1. p is a product distribution
2. If a_i has positive probability under p ,
then a_i is a best response to p_{-i}
- 3.

p is a **correlated equilibrium** without condition 1 and a natural modification to 2.

A Learning Process

Predictions:

- Players know joint history X_1, X_2, \dots, X_{t-1} where each $X_k \in A$
- Players make forecast f_t which is a distribution over A_{-i}

Actions:

- Player i takes a best response to f_t

Calibration and Learning

Suppose players make calibrated predictions

Theorem: [Foster & Vohra] The frequency of empirical play converges into the set of correlated equilibria.

$$d\left(\frac{1}{T}\sum_{t=1}^T X_t, CE\right) \rightarrow 0$$

The “Public” Learning Process

Public Predictions:

- At time t , the joint forecast is $f_t = F(X_1, X_2, \dots, X_{t-1})$
- f_t is a distribution over A

Players' Actions:

- Player i marginalizes to get $(f_t)_{-i}$
- Player i then takes “continuous” ϵ -best response to $(f_t)_{-i}$ (which is a mixed strategy)
- Using randomized rounding is one scheme

Interpretation of the Process

- Players use the **same algorithm** (based on F).
- Predictions are guaranteed to be calibrated, **regardless** of how the other players act.
- Algorithm uses observable information and is uncoupled.
- Additional 'consistency' in the predictions.

Theorem 2

Assuming:

- F is any weakly calibrated
- the best response functions slowly sharpen

Then:

- The joint empirical frequency of play converges into the set of convex combinations, almost surely

- $$d\left(\frac{1}{T}\sum_{t=1}^T X_t, \text{Convex}(NE)\right) \rightarrow 0$$

- (Merging) The play and predictions are often close to some Nash Equilibrium:

The Proof

Say f used "often":

- Independence: All players act independently conditioned on of f . Call this constant product play p .
- Best Responses: All players take best responses to f .
- **Deterministic Calibration:** f is calibrated $\rightarrow f=p$
 - why we get NE not just CE

Convergence Rates and Practicality

- Come to Dean's Talk: 10:30
- Doubly exponential in #players
- What do we really **need** to check?
 - much less, in general
 - even less in structure games
- We do really **want** to check ourselves on opponents decisions?