A simple adaptive procedure leading to correlated equilibrium

Krishna Calindi

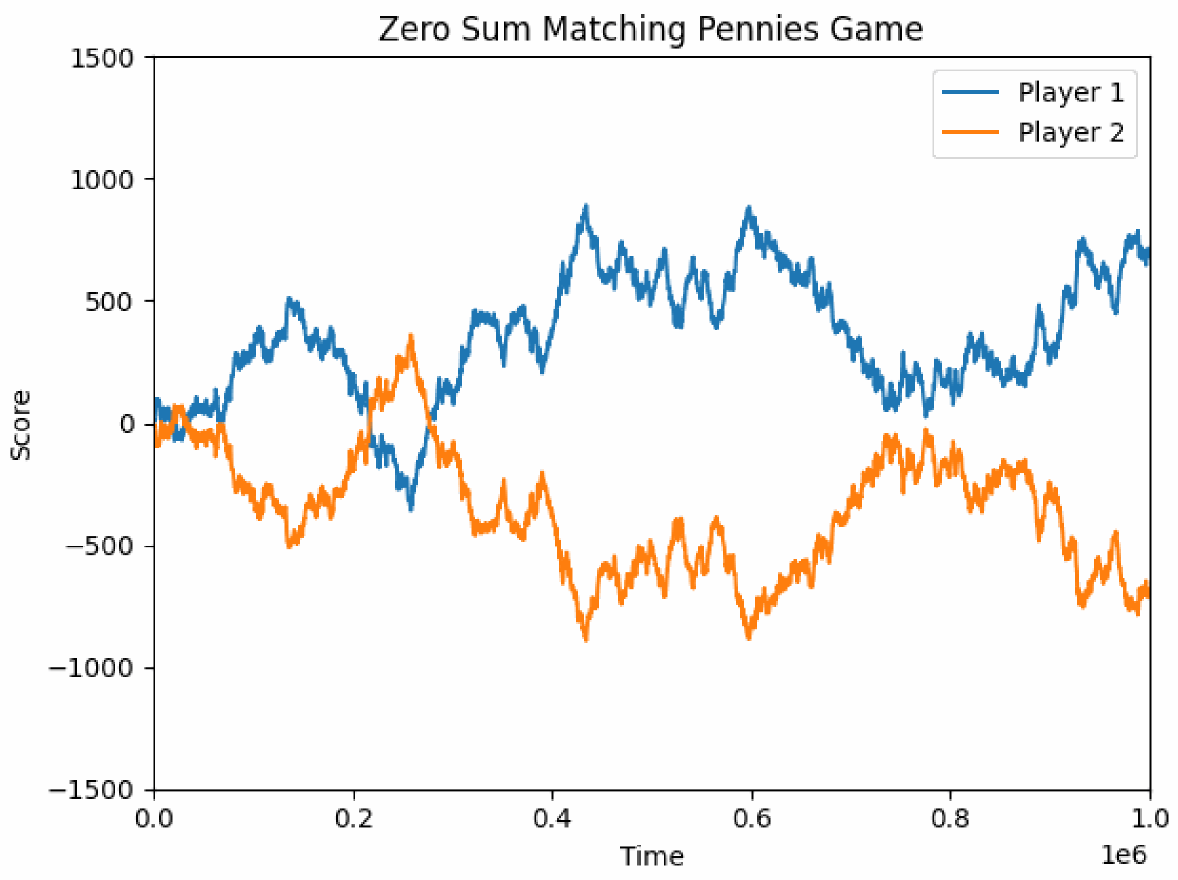
Paper

- Title: A simple adaptive procedure leading to correlated equilibrium
- Authors: Sergiu Hart and Andreu Mas-Colell
- Time: September 2000
- Journal: Econometrica
- Published by: The Econometric Society
- Vol. 68, No. 5, pp 1127-1150

Background

Aspect	Nash Equilibrium	Correlated Equilibrium
Coordination	No external coordination	Coordination via third party correlation device
Strategy	Chosen independently	Condition their actions on input signal
Outcome	Limited to stable outcomes	Can achieve diverse outcomes, may include higher payoffs
Incentive	No incentive to deviate unilaterally (cannot improve payoff unilaterally)	No incentive to deviate from recommended action

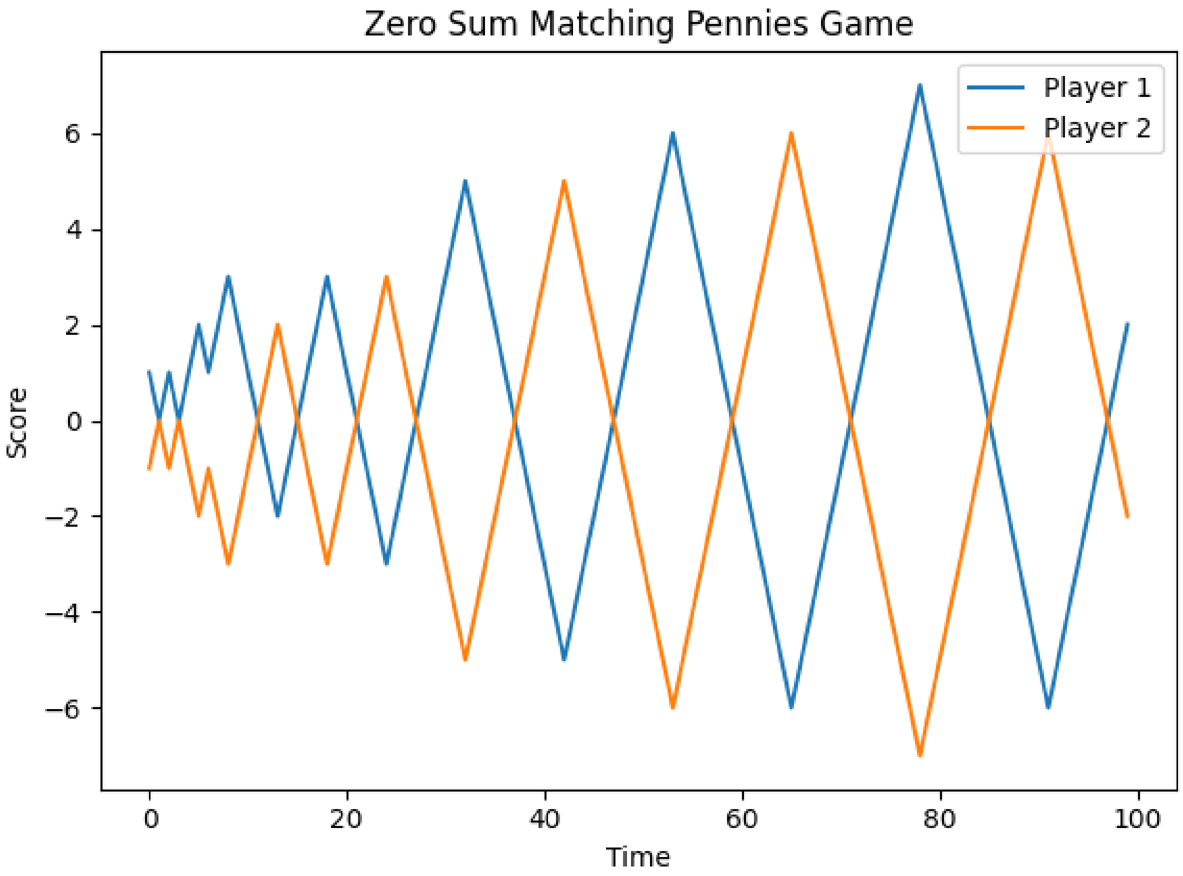
Background – In Action



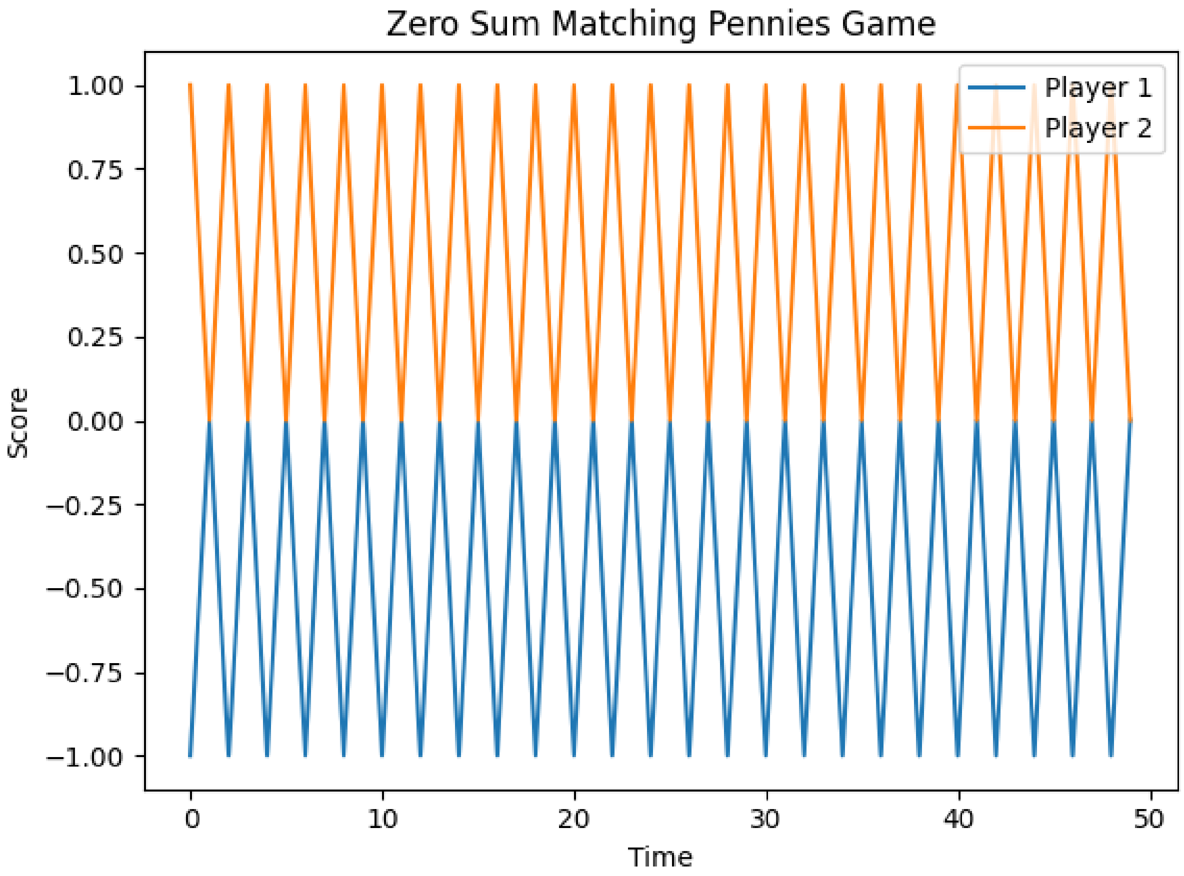
Adaptive procedures

- Adaptive procedures: used to adjust actions based on past experiences, observations or interactions with other players.
- Previous procedures:
  - Foster and Vohra (1997): myopic best response to a calibrated forecast
  - Fudenberg and Levine (1999): smooth fictitious play

Adaptive procedure – In Action



Players chose action based on most frequently played opponent action



Players chose action based on most recently played opponent action

Model – Motivation

This	Not	That
Simple	Too complex	
Past	Future	Foster and Vohra (1997)
Better	Best	fictitious play, smooth fictitious play, calibrated learning
Inertia	Motion	

Adaptive procedure

- At each period, a player can continue with same strategy or switch
- Probability of switching is based on regret
- Regret is calculated as the difference in expected and current payoff

Main Result

If every player plays according to the adaptive procedure
then the empirical distributions of play z converge
almost surely as $t \to \infty$
to the set of correlated equilibrium distributions of the game.

Model

- Game

$$\Gamma = (N, \{S_i\}_{i \in N}, \{u_i\}_{i \in N})$$

where:

- N: The set of players
- S_i: The set of strategies available to player i
- $u_i : S_i \times S_{-i} \to \mathbf{R}$: The payoff function for player i

Model

- Definition

A probability definition of ψ on S is a correlated equilibrium if

$$\sum_{\substack{s \in S \\ s_i = j}} \psi(s)[u^i(k, s^{-1}) - u^i(s)] \leq 0$$

- $i \in N$
- $j \in S^i$
- $k \in S^i$

Model

- Regret

$$W^i_\tau(j,k) := \begin{cases} u^i(k, s^{-i}_\tau) & \text{if } s^i_\tau = j \\ u^i(s^i_\tau) & \text{otherwise.} \end{cases}$$

$$D^i_t(j,k) := \frac{1}{t}\sum_{\tau=1}^{t} W^i_\tau(j,k) - \frac{1}{t}\sum_{\tau=1}^{t} u^i(s_\tau)$$

$$= \frac{1}{t}\sum_{\tau \le t : s^i_\tau = j} \left[u^i(k, s^{-i}_\tau) - u^i(s_\tau)\right]$$

$$R^i_j(j,k) := \left[D^i_j(j,k)\right]^+ = \max\left(D^i_j(j,k), 0\right)$$

- j: Current strategy for player i
- k: Possible strategy for player i
- $W^i_\tau(j,k)$: Expected payoff
- $D^i_\tau(j,k)$: Difference in average payoff
- $R^i_\tau(j,k)$: Measure of regret

Model

- Probability distribution

$$\begin{cases} p^i_{t+1} := \frac{1}{\mu} R^i_t(j,k) & \text{for all } k \neq j \\ p^i_{t+1} := 1 - \sum\limits_{k \in S^i, k \neq j} p^i_{t+1}(k) \end{cases}$$

- $p^i_{t+1} \in \triangle(S^i)$: Probability distribution of S at $t+1$
- μ: Measure of intertia

- Empirical distribution

$$z_t(s) := \frac{1}{t} |\{\tau \leq t : s_\tau = s\}|$$

- $z_t \in \triangle(S)$: Empirical distribution of N-tupes of strategies
- $z_t(s) fors \in S$: relative frequency that s has been played

Model

- Theorem:

Suppose that at every period $t+1$, player i chooses strategies according to a probability vector q_t^i that satisfies 1, then player i's regrets $R_i(j,k)$ converge to zero almost surely for every $j,k \in S$ with $j \neq k$.

- Preposition:

Let s_t be a sequence of plays ($s_t \in S$ for all t) and let $\epsilon \geq 0$. Then:

$$\limsup_{t\to\infty} R_t^i(j,k) \leq \epsilon \quad \text{for every } i \in N \text{ and every } j,k \in S \text{ with } j \neq k,$$

if and only if the sequence of empirical distributions z_t converges to the set of correlated ϵ-equilibria.

1 $\displaystyle\sum_{k\in S^i} q_t^i(k) R_t^i(k,j) = q_t^i(j) \sum_{k\in S^i} R_t^i(j,k),$

Proof

- Proposition

$$D_t^i(j,k) = \frac{1}{t} \sum_{\tau \le t : s_\tau^i = j} [u^i(k, s_\tau^{-i}) - u^i(j, s_\tau^{-i})]$$

$$= \sum_{s \in S : s^i = j} z_t(s)[u^i(k, s^{-i}) - u^i(j, s^{-i})]$$

$$\text{when } z_t \text{ converges as } z_t' \to \psi \in \triangle(S)$$

$$D_t^i(j,k) \to \sum_{s \in S : s^i = j} \psi(s) \left[u^i(k, s^{-i}) - u^i(j, s^{-i})\right]$$

Proof

- Main theorem

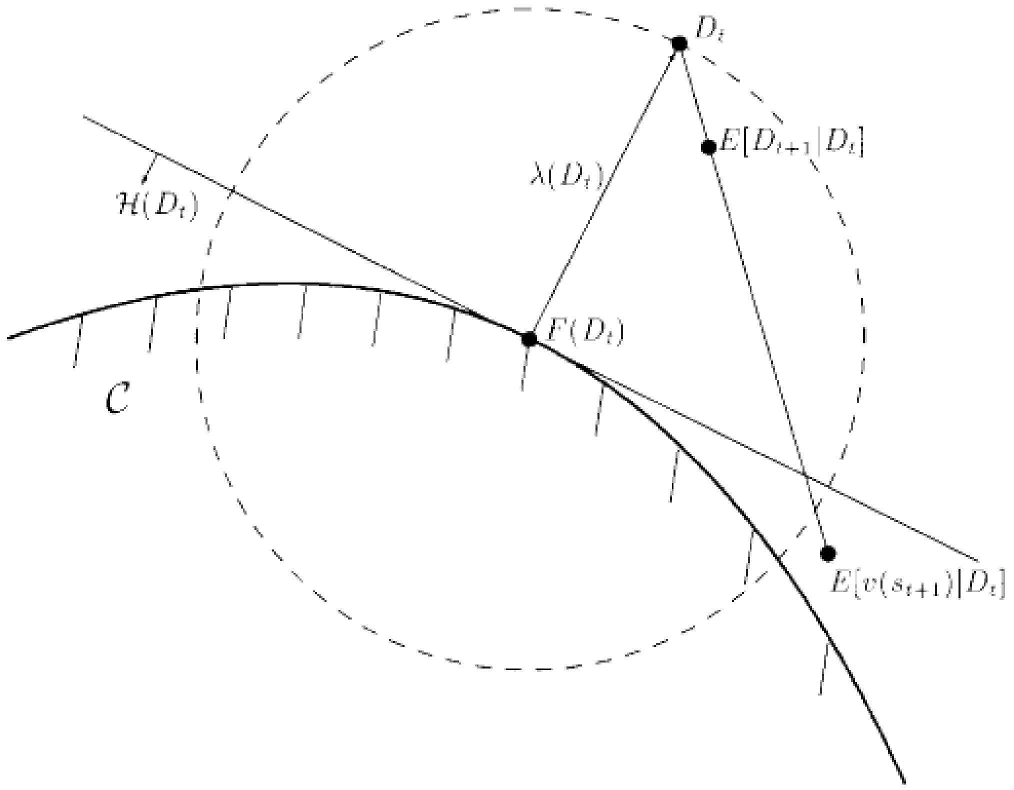


FIGURE 1.—Approaching the set $\mathscr{C}$.

- Convergence

The expectation $E[R^i_j(j,k)]$ of the regrets is of the order of $\frac{1}{\sqrt{t}}$

Assumptions

- Finite N – person strategic form game
- Game occurs in discrete time periods
- Players have access to their own payoff matrix
- Players have access to the history of all past plays
- No attempts to predict the future

Interpretation

- Regret matching is a path to equilibrium
- Non cooperative setting with simple adaptive procedure can yield stable outcomes
- Inertia as a stabilizing factor

Summary

Aspect	How
Simple	No computation of an invariant vector
Adaptive procedure	Based on regret matching
Correlated equilibrium	$t \to \infty$

Thank You

Email: kxc@tamu.edu