A Probabilistic Analysis of Prisoner’s Dilemma with an Adaptive Population

Yao Chou, Craig Wilson
Department of Electrical and Computer Engineering
Brigham Young University
Provo 84602, United States

Abstract—A probabilistic approach of analyzing successive round prisoner’s dilemma in a large populations. We introduce a unique model capable of handling multiple adaptive player strategies and how to compute expected distributions after several rounds from an initial distribution. We examine three cases studies, and their convergence properties.

I. INTRODUCTION

Prisoner’s Dilemma is not only a popular topic in daily life, but also very important in numerical analysis. This problem was a thought experiment originally proposed by economists in the 1950’s and has been the subject of countless papers and applied across several fields [1], presenting it as follows: Two members of a criminal gang are arrested and imprisoned. Each prisoner is in solitary confinement with no means of speaking to or exchanging messages with the other. The police admit they don’t have enough evidence to convict the pair on the principal charge. They plan to sentence both to a year in prison on a lesser charge. Simultaneously, the police offer each prisoner a Faustian bargain. Each prisoner is given the opportunity either to betray the other, by testifying that the other committed the crime, or to cooperate with the other by remaining silent. Here’s how it goes [1]:

1) If A and B both betray the other, each of them serves 2 years in prison
2) If A betrays but B remains silent, A will be set free and B will serve 3 years in prison (and vice versa)
3) If A and B both remain silent, both of them will only serve 1 year in prison (on the lesser charge)

The prisoner’s dilemma is a well know game that has been extensively studied in economics, political science, machine learning [2], [3] and evolutionary biology [4]. In this paper, in order to analyze this problem better, we make this problem as a mathematics model:

<table>
<thead>
<tr>
<th></th>
<th>A</th>
<th>B</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Split</td>
<td>500,500</td>
</tr>
<tr>
<td>2</td>
<td>Steal</td>
<td>1000,0</td>
</tr>
<tr>
<td>3</td>
<td></td>
<td>0,0</td>
</tr>
</tbody>
</table>

The main contributions in this paper are as follows.

1) This paper presents and further investigates the prisoner’s dilemma, and formulate this problem by mathematics, make it easy to further analyze,
2) Apply a probabilistic approach to analysis,
3) Analyze whether given distributions will converge.

II. THEORY

A. Overview

Our experiment will be played out over a series of rounds between a population of N people. Each person has two important traits:

1) a probability of choosing to Split or to Steal and
2) a playing strategy which will alter their probability based on interactions with other players. we will use players with two different playing strategies in this paper, type A and type B.

In a round each person in the population will be paired-with equal probability-with exactly one other person. These pairs then play a game of the prisoner’s dilemma with each other, afterwards a player’s p value can change depending on the results of the game.

B. Definitions

Let the vector $X_t$ represent a population of $N$ people where

$$X_t = X_{t1}, X_{t2}, X_{t3}, \ldots , X_{tN}$$

$t$ is some time index, and each $X_{it}$ represents an $i_{th}$ individual person in t round.
Let $\text{Split} \triangleq 1$ and $\text{Steal} \triangleq 0$. For each person $X_{it}$,
\[
P_r(X_{it} = x) = \begin{cases} 
p_{it}, & \text{for } x = \text{Split}, \\
1 - p_{it}, & \text{for } x = \text{Steal}.
\end{cases}
\] (1)

That is to say each individual has their own $p$ value which can vary with time. $p_{it}$ is a given initial values, and subsequent values are determined by the equation:
\[
p_{i(t+1)} = p_{it} + r \times D_{ijt}.
\]

where $r$ is a given static learning rate and is greater than 0 and less than 1.

We call $D_{ijt}$ the decision matrix, it represents how person $X_i$’s $p$ value is going to shift in response to a single round of the prisoner’s dilemma between themselves and person $X_j$ (note $i \neq j$).

$D_{ijt}$ differs depending on what type of person $X_i$ is. For people of type A:
\[
D_{ijt} = \begin{cases} 
1, & \text{for } X_{jt} = \text{Split}, \\
-1, & \text{for } X_{jt} = \text{Steal}.
\end{cases}
\] (2)

For people of type B:
\[
D_{ijt} = \begin{cases} 
1, & \text{for } X_{jt} = \text{Split}, X_{it} = \text{Split}, \\
-1, & \text{for } X_{it} = \text{Steal}, X_{jt} = \text{Split}, \\
-1, & \text{for } X_{it} = \text{Steal}, X_{jt} = \text{Steal}, \\
1, & \text{for } X_{jt} = \text{Steal}, X_{it} = \text{Steal}.
\end{cases}
\] (3)

From these definitions we can see type A people tend the mimic the behavior for who they’ve been paired against, while people of type B tend to reinforce choices that made them money, or lean away from choices that didn’t gain them money.

Finally we define the pdf
\[
f_{X_i}(x) = \frac{1}{N} \sum \delta[p_{it} - x],
\]
which shows us the probability of randomly choosing an individual whose $p_{it} = x$.

Our goal is given some initial distribution $X_1$ is to estimate $f_{X_i}(x)$ as $t \rightarrow \infty$ and determine if $\lim_{t \rightarrow \infty} f_{X_i}(x) = f_{SS}(x)$, where $f_{SS}(x)$ is some steady state pdf.

C. Expectations

With each round dependent on previous rounds this problem is iterative in nature, so we approach it first by solving for the $E[f_{X_{t+1}X_i}(x)]$. We will use this to iteratively see if the expectation converges as $t \rightarrow \infty$.

We will use the following derived expectations:
\[
E[f_{X_{t+1}X_i}(x)] = \frac{1}{N} \sum \delta[E[p_{i(t+1)}|X_i] - x].
\] (4)

\[
E[p_{i(t+1)}|X_i] = p_{it} + r \times E[D_{ijt}].
\] (5)

\[
E[X_i] = \frac{1}{N} \sum_{i=1}^{N} p_{it}.
\] (6)

Type A:
\[
E[D_{ijt}|X_i,i \neq j] = 2 \times E[X_{jt}] - 1.
\] (7)

Type B:
\[
E[D_{ijt}|X_i] = 4 \times p_{it} \times E[X_{jt}] - 2 \times p_{it} - 2 \times E[X_{jt}] + 1.
\] (8)

For large values $N$, $E[X_{jt}] \approx E[X_i]$ so we further simplify

Type A:
\[
E[D_{ijt}|X_i] = 2 \times E[X_t] - 1
\] (9)

Type B:
\[
E[D_{ijt}|X_i] = 4 \times p_{it} \times E[X_i] - 2 \times p_{it} - 2 \times E[X_i] + 1.
\] (10)

III. Experiment

There are two parts to our experiment, a calculation of the expectation, and a simulation. Both of these are run for a large number of rounds and the results are observed. Algorithm 1 shows our method for calculating $E[f_{X_{t}X_i}(x)]$ for each round we first calculate the average $p$ of all the players $E[X_i]$. Then depending on each person’s type the $E[D_{ijt}]$ is calculated and applied. Once all rounds are run and we have $E[X_{it}X_i]$ we can directly calculate $E[f_{X_{t}X_i}(x)]$.

<table>
<thead>
<tr>
<th>Algorithm 1 Estimation</th>
</tr>
</thead>
<tbody>
<tr>
<td>for $r=2$ rounds</td>
</tr>
<tr>
<td>$EX(t)=SUM(X*p)/N$;</td>
</tr>
<tr>
<td>for $k=1:N$</td>
</tr>
<tr>
<td>$X_i=X(r); \text{if}(X_i \text{type == A})$</td>
</tr>
<tr>
<td>$ED(r,k)=2*EX(r)−1$;</td>
</tr>
<tr>
<td>else</td>
</tr>
<tr>
<td>$ED(r,k)=4<em>X_i</em>p<em>EX(r)−2</em>EX(r)−2<em>X_i</em>p+1$;</td>
</tr>
<tr>
<td>end</td>
</tr>
<tr>
<td>$p = max(p,0)$;</td>
</tr>
<tr>
<td>$X_i.p = p$;</td>
</tr>
<tr>
<td>end</td>
</tr>
</tbody>
</table>

Algorithm 2 shows our simulation methodology. It is very similar to the Algorithm 1, but we go through the process of matching the players together in pairs. For each person a choice is generated and their $p$ value is updated according to the equations given for $p_{i(t+1)}$, and $D_{ijt}$ given above.

<table>
<thead>
<tr>
<th>Algorithm 2 Simulation</th>
</tr>
</thead>
<tbody>
<tr>
<td>for $r=1$ rounds</td>
</tr>
<tr>
<td>match = randperm( population_size)</td>
</tr>
<tr>
<td>for $i = 1:2*population_size$</td>
</tr>
<tr>
<td>$X_i = X(\text{match}(i));$</td>
</tr>
<tr>
<td>$X_j = X(\text{match}(i+1));$</td>
</tr>
<tr>
<td>ChoiceA = $X_i, \text{GetChoice}();$</td>
</tr>
<tr>
<td>ChoiceB = $X_j, \text{GetChoice}();$</td>
</tr>
<tr>
<td>$X_i, \text{update}_p(\text{ChoiceA, ChoiceB});$</td>
</tr>
<tr>
<td>$X_j, \text{update}_p(\text{ChoiceA, ChoiceB});$</td>
</tr>
<tr>
<td>end</td>
</tr>
</tbody>
</table>

We have run our experiment with three different scenario cases. In case 1 all players use the type A strategy. In case 2
all players use the type B strategy. And in case 3 we use a 70-30% mix of type A and type B strategies. For all cases we use the same population(1000), number of rounds(500), learning rate(0.05) and initial distribution of \( p \) (approx gaussian with mean of 0.6 and std of 0.1) showed in the figure 1.

IV. RESULTS AND THE ANALYSIS

A. Case 1

\( N = 1000, 100\% \) of type A, Rounds = 500, learning_rate =0.05, \( f_{X_1} = \text{approx.gaussian} \) with \( \mu = 0.6, \sigma^2 = 0.1 \) For our first case the expectation and simulation matched perfectly. This is because the way \( D_{ijt} \) was defined for type A players create absorbing states when \( p \) equals 0 or 1 for all players. After running the simulation 1000 times the same result was obtained, however by inspection we can see there is non-zero probability that \( \lim_{t \to \infty} f_{X_t}(0) = 1 \) (that is all players would shift completely to the left). We therefore cannot say this surely converges. The results are showed in Figure 2 and 3.

B. Case 2

\( N = 1000, 100\% \) of type B, Rounds = 500, learning_rate =0.05, \( f_{X_1} = \text{approx.gaussian} \) with \( \mu = 0.6, \sigma^2 = 0.1 \) Case 2 had the largest discrepancy between the simulation and expectation of the three cases. \( E[f_{X_t}(x)] \) did converge, but the simulation never did. This is to be expected because \( \lim_{t \to \infty} E[f_{X_t}(x)] \) is not an absorbing state. \( D_{ijt} \) inherently has higher variance for type A than type B. Ironically, \( \lim_{t \to \infty} f_{X_t}(0) = 1 \) for case 2 does converge. The only absorbing state is \( f_{X_{ss}}(1) = 1 \), furthermore \( P_r(\lim_{t \to \infty} f_{X_{t+\tau}} | X_t = X_{ss}) > 0 \) for any \( X_t \). So though seemingly unlikely if we were able to run our simulation for an infinite time it would almost surely converge. Figure 4,5 show the results.

C. Case 3

\( N = 1000, 70\% \) of type A 30\% of type B, Rounds = 500, learning_rate =0.05, \( f_{X_1} = \text{approx.gaussian} \) with \( \mu = 0.6, \sigma^2 = 0.1 \) Our expectation closely predicted the simulation results (note \( f_{X_t}(0) \) on Fig 6. is non-zero). Once again the \( \lim_{t \to \infty} E[f_{X_t}(x)] \) converged, but the simulation did not. By having a high concentration of type A people, the variance of the system was greatly decreases from case 2 so the simulation was a closer match. Using the same reasoning as case 2, this system will also almost surely converge to
\( f_{x,x}(1) = 1 \). The results are presents in Figure 6,7.

V. CONCLUSION

Unsurprisingly, when looking at systems with low variance we were very good at estimating the future values of \( f_{X_t}(x) \), but with higher variances, or as the initial mean approached 0 it became increasingly difficult. More research could be done on how the variance of the system changes with each round. But what was most interesting was the realization that, against our original assumptions, the cases that perfectly matched their expectations did not in fact converge.

REFERENCES
