Only Two Types of Strategies Enforce Linear Payoff Relationships
Under Observation Errors in Repeated Prisoner’s Dilemma Games

Azumi Mamiya\(^1\) and Genki Ichinose\(^1\)

\(^1\)Department of Mathematical and Systems Engineering, Shizuoka University,
3-5- 1 Johoku, Naka-ku, Hamamatsu, 432-8561, Japan
ichinose.genki@shizuoka.ac.jp

Abstract

The repeated prisoner’s dilemma (RPD) game has revealed how cooperation and competition arise among competitive players in long-run relationships. In the RPD game with no errors, zero-determinant (ZD) strategies allow a player to unilaterally set a linear relationship between the player’s own payoff and the opponent’s payoff regardless of the strategy of the opponent. On the other hand, unconditional strategies such as ALLD and ALLC also unilaterally set a linear relationship. However, little is known about the existence of such strategies in the RPD game with errors. Here, we analytically search for the strategies that enforce a linear payoff relationship under observation error in the RPD game. As a result, we found that, even in the case with observation errors, the only strategy sets that enforce a linear payoff relationship are either ZD strategies or unconditional strategies.

The two-player repeated prisoner’s dilemma (RPD) game is a model for exploring the players’ long-run relationships, which mathematically reveals how cooperation and competition arise among competitive players. In the long history of the RPD game, strategies that can unilaterally control opponent’s payoff have been unknown. However, in 2012, Press and Dyson found a novel class of strategies which contain such ultimate strategies, called zero-determinant (ZD) strategies (Press and Dyson, 2012). ZD strategies impose a linear relationship between the payoffs for a focal player and his opponent regardless of the strategy that the opponent implements. The discovery of ZD strategies inspired various relevant studies, which promote an understanding of the nature of human cooperation. In contrast, unconditional strategies such as ALLC and ALLD can also unilaterally set a linear payoff relationship to the opponent. A previous study revealed that those two types of strategies are the only sets which enforce a linear payoff relationship (Ichinose and Masuda, 2018).

These two strategies were found in the case of no errors. Errors are unavoidable in biological organisms and it may lead to the collapse of cooperation due to negative effects. Thus, the effect of errors has been considered in the literature of the RPD game. However, except for Hao et al. (2015), the effect of errors has not been considered for strategies that enforce a linear relationship. In this study, we incorporate observation errors within our model. In human interactions, observation errors refer to the misunderstanding of the opponent’s action, which often happens in a real society. The purpose of this study is how those errors affect the strategies that enforce a linear payoff relationship. Specifically, we aim to investigate whether those strategies can exist even in the case where observation errors are considered.

Here, we consider the symmetric two-person prisoner’s dilemma game. Each player \(i \in \{X, Y\}\) chooses an action \(a_i \in \{C, D\}\). Each player cannot directly see what action the opponent chose. Instead, they can only observe a signal \(\omega_i \in \{g, b\}\), where \(g\) and \(b\) denote good and bad signals, respectively. The signal cannot be observed by the other player, meaning that the signal is private information. Each player’s signal \(\omega_i\) basically depends on the opponent’s action but is also affected by noise from the environment, which is a stochastic variable. In other words, a player observes \(g\) (or \(b\)) when the other player chooses an action \(C\) (or \(D\)). However, when an error occurs, a player observes \(b\) (or \(g\)) although the other player chooses an action \(C\) (or \(D\)). Let \(\epsilon\) be the probability that an error happens to one particular player but not to the other and \(\xi\) be the probability that an error happens to both players. Then, the probability that an error occurs to neither player is \(1 - 2\epsilon - \xi\). The realized payoff for each player depends only on the action he chose and the signal he received. Then the payoff matrix is given by

\[
\begin{pmatrix}
g & b \\ C & R & S \\ D & T & P
\end{pmatrix}
\]

The entries represent the payoffs that the focal player gains in a single round of a repeated game. Each row and column represents the action that the focal player chose and the signal he observed, respectively. The expected payoffs under different action profiles \((C, C), (C, D), (D, C)\) and \((D, D)\) are denoted as \(R_E, S_E, T_E\) and \(P_E\), where \(R_E = R(1 - \epsilon - \xi) + S(\epsilon + \xi), S_E = S(1 - \epsilon - \xi) + R(\epsilon + \xi), T_E = T(1 - \epsilon - \xi) + P(\epsilon + \xi), P_E = P(1 - \epsilon - \xi) + T(\epsilon + \xi)\), respectively. We assume that \(T_E > R_E > P_E > S_E\) and \(2R_E > T_E + S_E\), which dictates the repeated prisoner’s dilemma game.

Consider two players \(X\) and \(Y\) that adopt memory-one strategies, with which they use only the outcomes of the last round to decide the action to be submitted in the current
round. A memory-one strategy is specified by a 4-tuple; X’s strategy is given by a combination of \( p = (p_1, p_2, p_3, p_4) \), where \( 0 \leq p_1, p_2, p_3, p_4 \leq 1 \). The subscripts 1, 2, 3 and 4 of \( p \) mean previous outcome C( or C), D( or D), and C( or D), respectively. \( p_1 \) is the conditional probability that X cooperates when X cooperated and observed signal \( g \) in the last round, \( p_2 \) is the conditional probability that X cooperates when X cooperated and observed signal \( b \) in the last round, and so forth. Similarly, Y’s strategy is specified by a combination of \( q = (q_1, q_2, q_3, q_4) \), where \( 0 \leq q_1, q_2, q_3, q_4 \leq 1 \). Because both players adopt a memory-one strategy, the stochastic state of the two players in round \( t \) is described by \( v(t) = (v_1(t), v_2(t), v_3(t), v_4(t)) \), where the subscripts 1, 2, 3 and 4 of \( v \) mean the stochastic state (C, C), (C, D), (D, C) and (D, D), respectively. \( v_1(t) \) is the probability that both players cooperate in round \( t \), \( v_2(t) \) is the probability that X cooperates and Y defected in round \( t \), and so forth.

The state transition matrix for the game can be described by \( M \). Then, the stochastic state of the two players in round \( t + 1 \) is calculated by \( v(t + 1) = v(t)M \). The stationary distribution for \( M \) is a vector \( v \) such that \( v = vM \). The dot product of an arbitrary vector \( f = (f_1, f_2, f_3, f_4) \) with the stationary distribution vector \( v \) is \( D(p, q, f) = (\sum_{i=1}^{4} f_i) = (\sum_{i=1}^{4} f_i)v_i \).

\[
\begin{align*}
| \tau p_1 q_1 + \tau p_2 q_1 + \tau p_2 q_1 + \tau p_2 q_1 - 1 | & \quad \mu p_1 + \mu p_2 - 1 | \mu p_1 + \mu p_2 - 1 |
\end{align*}
\]

where \( \mu = 1 - \epsilon - \xi, \eta = \epsilon + \xi \). Then player X’s expected payoff in the stationary state is represented by

\[
s_X = \frac{v \cdot S_X}{v \cdot 1} = \frac{D(p, q, S_X)}{D(p, q, 1)},
\]

where \( S_X = (R_E, S_E, T_E, P_E) \) is X’s payoff vector. Similarly, player Y’s expected payoff is represented by

\[
s_Y = \frac{v \cdot S_Y}{v \cdot 1} = \frac{D(p, q, S_Y)}{D(p, q, 1)},
\]

where \( S_Y = (R_E, T_E, S_E, P_E) \) is Y’s payoff vector.

We search for player X’s strategies which impose a linear payoff relationship between the two players, i.e.,

\[
\alpha s_X + \beta s_Y + \gamma = 0.
\]

Here, the linear combination of \( s_X \) and \( s_Y \) is given by

\[
\alpha s_X + \beta s_Y + \gamma = \frac{D(p, q, \alpha S_X + \beta S_Y + \gamma 1)}{D(p, q, 1)}.
\]

If the numerator of the right side of Eq. (6) is zero, that is, \( D(p, q, \alpha S_X + \beta S_Y + \gamma 1) = 0 \), the right side of Eq.(6) becomes zero and it holds Eq.(5). We use the following determinant theorem for the analysis: For an \( n \times n \) matrix \( A \), the following holds: \( det(A) = 0 \iff \) The columns of the matrix \( A \) are dependent vectors. Under this condition, we searched for X’s strategies which impose a linear relationship between the two players’ payoffs. The result showed that the only strategies that impose a linear relationship between the two players’ payoffs are either

\[
\begin{align*}
\mu p_1 + \eta p_2 - 1 &= \alpha R_E + \beta T_E + \gamma \\
\eta p_1 + \mu p_2 - 1 &= \alpha S_E + \beta P_E + \gamma \\
\mu p_3 + \eta p_4 &= \alpha T_E + \beta S_E + \gamma \\
\eta p_3 + \mu p_4 &= \alpha P_E + \beta E_E + \gamma,
\end{align*}
\]

or

\[
\begin{align*}
p_1 = p_2 = p_3 = p_4.
\end{align*}
\]

The former corresponds to the ZD strategies and the latter corresponds to the unconditional strategies, respectively.

Figure 1 shows the numerical realizations of Extortioner (example of ZD; left panel) and ALLD (example of unconditional strategies; right panel). In each panel, player X (vertical axis) adopts Extortioner (left) or ALLD (right), respectively. We randomly generated 1,000 strategies for player Y (horizontal axis) for each panel. We numerically confirmed that Extortioner and ALLD can enforce a linear payoff relationship to player Y even in the case that errors are considered (\( \epsilon + \xi = 0.1, 0.2, 0.3 \)).

In conclusion, we analytically found that, in the RPD game with observation errors, the only strategy sets that enforce a linear payoff relationship are either the ZD strategies or the unconditional strategies, which are consistent with the case of no errors. This result suggests that, in any case, those two sets are only types of strategies that enforce a linear payoff relationship between two players.

References

