

Economics 209A
Theory and Application of Non-Cooperative Games
(Fall 2013)

Repeated games
OR 8 and 9, and FT 5

The basic idea – prisoner's dilemma

The prisoner's dilemma game with one-shot payoffs

	<i>C</i>	<i>D</i>
<i>C</i>	2, 2	0, 3
<i>D</i>	3, 0	1, 1

has a unique Nash equilibrium in which each player chooses *D* (defection), but both player are better if they choose *C* (cooperation).

If the game is played repeatedly, then (C, C) accrues in every period if each player believes that choosing *D* will end cooperation (D, D) , and subsequent losses outweigh the immediate gain.

Strategies

Grim trigger strategy

$$\boxed{C : C} \xrightarrow{(\cdot, D)} \boxed{D : D}$$

Limited punishment

$$\begin{array}{ccccccc} \dashrightarrow & \boxed{P_0 : C} & \longrightarrow & \boxed{P_1 : D} & \longrightarrow & \boxed{P_2 : D} & \longrightarrow & \boxed{P_3 : D} & \dashrightarrow \\ & & (\cdot, D) & & (\cdot, \cdot) & & (\cdot, \cdot) & & (\cdot, \cdot) \end{array}$$

Tit-for-tat

$$\begin{array}{ccc} \dashrightarrow & \boxed{C : D} & \longrightarrow & \boxed{D : D} & \dashrightarrow \\ & & (\cdot, D) & & (\cdot, C) \end{array}$$

Payoffs

Suppose that each player's preferences over streams $(\omega^1, \omega^2, \dots)$ of payoffs are represented by the discounted sum

$$V = \sum_{t=1}^{\infty} \delta^{t-1} \omega^t,$$

where $0 < \delta < 1$.

The discounted sum of stream (c, c, \dots) is $c/(1 - \delta)$, so a player is indifferent between the two streams if

$$c = (1 - \delta)V.$$

Hence, we call $(1 - \delta)V$ the discounted average of stream $(\omega^1, \omega^2, \dots)$, which represent the same preferences.

Let

$$V_T = c + \delta c + \delta^2 c + \dots + \delta^T c.$$

Then,

$$V_T - \delta V_T = c - \delta^{T+1} c$$

and

$$V_T = \frac{1 - \delta^{T+1}}{1 - \delta} c.$$

Nash equilibria

Grim trigger strategy

$$(1 - \delta)(3 + \delta + \delta^2 + \dots) = (1 - \delta) \left[3 + \frac{\delta}{(1 - \delta)} \right] = 3(1 - \delta) + \delta$$

Thus, a player cannot increase her payoff by deviating if and only if

$$3(1 - \delta) + \delta \leq 2,$$

or $\delta \geq 1/2$.

If $\delta \geq 1/2$, then the strategy pair in which each player's strategy is grim strategy is a Nash equilibrium which generates the outcome (C, C) in every period.

Limited punishment (k periods)

$$(1-\delta)(3+\delta+\delta^2+\dots+\delta^k) = (1-\delta) \left[3 + \delta \frac{(1-\delta^k)}{(1-\delta)} \right] = 3(1-\delta) + \delta(1-\delta^k)$$

Note that after deviating at period t a player should choose D from period $t+1$ through $t+k$.

Thus, a player cannot increase her payoff by deviating if and only if

$$3(1-\delta) + \delta(1-\delta^k) \leq 2(1-\delta^{k+1}).$$

Note that for $k=1$, then no $\delta < 1$ satisfies the inequality.

Tit-for-tat

A deviator's best-reply to tit-for-tat is to alternate between D and C or to always choose D , so tit-for-tat is a best-reply to tit-for-tat if and only if

$$(1 - \delta)(3 + 0 + 3\delta^2 + 0 + \dots) = (1 - \delta)\frac{3}{1 - \delta^2} = \frac{3}{1 + \delta} \leq 2$$

and

$$(1 - \delta)(3 + \delta + \delta^2 + \dots) = (1 - \delta) \left[3 + \frac{\delta}{(1 - \delta)} \right] = 3 - 2\delta \leq 2.$$

Both conditions yield $\delta \geq 1/2$.

Subgame perfect equilibria

Grim trigger strategy

For the Nash equilibria to be subgame perfect, "threats" must be credible: punishing the other player if she deviates must be optimal.

Consider the subgame following the outcome (C, D) in period 1 and suppose player 1 adheres to the grim strategy.

Claim: It is not optimal for player 2 to adhere to his grim strategy in period 2.

If player 2 adheres to the grim strategy, then the outcome in period 2 is (D, C) and (D, D) in every subsequent period, so her discounted average payoff in the subgame is

$$(1 - \delta)(0 + \delta + \delta^2 + \dots) = \delta,$$

where as her discounted average payoff is 1 if she choose D already in period 2.

But, the "modified" grim trigger strategy for an infinitely repeated prisoner's dilemma

$$\boxed{C : C} \rightarrow \boxed{D : D}$$

$$(\cdot, \cdot) / (C, C)$$

is a subgame perfect equilibrium strategy if $\delta \geq 1/2$.

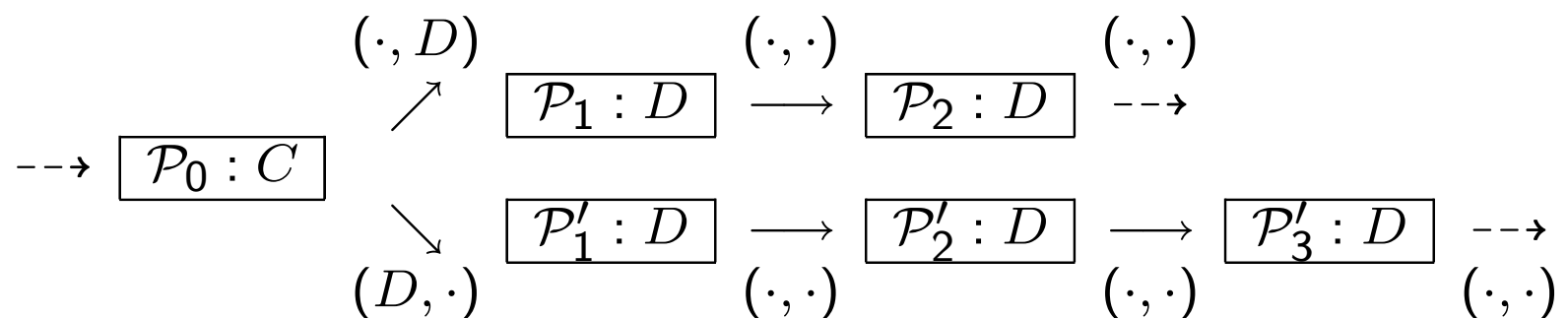
Limited punishment

The game does not have such subgame perfect equilibria from the same reason that a pair of grim strategies is never subgame perfect.

But, we can modify the limited punishment strategy in the same way that we modified the grim strategy to obtain subgame perfect equilibrium for δ sufficiently high.

The number of periods for which a player chooses D after a history in which not all the outcomes were (C, C) must depend on the identity of the deviator.

Consider the strategy of player 2, where the top part entails her reaction to her own deviation



Tit-for-tat

The optimality of tit-for-tat after histories ending in (C, C) is covered by our analysis of Nash equilibrium.

If both players adhere to tit-for-tat after histories ending in (C, D) : then the outcome alternates between (D, C) and (C, D) .

(The analysis is the same for histories ending in (D, C) , except that the roles of the players are reversed.)

Then, player 1's discounted average payoff in the subgame is

$$(1 - \delta)(3 + 3\delta^2 + 3\delta^4 + \dots) = \frac{3}{1 + \delta},$$

and player 2's discounted average payoff in the subgame is

$$(1 - \delta)(3\delta + 3\delta^3 + 3\delta^5 + \dots) = \frac{3\delta}{1 + \delta}.$$

Next, we check if tit-for-tat satisfies the one-deviation property of subgame perfection.

If player 1 (2) chooses C (D) in the first period of the subgame, and subsequently adheres to tit-for-tat, then the outcome is (C, C) ((D, D)) in every subsequent period. Such a deviation is profitable for player 1 (2) if and only if

$$\frac{3}{(1 + \delta)} \geq 2, \text{ or } \delta \leq 1/2$$

and

$$\frac{3\delta}{(1 + \delta)} \geq 1, \text{ or } \delta \geq 1/2,$$

respectively.

Finally, after histories ending in (D, D) , if both players adhere to tit-for-tat, then the outcome is (D, D) in every subsequent period.

On the other hand, if either player deviates to C , then the outcome alternates between (D, C) and (C, D) (see above).

Thus, a pair of tit-for-tat strategies is a subgame perfect equilibrium if and only if $\delta = 1/2$.

Nash equilibria discounted average payoffs

Nash folk theorem

- For any $\delta \in [0, 1]$, the discounted average payoffs of each player i in any Nash equilibrium is at least $u_i(D, D)$.
- Let $x = (x_1, x_2)$ be a feasible payoff pair for which $x_i > u_i(D, D)$ for $i = 1, 2$. There exists $\bar{\delta}$ such that for any $\delta > \bar{\delta}$ there exists a Nash equilibrium in which the discounted average payoffs of each player i is x_i .
- For any $\delta \in [0, 1]$, there is a Nash equilibrium in which the discounted average payoffs of each player i is $u_i(D, D)$.

Every subgame perfect equilibrium is also a Nash equilibrium, so the set of subgame perfect equilibrium payoff pairs is a subset of the set of Nash equilibrium payoff pairs.

But, strategies that are not subgame perfect equilibrium strategies, like grim, can be modified to make the punishment it imposes credible.

Thus, the set of subgame perfect equilibrium payoff pairs is the same as of the set of Nash equilibrium payoff pairs.

Subgame perfect equilibria discounted average payoffs

Subgame perfect folk theorem

- For any $\delta \in [0, 1]$, the discounted average payoffs of each player i in any subgame perfect equilibrium is at least $u_i(D, D)$.
- Let $x = (x_1, x_2)$ be a feasible payoff pair for which $x_i > u_i(D, D)$ for $i = 1, 2$. There exists $\bar{\delta}$ such that for any $\delta > \bar{\delta}$ there exists a subgame perfect equilibrium in which the discounted average payoffs of each player i is x_i .
- For any $\delta \in [0, 1]$, there is a subgame perfect equilibrium in which the discounted average payoffs of each player i is $u_i(D, D)$.

Folk theorems

The set of payoffs (not outcomes) that can be sustained by equilibria. Socially desirable outcomes can be sustained if players have long-term objectives. But, The set of equilibrium outcome is huge, so it lacks predictive power.

Infinitely vs. finitely repeated games: finite and infinite horizon yield different results. In the prisoner's dilemma (D, D) in every period is the only Nash equilibrium of any finite repetition.

Preliminaries

An N -player game with perfect monitoring

$$G = \langle N, (A_i), (u_i) \rangle$$

repeated T times with discount factor $\delta \in (0, 1)$ is denoted by

$$G^\delta(T)$$

and

$$G^\delta(\infty)$$

denotes the infinitely repeated version of game G with discount factor $\delta \in (0, 1)$.

An N -player game with perfect monitoring

$$G = \langle N, (A_i), (u_i) \rangle$$

repeated T times with no discounting is denoted by

$$G(T),$$

and

$$G(\infty)$$

denotes the infinitely repeated version of game G .

Strategies

A strategy of player i in $G^\delta(T)$ or $G(T)$ is a behavioral strategy

$$\sigma_i = (\sigma_i^1, \sigma_i^2, \dots, \sigma_i^T)$$

where $\sigma_i^1 \in \Delta(A_i)$ and for any $t > 1$

$$\sigma_i^t : (\times_{j \in N} \Delta(A_j))^{t-1} \rightarrow \Delta(A_i),$$

and a similar definition applies to $G^\delta(\infty)$ or $G(\infty)$.

Outcomes

An N -tuple of pure strategies in $G^\delta(T)$ or or $G(T)$ inductively defines an outcome path

$$(a^1, a^2, \dots, a^T)$$

where for any $t > 1$

$$a^t = a^t(a^1, a^2, \dots, a^{t-1}),$$

and a similar definition applies to $G^\delta(\infty)$ or $G(\infty)$.

An N -tuple of behavioral strategies probabilistically defines outcome paths.

Payoffs

Player i 's payoffs in $G^\delta(T)$ and $G^\delta(\infty)$ are given respectively by

$$\sum_{t=0}^T \delta^t u_i(a^t) \text{ and } \sum_{t=0}^{\infty} \delta^t u_i(a^t).$$

Player i 's discounted average payoffs in $G^\delta(T)$ and $G^\delta(\infty)$ are given respectively by

$$\frac{1 - \delta}{1 - \delta^{T+1}} \sum_{t=0}^T \delta^t u_i(a^t) \text{ and } (1 - \delta) \sum_{t=0}^{\infty} \delta^t u_i(a^t).$$

Player i 's payoffs in $G(T)$ and $G(\infty)$ are given respectively by

$$\frac{1}{T} \sum_{t=0}^{T-1} u_i(a^t)$$

and what so-called limit of the means

$$\lim_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} u_i(a^t)$$

which may not exist as defined for a particular path. So, two possible payoff criteria are

$$\liminf_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} u_i(a^t) \text{ or } \limsup_{T \rightarrow \infty} \frac{1}{T} \sum_{t=0}^{T-1} u_i(a^t).$$

Equilibrium

The sets of Nash and subgame perfect equilibria discounted average payoffs of $G^\delta(T)$ are denoted respectively by $N^\delta(T)$ and $P^\delta(T)$, repetitively.

The set of discounted average payoffs of $G^\delta(T)$ which can be sustained using grim trigger strategies is denoted by $TR^\delta(T)$.

The corresponding sets in a game consists of $G(T)$ ($\delta = 1$) are denoted by $N(T)$, $P(T)$ and $TR(T)$. Similar definitions apply to $G^\delta(\infty)$ and $G(\infty)$.

Nash equilibrium

A Nash equilibrium of $G^\delta(T)$ or $G(T)$ is an N -tuple of strategies

$$\sigma = \times_{i \in N} \sigma_i$$

such that for any player i

$$H_i(\sigma) \geq H_i(\sigma'_i, \sigma_{-i})$$

for any $\sigma'_i \neq \sigma_i$ and H_i is the payoff criterion.

Subgame perfect equilibrium

Let $h(K) = (a^1, \dots, a^K)$ and define $\sigma/h(K)$ such that

$$\sigma^1/h(K) = \sigma^{K+1}(a^1, \dots, a^K)$$

and for all $t > 1$ and $(\bar{a}^1, \dots, \bar{a}^{t-1})$

$$\sigma^t/(h(K), (\bar{a}^1, \dots, \bar{a}^{t-1})) = \sigma^{K+1}(a^1, \dots, a^K, \bar{a}^1, \dots, \bar{a}^{t-1}).$$

Hence, $\sigma/h(K)$ is the strategy profile induced by the history $h(K)$ in the remaining of the game.

A subgame perfect equilibrium of $G^\delta(T)$ or $G(T)$ is σ which is a Nash equilibrium of $G^\delta(T - K)$ or $G(T - K)$ for any $0 \leq K < T$.

Strategies as machines

A machine (finite automaton) for player i in an infinitely repeated game of G is a four-tuple $(Q_i, q_i^0, f_i, \tau_i)$ where

- Q_i is a finite set of states,
- $q_i^0 \in Q_i$ is the initial state,
- $f_i : Q_i \rightarrow A_i$ is an output function that assigns an action to every state, and
- $\tau_i : Q_i \times A_j \rightarrow Q_i$ is a transition function (depends only on the present state and the other player's action).

A machine game

The machine game is a two-player strategic game in which each player chooses a (finite) machine to play the infinitely repeated game.

The set of strategies for player i in the machine game is the set of all finite machines for player i , denoted by M_i .

Player i prefers the pair of machines (M_i, M_j) to (M'_i, M'_j) if in the repeated game

$$(a^t(M_i, M_j))_{t=1,2,\dots} \succsim_i a^t(M'_i, M'_j)_{t=1,2,\dots}$$

By moving from strategies to machines, we have restricted the set of strategies to those that can be executed by finite machines.

The restriction of the set of strategies to those implemented by machines does not affect the content of the Folk theorems.

Whatever player j 's machine is, player i can design a machine such that the induced sequence of payoffs is at least as good as the constant sequence of the minmax level.

Complexity

A very naive approach: the complexity, $c(M)$, of the machine $M = (Q, q^0, f, \tau)$ is taken to be its number of states (the cardinality of Q).

A machine game of an infinitely repeated game $G = \langle \{1, 2\}, (A_i), (u_i) \rangle$ with the complexity measure $c(M)$, is a strategic game in which for each player i , the set of strategies is M_i , and $(M_i, M_j) \succ_i (M'_i, M'_j)$ whenever

$$U_i(M_i, M_j) > U_i(M'_i, M'_j) \text{ and } c(M_i) = c(M'_i),$$

or

$$U_i(M_i, M_j) = U_i(M'_i, M'_j) \text{ and } c(M_i) < c(M'_i).$$

Example: Consider the two-state machine M above that implements the grim strategy in the PD .

If δ is not too small, then M is a best response to the other player using M in the δ -discounted repeated game.

Given that player 2 uses M , the machine with one state in which C is played yields player 1 the same payoff and is less complex.

Either player can drop a state without affecting the outcome so (M, M) is not a Nash equilibrium.

The structure of machine game Nash equilibria

Result I If (M_1^*, M_2^*) is a Nash equilibrium of a machine game, then $c(M_1^*) = c(M_2^*)$ and the pair of strategies in the repeated game associated with (M_1^*, M_2^*) is a Nash equilibrium of the repeated game.

- For each j , the solution to the problem $\max_{M_j} U_j(M_i^*, M_j^*)$ (when the complexity element is ignored) does not involve more than $c(M_i^*)$ states.
- Player j can use a machine with $c(M_i^*)$ states to achieve a payoff in the repeated game equal to $\max_{M_j} U_j(M_i^*, s_j)$ where s_j is his strategy in the repeated game.

Result II If (M_1^*, M_2^*) is a Nash equilibrium of a machine game, then there is a one-to-one correspondence between the actions of player 1 and player 2, prescribed by M_1^* and M_2^* : for some $t \neq s$,

$$\text{if } a_i^t(M_1^*, M_2^*) = a_i^s(M_1^*, M_2^*) \text{ then } a_j^t(M_1^*, M_2^*) = a_j^s(M_1^*, M_2^*).$$

- The proof is by contradiction – there is a machine which carries out an optimal strategy for player j using only $c(M_i^*) - 1$ states.
- In the *PD*, the *G*-outcomes played in equilibrium must be either in the set $\{(C, C), (D, D)\}$ or in the set $\{(C, D), (D, C)\}$.

Result III If (M_1^*, M_2^*) is an equilibrium of a machine game, then there exists a period t^* and an integer $l < t^*$ such that for $i = 1, 2$, the states in the sequence

$$(q_i^t(M_1^*, M_2^*))_{t=1, \dots, t^*}$$

are distinct and

$$q_i^t(M_1^*, M_2^*) = q_i^{t-l}(M_1^*, M_2^*) \text{ for } t > t^*,$$

i.e. an introductory phase and a cycling phase.

Concluding, the set of equilibria of the machine game is much smaller than that of the repeated game. To restrict the set even more we need to specify the tradeoff between $U_i(M_i, M_j)$ and $c(M_i)$. For example, by using lexicographic preferences (see OR 9.4).