Repeated games-Correlated equilibria

Roberto Lucchetti

Politecnico di Milano

The reference game

Example

$$\begin{pmatrix} (6,6) & (0,10) & (-2,-2) \\ (10,0) & (1,1) & (-1,-1) \\ (-2,-2) & (-1,-1) & (-2,-2) \end{pmatrix}$$

Equilibrium in strictly dominated strategies.

Suppose the game is played once a day for N days, i.e. it is repeated N times

Nash equilibria

Playing all days the strictly dominant strategy for both, with outcome $(1, 1)$, is an obvious equilibrium

Are there other Nash equilibria? The more appealing, socially efficient, outcome $(6, 6)$ can be obtained by the players?

We shall show that, for every $a > 0$, if the game lasts enough days (if N is sufficiently big), the players can get at least $6 - a$ each on average

An interesting strategy

Suppose the game is played N days. Consider the following symmetric strategy

- Player one (two) plays the first row (column) at the first $N-k$ days and the second row (column) in the last k days, if the second (first) player uses the same strategy
- Otherwise, if at one day the second (first) deviates, from that stage on player one (two) plays the third row (column)

In other words, the players play "friendly" for $N-k$ days, and the strictly dominant strategy for the remaining days. In case one deviates, the other shifts to the third strategy, for ever.

It is a Nash equilibrium

What a player gets under the strategy profile

$$\frac{(N-k)6+k1}{N}$$

What the player gets by deviating the most convenient day, i.e. day $N-k$:

$$\frac{(N-k-1)6+10+k(-1)}{N}$$

Thus the strategy profile is a NEp if and only if

$$\frac{(N-k)6+k1}{N} \geq \frac{(N-k-1)6+10+k(-1)}{N}$$

True provided $k \geq 2$

Payoffs

The payoffs at the NEp

$$\frac{(N-k)6+k1}{N}$$

For every k

$$\lim_{N\to\infty}\frac{(N-k)6+k1}{N}=6$$

On average the players can get at least $6-a$ each per day, if they play a sufficiently large number of days

■

Remarks

- When a game is repeated, collaboration, even if dominated in the one shot game, can be based on rationality
- The (common) strategy of the NEp has a weakness: it is based on a mutual threat of the players, which is not completely credible
- In general, the number of the NEp, in the repetition is very large

Infinite repetitions of the game

A stage game (usually a finite game in strategic form) is played with infinite horizon (for infinite days) by the players

We need to specify:

1. strategies
2. payoffs

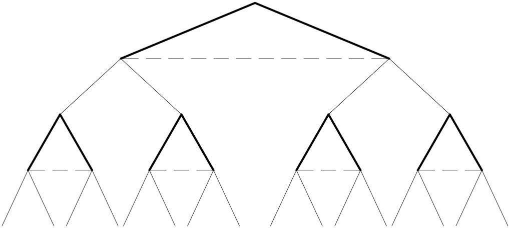

Figure: The prisoner dilemma repeated twice.

Strategy

Strategy for a player is

$$s = s(\tau), \tau = 0, \ldots$$

where for each τ $s(\tau)$ is a specification of moves of the stage game, in general function of the past choices of the players

In the Example a possible strategy $s = (s(0), s(1))$

- $s(0)$ do not confess
- $s(1)$ do not confess if the other player did not confess at stage zero, otherwise confess

Observe the simplification w.r.t. the general definition of strategy

Payoff

It is not possible to sum payoffs obtained at each stage since the sum will be infinite in general

Different possible choices, one standard is to use a discount factor δ

$$0 < \delta < 1$$

The utility of player i

$$u_i(s,t) = (1-\delta)\sum_{\tau=0}^{\infty} \delta^{\tau} u_i(s(\tau), t(\tau))$$

$u_i(s(\tau), t(\tau))$ is the stage-game payoff of player i at time τ given strategy profile $(s(\tau), t(\tau))$

$1-\delta$ is a normalizing factor: if $u_i(s(\tau), t(\tau)) = a$ for all τ then

$$u_i(s,t) = a$$

Threat values

The stage game is (A, B)

Definition

For the bimatrix game (A, B)

$$\underline{v}_1 = \min_j \max_i a_{ij}, \qquad \underline{v}_2 = \min_i \max_j b_{ij}$$

are called the threat values of the two players

Suppose $\underline{v}_1$ is obtained with $j = \bar{\jmath}$. If Pl2 wants to punish Pl1, $\underline{v}_1$ is the highest utility Pl 1 can get if Pl2 uses $\bar{\jmath}$.

The folk theorem

Suppose a bimatrix game (A, B) is given.

Theorem

For every outcome $v = (v_1, v_2) = (a_{\bar{\imath}\bar{\jmath}}, b_{\bar{\imath}\bar{\jmath}})$ such that $v_i > \underline{v}_i$ $i = 1, 2$, there exists $\bar{\delta} < 1$ such that for all $\delta > \bar{\delta}$ there is a Nash equilibrium of the repeated game with discounting factor δ, with payoffs v.

The proof

Proof $v = (v_1, v_2) = (a_{\overline{\text{IJ}}}, b_{\overline{\text{IJ}}})$ such that $v_i > \underline{v}_i$,
Define the following strategy s

Play the strategy leading to v at any stage, unless the opponent deviates at time t. In this case play the threat strategy from the stage $t + 1$ onwards

Need to prove

- s provides utility vector v
- s is a Nash equilibrium for all δ close to 1

Cont'd

Denote by s_t the strategy of deviating at time t.

$$u_1(s_t, s) \leq (1-\delta)\left(\sum_{\tau=0}^{t-1} \delta^\tau v_1 + \delta^t \max_{i,j} a_{ij} + \sum_{\tau=t+1}^{\infty} \delta^\tau \underline{v}_1\right) = \quad (1)$$

$$= (1-\delta^t)v_1 + (1-\delta)\delta^t \max_{i,j} a_{ij} + (\delta^{t+1})\underline{v}_1 \quad (2)$$

$$u_1(s, s) = (1-\delta)\sum_{\tau=0}^{\infty} \delta^\tau v_1 = v_1 \quad (3)$$

Cont'd

$$u_1(s,s) = v_1 \geq u_1(s_t,s) = (1-\delta^t)v_1 + (1-\delta)\delta^t \max_{i,j} a_{ij} + (\delta^{t+1})\underline{v}_1$$

if and only if

$$\begin{aligned}
(1-\delta^t)v_1 + \delta^t(1-\delta)\max_{i,j} a_{ij} + \delta^{t+1}\underline{v_1} &\leq v_1 \\
\delta^t(1-\delta)\max_{i,j} a_{ij} + \delta^{t+1}\underline{v_1} &\leq \delta^t v_1 \\
(1-\delta)\max_{i,j} a_{ij} + \delta\underline{v_1} &\leq v_1 \\
\delta(\max_{i,j} a_{ij} - \underline{v_1}) &\geq \max_{i,j} a_{ij} - v_1
\end{aligned}$$

We set

$$0 < \underline{\delta_1} = \frac{\max_{i,j} a_{ij} - v_1}{\max_{i,j} a_{ij} - \underline{v_1}} < 1$$

Conclusion

Thus setting

$$\underline{\delta_1} = \frac{\max_{i,j} a_{ij} - v_1}{\max_{i,j} a_{ij} - \underline{v_1}}$$

$$\underline{\delta_2} = \frac{\max_{i,j} b_{ij} - v_2}{\max_{i,j} b_{ij} - \underline{v_2}}$$

the theorem is proved with

$$\underline{\delta} = \max_{i=1,2} \underline{\delta_i} \quad \blacksquare$$

Correlated equilibria: the reference example

$$\begin{pmatrix} (6,6) & (2,7) \\ (7,2) & (0,0) \end{pmatrix}$$

There are three NEp:

$[(1,0)(0,1)]$ $[(0,1)(1,0)]$ $[(\frac{2}{3},\frac{1}{3}),(\frac{2}{3},\frac{1}{3})]$

with outcomes $(2,7)$ $(7,2)$, and $(\frac{14}{3},\frac{14}{3})$ respectively.

Is it possible to do better for the players?

Consider the following probability distribution over the outcomes

$$\begin{pmatrix} \frac{1}{3} & \frac{1}{3} \\ \frac{1}{3} & 0 \end{pmatrix}$$

This provides better outcome ($\frac{15}{3}$) for both than the mixed NEp, but how is it possible to convince the players to agree on this?

Partial information to the players

Suppose the players agree on the following mechanism. An external entity makes a random choice on the outcomes according to the probabilities on the outcomes given by the table, and tells the players what to do, privately

Given this private information, we see that the players do not have incentive to change strategy!

No incentive to change

1 The random choice selects outcome $(7,2)$. Pl1 is told to play second row, PL2 first column. PL1 now knows that PL2 is told to play first column: he does not deviate since the outcome in NEp. Pl2 knows that the probability Pl1 is told to play first row is $\frac{1}{2}$. Thus his expected value following the suggestion is $\frac{1}{2}(6+2)$. If he deviates his expected value is $\frac{1}{2}(7+0)$: no interest to deviate for both

2 The random choice selects outcome $(6,6)$. Pl1 is told to play first row, PL2 first column. Both players know that the other player will play the two strategies with the same probability. Thus the expected value following the suggestion is $\frac{1}{2}(6+2)$. If the player deviates his expected value is $\frac{1}{2}(7+0)$: no interest to deviate for both

3 The random choice selects outcome $(2,7)$. Just as in 1 (interchanging the role of the players): no interest to deviate for both

Toward the correlated equilibrium

In the above example the probability distribution over the outcomes is accepted by all players, since in any case they do not have interest to deviate, given the upgrade on the probability of the outcomes after the random choice and the information obtained (what to do)

Given the game $(A, B) = (a_{ij}, b_{ij})$, $i = 1, \dots, n$, $j = 1, \dots, m$, let $I = \{1, \dots, n\}$, $J = \{1, \dots, m\}$ and $X = I \times J$

Correlated equilibrium

Given the game $(A, B) = (a_{ij}, b_{ij})$, $i = 1, \dots, n$, $j = 1, \dots, m$, let $I = \{1, \dots, n\}$, $J = \{1, \dots, m\}$ and $X = I \times J$

Definition

A correlated equilibrium is a probability distribution $p = (p_{ij})$ on X such that, for all $\bar{\imath} \in I$

$$\sum_{j=1}^{m} p_{\bar{\imath}j} a_{\bar{\imath}j} \geq \sum_{j=1}^{m} p_{\bar{\imath}j} a_{ij} \qquad \forall i \in I$$

and for all $\bar{\jmath} \in J$

$$\sum_{i=1}^{n} p_{i\bar{\jmath}} b_{i\bar{\jmath}} \geq \sum_{i=1}^{n} p_{i\bar{\jmath}} b_{ij} \qquad \forall j \in J$$

The inequalities in the example

$$\begin{pmatrix} (6,6) & (2,7) \\ (7,2) & (0,0) \end{pmatrix} \quad \begin{pmatrix} x_1 & x_2 \\ x_3 & x_4 \end{pmatrix}$$

$$\begin{cases} 6x_1 + 2x_2 \geq 7x_1 \\ 7x_3 \geq 6x_3 + 2x_4 \\ 6x_1 + 2x_3 \geq 7x_1 \\ 7x_2 \geq 6x_2 + 2x_4 \\ x_1 + x_2 + x_3 + x_4 = 1 \\ x_i \geq 0 \qquad\qquad i = 1, \dots, 4 \end{cases}$$

Existence

The set of the correlated equilibria of a finite game is nonempty

Theorem

A NEp profile generates a Correlated equilibrium

Given the NEp $(\bar{x}, \bar{y})$ the probability distribution on the outcome matrix is $p = (p_{ij})$ with $p_{ij} = \bar{x}_i \bar{y}_j$

The proof

Proof

We have to prove that

$$\sum_{j=1}^{m} \bar{x}_{\bar{\imath}} \bar{y}_j a_{\bar{\imath}j} \geq \sum_{j=1}^{m} \bar{x}_{\bar{\imath}} \bar{y}_j a_{ij} \qquad \forall i \in I$$

Obvious if $\bar{x}_{\bar{\imath}} = 0$. If $\bar{x}_{\bar{\imath}} > 0$ we need to show that

$$\sum_{j=1}^{m} \bar{y}_j a_{\bar{\imath}j} \geq \sum_{j=1}^{m} \bar{y}_j a_{ij} \qquad \forall i \in I$$

The left (right) hand side is the expected utility of the first player if he plays row $\bar{\imath}$(row i) and the second his equilibrium strategy $\bar{y}$

The inequality holds since the pure strategy $\bar{\imath}$ is played with positive probability so $\bar{\imath}$ must be (one of) best reaction(s) to $\bar{y}$ ■

The set of the correlated equilibria

Theorem

The set of the correlated equilibria of a finite game is a nonempty convex polytope

Proof Remember that a convex polytope is a closed bounded convex set which is the smallest convex set containing a finite number of points. The set of the correlated equilibria is the solution set of a system of $n^2 + m^2$ linear inequalities (where n, m are the number of the pure strategies of the players), called incentive constraints, plus the conditions of being a probability distribution $(p_{ij} \geq 0, \sum p_{ij} = 1)$ ■

Dominated strategies

Proposition

If a row $\bar{\imath}$ is strictly dominated, then $p_{\bar{\imath}j} = 0$ for every j

Proof

Suppose $\bar{\imath}$ is strictly dominated by i. Since

$$\sum_{j=1}^{m} p_{\bar{\imath}j}(a_{\bar{\imath}j} - a_{ij}) \geq 0$$

it must be $p_{\bar{\imath}j} = 0$ for every j ■

Is the same true for a weakly dominated row?

Conclusion

This conclutes the part of the noncooperative game theory, according to Nash model.
The most important conclusion we can draw is that there is essentially a unique rationality paradigm in the whole theory:

this is the idea of best reaction

As we have seen, not always the idea of NE is fully convincing, still it remains the foundation of rationality in non cooperative theory