Infinitely repeated games

Marcin Pęski

March 11, 2020

Plan

Infinitely repeated prisoner's dilemma

Grim trigger equilibrium

Cooperation and threat of punishment

Infinitely repeated prisoner's dilemma

Payoffs in stage game

- Two actions C(Cooperate) and D(deviate)
- Per-period payoffs $g_i(a_i, a_{-i})$ from action profile (a_i, a_{-i}) are in the table

	C	D
C	x, x	$0, y$
D	$y, 0$	$1, 1$

where $y > x > 1$.

Infinitely repeated prisoner's dilemma

Payoffs in repeated game

- Players play for $T = \infty$ periods.
- History in period t
  $\left(a^1, ..., a^{t-1}\right) = \left(\left(a_1^1, a_2^1\right), ..., \left(a_1^{t-1}, a_2^{t-1}\right)\right)$.
  - terminal histories have infinitely many periods,
- Payoffs in discounted game along infinite history

$$g_i\left(a_i^1, a_{-i}^1\right) + \delta g_i\left(a_i^1, a_{-i}^1\right) + \delta^2 g_i\left(a_i^3, a_{-i}^3\right) + ...$$
$$= \sum_{t=1}^{\infty} \delta^{t-1} g_i\left(a_i^t, a_{-i}^t\right)$$

  - $g_i\left(a_i^t, a_{-i}^t\right)$ is the payoff in period t,
  - $\delta < 1$ - discount factor,
  - notice that the discount in period t is δ^{t-1}.

Infinitely repeated prisoner's dilemma
Payoffs

Different interpretations of discount factor

- delayed payoff (\$1 tomorrow is worth only \$$\delta$ today),
- probability that the game ends in each period is equal to $1-\delta$
  - probability that the game continues is δ in each period,
  - probability that game ends after period 2 (i.e., survives till after period 2) is δ,
  - probability that game ends after period 3 is δ^2,
  - probability that game ends after period 4 is δ^3, etc.
  - $\sum_{t=1}^{\infty} \delta^{t-1} g_i \left(a_i^t, a_{-i}^t\right)$ is the expected payoff until the game ends.
- High δ means that players are patient and value future interactions.

Infinitely repeated prisoner's dilemma

Payoffs

- The payoff along infinite histories
  - $(DD, DD, DD, ...)$ is

$$1 + \delta + \delta^2 + \delta^3 + ... = \frac{1}{1-\delta},$$

  - $(CC, CC, CC, ...)$ is

$$x + \delta x + \delta^2 x + \delta^3 x + ... = x\left(1 + \delta + \delta^2 + ...\right) = x\frac{1}{1-\delta},$$

  - $(CD, DC, CD, DC, ...)$ is

$$0 + \delta y + \delta^2 0 + \delta^3 y + \delta^4 0 + \delta^5 y = \delta y\left(1 + \delta^2 + \delta^4 + ...\right) = \delta y\frac{1}{1-\delta^2}.$$

Infinitely repeated prisoner's dilemma

Examples of strategies

- Always Cooperate $\sigma_{AC}(h) = C$ for each history h,
- Always Defect $\sigma_{AD}(h) = D$ for each h,
- Tit-for-Tat (of player i $\sigma_{GT}(\emptyset) = C$ and $\sigma_{GT}(h) = a_{-i}^{t-1}$ for any other t-period history,
    - player i starts with C,
    - and plays the previous period action of the other player in each period

Infinitely repeated prisoner's dilemma

Examples of strategies

- Grim Trigger strategy

$$\sigma_{GT}(h) = \begin{cases} C & \text{if } h = \emptyset \text{ or } h = (CC, ..., CC) \\ D & \text{otherwise} \end{cases},$$

  - player starts with C,
  - continues playing C after any history in which both players always played C
  - plays D otherwise.
  - *I will trust you first and you can cheat me only once.*

Infinitely repeated prisoner's dilemma

Strategies and best response

- Best response to Always Cooperate.
  - Payoff from Always Cooperate (against AC) is $x + \delta x + \delta^2 x + \ldots = \frac{1}{1-\delta} x$,
  - Payoff from Always Defect is $y + \delta y + \delta^2 y + \ldots = \frac{1}{1-\delta} y$,
    - that is a maximum possible payoff,
  - Hence, Always Defect is a best response and (AC, AC) is not a (Nash) equilibrium.
- OTOH, (AD, AD) is a Nash and Subgame Perfect Equilibrium
  - convince yourself that AD is a best response to AD in each subgame,
  - bad payoffs $\frac{1}{1-\delta}$.

Infinitely repeated prisoner's dilemma

Strategies and best response

Payoffs from strategy σ_i against Grim Trigger $\sigma_{-i} = GT_{-i}$

- if $\sigma_i = AC_i$, the payoff is equal to $x + \delta x + \delta^2 x + ... = \frac{1}{1-\delta}x$,
- if $\sigma_i = GT_i$, the payoff is equal to $x + \delta x + \delta^2 x + ... = \frac{1}{1-\delta}x$,
- if $\sigma_i = AD_i$, the payoff is equal to $y + \delta 1 + \delta^2 1 + ... = y + \frac{\delta}{1-\delta}$.
  - one successful defection with great payoff y, followed by eternal punishment with crappy payoff 1.

Infinitely repeated prisoner's dilemma

Strategies and best response

- Which one GT or AD is a better response against GT?
- GT is better if
$$y + \frac{\delta}{1-\delta} \leq \frac{1}{1-\delta}x.$$
- After some algebra, we can check that the inequality is equivalent to
$$\delta \geq \frac{y-x}{y-1}.$$
(check! Also convince yourself that the right-hand side is strictly smaller than 1.)
- The last inequality means that the players must be suficiently patient and value future interactions sufficiently strongly.

Plan

Infinitely repeated prisoner's dilemma

Grim trigger equilibrium

Cooperation and threat of punishment

Grim trigger equilibrium

Theorem

The profile of grim trigger strategies (GT_i, GT_{-i}) *is a subgame perfect equilibrium if and only if*

$$y + \delta \frac{1}{1-\delta} \leq \frac{1}{1-\delta} x. \tag{1}$$

- There is an equilibrium in which players play (C, C) if and only if
  - payoff from AD is smaller than the payoff from GT, or
  - players are sufficiently patient and value future interactions sufficiently strongly.

Grim trigger equilibrium

Proof

- If the inequality is not satisfied, then AD is a better response than GT and (GT, GT) cannot be an equilibrium.
- Suppose from now on that the inequality is satisfied and the other player plays GT .
- We are going to check that GT is a best response at any history h.
- Two types of histories:
  - h with some D played in the past,
  - $h = (CC, CC, ..., CC)$.

Grim trigger equilibrium

Proof

- Suppose that h with some D played in the past.
- After such a history, the GT strategy of the other player tells her to always play D.
- But then, one cannot do better than respond with D.
- But this is what GT says to do. Hence, GT is a best response after such a history.

Grim trigger equilibrium
Proof

- Suppose that $h = (CC, CC, ..., CC)$ is a t-period history.
- Consider a deviation to play D for the first time in period $s \geq t$.
  - the payoff from such a deviation in periods $t, ..., s-1$ is x,
  - in period s is y,
  - in periods $s+1, s+2, ...$ is at most 1.
- Hence, the overall payoff from the deviation is not higher than

$$\underbrace{\delta^{t-1}x + ... + \delta^{s-2}x}_{\text{before period } s} + \underbrace{\delta^{s-1}y}_{\text{in period } s} + \underbrace{\delta^{s}1 + \delta^{s+1}1 + ...}_{\text{after period } s}.$$

- The overall payoff from following GT is

$$\underbrace{\delta^{t-1}x + ... + \delta^{s-2}x}_{\text{before period } s} + \underbrace{\delta^{s-1}x}_{\text{in period } s} + \underbrace{\delta^{s}x + \delta^{s+1}x + ...}_{\text{after period } s}.$$

Grim trigger equilibrium

Proof

- We compare both payoffs. *GT* has a higher payoff if

$$\underbrace{\delta^{t-1}x+...+\delta^{s-2}x}_{\text{before period } s}+\underbrace{\delta^{s-1}y}_{\text{in period } s}+\underbrace{\delta^{s}1+\delta^{s+1}1+...}_{\text{after period } s}$$
$$\leq\underbrace{\delta^{t-1}x+...+\delta^{s-2}x}_{\text{before period } s}+\underbrace{\delta^{s-1}x}_{\text{in period } s}+\underbrace{\delta^{s}x+\delta^{s+1}x+....}_{\text{after period } s}$$

- The first terms are identical. Subtracting them from both sides yields

$$\underbrace{\delta^{s-1}y}_{\text{in period } s}+\underbrace{\delta^{s}1+\delta^{s+1}1+...}_{\text{after period } s}\leq\underbrace{\delta^{s-1}x}_{\text{in period } s}+\underbrace{\delta^{s}x+\delta^{s+1}x+....}_{\text{after period } s}$$

  - the left-hand is equal to $\delta^{s-1}\left(y+\frac{\delta}{1-\delta}\right)$.
  - the right hand is equal to $\delta^{s-1}\left(\frac{1}{1-\delta}x\right)$.
  - By dividing by δ^{s-1}, we obtain that the left-hand side is smaller because of inequality (1).

Plan

Infinitely repeated prisoner's dilemma

Grim trigger equilibrium

Cooperation and threat of punishment

Cooperation and threat of punishment

- The Theorem says that Cooperation can exists in equilibrium if and only if
$$y + \delta \frac{1}{1-\delta} \leq \frac{1}{1-\delta} x$$
- Let's look at the above inequality in a more detail.
It is equivalent to

$$\underbrace{y}_{\text{payoff today from defection}} + \underbrace{\delta \frac{1}{1-\delta}}_{\text{future payoff after defection}}$$

$$\leq \underbrace{x}_{\text{payoff today from Cooperation}} + \underbrace{\delta \frac{1}{1-\delta} x}_{\text{future payoff after Cooperation}}$$

Cooperation and threat of punishment

- Or, after moving some terms, Cooperation requires that

$$\underbrace{y - x}_{\text{today payoff gain from defection}} \quad \leq \quad \underbrace{\delta \frac{1}{1-\delta} x - \delta \frac{1}{1-\delta}}_{\text{future payoff loss from defection}} \quad .$$

- In other words, cooperation is possible, if the today's gain from defection is smaller than the tomorrow's payoff loss for defection.
  - threat of punishment stops the defection.

Cooperation and threat of punishment
General Idea

- More generally, Cooperation requires that

  Today's gain from defection $\leq \delta$ [future payoff loss after defection].

- Anything that increases the left-hand side and reduces the right hand-side is bad news for cooperation.
- Cooperation is less likely if
  - today's gain from defection is very tempting
  - discount factor is small,
  - future punishment is small.
- OTOH, to maximize cooperation, minimize left-hand and maximize right-hand

Cooperation and threat of punishment

General Idea

- Or, after moving some terms, Cooperation requires that

  $$\text{Today's gain from defection} \leq \delta \left[\text{future payoff loss from defection}\right].$$

- Applications
  - limited punishment,
  - entry costs.

Cooperation and threat of punishment

Applications: Limited punishment

- The problem with grim trigger is that it is very grim. It punishes forever. Shorter punishments?
- K-punishment strategy:
  - the strategy moves between the normal and punishment phases:
  - Normal phase: Play C. If there is no deviation, continue in the normal phase. If there is a deviation, go to punishment phase.
  - Punishment phase: Plac D for K periods. After that, move to normal phase.
  - the game starts in the normal phase.

Cooperation and threat of punishment

Applications: Limited punishment

- When is the profile of K-punishments strategies an equilibrium?
- Example: suppose that $y = 5, x = 2$. We show $1-$ or $2-$ period of punishments are not enough to sustain the coooperation.
- To see why, notice that if no player deviates deviation, expected payoff is

$$x + \left(\delta + \delta^2 + ...\delta^K\right) x + \delta^{k+1} x + \delta^{k+2} x + ...$$

If player i deviates in the first period, and follow the strategies from tomorrow on, his payoff is

$$y + \left(\delta + \delta^2 + ...\delta^K\right) 1 + \delta^{k+1} x + \delta^{k+2} x + ...$$

Cooperation and threat of punishment

Applications: Limited punishment

- For the 2-punishment strategies to be an equilibrium, it must be that the latter is smaller than the former, which implies that

$$y + \left(\delta + \delta^2 + ...\delta^K\right) 1 < x + \left(\delta + \delta^2 + ...\delta^K\right) x,$$

  If $y = 5, x = 2$ and $K = 1$ or $K = 2$, then the above inequality fails *for any discount factor*!
- Lesson: The limited punishment strategy reduces future payoff loss from defection.
  - In order to ensure the equilibrium behavior, you should punish as hard as you can for any deviation!

Cooperation and threat of punishment

Applications: Entry costs

- One way to punish a cheater is to break up the relationship.
- But anew problem appears:
  If the cheater can immediately start a new relationship with somebody else, then the punishment is not big.
  And a small punishment means no cooperation.
- One way to deal with the new problem is to put a barriers to entry or re-entry into a new relationship.
- Three examples:
  - long-distance Maghrebi trade,
  - diamond trade,
  - celebratory meals.

Cooperation and threat of punishment

Applications: Entry costs in long-distance medieval trade

- Middle Ages, Mediterranean Sea, long-distance trade (spices, gold) is extremely proftiable.
  The whole area is separated into multiple jurisdictions. No common legal authority.
- If I send a ship packed with wheat and spices from Alexandria in Egypt to Malaga in Andalusia, how can I trust that somebody will send it back with gold and wool? Instead of stealing it for themselves?
- A significant fraction of this trade was dominated by a very small group of 50-100 families, all interrelated, of Maghrebi Jews (probably of Iranian origin).
  - the members were spread around Mediterranean,
  - the trade was done only between members of the group,
  - infinite barrier of entry: nobody outside the group could trade,
  - if any member cheated, it would be punished by exclusion from the group,
  - because future trade opportunities were very profitable, cheating was very rare.

Cooperation and threat of punishment

Applications: Entry costs in long-distance medieval trade

- Similar phenomenon in diamond market in Netherlands late XX century or nor in XXI century Punjab.
  - in both cases, the trade requires lots of trust (that the other party does not misrepresent the diamond quality, etc.),
  - trade is restricted to small ethnic, interrelated groups (a tiny religious sect or a subcaste),
  - trust is guaranteed by exclusion threat and infinite barrier to entry.

Cooperation and threat of punishment

Applications: Celebratory meals as entry cost

- Apocryphical story: Any business relationship in Souq (market) in Middle-Easter cities starts with a large party. Lots of expensive food, drink, valuable plates are broken.
- Why?
  It makes entry costly.
  If you cheat and you want to restart the relationship, you have to go through the wasteful and costly party again.
  Better not to cheat.

- Can you think of another relationship that starts with extravagant, and very costly party?