Section Notes 4

Wonbin Kang, Sam Richardson

February 17, 2011

Agenda

1. Infinitely Repeated Games

2. Midterm Review

1 Infinitely Repeated Games

Normally when we think about repeated games, we think about infinitely repeated prisoner's dilemma games because we're interested in when and how cooperation can be maintained in equilibrium. Let's look at the following example:

1.1 Discounting

Recall from class that the value of \$1 in period t is equal to $\frac{\$1}{1+\beta} \cdot (1-z) = \delta \cdot \$1$ in period t-1. The discount factor $\delta = \frac{1-z}{1+\beta}$ is composed of both an interest rate (which captures the time value of money) and the probability that a game ends after a certain number of periods. Note the following characteristics:

1. If the probability of a game ending at a certain period is fixed at z, then $z_1 = 1-z$; $z_2 = (1-z)z$; $z_3 = (1-z)^2 z$; ... ; $z_t = (1-z)^{t-1} z$; ...; where $z_t = \Pr(\text{the game ends at time } t)$.

2. $\sum_{t=0}^{\infty} z_t = 1$

1.2 Infinitely Repeated PD with Deterministic Discount Factor

Example 1.1. Payoffs are:

	Cooperate	Defect
Cooperate	3,3	0,5
Defect	5,0	1,1

; discount factor $\delta_t = \frac{1-z_t}{1+r}$, where $z_{it} = \Pr(player\, i\, ends\, game\, at\, period\, t)$, $z_t = z_{1t} + z_{2t}$, and r is the interest rate.[1]

Let's first assume that the probability of a game ending at any given period is equal to z, or $z_1 = z_2 = ... = z_t = ... = z$. Also assume that $r = 0$ for this and the following game. Given these assumptions, what level of z will sustain cooperation when both players are playing a grim trigger strategy?

1. We assumed that both player 1 and 2 are playing a *grim trigger strategy.* Therefore, the candidate strategy profile is $\vec{\sigma}^* = [s_1, s_2]$, where:

$$s_1 = s_2 = \begin{cases} "C" & in\, first\, period \\ "C" & if\, (c,c)\, in\, all\, previous\, periods \\ "D" & otherwise \end{cases}$$

[2]

2. Single Deviation Principle: If player i has some profitable deviation, then there is a profitable deviation that differs from s_i in only one period. This means that we need to check each possible deviation only once.

3. The payoff for player 1 (and player 2) under the strategy profile above starting at any time period t[3] is:

$$u_1(s_1, s_2) = 3 + 3\delta + 3\delta^2 + ... = \sum_{t=0}^{\infty} 3\delta^t = 3 \cdot \frac{1}{1-\delta}$$

[4]

4. Now let's consider the payoff of player 1, if she decides to deviate in period t:

$$u_1(s_1^{dev}, s_2) = 5 + 1 \cdot \delta + 1 \cdot \delta^2 + ... = 5 + \frac{\delta}{1-\delta}$$

[1] The probability of the game ending at any given period t is equal to the probability that player 1 ends the game OR player 2 ends the game.

[2] The "otherwise" is the punishment strategy. Recall in the Cold War game from the first semester that we had "Nuke" as a punishment option. Here, the punishment is "D" for all remaining periods.

[3] The reason that I say player 1 deviates at time t is to emphasize the single deviation principle above.

[4] The sum of a geometric series where $\alpha \in (0,1)$ can be solved by:

$$\begin{aligned} V &= 1 + \alpha + \alpha^2 + \alpha^3 + ... \quad (1) \\ \alpha V &= \alpha + \alpha^2 + \alpha^3 + \alpha^4 + ... \quad (2) \end{aligned}$$

Subtracting equations 1 and 2, we can find the sum of a geometric series:

$$(1-\alpha)V = 1 \Rightarrow V = \frac{1}{1-\alpha}$$

Therefore, in order to sustain cooperation, we need to have $u_1(s_1, s_2) \geq u_1(s_1^{dev}, s_2)$:

$$\frac{3}{1-\delta} \geq 5 + \frac{\delta}{1-\delta} \Rightarrow \delta = 1 - z \geq \frac{1}{2} \Rightarrow z \leq \frac{1}{2}.$$

What have we just shown? Have we shown that the Grim Trigger Strategy is a NE or an SPNE? We've just shown that this is a NE. Now we want to show that this is a SPNE. How? Based on the definition of SPNE, we need to show that the Grim Trigger strategy is a NE in each of the subgames of the infinitely repeated stage game: here the stage game is a Prisoners' Dilemma. What are the subgames of the infinitely repeated Prisoners' Dilemma? Note that each subgame of the infinitely repeated Prisoners' Dilemma is equal to the entire game.

Pick a stage t. The strategies that precede t are either "Cooperate" has been played in all stages by both players, in which case the Grim Trigger strategy is in play at stage t, which we know is a NE; or there was a "Deviate" somewhere before stage t, in which case the NE punishment strategy is in play, which we know is also a NE for that subgame. The conclusion is that a NE strategy is being used by both players in each and every subgame of the infinitely repeated game.

1.3 Infinitely Repeated PD with Stochastic Discount Factor

Example 1.2. Now let's assume that the interest rate is equal to zero (0) and the probability of a game ending at any given period t is equal to $z_t = z_{1t} + z_{2t}$, where $z_{it} \sim^{iid} U\left[0, \frac{1}{4}\right]$. Given these assumptions, what level of z_{1t}[5] will sustain cooperation when both players are playing a grim trigger strategy? Let's further assume that at the beginning of period t, each player i learns z_{it}, but only knows the distribution of the other player's probability of ending the game.

1. What is the expected discount factor at period t?

$$\mathbb{E}\left[\delta_t\right] = \mathbb{E}\left[1 - z_t\right] = 1 - \mathbb{E}\left[z_{1t}\right] - \mathbb{E}\left[z_{2t}\right] = 1 - \frac{1}{4} = \frac{3}{4}.$$

   This is also the probability that the game does not end at period t.

2. For each player what is the expected probability of the game ending given player 1's own draw of z_{1t}? In other words, what is the expected discount factor

$$\mathbb{E}\left[z_t|z_{1t}\right] = \mathbb{E}\left[z_{1t} + z_{2t}|z_{1t}\right] = z_{1t} + \mathbb{E}\left[z_{2t}|z_{1t}\right] = z_{1t} + \mathbb{E}\left[z_{2t}\right] = z_{1t} + \frac{1}{8}$$

$$\therefore \mathbb{E}\left[\delta_t|z_{1t}\right] = \mathbb{E}\left[1 - z_t|z_{1t}\right] = 1 - \mathbb{E}\left[z_t|z_{1t}\right] = \frac{7}{8} - z_{1t}$$

[5] This will generalize to player 2.

3. The payoff for player 1 (and player 2) under the strategy profile above starting at any time period t is:

$$\begin{aligned} U_1(s_1, s_2|z_{1t}) &= 3 + 3 \cdot \mathbb{E}[\delta_t|z_{1t}] + 3 \cdot \mathbb{E}[\delta_t|z_{1t}] \cdot \mathbb{E}[\delta_t] + 3 \cdot \mathbb{E}[\delta_t|z_{1t}] \cdot \mathbb{E}[\delta_t]^2 + ... \\ &= 3 + \frac{3 \cdot \mathbb{E}[\delta_t|z_{1t}]}{1 - \mathbb{E}[\delta_t]} = 3 + \frac{3\left(\frac{7}{8} - z_{1t}\right)}{\frac{1}{4}} \end{aligned}$$

4. Now let's consider the payoff of player 1, if she decides to deviate in period t:

$$\begin{aligned} U_1(s_1^{dev}, s_2|z_{1t}) &= 5 + 1 \cdot \mathbb{E}[\delta_t|z_{1t}] + 1 \cdot \mathbb{E}[\delta_t|z_{1t}] \cdot \mathbb{E}[\delta_t] + 1 \cdot \mathbb{E}[\delta_t|z_{1t}] \cdot \mathbb{E}[\delta_t]^2 + ... \\ &= 5 + \frac{\mathbb{E}[\delta_t|z_{1t}]}{1 - \mathbb{E}[\delta_t]} = 5 + \frac{\left(\frac{7}{8} - z_{1t}\right)}{\frac{1}{4}} \end{aligned}$$

5. Therefore, cooperation is sustainable if:

$$3 + \frac{3\left(\frac{7}{8} - z_{1t}\right)}{\frac{1}{4}} \geq 5 + \frac{\left(\frac{7}{8} - z_{1t}\right)}{\frac{1}{4}} \Rightarrow z_{1t} \leq \frac{5}{8}$$

6. But what did we say about the initial distribution? $z_{it} \sim^{iid} U\left[0, \frac{1}{4}\right]$. Therefore, cooperation is always sustainable. For a more interesting case where cooperation can unravel after a certain draw, consider the case where $z_{it} \sim^{iid} U[0, 0.42]$. Try it and see what happens.

1.4 The Grim Trigger Strategy[6]

Why do we care so much about Grim Trigger strategies? The reason is that they always exist in dynamic or repeated games. Consider the following theorem (the proof of which can be found in Fudenberg and Tirole):

Theorem 1.3. *Suppose $\sigma^* \in S$ is a pure strategy SPNE, with the distribution of on the equilibrium path actions (a) given by $p(a|\sigma^*)$. Then, there exists a discount factor $\bar{\delta} \in (0, 1)$ such that for all $\delta > \bar{\delta}$, there exists a SPNE σ^{**} with $p(a|\sigma^*) = p(a|\sigma^{**})$ and σ^{**} involves a continuation payoff of $U_i^{minmax}(\sigma_i, \sigma_{-i}^*)$ to player i, if player i is the first to deviate from σ^{**} at date t after some history $h^{t-1} \in H^{t-1}$.*

Where,

$$U_i^{minmax}(\sigma_i, \sigma_{-i}^*) = \min_{\sigma_{-i}^* \in S_{-i}} \max_{\sigma_i \in S_i} \sum_{s=0}^{\infty} \delta^s u_i(\sigma_{i,t+s}, \sigma_{-i,t+s}^*).$$

[6] Only for those interested. For those not interested, just remember that in an infinitely repeated game, you'll be able to find a grim trigger strategy which is SPNE.

In short, if there exists a SPNE in an infinitely repeated game, then for a high enough discount factor, there exists a strategy profile where the first player to deviate receives her minmax continuation payoff which is also SPNE. Note that it could be the case that $\sigma^* = \sigma^{**}$.

This is why we look at Grim Trigger Strategies. Note that the theorem above was written for a general dynamic model. It can be simplified for purposes of the infinitely repeated game.

2 Recursive Methods

There is no way to provide you with a thorough review of recursive methods and Bellman equations, however, the key insights are fairly simple and may be helpful to you in solving some of the practice problems.

Assume that the discounted value of a stream of payoffs at time t given some domain is equal to $V(\cdot|t)$; let the instantaneous payoff be equal to $u(\cdot)$; and the discount rate is equal to δ. Further assume that the relevant domain of the function $V(\cdot|t)$ is the strategy space of all of the players playing the infinite game, therefore,$V : \times_{i \in N} \Sigma_i \rightarrow \mathbb{R}$. Then we can use the following recursive equation:[7]

$$V(\sigma_i, \sigma_{-i}|t) = u_i(\sigma_i, \sigma_{-i}) + \delta \mathbb{E}_{t+1}\left[V(\sigma_i, \sigma_{-i}|t+1)\right] \Rightarrow V = u + \delta \mathbb{E}_{t+1}\left[V\right].$$

The key is understanding that the value function today is the same as the expected value function tomorrow. Using the terminology of recursive methods: The value function today is equal to the instantaneous payoff today and the discounted continuation value.

[7]Note that we could write this as: $V(\sigma_i, \sigma_{-i}|t) = u_i(\sigma_i, \sigma_{-i}) + \delta V(\sigma_i, \sigma_{-i}|t+1) \Rightarrow V = u + \delta V$ if there are no stochastic elements that would determine the value function in the future.