Economic Theory: Final Exam

Repeated Games, Signaling, and Reputations

Niccolò Lomys*

April 8, 2019

Question 1 [17 points]

1. [7 points] Consider the infinitely repeated prisoners' dilemma for which the reward matrix of the stage game is given by

		Player 2: C	Player 2: D
Player 1	C	$2, 2$	$-c, b$
	D	$b, -c$	$0, 0$

   where $b > 2$, $c > 0$, and $b - c < 4$. Suppose $\delta \geq (b-2)/2$. Show that the set $\{(0, 0), (2, 2)\}$ is pure-action self-generating.

2. [10 points] Consider the infinitely repeated prisoners' dilemma for which the reward matrix of the stage game is given by

		Player 2: C	Player 2: D
Player 1	C	$1, 1$	$-1, 2$
	D	$2, -1$	$0, 0$

   Find a discount factor $\delta \in [0, 1)$ for which the set $\{(v, v) \in \mathbb{R}^2 : v \in [0, 1]\}$ is pure-action self-generating. Explain your reasoning.

Solution

1. Set $W := \{(0, 0), (2, 2)\}$.

   (i) The payoff vector $(0, 0)$ corresponds to the payoff from the stage game Nash equilibrium (D, D). Thus, $(0, 0)$ is pure-action decomposable on W for any δ (as in the Lecture Notes, use a constant γ mapping each strategy profile into $(0, 0)$).

*Toulouse School of Economics, University of Toulouse Capitole; niccolo.lomys@tse-fr.eu.

(ii) The action profile (C, C) is enforceable on W for $\delta \geq (b-2)/2$. To see this, consider $\gamma \in W^A$ defined pointwise as $\gamma(C, C) := (2, 2)$ and $\gamma(a) := (0, 0)$ for $a \neq (C, C)$. Clearly, for $i = 1, 2$, we have $2 = (1-\delta)u_i(C, C) + \gamma_i(C, C)$. Moreover, (C, C) is enforceable on W if

$$\begin{aligned} 2 &\geq \max_i \left[(1-\delta)u_i(a_i = D, a_{-i} = C) + \delta\gamma_i(a_i = D, a_{-i} = C)\right] \\ &= (1-\delta)b, \end{aligned}$$

which is satisfied for any $\delta \geq (b-2)/2$. Therefore, for any $\delta \geq (b-2)/2$, the action profile (C, C) is enforceable on W and the payoff vector $(2, 2)$ is pure-action decomposable on W.

From the arguments in (i) and (ii) it follows immediately that W is a pure-action self-generating set of payoffs for $\delta \geq (b-2)/2$, as was to be shown.

2. Claim. For $\delta = 2/3$, the set $W := \{(v, v) \in \mathbb{R}^2 : v \in [0, 1]\}$ is pure-action self-generating.

Proof. We distinguish two cases.

(i) $v \in [0, 2/3]$.
Pick any $v \in [0, 2/3]$. Such v is pure-action decomposed by (D, D) and $\gamma \in W^A$ defined pointwise as $\gamma(D, D) := (3v/2, 3v/2)$ and $\gamma(a) := (0, 0)$ if $a \neq (D, D)$ (Note that $\gamma(D, D) \in W$ because $v \in [0, 2/3]$). Too see this, note that

$$(1-\delta)u_i(D, D) + \delta\gamma_i(D, D) = 0 + \frac{2}{3}\frac{3v}{2} = v.$$

Moreover,

$$v \geq (1-\delta)u_i(a_i = C, a_{-i} = D) + \delta\gamma_i(a_i = C, a_{-i} = D) = \frac{1}{3}(-1) + 0 = -\frac{1}{3}.$$

(ii) $v \in (2/3, 1]$.
Pick any $v \in (2/3, 1]$. Such v is pure-action decomposed by (C, C) and $\gamma \in W^A$ defined pointwise as $\gamma(C, C) := ((3v-1)/2, (3v-1)/2)$ and $\gamma(a) := (0, 0)$ if $a \neq (C, C)$ (Note that $\gamma(C, C) \in W$ because $v \in (2/3, 1]$). Too see this, note that

$$(1-\delta)u_i(C, C) + \delta\gamma_i(C, C) = \frac{1}{3} + \frac{2}{3}\frac{3v-1}{2} = v.$$

Moreover,

$$v \geq (1-\delta)u_i(a_i = D, a_{-i} = C) + \delta\gamma_i(a_i = D, a_{-i} = C) = \frac{1}{3}2 + 0 = \frac{2}{3},$$

since, by assumption, $v \in (2/3, 1]$.

Question 2 [17 points]

1. [9 points] Consider a repeated game with imperfect public monitoring G^δ and restrict attention to public strategies. Define the notions of one-shot deviation and profitable one-shot deviation. State and prove the One-Shot Deviation Principle.

2. [8 points] State the Folk Theorem for repeated games with imperfect public monitoring. Explain in words the role of all its assumptions.

Solution

1. <u>One-Shot Deviation.</u> Let H be the set of public histories. A one-shot deviation for player i from the public strategy σ_i is a public strategy σ_i' such that $\sigma_i'(h') = \sigma_i(h')$ for all $h' \in H$ with $h' \neq h$, and $\sigma_i'(h) \neq \sigma_i(h)$.

   <u>Profitable One-Shot Deviation.</u> Fix a public strategy profile σ. A one-shot deviation σ_i' from strategy σ_i is profitable if, at the public history $h \in H$ for which $\sigma_i'(h) \neq \sigma_i(h)$,

   $$v_i(\sigma_i|_h, \sigma_{-i}|_h) > v_i(\sigma|_h).$$

   <u>One-Shot Deviation Principle.</u> A public strategy profile σ is a perfect public equilibrium of G^δ if and only if no player has a profitable one-shot deviation.

   <u>Proof.</u> Necessity ("only if") is immediate: if σ is a PPE, then there are no profitable deviations in public strategies, whether one-shot or not. For sufficiency ("if") we need to show that, whenever a profitable deviation exists, we can define a profitable one-shot deviation (all of this in public strategies). Suppose then that the public strategy profile σ is not a PPE. Then, there exists $i \in N$, a public history $h^t \in H$, and a public strategy σ_i' such that

   $$v_i(\sigma_i'|_{h^t}, \sigma_{-i}|_{h^t}) - v_i(\sigma|_{h^t}) > 0.$$

   Because payoffs are discounted, it follows that there exists $T \in \mathbb{N}$ such that, defining σ_i'' by $\sigma_i''(h^\tau) := \sigma_i(h^\tau)$ for all public histories $h^\tau \in H$ with $\tau \geq t + T$ and $\sigma_i''(h^\tau) := \sigma_i'(h^\tau)$ otherwise, we have

   $$v_i(\sigma_i''|_{h^t}, \sigma_{-i}|_{h^t}) - v_i(\sigma|_{h^t}) > 0.$$

   That is, the public strategy σ_i'' is a profitable deviation that only differs from σ_i at finitely many public histories. We now proceed by induction. Consider all public histories $h \in H^{t+T-1}$. We have two possibilities:

   - Either $v_i(\sigma_i''|_{h^t}, \sigma_{-i}|_{h^t}) - v_i(\sigma|_{h^t}) > 0$ for some $h \in H^{t+T-1}$. In this case, define σ_i''' as the one-shot deviation from σ_i at h, with $\sigma_i'''(h) := \sigma_i''(h)$. This is a profitable one-shot deviation, and so we are done.
   - Or $v_i(\sigma_i''|_{h^t}, \sigma_{-i}|_{h^t}) - v_i(\sigma|_{h^t}) \leq 0$. In this case, define σ_i''' by $\sigma_i'''(h^\tau) = \sigma_i(h^\tau)$ for all $h^\tau \in H$ with $\tau \geq t + T - 1$ and $\sigma_i'''(h^\tau) := \sigma_i'(h^\tau)$ otherwise. Again, σ_i''' is a profitable deviation that only differs from σ_i at finitely many public histories.

Repeat the above process iteratively until the difference is positive.

2. Folk Theorem. Let $\mathcal{F}^*$ be the set of feasible and individually rational rewards. Suppose: (i) $\dim(\mathcal{F}^*) = n$; (ii) every pure action profile has individual full rank; (iii) for any $i, j \in N$ with $i \neq j$, there exists an action profile which has pairwise full-rank for i and j. Then, for any closed set $W \subseteq Int(\mathcal{F}^*)$, there exists $\underline{\delta} < 1$ such that $W \subseteq E_\delta$ for all $\delta \in (\underline{\delta}, 1)$.

   Discussion.

   - Assumption (i) ensures that players need not be simultaneously minmaxed. More specifically, it ensures that player-specific punishments are possible and that, at the same time, punishers can be rewarded for punishing. When the assumption fails, punishing a deviator may also punishes someone else (i.e. a punisher). In such a circumstance, it would be difficult to incentivize someone else to punish the deviator.
   - Assumption (ii) ensures that the signals generated by any (possibly mixed) action of player i, say action α_i, are statistically distinguishable from those generated by any other of his actions, say $\alpha_i' \neq \alpha_i$. That is, player i's actions can be distinguished from one another.
   - Assumption (iii) ensures that a deviation by player i leads to a distribution over signals that is different from that induced by any deviation by player $j \neq i$. Thus, if everyone is playing action profile α, then not only can i's actions be distinguished and j's actions be distinguished (by individual full rank), but i's actions can be distinguished from j's actions (by pairwise full rank). You can think of this just like statistical identification. Here, i's action is the parameter. To identify it, you need the probability distribution over observables to change when it changes. Pairwise full rank is what you need to identify both a_i and a_j at the same time—we only require this for some profile played by the others, $k \neq i, j$.

Question 3 [16 points]

Consider the *product choice* game for which the payoff matrix is given by

		Consumer	
		h	ℓ
Firm	H	$2, 4$	$0, 2$
	L	$3, 0$	$1, 1$

The narrative is as follows. The firm can exert either high effort, H, or low effort, L, in the production of its output. The consumer can buy either a high-priced product, h, or a low-priced product, ℓ. Moves are simultaneous. The consumer prefers the high-priced product if the firm has exerted high effort, but prefers the low-priced product if the firm has not. The firm prefers that consumers purchase the high-priced product. In a simultaneous move game, however, the firm cannot observably choose effort before the consumer chooses the product. Because high effort is costly, the firm prefers low effort, no matter the choice of the consumer. The stage game has a unique NE, (L, ℓ).

As in the canonical reputation model with behavioral types seen in class, suppose the firm is a long-lived player, playing the game at times $t = 0, 1, 2, \ldots$ with discount factor $\delta \in (0,1)$. The firm faces a successions of short-lived consumers, with a new consumer in each period. Each consumer observes the actions in all prior periods. The firm can be of two types, ω_n or ω_0: ω_n is the normal type; ω_0 is the Stackelberg type, who always plays H. Let $\mu(\omega_0) \in (0,1)$ be the prior probability that the firm is of type ω_0 and $\mu(\omega_n) = 1 - \mu(\omega_0)$ the prior probability that the firm is of type ω_n.

(a) [9 points] Fix a Nash equilibrium of this repeated game with incomplete information. Let h^t be a positive probability period-t history in which the firm always plays H. Provide an upper bound for the number of periods in h^t in which a consumer played ℓ. Prove your statement.

(b) [7 points] Show that, for all $\varepsilon > 0$, there exists $\overline{\delta} \in (0,1)$ such that for all $\delta \in (\overline{\delta}, 1)$, the normal type's payoff in any Nash equilibrium is at least $2 - \varepsilon$.

Solution

(a) <u>Claim.</u> The number of periods in h^t in which a consumer played ℓ is no larger than

$$k^* := -\frac{\log \mu(\omega_0)}{\log 3}.$$

<u>Proof.</u> Let q_τ be the probability that the firm plays H in period τ condition on h^τ. In equilibrium, if consumer τ plays ℓ, then $q_\tau \leq 1/3$ (if $q_\tau > 1/3$, it is not a best reply for the consumer to play ℓ). An upper bound on the number of periods in h^t in which a consumer played ℓ is thus

$$k(t) := \#\{\tau \in \mathbb{N} : \tau < t \text{ and } q_\tau \leq 1/3\},$$

the number of periods in h^t where $q_\tau \leq 1/3$. (This is an upper bound, and not the actual number, since the consumer is indifferent if $q_\tau = 1/2$).

Let $\mu_\tau := \mathbb{P}(\omega_0 \mid h^\tau)$ be the posterior probability assigned to ω_0 after h^τ, where $\tau < t$ (so that h^τ is an initial "segment" of h^t). Since h^t is a history in which the firm always plays H, the firm plays H is period $\tau + 1$ and μ_τ is updated to $\mu_{\tau+1}$ as follows:

$$\begin{aligned}\mu_{\tau+1} = \mathbb{P}(\omega_0 \mid h^\tau, H) &= \frac{\mathbb{P}(\omega_0, H \mid h^\tau)}{\mathbb{P}(H \mid h^\tau)} \\ &= \frac{\mathbb{P}(H \mid \omega_0, h^\tau)\mathbb{P}(\omega_0 \mid h^\tau)}{\mathbb{P}(H \mid h^\tau)} \\ &= \frac{\mu_\tau}{q_\tau}.\end{aligned}$$

Thus, for all $\tau \leq t$, we have

$$\mu_\tau = q_\tau \mu_{\tau+1}.$$

Then,

$$\begin{aligned}
\mu(\omega_0) = q_0\mu_1 &= q_0q_1\mu_1 \\
&= \cdots \\
&= \mu_t \prod_{\tau=0}^{t-1} q_\tau \\
&= \left(\mu_t \prod_{\{\tau:\tau<t \text{ and } q_\tau \le 1/3\}} q_\tau\right) \times \left(\mu_t \prod_{\{\tau:\tau<t \text{ and } q_\tau > 1/3\}} q_\tau\right) \\
&\le \left(\frac{1}{3}\right)^{k(t)} \times 1^{t-k(t)} \\
&= \left(\frac{1}{3}\right)^{k(t)}.
\end{aligned}$$

Taking logs,

$$\log \mu(\omega_0) \le k(t) \log \frac{1}{3},$$

and so

$$k(t) \le -\frac{\log \mu(\omega_0)}{\log 3},$$

as was to be shown.

(b) The normal type can guarantee himself histories in which he always plays H. Such behavior yields payoffs that are no larger than the firm's NE payoffs in any equilibrium (if not, the firm has an incentive to deviate). Since there is positive probability that the firm is the Stackelberg type, the history resulting from always playing H has positive probability. Applying the result in part (a) yields a lower bound on the normal types payoff of

$$\sum_{t=0}^{k^*-1}(1-\delta)\delta^\tau(0) + \sum_{t=k^*}^{\infty}(1-\delta)\delta^t 2 = 2\delta^{k^*}.$$

This can be made arbitrarily close to 2 by choosing δ sufficiently close to 1.