ECON0059: **Advanced Microeconomic Theory: Part 2**

A Summary

Duarte Gonçalves*

University College London

1. Normal-Form and Extensive-Form Games

- **Normal-Form Games**: $\Gamma = \langle I, S, u \rangle$

  1. Who is playing: I, set of players; a specific player i, $i \in I$;
     $-i := I \setminus \{i\}$, players other than player i (player i's "opponents")
  2. What each player can do: S_i, set of feasible strategies for each player;
     $s_i \in S_i$, one particular strategy of player i;
     $S := \times_{i \in I} S_i$ set of feasible strategy profiles; a particular strategy profile $s \in S$, $s = (s_i)_{i \in I}$, specifying one strategy for each player
     $s_{-i} \in S_{-i} := \times_{j \in -i} S_j$ a specific strategy profile of player i's opponents
     A strategy profile s determines an outcome of the game.
  3. What are each player's incentives: $u_i : S \to \mathbb{R}$, player i's payoff function, determining how player i evaluates a specific outcome/strategy profile s;
     $u = (u_i)_{i \in I}$, all players' payoff functions

- **Extensive-Form Games**: $\Gamma = \langle I, \mathscr{A}, H, \mathscr{I}, \rho, u \rangle$ where

  1. I denotes a **set of players**
  2. $\mathscr{A}$ denotes the overall **set of actions**
     These are all the actions that some player (or nature) can take at some point.
  3. H denotes the **set of histories**
     (i) The *empty history* $\emptyset$ is a member of H (the 'starting point' of the game)
     (ii) A *nonempty history* $h \in H$ consists of a (possibly infinite) sequence of actions, $h = (a^1, \dots, a^t) \in \mathscr{A}^t$ for some $t \in \mathbb{N} \cup \{\infty\}$
     (what has happened thus far)

* Department of Economics, University College London; duarte.goncalves@ucl.ac.uk.

(iii) If a sequence of actions $h^{t+1} = (a^1, ..., a^t, a^{t+1})$ is a possible history, then so is the sequence $h^t = (a^1, ..., a^t)$; h^t is called a *proper subhistory* of h^{t+1}[1]

(iv) A history $h \in H$ is a *terminal history* if there is no $a \in \mathscr{A}$ such that $(h, a) \in H$ or if it is an infinite sequence of actions.

The *set of terminal histories* is denoted by $T \subset H$.

A history which is not terminal ($h \in H \setminus T$) is called a *nonterminal history*.

(The intuition is that, following a nonterminal history h, there is some player who can still take some action $a \in \mathscr{A}$)

(v) The *set of feasible actions* following nonterminal history h is defined as $A(h) := \{a \in \mathscr{A} \mid (h, a) \in H\}$. That is, if nonterminal history h occurs, then some player is called upon to play and the player can choose from actions $A(h)$

4. $\mathscr{I} := \{\mathscr{I}_i\}_{i \in I}$, where $\mathscr{I}_i$ denotes player i's **information sets** or *information partition*

   (i) A player's information set corresponds to the set of histories the player called to play cannot distinguish between.

   That is, if $\{h, h'\}$ correspond to an information set of player i, then if either h or h' occur, player i *knows* that one of the two histories occurred (and not some other history), but not which one.

   (ii) For any two histories h, h' belonging to the same information set I_i, the *set of feasible actions is the same*, $A(h) = A(h') =: A(I_i)$.

   The reason for this is that if $A(h) \neq A(h')$, then player i could infer which history occurred just by knowing which actions are available to choose from.

   (iii) A game is said to be of *perfect information* if all information sets are singletons (contain a single history). Otherwise, it is called of *imperfect information*.

5. ρ is a function that determines the **probability with which nature "takes an action"** whenever nature is called upon to move. This is known to the players.

6. $u := (u_i)_{i \in I}$, where each u_i represents player i's **payoff function**, $u_i : T \to \mathbb{R}$

   Payoffs realize after terminal histories

   We will assume that u_i corresponds to a von-Neumann–Morgenstern utility function (Bernoulli index) representing preferences of player i over terminal histories.

- **Strategies in Extensive-Form Games**

[1] It is a *proper* subhistory because a history is also a (nonproper) subhistory of itself.

A *pure strategy* for player i is full contingent plan.

A pure strategy s_i specifies a feasible action *for every information set* I_i at which player i is called upon to play.

The number of pure strategies of a player = product of the number of feasible actions at each information set.

For example: player i chooses at 2 information sets; 3 feasible actions in one, 5 in other, the total number of player i's pure strategies is $3 \cdot 5 = 15$.

A *mixed strategy* for player i is a distribution of a player i's pure strategies.

A *behavioral strategy* for player i specifies a distribution over feasible actions *for every information set* I_i at which player i is called upon to play. In general, it is equivalent to think in terms of mixed and behavioral strategies (Kuhn's theorem).

2. Subgame Perfect Nash Equilibrium

- **Subgame**

(i) always starts at a history h^t corresponding to a singleton information set $\{h^t\}$;

(ii) includes all histories h^{t+h} following it ($h^{t+h} = (h^t, a_t, \ldots, a_{t+h-1})$);

(iii) never 'cuts across' information sets (players always know that are at the subgame).

The game is a subgame of itself. A proper subgame of a game is any subgame different from the game itself.

- **SPNE**: a strategy profile that induces a Nash equilibrium in every subgame.

If the game is finite, there is an SPNE.

An SPNE is a NE.

Any SPNE can be found by (generalized) backward induction; all strategy profiles resulting from (generalized) backward induction are SPNE.

Generalized backward induction: start by the 'smallest' subgames (closest to terminal nodes and not containing any proper subgame) and solve for NE in the the subgames; fix the strategies in the subgames and iterate.

If the game is of perfect information and no two players with the same payoffs at any two terminal histories, there is a unique SPNE. (Zermelo's theorem)

3. Weak Perfect Bayesian Equilibrium and Sequential Equilibrium

- **Belief System** μ: specifies beliefs players hold at each information set I_i about the probability with which a history h in that information set occurred, given that the information set is reached.

  e.g. $\mu(h|I_i)$: prob player i assigns to h having occurred given that the player is called upon to play at information set I_i.

  If $I_i = \{h\}$ is a singleton information set, then $\mu(h|I_i) = 1$.

- **Sequential rationality**: *given a belief system* μ, sequential rationality of σ requires that at every information set I_i, the player called upon to play chooses optimally considering only histories following I_i.

  The player only considers terminal histories that are compatible with being at I_i. For instance, suppose that player 1 moves first and chooses between A,B,C and then player 2 moves, but cannot distinguish between A,B. Sequential rationality requires that, at the information set $\{A,B\}$, player 2 chooses an action at information set $\{A,B\}$ in order to maximize the expected payoff, given a strategy profile that specifies the behavior of everyone else called upon to play after player 2, and given player 2's beliefs about whether A or B was chosen, and ignoring terminal histories in which player 1 chooses C. (The reason is that if player 2 knows that either A or B were chosen, it makes no sense to consider terminal histories in which C was chosen by player 1.)

  Beliefs matter! Different μ can lead to different strategy sequentially rational strategies.

- **Information sets on-path or reached**: an information set I_i is **reached given** σ if there is a positive probability that some history $h \in I_i$ is played with positive probability, $\mathbb{P}(I_i \mid \sigma) > 0$

- **Beliefs Derived through Bayes Rule**: a belief system is **derived through Bayes rule whenever possible given** σ if, for any information set I_i that is reached given σ, beliefs $\mu(I_i)$ over histories in I_i equal the distribution over histories in I_i conditional on I_i, as induced by σ.

Example: Player 1 chooses $a_1 \in \{A, B, C\}$ with prob $\sigma_1(a_1)$, and Player 2 chooses $a_2 \in \{A, B, C\}$ with prob $\sigma_2(a_2)$. Player 3 moves only following (A, B) and (C, A), and cannot distinguish between these, i.e. $I_3 = \{(A, B), (C, A)\}$ is Player 3's information set. Then the prob that I_3 is reached is $\mathbb{P}_\sigma(I_3) = \sigma_1(A)\sigma_2(B) + \sigma_1(C)\sigma_2(A)$.
If $\mathbb{P}_\sigma(I_3) = 0$, then beliefs at I_3 can**not** be derived through Bayes rule given σ.
If $\mathbb{P}_\sigma(I_3) > 0$, then beliefs at I_3 can be derived through Bayes rule given σ, and we would have $\mu((A, B)|I_3) = \mathbb{P}_\sigma((A, B)|I_3) = \frac{\sigma_1(A)\sigma_2(B)}{\sigma_1(A)\sigma_2(B)+\sigma_1(C)\sigma_2(A)}$.

- **Weak Perfect Bayesian Nash Equilibrium** (weak PBE): strategy profile σ *and* a belief system μ such that

  (i) σ is sequentially rational given μ;

  (ii) μ is derived through Bayes rule whenever possible given σ.

  Note: need to define both the strategy profile *and the belief system*.

  If (σ, μ) is wPBE, then σ is a NE.

- **Perfect Bayesian Nash Equilibrium** (PBE): strategy profile σ *and* a belief system μ such that it induces a wPBE in every subgame.

  If (σ, μ) is a PBE, then (σ, μ) is a wPBE and σ is an SPNE.

- **Sequential Equilibrium**: strategy profile σ *and* a belief system μ satisfying

  (i) σ is sequentially rational given μ;

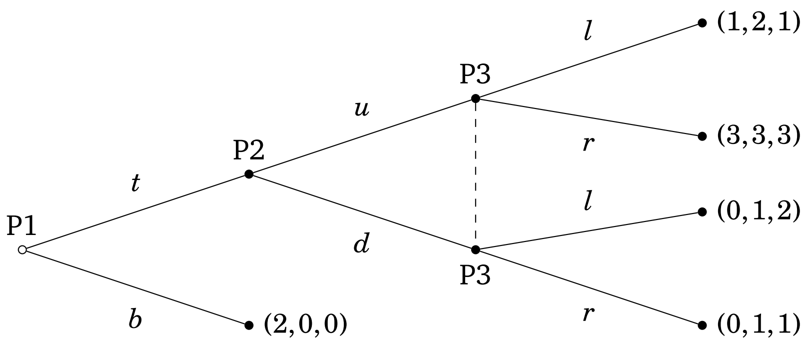
  (ii) there is a sequence of fully mixed strategy profiles[2] σ^n such that σ^n converges to σ ($\sigma^n \to \sigma$),
  and beliefs μ^n derived through Bayes rule given σ^n converge to μ ($\mu^n \to \mu$).

  Note: (ii) implies that μ is derived through Bayes rule whenever possible given σ

  (iii) (σ, μ) SE $\Longrightarrow$ (σ, μ) PBE $\Longrightarrow$ (σ, μ) wPBE and σ SPNE $\Longrightarrow$ σ NE

[2] You can think of this as every feasible action at every information set is played with positive prob.

3.1. A Worked-Out Example



For any belief system the only nontrivial belief is P3's belief at information set $I_3 = \{(tu),(td)\}$. Let μ denote the belief that P3 holds that (tu) is played given I_3.

Note that P3 prefers l to r given μ if $\mu 1 + (1-\mu)2 \geq \mu 3 + (1-\mu)1 \iff 1/3 \geq \mu$. We first solve for wPBE; we'll break the analysis down into three cases.

(1) $\mu > 1/3$.

Then $\sigma_3(r) = 1$ (as the preference for r over l is strict).

By sequential rationality, P2 always strictly prefers u over d, as for any $\sigma_3(l) \in [0,1]$, $\sigma_3(l)2 + (1-\sigma_3(l))3 > \sigma_3(l)1 + (1-\sigma_3(l))1 = 1$. Therefore, P2 always chooses $\sigma_2(u) = 1$.

Then, by sequential rationality, given σ_2, σ_3, P1 strictly prefers to choose t over b. Hence, $\sigma_1(t) = 1$.

As I_3 is reached given σ, we update by Bayes rule: $\mu = \mathbb{P}_\sigma((tu)|I_3) = \frac{\sigma_1(t)\sigma_2(u)}{\sigma_1(t)\sigma_2(u)+\sigma_1(t)\sigma_2(d)} = \frac{1\cdot 1}{1\cdot 1+1\cdot 0} = 1$.

We then have that the unique wPBE if $\mu > 1/3$ is given by $((\sigma_1(t), \sigma_2(u), \sigma_3(l)), \mu) = ((1,1,0),1)$.

(2) $\mu < 1/3$.

Then $\sigma_3(l) = 1$ (as the preference for l over r is strict).

Again, by the same reason, $\sigma_2(u) = 1$. Finally, by sequential rationality, P1 strictly prefers b over t.

As I_3 is not reached given σ, μ cannot be derived through Bayes rule. Hence, if $\mu < 1/3$, then for any $p \in [0, 1/3)$, we have wPBE given by $((\sigma_1(t), \sigma_2(u), \sigma_3(l)), \mu) = ((0,1,1), p)$.

(3) $\mu = 1/3$. By definition, P3 is indifferent between l and r. It is immediate that we have a wPBE given by $((\sigma_1(t), \sigma_2(u), \sigma_3(l)), \mu) = ((0,1,1), 1/3)$.

We also need to check for wPBE in which P3 is mixing.

As, by sequential rationality, $\sigma_2(u) = 1$, then for any $\sigma_1(t) > 0$, I_3 would be reached and μ would need to be derived through Bayes rule given σ_1, σ_2, implying $\mu = \mathbb{P}_\sigma((tu)|I_3) = \frac{\sigma_1(t)\sigma_2(u)}{\sigma_1(t)\sigma_2(u)+\sigma_1(t)\sigma_2(d)} = \sigma_2(u) = 1 \neq 1/3$, which is a contradiction. Hence, we need that $\sigma_1(t) = 0$.

In order for P1 to prefer b to t (and thus choose $\sigma_1(t) = 0$), by sequential rationality we need that $2 \geq \sigma_2(u)(\sigma_3(l)1 + (1-\sigma_3(l))3) + (1-\sigma_2(u))(\sigma_3(l)0 + (1-\sigma_3(l))0) = \sigma_2(u)(\sigma_3(l)1 + (1-\sigma_3(l))3)$.

Since, by sequential rationality $\sigma_2(u) = 1$, we need that $2 \geq \sigma_3(l)1+(1-\sigma_3(l))3 \iff \sigma_3(l) \geq 1/2$.

Again, if $\sigma_1(b) = 1$, I_3 is not reached and so we don't need to derive μ through Bayes rule.

Consequently, if $\mu = 1/3$, then for any $q \in [1/2, 1]$, we have wPBE given by $((\sigma_1(t), \sigma_2(u), \sigma_3(l)), \mu) = ((0, 1, q), 1/3)$.

The above fully describes the set of wPBE of the game. Let's now turn to SE.

Recall that an SE is a wPBE, so we need only to check whether the wPBE we found are SE. For a wPBE (σ, μ) to be SE, we need to find a sequence $\sigma^n = (\sigma_1^n(t), \sigma_2^n(u), \sigma_3^n(l))$ such that
(i) for every n, $\sigma_1^n(t), \sigma_2^n(u), \sigma_3^n(l) \in (0,1)$,
(ii) $\sigma^n \to \sigma$, and
(iii) $\mu^n := \mathbb{P}_{\sigma^n}((tu)|I_3) = \frac{\sigma_1^n(t)\sigma_2^n(u)}{\sigma_1^n(t)\sigma_2^n(u)+\sigma_1^n(t)\sigma_2^n(d)} \to \mu$.

Note that for any such sequence σ^n, we have $\mu^n = \frac{\sigma_1^n(t)\sigma_2^n(u)}{\sigma_1^n(t)\sigma_2^n(u)+\sigma_1^n(t)\sigma_2^n(d)} = \frac{\sigma_2^n(u)}{\sigma_2^n(u)+\sigma_2^n(d)} = \frac{\sigma_2^n(u)}{\sigma_2^n(u)+(1-\sigma_2^n(u))} = \frac{\sigma_2^n(u)}{1} = \sigma_2^n(u)$.

We number the sets of wPBE as above.

(1) As $\sigma_2(u) = 1 = \mu$, then take e.g. $\sigma_1^n(t) = \sigma_2^n(u) = \frac{n}{n+1}$. $((\sigma_1(t), \sigma_2(u), \sigma_3(l)), \mu) = ((1, 1, 0), 1)$ is an SE.

(2) As $\mu^n = \sigma_2^n(u) \to \sigma_2(u) = 1 > 1/3 > \mu$, then for any $p \in [0, 1/3]$, the wPBE given by $((\sigma_1(t), \sigma_2(u), \sigma_3(l)), \mu) = ((0, 1, 1), p)$ is not an SE.

(3) Similar to (2), as $\mu^n = \sigma_2^n(u) \to \sigma_2(u) = 1 > 1/3 = \mu$, then the wPBE described in (3) above are not SE.

We conclude that the unique SE is $((\sigma_1(t), \sigma_2(u), \sigma_3(l)), \mu) = ((1, 1, 0), 1)$.

4. Repeated Games

- **Repeated game** (G^T, δ): stage game G repeated T times (T can be ∞) and δ is the discount factor ($\delta \in [0,1)$ if $T = \infty$).

  (i) Players take actions in the stage game. Action profile of the stage game is denoted by a and vector of *stage-game payoffs* by $\pi(a) = (\pi_i(a))_{i \in I}$.
  We write $\alpha = (\alpha_i)_{i \in I}$ where α_i is a distribution over player i's actions A_i in the stage game.

  In general, the stage-game can be either a normal-form game or an extensive-form game (with the due adjustments to the definitions).

  (ii) We assume *perfect monitoring*: at period t players observe actions played in periods $1, \ldots, t-1$. Histories are then $h^t = (a^1, \ldots, a^{t-1})$, where $a^t = (a_i^t)_{i \in I}$.

  (iii) A repeated game is nothing but an extensive-form game of a special kind.

  For simplicity, we tend to work with behavioral strategies instead of mixed strategies.

  (iv) Given (behavioral) strategies, we consider the *(expected) discounted payoffs* for the repeated game u given δ. For instance, consider the pure strategy profile s where $s_i(h^t) = a_i^t$. Then $u_i(s) = \sum_{t=1}^{T} \delta^{t-1} \pi_i(a^t)$.
  The *average (expected) discounted payoffs* for the repeated game is given by normalizing $\tilde{u}_i(s) := (1-\delta) u_i(s)$. This is done so that if the same action profile is played, $a^t = a$, and $T = \infty$, then $\tilde{u}_i(a) = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} \pi_i(a) = \pi_i(a)$. In this case, we require $\delta \in [0,1)$.

- **Payoffs**

  - **Feasible payoffs**

    Payoff vector $v = (v_i)_{i \in I}$ is feasible if there is a distribution p over action profiles A such that $v_i = \sum_{a \in A} p(a) \pi_i(a)$ for every player i.

  - **Nash-threat payoff** for player i:
    $\underline{v}_i := \inf\{\pi_i(\alpha) \mid \alpha \text{ is a (possibly mixed) stage-game Nash equilibrium}\}$

    Worst stage-game Nash equilibrium payoff for player i.

  - **Minmax payoff** for player i: $\underline{\underline{v}}_i := \min_{\alpha_{-i}} \max_{\alpha_i} \pi_i(\alpha_i, \alpha_{-i})$

    The best that player i can do if everyone else decides to punish player i.

Minmax actions can be used to punish an agent who deviates from equilibrium path play

Minmax payoffs set bounds for average payoffs that we can expect: player i can always secure at least $\underline{\underline{v}}_i$

– (Strictly) **Individually Rational payoffs** v

Payoff vector $v = (v_i)_{i \in I}$ is (strictly) individually rational if for any player $i \in I$, $v_i \geq (>) \underline{\underline{v}}_i$

- **Important Results**

  – **One-Shot Deviation Principle**: A strategy profile σ is an SPNE if and only if there are no profitable one-shot deviations.

  This means that to check if a strategy profile is an SPNE it is sufficient to check that for every player i there is no other strategy σ'_i and no history h such that the alternative strategy profile

  (i) is different only at a particular history corresponding to a root of a subgame, $\sigma'_i(h) \neq \sigma_i(h)$,

  (ii) is the same at every other history, $\sigma'_i(h') = \sigma_i(h')$ for all $h' \neq h$, and

  (iii) delivers a strictly higher payoff for player i in the subgame starting at h given that every other player is playing according to σ_{-i}

  – **Unique SPNE in Finitely Repeated Games**:

  In any finitely repeated game (G^T, δ) $(T < \infty)$ such that the stage game G has a unique (possibly mixed) Nash equilibrium α[3], the unique SPNE of the repeated game (G^T, δ) is the history-independent strategy profile σ such that for all $h^t \in H^t$ and all $t = 1, \dots, T$, $\sigma(h^t) = \alpha$

  In other words: if the stage game has a unique (SP)NE, then the unique SPNE of the finitely repeated game is given by having players playing the stage game equilibrium every period.

  – **Repeated Equilibrium is SPNE**: a strategy profile in which players play a history-independent (SP)NE of the stage-game every period is an SPNE of the repeated game.

  – **Folk Theorem for NE in Infinitely Repeated Games**

[3] Unique SPNE if the stage game is an extensive-form game.

Let v be a feasible and strictly individually rational payoff vector ($v_i > \underline{\underline{v}}_i\ \forall i$).
There is $\delta^* \in (0,1)$ such that if $\delta > \delta^*$, (G^∞, δ) has a Nash equilibrium σ that yields average discounted equilibrium payoffs of v.

Intuition:

Sticking to the strategy: get payoff v_i

Deviating: $(1-\delta)\max_{a_i} \pi_i(a_i, p_{-i}) + \delta \underline{\underline{v}}_i$

Need: $\delta \geq \frac{\max_{a_i} \pi_i(a_i, p_{-i}) - v_i}{\max_{a_i} \pi_i(a_i, p_{-i}) - \underline{\underline{v}}_i}$

- **Nash-threats Folk Theorem for SPNE in Infinitely Repeated Games**

Let v be a feasible payoff vector such that $v_i > \underline{v}_i\ \forall i$.
There is $\delta^* \in (0,1)$ such that if $\delta > \delta^*$, (G^∞, δ) has a SPNE σ that yields average discounted equilibrium payoffs of v.

Intuition:

Coordinate on prob. p that generates v; if player j deviates, revert to j-worst stage-game NE forever
If there are multiple players deviating, then choose one to be punished

OSDP: only need to check for one-shot profitable deviations

- **Folk Theorem for SPNE in Infinitely Repeated Games**

Let (i) v be a feasible and SIR payoff vector ($v_i > \underline{\underline{v}}_i\ \forall i$), and (ii) $\{v'$ is feasible and SIR$\}$ have a nonempty interior.
There is $\delta^* \in (0,1)$ such that if $\delta > \delta^*$, (G^∞, δ) has a SPNE σ that yields average discounted equilibrium payoffs of v.

Intuition:

Can do away with (ii) for 2-player games.

Can have the stage game being an extensive-form game.

4.1. How to use the Folk Theorem and the OSDP: A Worked-Out Example

		Col		
		A	B	C
	A	4,4	0,0	0,5
Row	B	0,0	1,1	0,0
	C	5,0	0,0	3,3

Action A is strictly dominated by e.g. $1/10B + 9/10C$. A is never played at a stage-game NE.
PSNE of the stage game: (B,B), (C,C).
MSNE of the stage game: (3/4 B + 1/4 C, 3/4 B + 1/4 C)
Only B and C chosen with positive prob (A strictly dominated). For player i to be indifferent between choosing B and C we need $\sigma_j(B)1 = \sigma_j(C)3 = (1-\sigma_j(B))3 \Longleftrightarrow \sigma_j(B) = 1-\sigma_j(C) = 3/4$.

Stage game NE payoffs: (1,1), (3,3), (3/4,3/4)

- Nash-threat payoff: $\underline{v}_i = 3/4$
- Minmax payoff: $\underline{\underline{v}}_R := \min_{\sigma_C} \max_{\sigma_R} \pi_R(\sigma_C, \sigma_R)$

It is sufficient to consider σ_R such that $\sigma_R(A) = 0$ (due to being strictly dominated) and, therefore, for any $\sigma_R(A) = 0$, for Col, it is sufficient to minimize using B and C as $u_R(\sigma_R, A) \geq \pi_R(\sigma_R, C)$, with the inequality strict when $\sigma_R(C) > 0$.
Note that $\underline{\underline{v}}_R > 0$ (think that if $\sigma_R(B) = \sigma_R(C) = 1/2$, Row has to be getting a strictly positive payoff).
Let $\sigma_C^* \in \arg\min_{\sigma_C} \max_{\sigma_R} \pi_R(\sigma_R, \sigma_C)$ and $\sigma_R^*(\sigma_C) \in \arg\max_{\sigma_R} \pi_R(\sigma_R, \sigma_C)$.
First we show that Row needs to be indifferent given σ_C^*. If Row is not indifferent, then $\pi_R(B, \sigma_C^*) > \pi_R(C, \sigma_C^*)$ (or vice versa), and then the best-response to σ_C^* is $\sigma_R^*(\sigma_C^*)$ placing prob 1 on B (respectively C).
But if this were the case, then Col would punish Row harsher by playing C with prob 1 (respectively B) and make Row get a payoff of 0, which contradicts the fact that σ_C^* minimizes Row's payoff given σ_R^*.
Now, if Row is indifferent between B and C, we then need we need $\sigma_C(B)1 = \sigma_C(C)3 = (1-\sigma_C(B))3 \Longleftrightarrow \sigma_C(B) = 1-\sigma_C(C) = 3/4$.

• Suppose you want to find out the smallest δ^* such that a given strategy can be supported as an average discounted SPNE equilibrium payoff of an infinitely repeated game using Nash-reversion strategies. We focus on case where the strategy is as follows: play A; if at every previous period (A,A) was played, then continue playing A; otherwise punish the opponent by switching to strategy that gives the deviator the worst payoff forever.

In this case, the punishment strategy is to mix (3/4 B + 1/4 C).

Note that if at every previous period (A,A) was played, then Row is considering between playing A and deviating.

As $4 = \pi_R(A,A) < \max_{a_R} \pi_R(a_R, A) = \pi_R(C,A) = 5$, then Row would best deviate to playing C. By the OSDP (one-shot deviation principle), it suffices to consider a single deviation. This means that Row deviates at that period, but then the player goes back to the original strategy. Consequently, deviating to playing C would trigger both players playing (3/4 B + 1/4 C) forever.

Row would not want to deviate if

$$
\begin{aligned}
4 = \pi_R(A,A) &= (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1}\pi_R(A,A) \\
&\geq (1-\delta)\left(\pi_R(C,A) + \sum_{t=2}^{\infty} \delta^{t-1}\pi_R(3/4B + 1/4C, 3/4B + 1/4C)\right) = (1-\delta)5 + \delta 3/4 \\
\Longleftrightarrow & \delta 17/4 = \delta(5 - 3/4) \geq 5 - 4 = 1 \Longleftrightarrow \delta \geq 4/17
\end{aligned}
$$

As the game is symmetric, we need that $\delta \geq 4/17$ to support (A,A) being played in equilibrium at every period using Nash-reversion strategies.

• Suppose you are asked whether you can support payoff vector $v = (v_R, v_C)$ as an average discounted SPNE equilibrium payoff of an infinitely repeated game for some discount factor. The first thing to do is to check whether v is feasible, i.e. that there is a convex combination of $(\pi_R(a_R, a_C), \pi_C(a_R, a_C))$, $a_i \in \{A, B, C\}$, $i = R, C$, such that v is equal to the convex combination of the payoff vectors of the stage game; if the answer is no, then v cannot be supported as an average discounted SPNE equilibrium payoff of an infinitely repeated game for some discount factor. If the answer is yes, then check if v is strictly individually rational. And if it is strictly individually rational (for all players), then, as the game has 2 players, appeal to the Folk Theorem for SPNE in Infinitely Repeated Games to answer positively.