Felix Dosch, Andrey Tyukin

Exercise 11.1 (Iterated games with decaying payoffs) We consider the iterated prisoner's dilemma with a decay coefficient $\lambda \in (0,1)$. The "gross" payoffs in each round are given by the following table:

$$\begin{bmatrix} & \text{Defect} & \text{Cooperate} \\ \text{Defect} & (1,1) & (4,0) \\ \text{Cooperate} & (0,4) & (3,3) \end{bmatrix}$$

The total payoff of the player p is given by:

$$U_p(\sigma_1, \sigma_2) = \mathbb{E}\left[\sum_{t=0}^{\infty} \lambda^t u_p(\sigma_1^t, \sigma_2^t)\right].$$

Notice that we start in the round $t = 0$.

a) We want to investigate for which λ's the *Grim trigger* strategy is a Nash equilibrium.

Claim: (GT, GT) is a Nash equilibrium if and only if $\lambda \geq 1/3$.

Proof: Suppose $\lambda \geq 1/3$. Let σ be any other strategy. Define the *stopping time*[1] τ as follows:

$$\tau := \inf\left\{t \in \mathbb{N}_0 \,\middle|\, \sigma^t = D\right\},$$

that is, it is the index of the round when σ plays D for the first time. It holds:

$$\begin{aligned}
\sum_{t=0}^{\infty} \lambda^t u_1(\sigma^t, TT^t) &= \sum_{t=0}^{\tau-1} \lambda^t u_1(C,C) + \lambda^\tau u_1(D,C) + \sum_{t=\tau+1}^{\infty} \lambda^t u_1(\sigma^t, D) \\
&\leq \sum_{t=0}^{\tau-1} \lambda^t \cdot 3 + 4\lambda^\tau + \sum_{t=\tau+1}^{\infty} \lambda^t \cdot 1 \\
&= 3\frac{1-\lambda^\tau}{1-\lambda} + 4\lambda^\tau + \frac{\lambda^{\tau+1}}{1-\lambda}.
\end{aligned}$$

Compare this expression to the payoff for playing GT:

$$U_1(GT, GT) = \frac{3}{1-\lambda}.$$

It holds:

$$\begin{array}{rrcl}
& 3\frac{1-\lambda^\tau}{1-\lambda} + 4\lambda^\tau + \frac{\lambda^{\tau+1}}{1-\lambda} & \leq & \frac{3}{1-\lambda} \\
\Leftrightarrow & 3 + \lambda^\tau - 3\lambda^{\tau+1} & \leq & 3 \\
\Leftrightarrow & \lambda^\tau(1-3\lambda) & \leq & 0 \\
\Leftrightarrow & \lambda & \geq & \frac{1}{3},
\end{array}$$

[1] See next footnote

therefore by monotonicity of the expected value:

$$U_1(\sigma, GT) = \mathbb{E}\left[\sum_{t=0}^{\infty} \lambda^t u_1(\sigma^t, GT^t)\right] \leq \mathbb{E}\left[\sum_{t=0}^{\infty} \lambda^t u_1(GT^t, GT^t)\right] = U_1(GT, GT),$$

hence $\lambda \geq 1/3$ implies that there are no advantageous deviations from GT, i.e. (GT, GT) is a Nash equilibrium.

To show the opposite implication, assume that $\lambda < 1/3$. Consider the strategy *Always-Defect*. It holds:

$$U_1(D, GT) = 4 + \frac{\lambda}{1-\lambda} = \frac{4-3\lambda}{1-\lambda}.$$

Since we assumed $\lambda < 1/3$ it holds:

$$U_1(D, GT) = \frac{4-3\lambda}{1-\lambda} > \frac{3}{1-\lambda} = U_1(GT, GT),$$

therefore we have found an advantageous deviation from GT. This means that (GT, GT) is not a Nash equilibrium.

Both implications together give the claimed equivalence. ■

b) Now we wish to understand for what λ's the Tit-for-Tat strategy is a Nash-equilibrium. Before giving a formal proof, we give an ansatz: an informal heuristic that shows how one could come up with the critical value $\lambda_{crit} = 1/3$.

One simple "exploit" of the Tit-for-Tat strategy the player 1 could attempt is the following. Player 1 can try to deviate from C to D, then continue to play D for $(k-1)$ rounds, then change back to C. The game then looks like this:

σ_1	D	D	D	D	...	D	C
TT	C	D	D	D	...	D	D
u_1	4	1	1	1	...	1	0

The ansatz is that this maneuvre by the player 1 should give him no advantage for the critical λ_{crit}, that is, the outcome for this block should be the same as if he would just play C all the time, and get a gross payoff of 3 per round:

$$4 + \lambda + \lambda^2 + \cdots + \lambda^{k-1} + 0 \stackrel{!}{=} 3(1 + \lambda + \lambda^2 + \cdots + \lambda^{k-1} + \lambda^k)$$

Simplifying both expressions using the formula for geometric series we obtain:

$$3 + \frac{1-\lambda^k}{1-\lambda} \stackrel{!}{=} 3\frac{1-\lambda^{k+1}}{1-\lambda}$$

After some shuffling of the summands we get:

$$(1-3\lambda) \stackrel{!}{=} (1-3\lambda)\lambda^k.$$

The only way this can hold for all k is that $\lambda_{crit} \equiv \lambda = \frac{1}{3}$.

Now we have some intuition and a plausible claim. Before we give a formal proof, let's take a quick look at some potential pitfalls. Notice that:

- Coming up with some exploit that is better than TT for some λ gives only a lower bound for the critical lambda, it does not show that there are no advantageous deviations for larger λ's;
- If one tries to prove something considering only hard-coded non-randomized strategies, one needs some argument that guarantees that there are no randomized strategies that are better;
- If one tries to prove something considering only "the previous step" or some finite number of "previous steps" (kind of Markovian property), one needs some extra argument that guarantees that there are no better strategies with built-in long-term memory.

Claim: The strategy (TT, TT) is a Nash equilibrium if and only if $\lambda \geq \frac{1}{3}$.

Proof: First, let's introduce a notation for the payoff of a block that starts at time 0 and goes until time ζ, with player 1 starting to play D at time δ and switching back to C at time ζ. Here a sketch of what we are trying to compute:

$$\left[\begin{array}{c|cccccccccccc} t & 0 & 1 & & & \dots & & \delta & & & \dots & & \zeta \\ \sigma_1 & C & C & C & C & \dots & C & D & D & D & \dots & D & C \\ TT & C & C & C & C & \dots & C & C & D & D & \dots & D & D \\ u_1 & 3 & 3 & 3 & 3 & \dots & 3 & 4 & 1 & 1 & \dots & 1 & 0 \end{array}\right]$$

Here is the definition, combined with an immediate simplification:

$$\begin{aligned} block_\lambda(\delta, \zeta) &:= \sum_{t=0}^{\delta-1} \lambda^t u_1(C, C) + \lambda^\delta u_1(D, C) + \sum_{t=\delta+1}^{\zeta-1} \lambda^t u_1(D, D) + \lambda^\zeta u_1(C, D) \\ &= \frac{1-\lambda^\delta}{1-\lambda} \cdot 3 + \frac{\lambda^\delta - \lambda^{\delta+1}}{1-\lambda} \cdot 4 + \frac{\lambda^{\delta+1} - \lambda^\zeta}{1-\lambda} \cdot 1 + 0 \\ &= \frac{1}{1-\lambda}\left(3 + \lambda^\delta - 3\lambda^{\delta+1} - \lambda^\zeta\right). \end{aligned}$$

We want to compare this block to the outcome of just playing C, that is, we want to show that for all δ and ζ it holds:

$$block_\lambda(\delta, \zeta) \leq \sum_{t=0}^{\zeta} \lambda^t \cdot 3 = \frac{3 - 3\lambda^{\zeta+1}}{1-\lambda},$$

if $\lambda \geq 1/3$. For this, consider the following equivalences:

$$\begin{array}{crl} & \frac{1}{3} & \leq \lambda \\ \Leftrightarrow & 1 & \leq 3\lambda \\ \Leftrightarrow & \lambda^\delta(1-\lambda^{\zeta-\delta}) & \leq 3\lambda^{\delta+1}(1-\lambda^{\zeta-\delta}) \\ \Leftrightarrow & \lambda^\delta - 3\lambda^{\delta+1} - \lambda^\zeta & \leq -3\lambda^{\zeta+1} \\ \Leftrightarrow & \frac{3+\lambda^\delta-3\lambda^{\delta+1}-\lambda^\zeta}{1-\lambda} & \leq \frac{3-3\lambda^{\delta+1}}{1-\lambda} \\ \Leftrightarrow & block_\lambda(\delta, \zeta) & \leq \sum_{t=0}^{\zeta} \lambda^t \cdot 3. \end{array}$$

Now let σ be an arbitrary strategy of player 1. We want to show that for $\lambda \geq 1/3$, the expected payoff is less or equal to the payoff for TT. For this, we will "cut" the game into blocks like the one above, and treat each block separately. For convenience, set $\sigma^{-1} := C$. Define a sequence of *stopping times*[2] δ_i, ζ_i as follows:

$$\begin{aligned}
\zeta_0 &:= -1 \\
\delta_i &:= \inf \left\{ t > \zeta_{i-1} \,\middle|\, \sigma^t = D \right\} \\
\zeta_i &:= \inf \left\{ t > \delta_i \,\middle|\, \sigma^t = C \right\}
\end{aligned}$$

The ζ_0 together with $\sigma^{-1} := C$ are set this way just for convenience, so that we do not need two cases in the definition of δ_i. Notice that $\inf$ includes the possibility that δ_i or ζ_i remain infinite, that is, at some point the player stops switching between C and D. To keep the notation simple yet consistent, we henceforth treat expressions like $\sum_{t=\infty}^{\infty} f(t)$ as empty sums with value 0. Also notice that the definition of $block_\lambda$ is compatible with this convention and can accept ∞ as arguments.

The stopping times now let us express the payoff as a sum (or series) of payoffs for simple blocks, contribution of each block can be bounded using the previous estimations:

$$\begin{aligned}
\sum_{t=0}^{\infty} \lambda^t u_1(\sigma^t, TT^t) \quad &= \quad \sum_{i=1}^{\infty} \Bigg(\sum_{t=\zeta_{i-1}+1}^{\delta_i - 1} \lambda^t u_1(\sigma^t, \sigma^{t-1}) + \lambda^{\delta_i} u_1(\sigma^{\delta_i}, \sigma^{\delta_i - 1}) + \\
&\qquad\qquad \sum_{t=\delta_i+1}^{\zeta_i - 1} \lambda^t u_1(\sigma^t, \sigma^{t-1}) + \lambda^{\zeta_i} u_1(\sigma^{\zeta_i}, \sigma^{\zeta_i - 1}) \Bigg) \\
&\overset{\text{def}_{\delta_i,\zeta_i}}{=} \quad \sum_{i=1}^{\infty} \Bigg(\sum_{t=\zeta_{i-1}+1}^{\delta_i - 1} \lambda^t u_1(C, C) + \lambda^{\delta_i} u_1(D, C) + \\
&\qquad\qquad \sum_{t=\delta_i+1}^{\zeta_i - 1} \lambda^t u_1(D, D) + \lambda^{\zeta_i} u_1(C, D) \Bigg) \\
&\overset{\text{def}_{block}}{=} \quad \sum_{i=1}^{\infty} \lambda^{\zeta_{i-1}+1} block_\lambda \left(\delta_i - (\zeta_{i-1}+1), \zeta_i - (\zeta_{i-1}+1) \right) \\
&\leq \quad \sum_{i=1}^{\infty} \lambda^{\zeta_{i-1}+1} \sum_{t=0}^{\zeta_i - (\zeta_{i-1}+1)} \lambda^t \cdot 3 \\
&= \quad \sum_{i=1}^{\infty} \sum_{t=\zeta_{i-1}+1}^{\zeta_i} \lambda^t \cdot 3 \; = \; \sum_{t=0}^{\infty} \lambda^t \cdot 3 \\
&= \quad \sum_{t=0}^{\infty} \lambda^t u_1(TT^t, TT^t).
\end{aligned}$$

[2] Although "*stopping time*" has a precise meaning, we do not want to go too much into details here. Just think of δ_i and ζ_i (for $i \in \mathbb{N}_0$) as of $\mathbb{N}_0 \cup \{\infty\}$-valued random variables that tell us when player 1 switches from C to D resp. from D to C. These quantities have to be random, because σ can be randomized.

Now the payoff for σ is bounded by the payoff for TT, by monotonicity of the expected value it holds:

$$U_1(\sigma, TT) = \mathbb{E}\left[\sum_{t=0}^{\infty} \lambda^t u_1(\sigma^t, TT^t)\right] \leq \mathbb{E}\left[\sum_{t=0}^{\infty} \lambda^t u_1(TT^t, TT^t)\right] = U_1(TT, TT).$$

This shows: for all $\lambda \geq 1/3$ there are no advantageous deviations from the strategy TT, so that (TT, TT) is a Nash equilibrium. Now we have to show the opposite direction by finding an advantageous deviation in the case that $\lambda < 1/3$. For this, consider the following strategy:

$$\sigma^t := \begin{cases} D & \text{for even } t \\ C & \text{for odd } t. \end{cases}$$

It holds:

$$\begin{aligned}
\sum_{t=0}^{\infty} \lambda^t u_1(\sigma^t, TT^t) &= \sum_{i=0}^{\infty} \lambda^{2i} \left(u_1(D, C) + \lambda u_1(C, D)\right) = \sum_{i=0}^{\infty} \lambda^{2i} \left(4 + \lambda \cdot 0\right) \\
&> \sum_{i=0}^{\infty} \lambda^{2i} \left(3 + \lambda \cdot 3\right) = \sum_{i=0}^{\infty} \lambda^{2i} \left(u_1(C, C) + \lambda u_1(C, C)\right) \\
&= \sum_{t=0}^{\infty} \lambda^t u_1(TT^t, TT^t),
\end{aligned}$$

that is, if $\lambda < 1/3$, player 1 can deviate from TT and gain an edge over player 2. In particular, (TT, TT) is not a Nash equilibrium.

Both directions together show the equivalence of statements "$\lambda \geq 1/3$" and "(TT, TT) is a Nash equilibrium". ∎

Exercise 11.2 (Auktionen)

a) Wir versuchen, den erwarteten Nutzen für jeden Spieler zu maximieren. Wir nehmen dabei an, dass $b_i \leq 2$, da es nicht sinnvoll ist, ein Gebot abzugeben, welches größer ist als der mögliche Wert des Gegenstandes für einen Spieler.

Wir zeigen, dass folgende Strategie für jeden Spieler ein symmetrisches Gleichgewicht ist:

$$b_i(v_i) = \frac{v_i}{2}, \quad i = 1, 2, 3$$

Wir betrachten ohne Einschränkung den Nutzen für Spieler I und bezeichnen mit b^* das Gebot, welches abgesehen vom Gebot von Spieler I, das höchste Gebot sei (unter der Annahme, dass diese auch diese Strategie spielen - ihre Gebote seien daher b_2^* und b_3^*):

$$b^* = \frac{v^*}{2} = max\{b_2^*, b_3^*\} = max\{\frac{v_2}{2}, \frac{v_3}{2}\}$$

Damit ist der Nutzen für Spieler 1:

$$\begin{aligned}
u_1(b_1, b^*, v_1) &= u_1(b_1, \frac{v^*}{2}, v_1) \\
&= P(b_1 > \frac{v^*}{2}) \cdot (v_1 - b_1) \\
&= P(2b_1 > v^*) \cdot (v_1 - b_1) \\
&= min\{2b_1, 1\} \cdot (v_1 - b_1)
\end{aligned}$$

Diese Funktion ist quadratisch auf dem Intervall $b_1 \in [0, \frac{1}{2}]$ (erreicht das Maximum bei $b_1 = \frac{v_1}{2}$), und linear mit negativer Steigung, wenn $b_1 > \frac{1}{2}$.

Fazit: Wenn die beiden anderen Spieler die Strategie $b_i = \frac{v_i}{2}$ spielen, so ist für Spieler I $b_1 = \frac{v_1}{2}$ die beste Antwort. Damit ist $b^* = b_i(v_i) = \frac{v_i}{2}$ ein symmetrisches Gleichgewicht.

b) [omitted]