Repeated Games

Moty Katzman

August 22, 2017

Example: Repeated Prisoners' Dilemma

Consider the following version of Prisoners' Dilemma

	n	c
N	2, 2	-1, 3
C	3, -1	0, 0

How do we model playing Prisoners' Dilemma twice?

Let $S_A = \{N, C\}$ and $S_B = \{n, c\}$.

If the game is repeated n times, Alice and Bob have

$$2^{4^0} \times 2^{4^1} \times \dots 2^{4^{n-1}}$$

strategies available to them, and for $n = 3$ this is already quite big ($2^{21} = 2,097,152$).

In each round of the game there are 4 possible strategy profiles $(N, n), (N, c), (C, n), (C, c)$. When Alice makes her choice in the ith stage of the game, there are 4^{i-1} possible histories, and she has to choose between N and C for each of them, and there are $K_i = 2^{4^{i-1}}$ possible choices.The total number of choices is

$$K_1 \times K_2 \times \dots K_n = 2^{4^0} \times 2^{4^1} \times \dots 2^{4^{n-1}}.$$

Payoffs of infinitely repeated games

If the game is played finitely many times, we can add up the payoffs of each stage.

What if the game is played infinitely many times, or potentially infinitely many times (e.g., after each repetition a coin is tossed and the game is stopped if Heads occurs.)

We'll use two methods for handling potentially infinite streams of payoffs. Given a sequence of intermediate payoffs $r_0, r_1, \ldots$ we can define

the average payoff: $\lim_{m\to\infty}(r_0 + \cdots + r_{m-1})/m$, and

the future discounted payoff: $r_0 + \beta r_1 + \ldots$ where $0 \le \beta < 1$ (notice that this series converges!)

There are two main motivations behind the second definition: (a) the game repeated every fixed period of time, β is the *present value* of a payoff of 1 occurring one period of time into the future, (b) the game is repeated at each stage with probability β.

We refer to games with average reward as being *infinite repetitions,* and those which use the the future discounted reward as *indefinite repetitions*.

A finitely repeated game

Use backward induction for finite repetition.

Proposition. Consider Prisoner's Dilemma repeated a finite number of times k and we add up the payoffs at each stage. There is a unique subgame-perfect Nash equilibrium in which the players play (C, c) at each stage.

Proof. Induction on k. The case $k = 1$ follows from the elimination of the dominated strategies n and N.

Assume that $k > 1$ and that the statement holds for $k - 1$.

Consider the kth and last stage of the game: regardless of previous history, Alice and Bob find themselves playing the case $k = 1$ and hence they play the dominant strategies (C, c) and both receive a payoff of 0.

Since the kth stage of the game does not affect previous payoffs, there is no advantage in deviating from the Nash equilibria of the first $k - 1$ repetitions of the game.

This shorter game has repeated (C, c) as a unique subgame perfect Nash equilibrium. $\square$

Nash equilibria of repeated games

Theorem 1. Let $G = (S, T, u_1, u_2)$ be a game with (pure- or mixed-strategy) Nash equilibrium (s, t). In the finite, infinite and indefinite repetitions of the game, playing (s, t) repeatedly is a subgame perfect Nash equilibrium.

Proof. If the game is played a finite number of times k, and Alice plays $s_1, \ldots, s_k$, against Bob's t, her payoff is $u_1(s_1, t) + \cdots + u_1(s_k, t) \leq u_1(s, t) + \cdots + u_1(s, t)$. Similarly, Bob does not gain from deviating from t, and we conclude that playing (s, t) repeatedly is a Nash Equilibrium. This argument works for all subgames, so it is a subgame perfect Nash equilibrium.

Subgames of the infinite and indefinite repeated games are identical to the whole game, so the restriction of the strategy of repeated (s, t) gives the same strategy on all subgames. So if the proposed strategy is a Nash equilibrium, it is automatically subgame perfect. Suppose that the first player deviates and plays s_k at the kth stage of the game. In the infinitely repeated game her payoff is

$$\lim_{k\to\infty} \frac{u_1(s_0, t) + \cdots + u_1(s_{k-1}, t)}{k} \leq \lim_{k\to\infty} \frac{u_1(s, t) + \cdots + u_1(s, t)}{k}$$

and in the indefinite repeated game the payoff is
$\displaystyle\sum_{k=0}^{\infty} \beta^k u_1(s_k, t) \leq \sum_{k=0}^{\infty} \beta^k u_1(s, t).$ $\square$

Consider an indefinitely repeated game of Prisoner's Dilemma where after each repetition the game is stopped with probability $1/3$ (and we measure payoffs with the the future discounted reward method with $\beta = 2/3$). The sets of strategies of Alice and Bob are now infinite; nevertheless we find some Nash equilibria.
Consider the following strategies:

- **HAWK** Always confess.
- **DOVE** Never confess.
- **GRIM** Don't confess until other person confesses, after that always confess.

Not surprisingly (HAWK, HAWK) is a (subgame-perfect) Nash equilibrium. (DOVE, DOVE) is not.

(GRIM,GRIM)

Proposition. The strategy profile (GRIM, GRIM) is a (subgame-perfect) Nash equilibrium.

Proof. If both players play GRIM, their payoffs are $2 + 2\beta^1 + \cdots = 2/(1-\beta) = 6$. If Alice deviates from GRIM for the first time at the kth stage of the game, her payoff is at most

$$2(1 + \beta + \cdots + \beta^{k-2}) + 3\beta^{k-1} = 2\frac{1-\beta^{k-1}}{1-\beta} + 3\beta^{k-1} =$$

$$6\left(1 - \left(\frac{2}{3}\right)^{k-1}\right) + 3\left(\frac{2}{3}\right)^{k-1} = -3\left(\frac{2}{3}\right)^{k-1} + 6 < 6.$$

□

The “folk theorem”– first version (indefinite games)

Given a game $G = (S, T, u_1, u_2)$ and $0 < \beta < 1$ we define $G^\infty(\beta)$ to be the indefinitely repeated game G with discount factor β.

Theorem 2. Let $G = (S, T, u_1, u_2)$ be a game with (pure- or mixed-strategy) Nash equilibrium (p, q). If (s, t) is a pure strategy profile with $u_1(s, t) > u_1(p, q)$ and $u_2(s, t) > u_2(p, q)$, then there exists a $0 \leq \beta_0 < 1$ such that for all $\beta_0 \leq \beta < 1$ there is subgame perfect Nash equilibrium of $G^\infty(\beta)$ with same payoff as that of playing (s, t) repeatedly.

Proof. Let $\mathcal{G}_1$ be the strategy for player 1 in which, if player 2 ever deviated from t she plays p and she plays s as long as player 2 sticks to t. Let $\mathcal{G}_2$ be the strategy for player 2 in which, if player 1 ever deviated from s he plays q and he plays t as long as player 1 sticks to s.
We first show that if β is close enough to 1, then $(\mathcal{G}_1, \mathcal{G}_2)$ is a Nash equilibrium for $G^\infty(\beta)$.
If $(\mathcal{G}_1, \mathcal{G}_2)$ is played, the players get a payoff of
$\sum_{i=0}^{\infty} \beta^i u_1(s,t) = u_1(s,t)/(1-\beta)$ and
$\sum_{i=0}^{\infty} \beta^i u_2(s,t) = u_2(s,t)/(1-\beta)$, respectively.

If player 1 deviates from playing s by playing s' at the k stage of the game and s_i $(i > k)$ thereafter, she gets the payoff

$$\sum_{i=0}^{k-1} \beta^i u_1(s,t) + \beta^k u_1(s',t) + \sum_{i=k+1}^{\infty} \beta^i u_1(s_i,q)$$

$$\leq \sum_{i=0}^{k-1} \beta^i u_1(s,t) + \beta^k u_1(s',t) + \sum_{i=k+1}^{\infty} \beta^i u_1(p,q)$$

$$= \sum_{i=0}^{k-1} \beta^i u_1(s,t) + \beta^k u_1(s',t) + u_1(p,q)\frac{\beta^{k+1}}{1-\beta}$$

We need to consider values of β for which

$$\sum_{i=0}^{\infty} \beta^i u_1(s,t) > \sum_{i=0}^{k-1} \beta^i u_1(s,t) + \beta^k u_1(s',t) + u_1(p,q) \frac{\beta^{k+1}}{1-\beta}$$

i.e.,

$$u_1(s,t) \frac{\beta^k}{1-\beta} > \beta^k u_1(s',t) + u_1(p,q) \beta^{k+1} \frac{1}{1-\beta},$$

i.e.,

$$u_1(s,t) \frac{1}{1-\beta} > u_1(s',t) + u_1(p,q) \frac{\beta}{1-\beta}$$

which simplifies to $u_1(s,t) > (1-\beta) u_1(s',t) + \beta u_1(p,q)$. As $\beta \to 1$, the right hand side converges to $u_1(p,q) < u_1(s,t)$, hence a continuity argument shows that there exists a $0 \leq \beta_1 < 1$ for which the inequality holds for all $\beta_1 < \beta < 1$.

A similar argument shows that there exists a $0 \leq \beta_2 < 1$ such that whenever $\beta_2 < \beta < 1$, player 2 would not deviate from playing $\mathcal{G}_2$.

We deduce that with $\beta_0 = \max\{\beta_1, \beta_2\}$ and $\beta \in (\beta_0, 1)$, $(\mathcal{G}_1, \mathcal{G}_2)$ is a Nash equilibrium whose payoff is identical to the payoff of playing repeated (s,t) repeatedly.

To show that $(\mathcal{G}_1, \mathcal{G}_2)$ is a *subgame-perfect* Nash equilibrium we need to show that it is a Nash equilibrium of its subgames: such a subgame occurs either after a defection or not. After a defection $(\mathcal{G}_1, \mathcal{G}_2)$ calls for playing (p, q) in every repetition of the game, which is a Nash equilibrium by Theorem 1, and if no defection occurs the subgame is identical to the original game. $\square$

Notation: G^∞

Given a game $G = (S, T, u_1, u_2)$ we define G^∞ to be the infinitely repeated game G with average reward payoff.

Minimax values

Definition. Given a finite game $G = (S, T, u_1, u_2)$ we define the *minimax values* of players 1 and 2 to be $\min_{t \in T} \max_{s \in S} u_1(s, t)$ and $\min_{s \in S} \max_{t \in T} u_2(s, t)$. Thus the minimax value of a player is the worst possible payoff the other player can inflict.

Example. The minimax values in

	a	b	c
I	1, 0	6, 4	0, 9
II	2, 1	0, 2	3, 0
III	3, 7	2, 3	4, 0

are $3 = u_1(III, a)$ and $2 = u_2(II, b)$.

Convex sets

Definition. A *convex* set in $\mathbb{R}^d$ is a subset S of $\mathbb{R}^d$ with the property that for any $v, w \in S$, $tv + (1-t)w \in S$ for all $0 \le t \le 1$.
Examples. Δ_d, vector-subspaces of $\mathbb{R}^d$, cones, intersections of convex sets, etc.
Definition. The *convex hull* of a set $A \subseteq \mathbb{R}^d$ is, equivalently,

(a) the intersection of all convex sets containing A,

(b) the smallest convex set containing A,

(c) the subset C of $\mathbb{R}^d$ defined as the set of all points of the form $\lambda_1 a_1 + \cdots + \lambda_k a_k$ where $k \ge 1$, $\lambda_1, \ldots, \lambda_k$ are non-negative real numbers, $\lambda_1 + \cdots + \lambda_k = 1$ and $a_1, \ldots, a_k \in A$.

The cooperative payoff region of games

Definition. Let $G = (S, T, u_1, u_2)$ be a game. The *cooperative payoff region* of G is the convex hull of $\{(u_1(s,t), u_2(s,t)) \mid s \in S, t \in T\} \subseteq \mathbb{R}^2$.

Definition. Let $A, B \subseteq \mathbb{R}^n$. We say that A is dense in B if for all $\epsilon > 0$ and all $b \in B$ there is an $a \in A$ whose distance to b is less than ϵ.

Example. $\mathbb{Q}$ is dense in $\mathbb{R}$ and more generally $\mathbb{Q}^n$ is dense in any subset of $\mathbb{R}^n$.

The "folk theorem"– second version (infinite games)

Theorem. Let $G = (S, T, u_1, u_2)$ be a game with minimax values $\mu_1 = u_1(\sigma_1, \tau_1)$ and $\mu_2 = u_2(\sigma_2, \tau_2)$. Let R be the cooperative payoff region of G and let $A = \{(x, y) \in R \,|\, x > \mu_1, y > \mu_2\}$. Let B be the set of payoffs of Nash equilibria of G^∞. Then B is dense in A.

Proof.
Consider $(x, y) \in A$ of the form
$(x, y) = \lambda_1(u_1(s_1, t_1), u_2(s_1, t_1)) + \cdots + \lambda_k(u_1(s_k, t_k), u_2(s_k, t_k))$
where $k \geq 1$, $s_1, \ldots, s_k \in S$, $t_1, \ldots, t_k \in T$ and $\lambda_1, \ldots \lambda_k$ are non-negative *rational* numbers in $[0, 1]$ adding up to 1. The set of all such points is dense in A. Also, we can find a positive common denominator N such that $\lambda_1 = m_1/N, \ldots, \lambda_k = m_k/N$ for integers $m_1, \ldots, m_k$.
Let s be the strategy for player 1 which consists of playing s_1 for m_1 turns followed by s_2 for m_2 turns, etc., ending with playing s_k for m_k turns and repeating this pattern cyclically.
Let t be the strategy for player 2 which consists of playing t_1 for m_1 turns followed by t_2 for m_2 turns, etc., ending with playing t_k for m_k turns and repeating this pattern cyclically.
Let $\mathcal{G}_1$ be the strategy for player 1 in which, if player 2 ever deviated from t she plays σ_2 at every turn and she plays s as long as player 2 sticks to t. Let $\mathcal{G}_2$ be the strategy for player 2 in which, if player 1 ever deviated from s he plays τ_1 at every turn and he plays t as long as player 1 sticks to s.

We show $(\mathcal{G}_1, \mathcal{G}_2)$ is a Nash equilibrium for G^∞.
If $(\mathcal{G}_1, \mathcal{G}_2)$ is played, the players get a payoff of
$\frac{1}{N}\sum_{i=1}^{k} m_i u_1(s_i, t_i) = x$ and $\frac{1}{N}\sum_{i=1}^{k} m_i u_2(s_i, t_i) = y$, respectively.
If player 1 deviates from playing s, player 2 punishes her and her subsequent payoffs are at most $\mu_1 < x$ and thus in the limit as the number of plays increases her average score converges to a limit below x.
If player 2 deviates from playing t, player 1 punishes him and his subsequent payoffs are at most $\mu_2 < y$ and thus in the limit as the number of plays increases his average score converges to a limit below y. $\square$.
The proof above does not show that the dense set of payoffs occur as payoffs of *subgame perfect* Nash equilibria, but in fact they are.