Repeated Games Step-by-Step

David K. Levine, 5/22/02

A Sample Game

To keep things concrete we will focus on a specific example, the normal form game in the matrix below:

	a	b	c	d
A	3,4	3*,2	-1,5*	0*,1
B	1,3	2,4*	0,3	0*,1
C	4*,0	3*,1	1*,2*	0*,1
D	0,1*	0,1*	0,1*	-1,0

The best responses for each player are marked with asterisks.

Static Benchmarks

The place to begin analyzing a repeated game is to ignore the fact that it is repeated, and focus on what happens when the game is played once. To clearly distinguish this from the repeated game, we refer to this as the *static game*. Most of the interesting information about the static game can be found directly from the best responses.

Nash equilibrium: The static Nash equilibrium is (equivalently – the static Nash equilibrium strategies are) Cc. The static Nash equilibrium payoff is (1,2).

Stackelberg equilibrium with player 1 leader:

strategy of 1	best response of 2	payoff to 1
A	c	-1
B	b	2*
C	c	1
D	a,b,c	0

The Stackelberg strategy is to play B. The Stackelberg payoff is the most that player 1 can get when player 2 plays a best response, that is, 2.

Problem: Show that the Stackelberg payoff to player 2 as the leader is 2.

Note: If there is a tie, such as when player 2 plays b we assume the Stackelberg follower plays the strategy most favorable to the leader. That is, since 1 is indifferent between A and C, we assume he plays A since that is better for the leader, player 2.

Minmax for player 1:

strategy of 2	best response of 1	payoff to 1
a	C	4
b	A,C	3
c	C	1
d	A,B,C	0*

The minmax payoff to player 1 is the *least amount* he gets when he plays a best-response, that is, 0.

Problem: Show that the minmax payoff for player 2 is 1.

Find the strategies that strictly Pareto dominate the static Nash equilibrium: The static Nash equilibrium has payoffs (1,2). Strict Pareto dominance means both players must be better off (no ties). This leaves only Aa payoffs (3,4) and Bb payoff (2,4).

Problem: Show that the only strategy profile (pair of strategies, one for each player) that Pareto dominates the static Nash equilibrium and is Pareto efficient is Aa.

Types of Repetition

There are many way in which a game can be repeated. It can be repeated a fixed number of times or an infinite number of times. Payoffs between different periods can be aggregated by several methods: adding them together, averaging them, taking the present value, or the average present value. Regardless, the repeated game strategies in which each player plays their static Nash equilibrium strategy no matter what the circumstances is always a subgame-perfect equilibrium of the repeated game. In the example, the strategies player 1 plays C no matter what, and player 2 plays c no matter what are a subgame perfect equilibrium. Notice that a strategy in a repeated game must not only say what to play (that is, Cc), but also under what circumstances to play it (in this case, always).

We will limit attention to two types of repetition. In both cases, the game is repeated infinitely (no definite ending). In the first case, which we refer to as patient players, both players use average present value, and a common (the same for both players) discount factor δ. In the second case, which we refer to as long run versus short-run player, one player uses average present value with discount factor δ and the other discount the future with discount factor 0 – that is, they don't care about the future at all.

Before examining the actual repetition of the games, we review the notion of average present value. If a player has a discount factor δ, and receives u_1 in period 1, u_2 in period 2, and so forth, the average present value is defined to be

$$(1 - \delta)(u_1 + \delta u_2 + \delta^2 u_3 \ldots).$$

The key to computations is the identity

$$1 + \delta + \delta^2 + \ldots = \frac{1}{1 - \delta}.$$

For example, suppose that in period 1 in the example, play is Cc, in period 2 again Cc, and forever Cc. That is, the static Nash equilibrium in every period. Then player 1 receives 1 in period 1, 1 in period 2 and so forth. His average present value is

$$(1 - \delta)(1 \cdot 1 + \delta \cdot 1 + \delta^2 \cdot 1 + \ldots) = (1 - \delta)\frac{1}{1 - \delta} = 1.$$

This is basically the point of average present value – if the same fixed utility is received every period, the average present value is equal to that same amount.

Problem: Show the average present value for player 2 is 2.

Now suppose that in period 1 play is Cc, in period 2 Dd, in period 3 again Dd, in period 4 Cc, period 5 Cc, and then Cc forever. Then player 1 gets 1 in period 1, 2 in period 2, 2 in period 3, then 1 in period 4 and after. The average present value is

$$(1 - \delta)(1 \cdot 1 + \delta \cdot 2 + \delta^2 \cdot 2 + \delta^3 \cdot 1 + \delta^4 \cdot 1 \ldots) =$$
$$(1 - \delta)\left(\left(1 \cdot 1 + \delta \cdot 2 + \delta^2 \cdot 2\right) + \left(\delta^3 \cdot 1 + \delta^4 \cdot 1 \ldots\right)\right)$$

The trick do doing computations like this is to rearrange the sum so that the final piece with the infinite part is a sum of constant utilities (in this case 1). We can then factor out the discount factor (in this example δ^3) and apply our formula.

$$(1 - \delta)\left(\left(1 \cdot 1 + \delta \cdot 2 + \delta^2 \cdot 2\right) + \left(\delta^3 \cdot 1 + \delta^4 \cdot 1 \ldots\right)\right) =$$
$$(1 - \delta)\left(\left(1 \cdot 1 + \delta \cdot 2 + \delta^2 \cdot 2\right) + \delta^3\left(1 + \delta \cdot 1 \ldots\right)\right) =$$
$$(1 - \delta)\left(\left(1 \cdot 1 + \delta \cdot 2 + \delta^2 \cdot 2\right) + \delta^3\frac{1}{1 - \delta}\right) =$$
$$(1 - \delta)(1 + 2\delta + 2\delta^2) + \delta^3$$

Patient Players

We deal first with the case of two equally patient players who discount the future with common discount factor δ with an infinite number of repetitions.

The Folk Theorem

You are expected to know the Folk Theorem: that when player are equally and sufficiently patient, all payoffs that are socially feasible and Pareto dominate the minmax are subgame perfect equilibrium payoffs. A good place to start understanding the
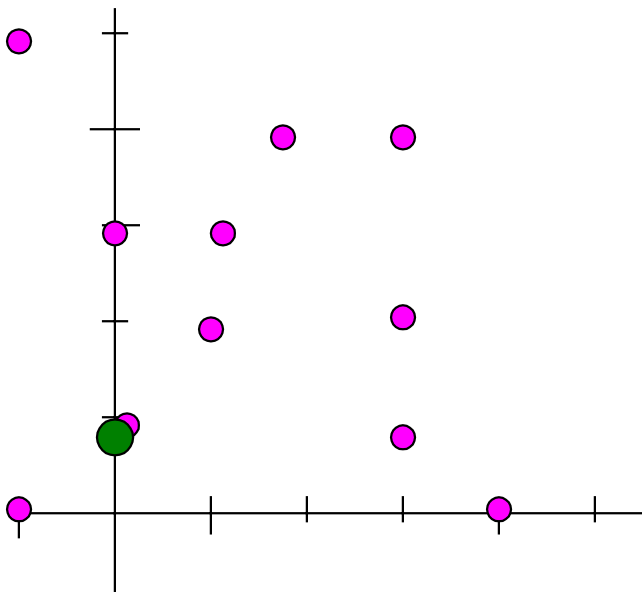
repeated game, is simply by plotting the socially feasible, individually rational (= Pareto dominates the minmax) region.
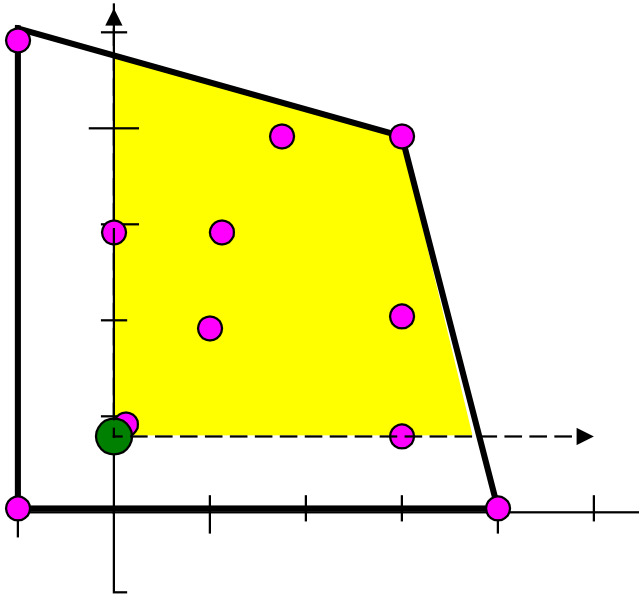
Step 1: plot the payoffs from the matrix
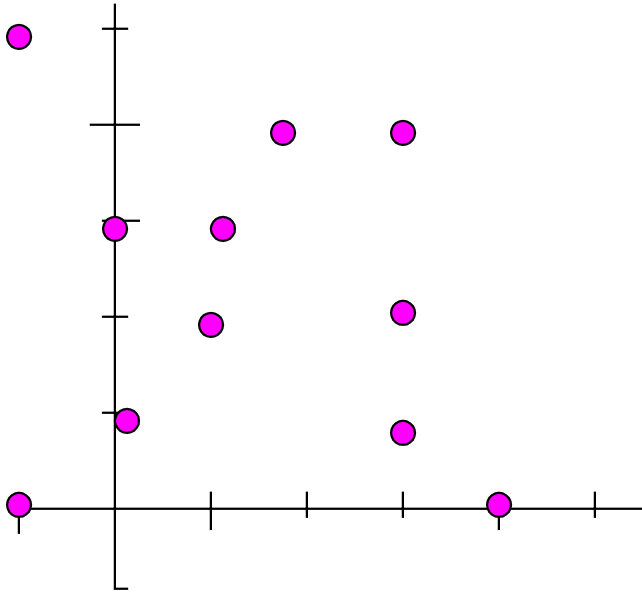
Step 2: find the minmax and plot it

minmax is (0,1) from above, marked in green in the figure

Step 3: shade in the area bounded by the matrix payoffs and above the minmax

Finding Grim Strategy Equilibria

You are not expected to be able to prove the Folk Theorem, or to be able to construct equilibrium strategies for arbitrary socially feasible individually rational payoffs. Given a strategy profile that strictly Pareto dominates the static Nash equilibrium, you are expected to be able to construct a *grim-strategy* equilibrium that sustains that outcome on the *equilibrium path*. Consider, specifically, in the example Aa, with payoffs (3,4) which strictly Pareto dominate the static Nash equilibrium at Cc with Payoff (1,2).

The structure of *grim strategies* is relatively simple. On the equilibrium path players "play the way they are supposed to." That means, they play Aa in the first period, and continue to play that way as long as "no player has deviated," that is, as long as they have seen Aa in every previous period. That means that the equilibrium path is Aa, and that players get an average present value of (3,4) on the equilibrium path. However, to prevent "cheating" there must be punishment for "deviation." With grim strategies, the punishment is always the static Nash equilibrium, that is, Cc. Specifically, player 1's strategy is: play A in period 1. Play A as long as Aa has been seen in every previous period. Play C if every anything other than Aa has ever occurred.

Grim strategies present a player with a very stark choice. They can go along with the program, in which case Aa is always played, and player 1 gets 3 and player 2 gets 4. Or they can deviate doing the best they can for one period, but recognizing that in every future period they will fact the static Nash equilibrium payoffs of (1,2). For the grim strategies to be an equilibrium, the discount factor must be high enough that remaining on the equilibrium path is better than deviating. Here are the computations involved:

Step 1: Find the average present value on the equilibrium path.
On the equilibrium path players play Aa and get (3,4) each period. The average present value is just (3,4).

Step 2: Find the best response of each player when his opponent remains on the equilibrium path
On the equilibrium path Player 2 plays a, so the most that player 1 can get is 4, by playing C.

Problem: Show that 5 is the most that Player 2 can get when Player 1 remains on the equilibrium path.

Step 3: Calculate the average present value from deviating.
This is the most a player can get when his opponent remains on the equilibrium path in the first period, followed by the static Nash equilibrium (the punishment) in the second and subsequent periods. For player 1 we found in step 2 that he could get 4 for one period. His static Nash equilibrium payoff is 1, so by deviating he gets gets 4 for one period, and 1 forever after. The average present value is

$$(1-\delta)\left(4 + \delta + \delta^2 + \ldots\right) = (1-\delta)(4 + \delta\frac{1}{1-\delta}) = (1-\delta)4 + \delta.$$

Problem: What is the average present value to deviating for Player 2?

Step 4: Compare the average present value to remaining on the equilibrium path to that from deviating.

If player 1 had elected not to cheat, he would have received 3 in every period, for an average present value of 3. So for the grim strategies to be an equilibrium, it must be that player 1 gets at least as much from not cheating as from cheating. In other words, we compare the answer from step 2 with that in step 3.

$$3 \geq (1 - \delta)4 + \delta.$$

Rearranging terms a bit, this means that $\delta \geq 1/3$. We refer to $1/3$ as *the critical discount factor for player 1*. The condition for equilibrium is that the discount factor must exceed the critical value for both players.

Problem: Show that the critical discount factor for player 2 is also $1/3$.

Problem: Find the grim-strategies and critical discount factors that support Bb and the corresponding payoff (2,4).

Long-Run versus Short-Run

Our second case is that of an infinitely repeated game in which one player, the long-run player, uses average present value with discount factor δ and the other player, the short-run player, is completely myopic, that is, discounts the future with discount factor 0. The crucial difference between this case and the previous one is that the short-run player behaves relatively passively – simply playing a best-response to whatever the long-run player is expected to do. This is sometime called *rational expectations*. Because the short-run player is always playing a best-response, the best the long-run player can hope to get in the repeated game is his Stackelberg payoff.

In the long-run versus short-run player case, we focus on what the long-run player can get. As is always the case in a repeated game, the repeated static Nash equilibrium played unconditionally is a subgame perfect equilibrium. If the long-run player is player 1, this means that the long-run player gets 1. In addition, if the long-run player is sufficiently patient, then there is a subgame perfect equilibrium in which he gets his Stackelberg payoff. In the case of Player 1 in the example, this is 2.

You should be able to find the grim strategies that support the Stackelberg equilibrium, and to calculate the critical discount factor for the long-run player for which these grim strategies are an equilibrium. The equilibrium path consists of the long-run

player playing his Stackelberg strategy (B) and the short-run player playing a best-response (b). The punishment is the static Nash equilibrium.

The grim strategies should be described as follows: For the long-run player (Player 1) play the Stackelberg strategy B in the first period and as long as Bb has always been played in the past. If ever Bb has not been played in the past, play the static Nash strategy C. For the short-run player (Player 2), play the best-response to the Stackelberg strategy b in the first period and as long as Bb has always been played in the past. If ever Bb has not been played in the past, play the static Nash strategy c.

To find the critical discount factor for the long-run player for which the grim strategies are an equilibrium:

Step 1: Find the long-run player Stackelberg payoff
For player 1 we calculated above this is 2.

Step 2: Find the long-run player static Nash payoff
For Player 1 we calculated above this is 1.

Step 3: Find the best-response of the long-run player to the short-run player play on the equilibrium path, and the payoff to the long-run player from that best-response.
The equilibrium path is the Stackleberg equilibrium Bb – that is, the short-run player plays b on the equilibrium path. The long-run player best-response to b is A or C, both of which give him a payoff of 3.

Step 4: Calculate the average present value to the long-run Player from deviating.
This is the amount from step 3 in the first period, and the static Nash payoff from step 3 subsequently. For player 1, it is 3 followed by 1 forever. The average present value

$$(1 - \delta)\left(3 + \delta + \delta^2 + ...\right) = (1 - \delta)\,(3 + \delta\,\frac{1}{1 - \delta}) = (1 - \delta)3 + \delta$$

Step 5: For the long-run player compare the average present value to remaining on the equilibrium path to that from deviating.

$$2 \geq (1 - \delta)3 + \delta$$

From which we easily find the critical discount factor of $\delta \geq 1/2$.

Problem: Find the grim strategies and critical discount factor that support the Stackelberg payoff when Player 2 is the long-run player and Player 1 is short-run.