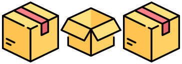
BDP 509: Applied Game Theory



Instructor: Nawaaz Khalfan

Summer Session II, 2022

Lecture Six: Repeated Games

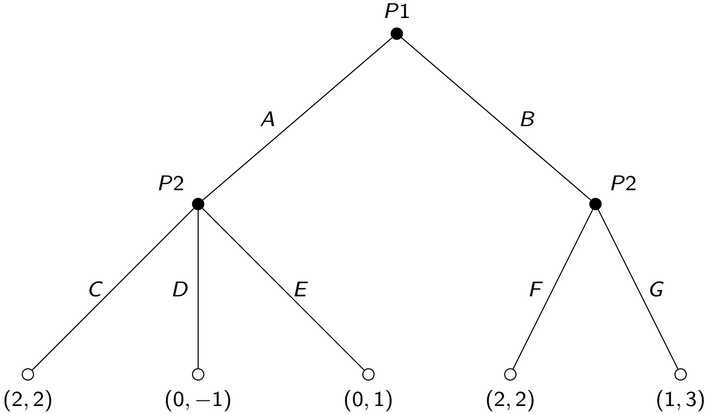
July 27, 2022

Today's Tasks

1. Notifications
2. Review question
3. Reputation and punishment
4. Finitely repeated games
   - A two period repeated game
   - Stage Nash and unraveling
5. Infinitely repeated games
   - Discounting and infinite sums
   - Revisiting the prisoners dilemma

Review

We saw that we can represent sequential games, that is games that have an order of play, with a game tree. For example, consider the game:



Review question

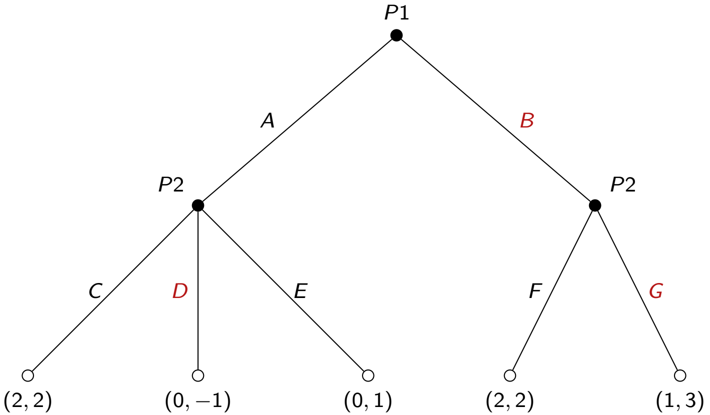
To find the Nash equilibria, we saw that we could translate these games back into strategic form. When we did this, it was important that we listed each players complete, contingent strategies even though some actions in the strategies wouldn't effect the equilibrium path:

		P2					
		CF	CG	DF	DG	EF	EG
$P1$	A	$2, 2$	$2, 2$	$0, -1$	$0, -1$	$0, 1$	$0, 1$
	B	$2, 2$	$1, 3$	$2, 2$	$1, 3$	$2, 2$	$1, 3$

As such the Nash equilibria for this game are: (A, CF), (A, CG), (B, DG) and (B, EG).

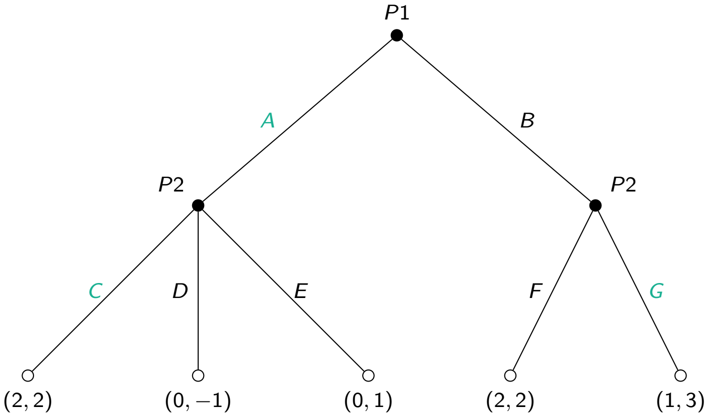
Review

When we analysed these equilibria we saw that some seemed implausible as they involved sub optimal choices being made off the equilibrium path. Nevertheless, they were Nash equilibrium because these choices never had to be made. Consider (B, DG):



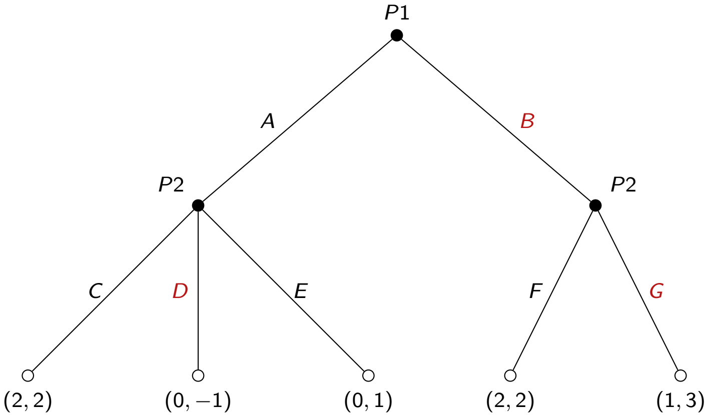
Review

We ruled such equilibrium out with the notion of subgame Perfect Nash equilibria which required all players to play their best strategy at all nodes, acting as if these nodes were reached. We saw that we can find this equilibrium using backwards induction. Consider (A, CG):



A question ...

Consider again (B, DG):



If you were P1, what would make you believe that D was going to be played if you played A?

Reputation and punishment

One answer might be that P2 has a reputation for playing D. Another answer might be the payoffs associate with (A, D) are derived from some commitment to future punishment.

Both of these suggests that P1 and P2 have some kind of long standing relationship: they've played games together before and they'll possibly play games together in the future. This is the case in many social settings:

- individuals interact in the workplace every day,
- voters select political parties each election cycle, and
- countries negotiate trade agreements on an ongoing basis.

We have seen that we can represent multiple staged games with game trees but there are two immediate drawbacks:

1. they're cumbersome to represent and describe, and
2. they're restricted to having finite horizons.

Repeated games

Instead, to get a grasp on topics such as reputation and punishment, we can directly analyse a type of game that may have a simple structure but is repeated either finitely many times or infinitely. Games of this nature are called repeated games and we refer to the simple structure at the core of the game as a stage game.

A two period repeated game

Consider for example, the following two player, simultaneous stage game:

		2		
		L	M	R
	U	$8,8$	$0,9$	$0,0$
1	C	$9,0$	$1,1$	$3,0$
	D	$0,0$	$0,3$	$3,3$

and suppose P1 and P2 play this game twice.

How many strategies does P1 have to choose from? P1 actually has 81 strategies to chose from: 3 actions in the first game, and then, contingent on what P2 chooses, P1 can choose to play each of their 3 actions. That's $3 \times 3 \times 3 \times 3$. This makes drawing up the payoff matrix for the entire game too cumbersome!

Stage Nash

A stage Nash profile is just a Nash equilibrium of the stage game.

			2	
		L	M	R
	U	$8, 8$	$0, \underline{9}$	$0, 0$
1	C	$\underline{9}, 0$	$\underline{1}, \underline{1}$	$\underline{3}, 0$
	D	$0, 0$	$0, \underline{3}$	$\underline{3}, \underline{3}$

Theorem: Any sequence of stage Nash profiles can be supported as the outcome of a subgame perfect Nash equilibrium.

As such, we can immediately find two strategy profiles that are subgame perfect Nash equilibrium:

1. (D in game 1 and D in game 2, R in game 1 and R in game 2)
2. (C in game 1 and C in game 2, M in game 1 and M in game 2)

Co-operation

But can we do better? (U,L) is not a stage Nash profile but it can be sustained in a subgame Nash equilibrium!

$$\begin{array}{cc|c|c|c|}
 & & & 2 & \\
 & & L & M & R \\
\hline
 & U & 8,8 & 0,\underline{9} & 0,0 \\
\hline
1 & C & \underline{9},0 & \underline{1},\underline{1} & \underline{3},0 \\
\hline
 & D & 0,0 & 0,\underline{3} & \underline{3},\underline{3} \\
\hline
\end{array}$$

Consider the strategy profile:

1. U in game 1, D in game 2 if L was played and C otherwise
2. L in game 1, R in game 2 if U was played and M otherwise

If 2 plays M in the second game, C is a best response for 1 and if 2 plays R in the second game, D is a best response. If 1 plays U in the first game, they will get 8 and then 3, but if they deviate to C, they will get 9 and then 1, and: $8 + 3 = 11 > 10 = 9 + 1$. As such, this profile is a SPNE!

Unraveling

But this doesn't work in every game! Note that we used the two different stage Nash profiles to build our punishment and reward schemes. Let's consider the prisoners dilemma which only has the one stage Nash profile:

		Bob	
		deny	*confess*
Alice	*deny*	$2, 2$	$0, 3$
	confess	$3, 0$	$1, 1$

In any SPNE, (confess, confess) must be played in the second game as it's the only stage Nash profile. As such, there's nothing to stop them deviating away from (deny, deny) in the first game!

This is the case no matter how many times this game is played: if Alice and Bob try to co-ordinate on (deny, deny), there will always be an incentive to deviate by confessing in the final stage, and then the penultimate stage, and so on and so forth . . .

Infinitely repeated games

		Bob	
		deny	*confess*
Alice	*deny*	$2, 2$	$0, 3$
	confess	$3, 0$	$1, 1$

But what if the game was to be repeated infinitely many times? Then there wouldn't be a final game for the proposed equilibrium to unravel ...

As always when dealing with "infinities", we have to be careful! For example, is it definitely the case that co-operating on (deny, deny) is preferred to deviating and ending up with (confess, confess)?

Which series of payoffs would you prefer:

$$2,\ 2,\ 2,\ 2,\ \ldots \quad \text{or} \quad 3,\ 1,\ 1,\ 1,\ \ldots \quad ?$$

Discounting

When we think about the payoffs our players receive in the future, we typically discount them. That is we reduce the magnitude of the payoffs because we believe our players value consequences less in the future than they do today. There are three broad reasons we do this:

1. people typically display impatience,
2. there may be some probability of breakdown so that tomorrow or the next game won't occur, and
3. monetary and resource payoffs have some investment or use value today so that they return more in the future than their nominal value.

The discount factor

We can achieve this discounting by multiplying payoffs by a factor, δ, between 0 and 1 for every period we want to discount them. This is called a discount factor.

For example, if you receive \$10 next year and you have a discount factor of 0.8, we say that the net present value of that money is $\$10 \times 0.8 = \$8$ today. If you instead receive the \$10 the year after then we discount it by an additional amount and say that the net present value of that money is $\$10 \times 0.8 \times 0.8 = \$10 \times 0.8^2 = \$6.40$ today.

Observe that if the discount factor is:

- high, it is close to 1, we aren't discounting by much and claiming our player is very patient, and
- low, it is close to 0, we are discounting payoffs by a lot and claiming our player is quite impatient.

Infinite (geometric) sums

Now that we have our way of discounting future payoffs, let's review how to sum an infinite series of numbers. Suppose our player receives a payoff of a every period which we can represent by the following series:

$$a, \ a, \ a, \ a, \ \ldots$$

This means their total payoff after discounting is given the value S:

$$S = a + \delta a + \delta^2 a + \delta^3 a + \ldots$$

Notice that if we multiply this entire sum by δ:

$$\delta S = \delta a + \delta^2 a + \delta^3 a + \delta^4 a + \ldots$$

we get a term that appears in the original sum. As such, taking the second sum away from the first gives us:

$$S - \delta S = a$$

And rearranging and solving for S gives us a formula for the sum of an infinite (geometric) series:

$$S = \frac{a}{1 - \delta}$$

where a is the "flow" payoff and δ is the discount factor.

Example

For example, suppose our player gets \$4 each year and has a discount factor of 0.5. Then using our formula, the net present value of that payoff stream is \$8!

Revisiting the prisoners dilemma

Now we have the ability to sum an infinite series, let's see if we can sustain (deny, deny) on equilibrium path in a subgame Nash equilibrium for the prisoners dilemma:

		Bob	
		deny	*confess*
Alice	*deny*	$2, 2$	$0, 3$
	confess	$3, 0$	$1, 1$

In particular consider the following strategy for both Alice and Bob:

- play *deny* in the first period.
- for all following periods, if (*deny*, *deny*) has been played by in every period prior, play *deny*
- otherwise play *confess*

This strategy is called the grim trigger strategy as once a player deviates, the "grim" outcome is triggered for **all** future periods.

Grim trigger Nash equilibrium

Is the grim trigger strategy profile a subgame Nash equilibrium?

		Bob	
		deny	*confess*
Alice	*deny*	$2, 2$	$0, 3$
	confess	$3, 0$	$1, 1$

There are only two types of subgame: one where in all past periods (*deny*, *deny*) has been played, and one where in some past period *confess* was played by some player.

Suppose we're in the later, then the strategy profile indicates that both players play *confess*. This results in a stage Nash profile which we know is a subgame Nash equilibrium!

Co-operate or deviate?

		Bob	
		deny	*confess*
Alice	*deny*	$2, 2$	$0, 3$
	confess	$3, 0$	$1, 1$

Suppose we're in the former, when in all past periods (*deny*, *deny*) has been played. By continuing to play the grim trigger strategy, that is continuing to play *deny*, Alice receives the following payoff:

$$2 + \delta 2 + \delta^2 2 + \delta^3 2 + \ldots = \frac{2}{1-\delta}$$

If instead Alice chooses to deviate, she receives the following payoff:

$$3 + \delta 1 + \delta^2 1 + \delta^3 1 + \ldots = 3 + \frac{\delta}{1-\delta}$$

As such, if the grim trigger strategy profile is a Nash equilibrium, then it must be that deviating isn't profitable:

$$\frac{2}{1-\delta} \geq 3 + \frac{\delta}{1-\delta} \Leftrightarrow \delta \geq \frac{1}{2}$$

General result

As such, the grim trigger strategy profile allows Alice and Bob to sustain playing (*deny*, *deny*) **if** Alice and Bob are patient enough.

Note that in this equilibrium (*deny*, *deny*) is always played and we've escaped the prisoners dilemma!

More generally let C be the payoff from co-operating, D be the payoff from deviating and P be the payoff from the punishing stage Nash profile. Then, for this idea to work in other games we need:

$$C + \delta C + \delta^2 C + \ldots = \frac{C}{1-\delta} \geq D + \frac{P\delta}{1-\delta} = D + \delta P + \delta^2 P + \ldots$$
$$\delta \geq \frac{D-C}{D-P}$$

There are many different types of strategy profile that also guarantee (*deny*, *deny*) on equilibrium path, but grim trigger does this for the greatest range of discount factors.