14.12 Recitation 7

Margaret Zheng (mzheng01@mit.edu)
Fridays 3-4pm, E51-361

Infinitely Repeated Games

General setup:

- stage game G (what actions each player have and payoff at each stage)
- horizon/timing: $t = 0,1,2,\ldots$
- perfect monitoring: each player observes what everyone else does and knows what happened in the past, but not in the future

Discounting

The idea: future values are worth less than current.

- Which choice do you prefer: \$100 now, or \$100 in 1 month?
- Most people would choose to get \$100 now!
- Discounting factor δ is between 0 and 1
- Can interpret this in several ways: e.g., interest rate, impatience, etc

Strategies for infinitely repeated games:

- For each player i, there needs to be an action at each distinct <u>history</u>
    - History = "which time t am I at now" AND "what happened in the past"
    - Instead of "I will play A at t=0, B at t=1…", the strategies look more like "I will play A at t=0, and at t=1, if my opponent played C, I will play A; but if my opponent played D, I will play B…"
- Yes, there are a LOT of strategies. (even more than the ~Monty Hall problem in PSET 1).
- It is pretty long to write this in normal form, so we can now describe them qualitatively. For example:
    - "Each player sells 1 unit every day, but if one of the players sells more than 1 unit on any day, every other player will sell 2 units a day for forever" (Grim and Trigger)
    - "Each player plays what the other player played last time"

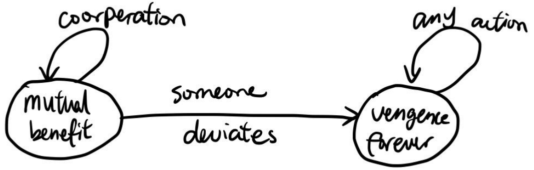
One-Shot Deviation (OSD)

- Start with some predetermined strategy profile (in which player i's strategy is s_i)
- "One shot" = at ONE specific time t, ONE specific player i decides to deviate from s_i
- We use one-shot deviation to find if a strategy profile is an SPNE
  - Idea: an SPNE should satisfy that FOR ALL histories h, at each step, there is no player who has a profitable one-shot deviation.
- The type of answer we look for is:
  - verifying some strategy profile is an SPNE
  - "Under what discount value is this strategy profile an SPNE?"

Common Types of OSD setups

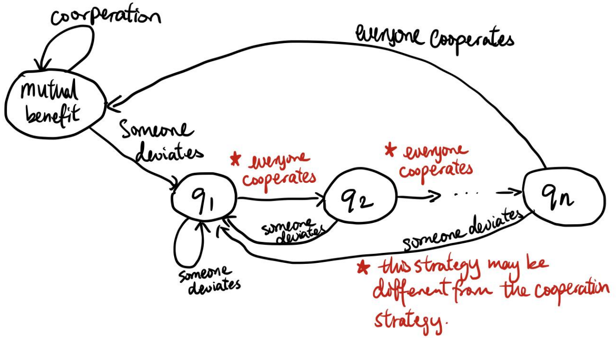
Grim-Trigger

- Punishment for deviation is eternal
- Application: Implicit cartel: n firms agreeing to produce at some quantity to maximize collective profit until someone betrays, then everyone acts in their own interest



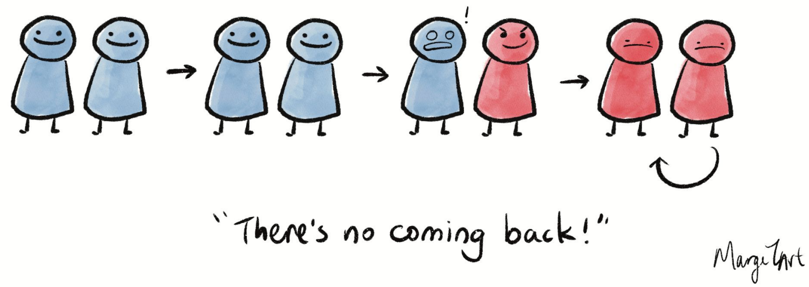
Carrot and Stick

- Punishment for deviation lasts a finite number of rounds; if no more deviation, then go back to normal
- Application: Trade wars

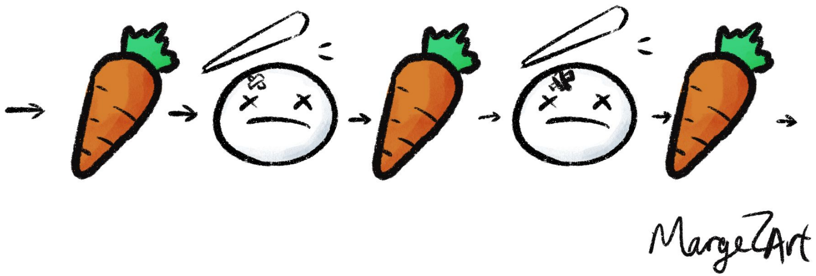


Common Types of OSD setups

Grim-Trigger



Carrot and Stick



Example problem: Walking through a general OSD set-up

Problem setting:

In an infinite repeated game, there are players, 1, 2, and 3. Each player has action {X,Y,Z} at every stage. The discount is δ (between 0 and 1, exclusive).

They agree to the strategy that **"everyone will play X in every stage, but if any player i deviates, then every player will play Z afterwards"**.

Assume that we have no information about the payoffs. How would we set up this question?

Example problem: Walking through a general OSD set-up

In an infinite repeated game, there are players, 1, 2, and 3. Each player has action {X,Y,Z} at every stage. They agree to the strategy that **"everyone will play X in every stage, but if any player i deviates, then every player will play Z afterwards"**.

Assume WLOG (without loss of generality) that **player 1 deviates** at period **t=0**.

- Why can we just pick a stage t=0?

Example problem: Walking through a general OSD set-up

In an infinite repeated game, there are players, 1, 2, and 3. Each player has action {X,Y,Z} at every stage. They agree to the strategy that "**everyone will play X in every stage, but if any player i deviates, then every player will play Z afterwards**".

Assume WLOG (without loss of generality) that **player 1 deviates** at period **t=0**.

Time (stage)	t=0	t=1	t=2	t=3	t=4
Stay	(X,X,X)	(X,X,X)	(X,X,X)	(X,X,X)	(X,X,X)
Deviate (option 1)	(Y,X,X)	(Z,Z,Z)	(Z,Z,Z)	(Z,Z,Z)	(Z,Z,Z)
Deviate (option 2)	(Z,X,X)	(Z,Z,Z)	(Z,Z,Z)	(Z,Z,Z)	(Z,Z,Z)

Example problem: Walking through a general OSD set-up

Time (stage)	t=0	t=1	t=2	t=3	t=4
Stay	(X,X,X)	(X,X,X)	(X,X,X)	(X,X,X)	(X,X,X)
Deviate (option 1)	(Y,X,X)	(Z,Z,Z)	(Z,Z,Z)	(Z,Z,Z)	(Z,Z,Z)
Deviate (option 2)	(Z,X,X)	(Z,Z,Z)	(Z,Z,Z)	(Z,Z,Z)	(Z,Z,Z)

$$\text{Stay: } u_1(X,X,X) + \delta u_1(X,X,X) + \delta^2 u_1(X,X,X) + \cdot = \frac{1}{1-\delta} u_1(X,X,X)$$

$$\text{D1: } u_1(Y,X,X) + \delta u_1(X,X,X) + \delta^2 u_1(X,X,X) + \cdot = u_1(Y,X,X) + \frac{\delta}{1-\delta} u_1(Z,Z,Z)$$

$$\text{D2: } u_1(Z,X,X) + \delta u_1(X,X,X) + \cdot = u_1(Z,X,X) + \frac{\delta}{1-\delta} u_1(Z,X,X) = \frac{1}{1-\delta} u_1(Z,X,X)$$

...Is this it?

Example problem: Walking through a general OSD set-up

We are not done yet!!

What we just showed was under the assumption that player 1 only deviate when no one has deviated. What if player 1 deviated after everyone played Z?

Answer:

- in theory, we would need to compare the utility of player 1 when staying/deviating under every possible history. Only when there is no incentive to deviate at any of those histories we can say we've found an SPNE.
- In practice, since many of those histories yield the exact same payoffs, we just compare the payoffs for deviating and not deviating at each possible state.
- In this game, the states are: "everyone played X" and "everyone plays Z"

Example problem: Walking through a general OSD set-up

More examples:

Time (stage)	**t=0**	t=1	t=2
Stay	(X,Y,Z)	(Z,Z,Z)	(Z,Z,Z)
Deviate (option 1)	(Y,Y,Z)	(Z,Z,Z)	(Z,Z,Z)
Deviate (option 2)	(Z,Y,Z)	(Z,Z,Z)	(Z,Z,Z)

Under the situation that in t=0, player 2 decides to deviate to Y, player 3 decides to deviate to Z, is deviating a good strategy for player 1?

Time (stage)	t=0	**t=1**	t=2
Stay	(Z,Z,Z)	(Z,Z,Z)	(Z,Z,Z)
Deviate (option 1)	(Z,Z,Z)	(Y,Z,Z)	(Z,Z,Z)
Deviate (option 2)	(Z,Z,Z)	(X,Z,Z)	(Z,Z,Z)

Under the situation that the game is already in the stage that everyone has played Z for 1 round, is it profitable for player 1 to deviate at t=1?

Example problem: Walking through a general OSD set-up

- Luckily, in practice, we usually don't have to think about many scenarios
  - More so: we could "group scenarios" smartly
- In many game formats, it is obvious that the game has 2 stages
  - Grim-Trigger
  - Carrot-stick
- In those scenarios - just need to check if it's profitable to deviate under each stage separately!

Remember:

- One shot deviation does not mean "in the whole game, only 1 player deviates at one point in time".
- OSD asks the question: **under <u>every</u> possible history (considering time period), is it possible for player 1 to deviate from the existing strategy for 1 step, and then resume the existing strategy forever?**
- We could be applying OSD even if someone has deviated from the strategy and the game entered "punishment stage", or whatever stage that isn't the same as what we started with!

Problem solving strategies & notes

- **Note the strategy profile we are trying to deviate from.**
  - Example 1: "everyone play C until one person plays D, then everyone plays D for forever"
  - Example 2: "player 1,2 will play (X,X). Then, each player will play X if they played the same thing last time, and play Y if they didn't"
- **Assume that player i deviates to what's best for them if they deviate.**
  - More formally: we say player 1 will not deviate to a strictly dominated strategy
- **Compare payoffs for staying/deviating for every history you need to consider - for the player i you are considering!**
  - Most times, it will be obvious how many states you will need to check
- **(Most times) you can normalize to t=0 for deviation stage**
  - ...because everything else that happened previously will have yielded the same output!
  - An exception: deviating after x turns in a trade war and then the "punishment" stage restarts - we need to shift time period smartly

Problem solving strategies & notes

For **Grim-Trigger**: think infinite sum, geometric series

For **Carrot-Stick**: we usually don't have to use infinite sum, can just add a few terms together

- after a while the payoff will be the same for either deviating or not deviating
- However, we might need to use finite geometric series

In either case...

(In)finite Geometric Series

$$\text{Infinite: } \sum_{i=0}^{\infty} \delta^i = \frac{1}{1-\delta} \text{ for } \delta \in (0,1)$$

$$\text{Finite: } \sum_{i=0}^{n} \delta^i = \frac{1-\delta^{n+1}}{1-\delta} \text{ for } \delta \in (0,1)$$

Example Problem 1 (part 1)

Exercise 12.23. Consider the infinitely repeated game with the following stage game

	Hawk	Dove
Hawk	$0, 0$	$4, 1$
Dove	$1, 4$	$3, 3$

and discount factor $\delta = 0.99$. For each strategy profile below check if it is a subgame-perfect Nash equilibrium.

1. There are two states: Cooperation and Fight. The game starts in the Cooperation state. In the Cooperation state, each player plays Dove. If both players play Dove, then they remain in the Cooperation state; otherwise they go to the Fight state in the next period. In the Fight state, both play Hawk, and they go back to the Cooperation state in the following period ~~(regardless of the actions).~~

Note from recitation: the last sentence question statement is a bit ambiguous. In the following solutions, we ignored the statement "(regardless of the actions)" above. If one were to solve it with this in mind, the question becomes easier and we should get the same conclusion regardless.

Answer (part 1)

Assume WLOG that player 1 thinks about deviating at time t=0 (normalized). We have to think about 2 cases:

Case 1: the game is in cooperation stage. We have

Time period	t=0	t=1	t=2
Stays	(3,3)	(3,3)	(3,3)
Deviates	(4,1)	(0,0)	(3,3)

Note that everything that happens after t=2 is the same!! Thus, we only need to add payoffs from t=0 and t=1:

Stay: $u_1 = 3+3\delta = 3+3*0.99 = 5.97$

Deviate: $u_1 = 4+0\delta = 4$

Thus, it is better to not deviate in cooperation stage.

Answer (part 1)

Assume WLOG that player 1 thinks about deviating at time t=0 (normalized). We have to think about 2 cases:

Case 2: player 1 deviates at fight stage. We have:

Time period	t=0	t=1	t=2
Stays	(0,0)	(3,3)	(3,3)
Deviates	(1,4)	(3,3)	(3,3)

Note that everything that happens after t=2 is the same!! Thus, we only need to add payoffs from t=0 and t=1:

Stay: $u_1 = 0+3\delta = 0+3*0.99 = 2.97$

Deviate: $u_1 = 1+0\delta = 1$

Thus, it is better to not deviate in fight stage, and we conclude that the strategy IS a SPNE.

Example Problem 1 (part 2)

Exercise 12.23. Consider the infinitely repeated game with the following stage game

	Hawk	Dove
Hawk	$0, 0$	$4, 1$
Dove	$1, 4$	$3, 3$

and discount factor $\delta = 0.99$. For each strategy profile below check if it is a subgame-perfect Nash equilibrium.

2. There are three states: Cooperation, P_1 and P_2. The game starts in the Cooperation state. In the Cooperation state, each player plays Dove. If they play (Dove, Dove) or (Hawk, Hawk), then they remain in the Cooperation state in the next period. If player i plays Hawk while the other player plays Dove, then in the next period they go to the state P_i. In state P_i, player i plays Dove while the other player plays Hawk; they then go back to Cooperation state (regardless of the actions).

Answer (part 2)

There are 3 cases to think about: deviating from stage cooperation, P_1, and P_2.

Case 1: player 1 deviates in cooperation stage. We have:

Time period	t=0	t=1	t=2
Stays	(3,3)	(3,3)	(3,3)
Deviates	(4,1)	(1,4)	(3,3)

Stay: $u_1 = 3+3\delta = 3+3*0.99 = 5.97$

Deviate: $u_1 = 4+1\delta = 4.99$

Thus, it is better to not deviate in cooperation stage.

Answer (part 2)

There are 3 cases to think about: deviating from stage cooperation, P_1, and P_2.

Case 2: player 1 deviates in P_1 stage (i.e., player 1 should play dove, and player 2 should play hawk). We have:

Time period	t=0	t=1	t=2
Stays	(**1**,4)	(**3**,3)	(3,3)
Deviates	(**0**,0)	(**1**,4)	(3,3)

It is pretty obvious that deviating is not a good idea for player 1 - deviating is strictly dominated by not deviating (in t=0 and t=1)!!

Answer (part 2)

There are 3 cases to think about: deviating from stage cooperation, P_1, and P_2.

Case 3: player 1 deviates in P_2 stage (i.e., player 2 should play dove, and player 1 should play hawk). We have:

Time period	t=0	t=1	t=2
Stays	(4,1)	(**3**,3)	(3,3)
Deviates	(**0**,0)	(**1**,4)	(3,3)

It is pretty obvious that deviating is not a good idea for player 1 - it is strictly dominated (in t=0 and t=1)!!

Answer: in this setting, the original strategy is also a SPNE.

Example Problem 2

Exercise 12.18. Consider the infinitely repeated game with discount factor δ and the stage game

	a	b	c
a	$3, 1$	$-1, 2$	$0, 0$
b	$-1, 2$	$2, -1$	$0, 0$
c	$0, 0$	$0, 0$	$-1, -1$

For each strategy profile below, determine the range of δ (and possibly the other parameters specified in the strategy profile) under which the strategy profile is a subgame-perfect Nash equilibrium; verify that it is indeed a SPNE for each such δ:

1. There are two states: a and c. In any state s, each player is to play s. The game starts at state a. At any state s, if both players play s, then the state in the next period is a. Otherwise, the state in the next period is c. (Determine δ.)

Answer

There are 4 cases to think about: (the game being in state a or state c) * (the deviating person is P1 or P2).

Note that the WLOG argument does not work here - the payoffs aren't symmetric!

Case 1: player 1 deviates in t=0, game state = a.

If the game state is a, then they should play (a,a).

Observe that assuming that player 2 plays a, it is not profitable for player 1 to deviate from a to b or c.

	a	b	c
a	$3,1$	$-1,2$	$0,0$
b	$-1,2$	$2,-1$	$0,0$
c	$0,0$	$0,0$	$-1,-1$

Answer

There are 4 cases to think about: (the game being in state a or state c) * (the deviating person is P1 or P2).

Note that the WLOG argument does not work here - the payoffs aren't symmetric!

Case 2: player 2 deviates in t=0, game state = a.

If the game state is a, then they should play (a,a). If player 2 deviates, she should play b (it will give her the greatest payoff). The stage payoffs can be written as:

Stay: $u_2 = 1+1\delta$
Deviate: $u_2 = 2-1\delta$

We solve: $1+1\delta = 2-1\delta \rightarrow \delta > 0.5$ would stop player 2 from deviating.

	a	b	c
a	$3, 1$	$-1, 2$	$0, 0$
b	$-1, 2$	$2, -1$	$0, 0$
c	$0, 0$	$0, 0$	$-1, -1$

Time period	t=0	t=1	t=2
Stays	(3,1)	(3,1)	(3,1)
Deviates	(-1,2)	(-1,-1)	(3,1)

Answer

There are 4 cases to think about: (the game being in state a or state c) * (the deviating person is P1 or P2).

Note that the WLOG argument does not work here - the payoffs aren't symmetric!

Case 3: player 1 deviates in t=0, game state = c.

If the game state is c, then they should play (c,c) and in t=1 the game will go back to stage a. Player 1's max payoff from deviating in stage t=0 is 0.

Stay: u_1 = -1+3δ

Deviate: u_1 = 0-1δ = -δ

We solve: -1+3δ =-δ -> δ>0.25 would stop player 1 from deviating.

	a	b	c
a	$3, 1$	$-1, 2$	$0, 0$
b	$-1, 2$	$2, -1$	$0, 0$
c	$0, 0$	$0, 0$	$-1, -1$

Time period	t=0	t=1	t=2
Stays	(-1,-1)	(3,1)	(3,1)
Deviates	(0,0)	(-1,-1)	(3,1)

Answer

There are 4 cases to think about: (the game being in state a or state c) * (the deviating person is P1 or P2).

Note that the WLOG argument does not work here - the payoffs aren't symmetric!

Case 3: player 2 deviates in t=0, game state = c.

If the game state is c, then they should play (c,c) and in t=1 the game will go back to stage a. Player 2's max payoff from deviating in stage t=0 is 0.

Stay: $u_2 = -1+1\delta=\delta-1$

Deviate: $u_2 = 0-1\delta = -\delta$

We solve: $\delta-1 =-\delta$ -> $\delta>0.5$ would stop player 2 from deviating.

Conclusion: when δ>0.5, the strategy we described is a SPNE.

	a	b	c
a	$3, 1$	$-1, 2$	$0, 0$
b	$-1, 2$	$2, -1$	$0, 0$
c	$0, 0$	$0, 0$	$-1, -1$

Time period	t=0	t=1	t=2
Stays	(-1,-1)	(3,1)	(3,1)
Deviates	(0,0)	(-1,-1)	(3,1)