Math 99R: Decision Theory

Sam Marks
smarks@math.harvard.edu
Office 431e (alcove just off the common room)

Office hours: TBD.

Grades based on:
- participation
- 30-min presentation
- Final paper

Game (Prisoner's dilemma)

Two criminals are arrested and imprisoned in separate rooms.
The police admit they don't have enough evidence to convict them on the principal charge (8 year sentence), but can convict them on a minor charge (2 year sentence).
To incentivize the prisoners to testify against each other, the police pose them the following game.

Prisoner 2

Prisoner 1	Remain silent	Testify
Remain silent	-2, -2	-10, 0
Testify	0, -10	-8, -8

"payoff matrix"

Analysis

Let's think about this from prisoner 1's perspective:
- if P2 is silent, then it's best to testify (payoff 0 vs -2)
- if P2 testifies, then it's best to testify (payoff -8 vs. -10)

So P1 should always testify.
Symmetrically, P2 should always testify.

Remarks
- Testify is a dominant strategy for each player
  - dominant strategy = a strategy which, no matter what strategy your opponent uses, is at least as good as any other strategy.
- Both players would prefer (silent, silent) to (testify, testify).
- "Remain silent" is traditionally called "cooperate," and "testify" is called "defect."
- Similar to games that naturally appear in real life.
  - How much to spend on advertising?
    - Back when cigarette advertising was legal, cigarette manufacturers endorsed the making of laws banning cigarette advertising.
  - Sell goods at above market prices
    - Cartels
  - Political donations
  - Arms races / disarmament
  - Doping in sports
  - Studying for exams graded on a curve
  - More examples?

Iterated PD

PD, but played multiple times against the same opponent.
In each round, the players' action may depend on the previous rounds.
Two variants:
- After each round, you stop playing with some probability p.
- You always play some fixed # of rounds, known to both players.

Let's analyze the second variant.

Claim In an IPD w/ finite time horizon both players should always defect.
("Always defect is the unique Nash equilibrium" — stay tuned!)

Argument. We'll argue by backwards induction.
Let N be the number of rounds.
In round N, a player's decision cannot affect the payoffs from the previous rounds, so we're in the setting of the one-off PD, where we already know defect is the dominant strategy.

In round $N-1$, a player's decision cannot affect the payoffs from the previous rounds. Moreover, we just determined that both players will defect in round N, so the decision in round $N-1$ cannot affect the decision in round N either. Thus, we are in the setting of one-off PD and the dominant strategy is defect.

Generally, let $1 \le n \le N$, and suppose that we've shown defect is the dominant strategy in rounds $n+1, \dots, N$. Then the decision in round n cannot affect the payoffs from rounds $1, \dots, n-1$ (because they already happened) nor from rounds $n+1, \dots, N$ (because we've shown that in these rounds both players will defect, regardless of what happens in round n). Thus we're in the setting of the one-off PD, so the dominant strategy is defect.

[Play some IPDs w/ following payoff matrix.]

	C	D
C	4, 4	0, 6
D	6, 0	1, 1

Why was the winner not someone who defected every time?

Possibilities:
- Luck; these were in fact suboptimal strategies, and by chance these bad players never matched with good players.
- Players being irrational systematically caused their opponents to be irrational (e.g. because of social pressure).

⋮

- The argument from before was flawed, or failed to properly model some part of the problem.

I claim it's this last option. Let's look at this part of the argument:
"a player's decision in round n cannot affect the payoff from previous rounds (because they already happened)"

Issues
1) What does this even mean? What is "affect"?
2) If this means something like "the function which outputs a player's move in round $m<n$ has no dependence on the choices of move in round n" then:
   2a) this *might* follow from a hypothesis that the choice of move in round n can only depend on the game transcript up to round n
   2b) But this hypothesis might be false; see (3).

(3) Counterexample: consider a player who:
- is a perfect predictor of its opponent
- plays cooperate in round 1 $\iff$ it predicts its opponent will play cooperate in round 2.

Then the claim is false for games involving this player.

Moral: if your opponent can predict your future moves, then future choices can affect present outcomes.

Q: How can you get your opponent to have useful predictions about your future moves?

Strategy 1: Track record
Strategy 2: Signaling
Strategy 3: Precommitments

Game (Chicken)

	Swerve	Drive
Swerve	0, 0	-2, 2
Drive	2, -2	-100, -100

How to win:
- Rip the steering wheel off your car (precommitment)
- Have always chosen drive in the past (track record)
- Make sure everyone knows you're a reckless sonofabitch (signaling)

Remark One can show that if you think the probability your opponent will drive is
- $<2\%$: you should drive.
- $>2\%$: you should swerve.
- $=2\%$: you're indifferent.

In fact, if both players plan on swerving w/ prob 2%, then neither has an incentive to change strategies. This is an example of a Nash Equilibrium (stay tuned!).

[In a moment, we'll play more IPD's. This time with longer to think about your strategy. Here are some suggestions.]

Grim reaper Cooperate until your opponent defects, then defect forever. (Works best if everyone knows this is your strategy.)
Tit-for-tat Do what your opponent did last round. (Empirically wins IPD tournaments.)

[If time, play:]

Game (smallest unique natural number)
Everyone picks a natural number ($n = 1, 2, 3, \dots$). The winner is the person who picks the smallest number that no one else picked.