# 14.770-Fall 2017
# Recitation 6 Notes

Arda Gitmez

October 20, 2017

Today:

- Markov Perfect Equilibrium.
- A discussion on why dynamic games is different from repeated games. Indirectly, this is a discussion on the difference between political economy and game theory.

Picking from where we left off last week... Recall that our focus was on *infinite horizon multistage games with observed actions.*

**Definition 1.** *An* ***infinite horizon multistage game with observed actions*** *is an extensive form game where:*

- *Set of players $\mathcal{I}$ is finite.*
- *At stage $t = 0, 1, 2, ...$ player $i \in \mathcal{I}$ chooses $a_{it} \in A_{it}$.*[1]
- *All players observe stage $t$ actions $a_t = (a_{1t}, \dots, a_{|\mathcal{I}|t})$ before choosing stage $t+1$ actions.*
- *Players' payoffs are some function of the action sequence: $u_i(a_0, a_1, ...)$. Typically, it is in the discounted sum form, i.e.*

$$u_i = \sum_{t=0}^{\infty} \delta^t u_{it}(a_t)$$

## Markov Perfect Equilibrium

The Markov Perfect Equilibrium (MPE) concept is a drastic refinement of SPE developed as a reaction to the multiplicity of equilibria in dynamic problems. (SPE doesn't suffer from this problem in the context of a bargaining game, but many other games -especially repeated games- contain a large number of SPE.) Essentially it reflects a desire to have some ability to pick out unique solutions in dynamic problems.

### Payoff-Relevant Variables

After seeing examples where SPE lacks predictive power, people sometimes start to complain about unreasonable "bootstrapping" of expectations. Suppose we want to rule out such things. One first step to developing such a concept is to think about the minimal set of things we must allow people to condition on. Often, there is a natural set of payoff-relevant variables.

**Example 1.** *Exploitation of Common Resources (or Common Pool Games, as Daron's notes put it).*

[1] I'm also allowing $A_{it} = \emptyset$, so it's possible that a player does not move at time $t$.

- *$N+1$ fishermen choose quantities/consumptions $(c_0^1, c_0^2, \dots, c_0^{N+1}), (c_1^1, c_1^2, c_1^{N+1}), \dots, (c_t^1, c_t^2, \dots, c_t^{N+1})$.*
- *Stock of fish grows by $K_{t+1} = AK_t - \sum_{i=1}^{N+1} c_t^i$ where $A > 0$ and $K_0$ is given.*
- *Player $i$'s utility is*

$$u_i = \sum_{t=0}^{\infty} \beta^t \log(c_t^i)$$

*With unrestricted strategy spaces, anything could happen with patient agents. This is because either fisherman can fish to extinction in the next period, and construct an equilibrium using it as a threat point.*
*On the other hand, Markov approach allow $c_t^i$ to depends only on $K_t$, and not on how we got there.*

**Example 2.** *Saving with Time Inconsistent Preferences.*

- *Suppose that the consumer chooses consumption $(c_t)_{t=0,1,\dots}$, such that*

$$A_{t+1} = (A_t - c_t)(1 + r_t)$$

  *Suppose*

$$u_t(c_t, c_{t+1}, \dots) = v(c_t) + \beta \sum_{s=1}^{\infty} \delta^s v(c_{t+s})$$

  *with $\delta, \beta \in (0, 1)$.*

- *We can think about this as a consumer playing a game against her future selves. (Makes it a game with infinitely many players, but results still go through.)*
- *In this game, we can construct SPE's where any consumption stream is possible. This is because any player can punish a deviation by consuming everything.*

Clearly we would like a solution concept that eliminates such punishments, and have a solution more like a "repeated static NE".

## Definition, Notation and Examples

Let $G$ be a multistage game with observed actions, i.e. at time $t = 0, 1, 2, \dots$ some subset of players simultaneously choose actions and all previous actions are observed before these choices are made.

**Definition 2.** *Period $t$ histories $h^t$ and $(h')^{t'}$ are said to be **Markov equivalent**, $h^t \sim (h')^{t'}$, if for any two action sequences $\{\alpha_s\}_{s=0}^{\infty}$ and $\{\beta_s\}_{s=0}^{\infty}$ of present and future action profiles and for all $i$:*

$$g_i(h^t, \{\alpha_s\}) \geq g_i(h^t, \{\beta_s\}) \iff g_i((h')^{t'}, \{\alpha_s\}) \geq g_i((h')^{t'}, \{\beta_s\}).$$

Intuitively, two histories are Markov equivalent when the relative preferences over future actions are identical. Note that we allow payoffs to be scaled differently but want decision making to be the same.

Some examples:

- In the common resource game, any two histories $h^t, (h')^{t'}$ with the same number of fish ($S_t = S'_{t'}$) are Markov equivalent.
- In the savings game, any two histories $h^t, (h')^{t'}$ with the same size of wealth ($A_t = A'_{t'}$) are Markov equivalent.
- Suppose in savings game, the action is $\alpha_t \in [0, 1]$: the fraction of wealth to consume. If $v(c) = \log(c)$, then any two histories are Markov equivalent.
- In a repeated game, any two histories are Markov Equivalent.

- In the divide-and-rule model, any two histories which end up with the same state ($K$ or $D$) are Markov equivalent. Indeed, in almost all of the dynamic game models we covered in class, the definition of what a *state* is will be obvious. (More on this later.)

Let $H = \cup_t H_t$ and $A_i = \cup_t A_{it}$. Recall that a **strategy** is a function $s_i : H \to A_i$.

**Definition 3.** *A **Markov strategy** is a function*

$$s_i : H \to A_i$$

*such that*

$$s_i(h^t) = s_i((h')^{t'})$$

*whenever* $h^t \sim (h')^{t'}$.

**Definition 4.** *A **Markov Perfect Equilibrium** of $G$ is a SPE of $G$ in which all players use Markov strategies.*

Note that an effect of this refinement is: it declares past to be irrelevant as long as present value of payoff-relevant variables to be the same. This essentially eliminates the effect of *punishments*.

Here's an existence result:

**Proposition 1.** *A MPE exists in any finite multistage game with observed actions, and in infinite horizon games whose payoffs are continuous at $\infty$.*

And here's another result which (in a loose sense) suggests that Markov strategies are self-enforcing:

**Proposition 2.** *If all other players use Markov strategies, then the remaining player has a Markov strategy among her (unrestricted) best responses.*

This result holds because when the other players are not conditioning on irrelevant things (past play), then you do not have to condition on irrelevant things either.

Does MPE fix the problems we started with?

- In some cases, yes. For instance, for a repeated game, Markov Perfect Equilibria are NE for the static game, repeated each period.
- Similarly, in the divide-and-rule model there are many SPE (where the ruler can condition punishments on the past behavior) but a unique MPE.
- In other cases, it may be tricky. For instance, in the common resources game, if the stock of fish grows in a deterministic manner, then one can still do punishments (e.g. "if number of fish is not exactly 137 we'll catch everything tomorrow.") Putting some noise in the stock of fish sometimes fixes this, but there is no general rule which asserts that it always will.
- In the consumption game, Markov + log utility is much more powerful than Markov alone. In the log utility, each self $t$ chooses an $\alpha_t \in [0,1]$, which is the fraction of wealth to consume. The only Markov Perfect Equilibrium in this game is $\alpha^*$, which solves:

  $$\alpha^* = \arg\max_{\alpha} \log(\alpha A_t) + \beta \sum_{s=1}^{\infty} \delta^s \log(\alpha^* A_{t+s})$$

  where the law of motion is given by:

  $$A_{t+s} = (1-\alpha) A_t (1+r)^s (1-\alpha^*)^{s-1}$$

substituting, rearranging and dropping constant terms, we have:

$$\alpha^* = \arg\max_{\alpha} \log(\alpha) + \beta \sum_{s=1}^{\infty} \delta^s \log(1-\alpha)$$

which implies:

$$\alpha^* = \frac{1-\delta}{1-\delta+\beta\delta}$$

# Dynamic Games

A good resource on dynamic games is the supplementary material by Daron I posted last week. You should really go check it, but let me briefly emphasize the concept of a *dynamic game*. From that document:

> The difference between dynamic games and infinitely repeated games is that in dynamic games, there is an underlying state, which evolves over time as a result of the actions by players and by nature.

So, long story short, a dynamic game makes our lives easier by explicitly giving us the state. Otherwise, we'd need to come up with the definition of a *state* and it's not always obvious. (There is a long discussion here on the possibility of defining the set of all histories as the set of states – this would allow us to make everything Markovian in an uninteresting way. I don't want to go into this discussion, partly because it's confusing and partly because it's not necessary for this class. For this class, if there's a *state* given, you should use it!)

So a dynamic game has a set of states $K$, a set of actions $A_i(k)$ for each $k \in K$, and a (per-period) utility function:

$$u_i : A \times K \to \mathbb{R}$$

along with a Markovian transition function:

$$q(k_{t+1}|a_t, k_t)$$

A history in this context will be the list of actions and states so far:

$$h^t = (a_0, k_0, a_1, k_1, \ldots, a_t, k_t)$$

Just for the sake of comparison with the definition made at the beginning, a *dynamic game* is a game where:

- Set of players $\mathcal{I}$ is finite.
- At stage $t = 0, 1, 2, ...$ player, if state is $k \in K$, $i \in \mathcal{I}$ chooses $a_{it} \in A_i(k)$.
- All players observe stage $t$ actions $a_t = (a_{1t}, \ldots, a_{|\mathcal{I}|t})$ before choosing stage $t+1$ actions.
- Players' payoffs are

$$u_i = \mathbb{E}[\sum_{t=0}^{\infty} \delta^t u_i(a_t, k_t)]$$

This definition is very similar to the definition of a multistage game, but there is more of a sense of stationarity here. (Everything that's time dependent is now taken care of via the states!)

A (pure) strategy for player $i$ in a dynamic game is:

$$s_i : H \times K \to A_i$$

Whereas a (pure) Markov strategy is:

$$s_i : K \to A_i$$

Now, with some contemplation, you should be able to realize that the Markov strategy I define here is really Markov in the sense of Definition 2. (This is because of the stationary nature of the game!)

## Dynamic Games vs Repeated Games

Existential question time: Why are we doing all these?

The analysis of games with discounting has been a focus of game theoretic analysis for years. They have tons of results (folk theorems, reputation results etc.) What are we learning *new* from all of these analysis?

I'd go on to argue that the broad implications are vastly different. A standard game theoretical model argues that in short term relationships (one-shot and finite horizon games) have a difficult time sustaining cooperation among players. A standard folk theorem result basically says that "In an infinite horizon game, with sufficiently patient players (i.e. with more frequent interactions of a longer expected life) we can sustain cooperation." This is a pretty robust insight which all of these game theoretical analysis builds upon. In a dynamic game, this insight can be overturned.

As an example, consider the "Constitutional Choice Game" example – a favorite example of Daron. There are three states: absolutism $a$, constitutional monarchy $c$, and full democracy $d$. There are two agents: elite $E$ and middle class $M$. The per period payoffs are:

$$u_E(d) < u_E(a) < u_E(c)$$
$$u_M(a) < u_M(c) < u_M(d)$$

Assume that $E$ rules in $a$ (i.e. they decide on which state to move to) and $M$ rules in $c$ and $d$. Suppose the initial state is $a$.

It is clear that a move from $a$ to $c$ is a Pareto improvement: both groups are better off. Consequently, $a$ is not Pareto efficient whereas $c$ and $d$ are. Do we obtain a Pareto improvement in equilibrium, are are we stuch with $a$? (Read: when do the elites extend the francise?)

The answer is: *whenever elites are **not** patient, we end up in a Pareto efficient state.* In particular, staying in $a$ yields a payoff of

$$\frac{1}{1-\beta}u_E(a)$$

to the elites. When they move to $c$, they realize that the next period $M$ will move to $d$ and they'll stay there forever. The payoff is:

$$u_E(c) + \frac{\beta}{1-\beta}u_E(d)$$

Clearly, a move to $c$ (and eventually $d$) occurs when

$$u_E(c) + \frac{\beta}{1-\beta}u_E(d) \geq \frac{1}{1-\beta}u_E(a) \Leftrightarrow u_E(c) \geq \frac{u_E(a) - \beta u_E(d)}{1-\beta}$$

which is true for sufficiently low $\beta$!

This is a simple and interesting "flip" on our intuition about the cooperation (i.e. obtaining the Pareto efficient outcome) in general. Indeed, Daron builds his 14.773 lectures based on this intuition – so if you want to see more, take 14.773!

Also, Daron has a paper titled "Why not a Political Coase Theorem?" (2013) investiagting this issue in depth – you can check it if interested.