Lecture 18
3/28/13

Multistage games with observed actions

These games are characterized by:

- The game progresses through a sequence of stages, possibly infinitely many. Often each stage is also a time period, as in repeated games
- Players move simultaneously within a stage
- Players know the actions of players in previous stages

For a finite number of stages, these games can be modeled as extensive form games with imperfect information (imperfect due to the simultaneous moves).

Notation

N players

$a_i(t)$ – action of player i at time t $\quad a_i(t) \in A_i$

$h_t = (a_i(s) : 1 \le i \le N, 0 \le s \le t-1)$ history through stage $t-1$.

A (pure strategy) policy for player i, μ_i,

is a mapping $\mu_i(t, h_t) \longrightarrow A_i$

stage $t \geq 0$ history

Payoff function:

time horizon could be $+\infty$

$$J_i(\mu_i, \mu_{-i}) = \sum_{t=1}^{K} u_i(t, \underbrace{\mu_i(t, h_t)}_{a_i(t)}, \underbrace{\mu_{-i}(t, h_t)}_{a_{-i}(t)})$$

where $u_i(t, a)$ is the payoff function for stage t.

For a given k with $0 \le k \le K$ and history h_k fixed, the subgame for (k, h_k) is the game with payoff function:

$$J_i^{(k)}(\mu_{-i}, \mu_i \mid h_k) = \sum_{t=k}^{K} u_i(t, \mu_i(t, h_t), \mu_{-i}(t, h_t))$$

(only depends on $\mu_j(t, h_t)$ for $t \ge k$)

The strategy profile $(\mu_j : 1 \le j \le N)$ is <u>subgame perfect</u> if it is a Nash equilibrium for every subgame (k, h_k).

Example

$(K=5)$

Six rounds of prisoner's dilema $t=0,1,\dots 5$

Player 2

Player 1	c (cooperate)	d
c	1,1	-1,2
d	2,-1	0,0

NE

In fact, d is a dominant strategy

Trigger strategy μ_i^T for player i: Play c if no player has ever played d. Otherwise play d.

Is (μ_1^T, μ_2^T) a NE?

No. In fact, no matter what happens in first 5 rounds, for any NE, both players play d in last round. So also in second to last round.

So for all rounds

i.e. $\mu_i(t, h_t) \equiv d$ for $i \in \{1,2\}$, $0 \le t \le K$ is a Nash equilibrium with payoff vector $(0,0)$.

Payoff vector is $(0,0)$.

$0+0+0+0+0+0$

If player i uses

$\mu_i(t, h_t) \equiv d$, player i will never encounter

a history such that player i himself played c

at some time -- even if other player deviates.

So $(\mu_1, \overset{\equiv d}{\mu_2})$ would still be a NE if for one i,

μ_i were redefined arbitrarily on histories

(yes, same i)

such that player i played c at least once.

For a subgame perfect equilibrium

need (μ_1, μ_2) to be a NE for every

subgame. (even ones that couldn't possibly occur --

think of nodes on tree of extensive form games.)

Can see $(\mu_1 \equiv d, \mu_2 \equiv d)$ is the <u>unique</u>

subgame perfect equilibrium for six rounds

of prisoner's dilemma.

Variation: After each round, roll a die. If 6 appears stop. Total number of rounds $X \sim Geo(1/6)$.

Is $(\mu_i(t, h_t) \equiv d \quad i=1,2)$ still a subgame perfect NE? Yes.

Any others?

How about if both players play trigger is strategy? "trigger"

$$J_i(\mu_i^T, \mu_{-i}^T) = E\left[\sum_{t=0}^{X-1} 1\right]$$

$$= E[X] = 6 = \sum_{t=0}^{\infty} \left(\frac{5}{6}\right)^t$$

probability game is played in period t.

expected payoff for period t $= P\{\text{game is played in period } t\}$.

Payoff vector is $(6,6)$.

If other player is known to be using the trigger strategy, and you use $\underbrace{\overbrace{c, c, c}^{L \text{ times}}, d}_{\text{start out this way}}$ – for an integer $L \geq 0$.

If you play d once, it's best to play d ever after.

payoff is $\overbrace{1, 1, 1}^{L \text{ times}}, 2, 0, 0, 0 \ldots$ up until $X-1$

i.e. $$\sum_{t=0}^{L-1} \left(\frac{5}{6}\right)^t + 2\left(\frac{5}{6}\right)^L = \frac{1-\left(\frac{5}{6}\right)^L}{1-\frac{5}{6}} + 2\left(\frac{5}{6}\right)^L$$

$$= 6 + \left(\frac{5}{6}\right)^L \{2-6\} < 6.$$

So (μ_i^T, μ^T) is a NE.

It's subgame perfect as well.

True for any continuation probability δ

with $\delta \geq \frac{1}{2}$ (payoff of "cooperate L times")

Strategy	payoff (vs. μ_{-i}^T)
μ_i^T	$\frac{1}{1-\delta}$ ← largest if $\frac{1}{2} \leq \delta < 1$
$d, d, \cdots$	2
$(\underbrace{c, c, \cdots, c}_{L \text{ times}}, d, d, \cdots)$	$\frac{1}{1-\delta} + \delta^L \left\{ 2 - \frac{1}{1-\delta} \right\}$

We can formalize the connection between backward induction and SPE as follows.

<u>Definition</u> A strategy profile μ^* satisfies the one-stage-deviation (OSD) condition if for any $i, \bar{k}, \bar{h}_{\bar{k}}$ with $0 \le \bar{k} \le K$, if $\tilde{\mu}$ agrees with μ^* except at $(i, \bar{k}, \bar{h}_{\bar{k}})$, then

$$J_i^{(\bar{k})}(\tilde{\mu}_i, \mu_{-i}^* \mid \bar{h}_{\bar{k}}) \le J_i^{(\bar{k})}(\mu_i^*, \mu_{-i}^* \mid \bar{h}_{\bar{k}}). \quad (*)$$

<u>Note</u> If K is finite, then μ^* satisfies OSD condition iff μ^* can be derived using backwards induction.

<u>Proposition</u> One-stage-deviation principle ($K < +\infty$) (for multistage games with observed moves)
A strategy profile μ^* is a subgame perfect equilibrium (SPE) iff μ^* satisfies the OSD condition.

<u>Proof</u> (only if) If μ^* is an SPE then, by definition, for a $i, \bar{k}, \bar{h}_{\bar{k}}$ with $0 \le k \le K$, (*) is true for any $\tilde{\mu}_i$, including $\tilde{\mu}_i$ that differs from μ^* only at stage $\bar{k}$ for history $\bar{h}_{\bar{k}}$. So μ^* satisfies OSD cond

(if) Suppose μ^* satisfies the OSD condition. Let i be an arbitrary player and let μ_i be an arbitrary strategy for i. To show that μ^* is an SPE, it suffices to show that for $1 \le k \le K$:

$$J_i^{(k)}(\mu_i, \mu_{-i}^* | h_k) \le J_i^{(k)}(\mu_i^*, \mu_{-i}^* | h_k) \text{ for all } h_k \quad (**)$$

We use proof by backwards induction on k. For the base case, note that (**) is true for $k=K$ because for each choice of h_K, μ_i enters only by its value for stage K and history h_K, $\mu_i(K, h_K)$.

Suppose (*) is true for $k+1$. Consider an arbitrary choice of h_k, and let h_{k+1} be the extension of h_k using μ_i and μ_{-i}^*:

$h_{k+1} = (h_k, \mu_i(k, h_k), \mu_{-i}^*(k, h_k))$. Then

$$J^{(k)}(\mu_i, \mu_{-i}^* | h_k) = u_i(k, \mu_i(k, h_k), \mu_{-i}^*(k, h_k)) + J_i^{(k+1)}(\mu_i, \mu_{-i}^* | h_{k+1})$$

by induction hypothesis

$$\leq u_i(k, \mu_i(k, h_k), \mu_{-i}^*(k, h_k)) + J_i^{(k+1)}(\mu_i^*, \mu_{-i}^* | h_{k+1})$$

$$= J_i^{(k)}(\tilde{\mu}_i, \mu_{-i}^* | h_k)$$

$$\leq J_i^{(k)}(\mu_i^*, \mu_{-i}^* | h_k)$$

(OSD condition)

where $\tilde{\mu}_i$ agrees with μ_i at (k, h_k) and with μ_i^* elsewhere. Hence (**) is true for $k+1$. Therefore, by proof by induction, (**) is true for $0 \leq k \leq K$, completing the proof. $\square$

The OSD principle can be extended to $K=+\infty$ under the following condition:

<u>Definition</u> The game is continuous at infinity if

$$\lim_{k\to\infty} \sup_{i,\mu,\tilde{\mu},h_k} \left| J_i^{(k)}(\mu|h_k) - J_i^{(k)}(\tilde{\mu}|h_k) \right| = 0$$

(True, for example, if stage payoffs have the form $g_i(\mu(t), h_k)\delta^k$ for discount factor δ with $0<\delta<1$.)

<u>Proposition</u> One-stage-deviation principle ($K=+\infty$)

If the game is continuous at infinity then μ^* is an SPE iff it satisfies the OSD condition.

<u>Proof</u> (only if) Same as case $K<+\infty$.

<u>Proof</u> (if) Suppose μ^* satisfies the OSD condition. Let $\varepsilon > 0$. Fix i and a policy μ_i. It suffices to show:

$$J_i^{(k)}(\mu_i, \mu_{-i}^* | h_k) \leq J_i^{(k)}(\mu_i^*, \mu_{-i}^* | h_k) + \varepsilon \text{ for all } h_k \quad (\diamond)$$

for all k. By the continuity at infinity condition, there exists $\tilde{k}$ so large that $(\diamond)$ is true for all $k > \tilde{k}$. The backwards induction proof used for $\underline{K} < +\infty$ can be used to show $(\diamond)$ for $0 \leq k \leq \tilde{k}$ as well, and hence for all k. $\square$

Feasibility theorems (aka folk theorems) for repeated games

Let $G = (I, (S_i : i \in I), (g_i : i \in I))$ be a finite strategic form game and let $0 < \delta < 1$. The repeated game with stage game G and discount factor δ is the strategic form game with payoff functions

$$J_i(\mu) = (1-\delta) \sum_{t=0}^{\infty} \delta^t g_i(\mu(t, h_t))$$

Here μ is a strategy profile

$$\mu = (\mu_i(t, h_t) : i \in I, t \geq 0)$$

Note that $(1-\delta) \sum_{t=0}^{\infty} \delta^t = 1$, so if each player uses a constant strategy $\mu_i(t) \equiv \alpha_i$ then $J_i(\mu) = g_i(\mu)$ for all i.

From example of prisoner's dilema, we've seen existence of multiple equilibria.

Feasibility theorems address what payoff vectors can be realized by the repeated game for δ sufficiently close to one.

Let $\underline{V} = (V_i : i \in I)$ be defined by

$$\underline{V}_i = \max_{\alpha_i} \min_{\alpha_{-i}} g_i(\alpha_i, \alpha_{-i})$$

$$= \text{min max value for player } i \text{ in } G.$$

By repeatedly playing $\hat{\alpha}_i \in \arg\max_{\alpha_i} \min_{\alpha_{-i}} g(\alpha_i, \alpha_{-i})$, player i can insure $J_i(\mu) \geq \underline{V}_i$, no matter what other players play. So

$J(\mu^{NE}) \geq \underline{V}$ for any NE μ^{NE}.

— A vector v is <u>individually rational</u> (IR) if $v_i \geq \underline{v}_i$ for $i \in I$, and strictly IR if $v_i > \underline{v}_i$ for all i.

— A vector v is a <u>feasible</u> payoff vector for game G if

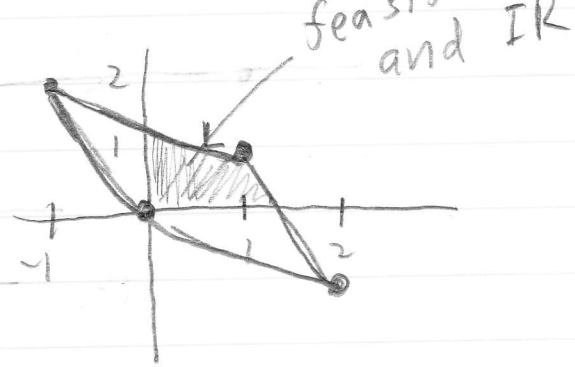
$$v \in \text{convex hull}\left\{ g(\alpha) : \alpha \in \bigtimes_{i \in I} S_i \right\}$$

Feasible payoffs are those that can be achieved for G using mixed strategies and randomization based on publicly available randomness.

eg. prisoners dilema

	d	c
d	1,1	-1,2
c	2,-1	0,0

$\underline{v} = (0,0) \leftarrow$ (Happens to be payoff for unique NE)



<u>Theorem</u> (Nash) If v is feasible and strictly IR then there exists $\bar{\delta} \in (0,1)$ so for any $\delta \in (\bar{\delta}, 1)$, there exists an <u>NE profile</u> μ so that $v = J(\mu)$

<u>Proof</u> (Assuming public randomness available -- can be simulated for $\bar{\delta}$ large enough)

There is a <u>random</u> strategy profile α for G so that $v = E[J(\alpha)]$. Use trigger strategies following α, which switch to punishing first player to violate this script. $\square$

Objection - the NE is typically not subgame perfect. Punishing one player could be costly to the others. (Not for prisoners diloma)

<u>Theorem</u> (Friedman) Let v^{NE} be a payoff vector for some NE α^{NE} of the stage game G. If v is a feasible vector such that $v_i > v_i^{NE}$ for all i, then there exist $\bar{\delta} \in (0,1)$ so that for any $\delta \in (\bar{\delta}, 1)$, there exists a <u>subgame perfect profile</u> μ with $v = J(\mu)$.

<u>Proof</u> (Again, under publicly observable randomness) Player i uses trigger strategy starting with α_i, switching to α_i^{NE} forever if one player deviates from using v. (α is random profile for G with $v = E[g(\alpha)]$.) $\square$

<u>Theorem</u> Under same conditions as
(Fudenberg + Maskin 86)
Nash theorem, if the dimension of the set of feasible vectors is equal to the number of players, then conclusion of Nash theorem holds with μ being sub game perfect. (Set F+T for proof.)

(Idea is to incentivize other players to punish first player to deviate from script for a trigger policy. Easier to prove under availability of public randomness, but still complicated.)

Lecture 20
4/3/13

<u>Bayesian games</u> (Bayes-Nash equilibrium for games with <u>incomplete</u> information

<u>Def</u> A Bayesian game is given by

$$G = (I, (S_i)_{i \in I}, (\Theta_i)_{i \in I}, u_i(s, \underline{\theta}), p(\underline{\theta}))$$

↑ space of types for player i

Marginals: $p(\theta_i) = \sum_{\theta_{-i}} p(\theta_i, \theta_{-i})$. Assume $p(\theta_i) > 0$

so $p(\theta_{-i} | \theta_i)$ is well defined. A pure strategy for a player i is a mapping $s_i : \Theta_i \to S_i$.
A strategy profile $(s_i)_{i \in I}$ is a Bayesian

(or Bayes-Nash) equilibrium if for each i

and $\theta_i \in \Theta_i$,

$$s_i(\theta_i) \in \arg\max_{a_i' \in S_i} \sum_{\theta_{-i}} p(\theta_{-i} | \theta_i) u_i(a_i', s_{-i}(\theta_{-i}), \theta_i, \theta_{-i})$$

In applications the type θ_i of player i is

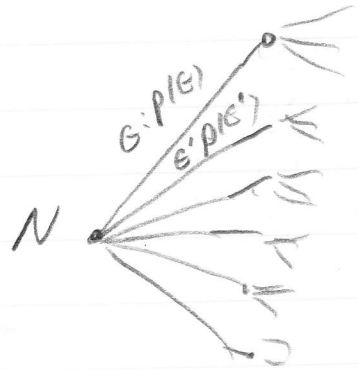
viewed as private information of player i,

determined before the game began.

Harsanyi (73) suggests viewing the situation

as a two stage game in which nature selects the vector of types $(\theta_i)_{i \in I}$ of the players in stage 1 using the known joint distribution p, each player learns his own type, and the players play the game with action spaces $(S_i)_{i \in I}$ and payoff functions $u_i(s, \theta)$. Player i doesn't know θ_{-i}, but knows its conditional distribution $p(\theta_{-i} | \theta_i) = \dfrac{p(\theta_i, \theta_{-i})}{\sum_{\theta'_{-i}} p(\theta_i, \theta'_{-i})}$.

If the game <u>following</u> assignment of θ's by nature has multiple stages, strategies could



involve learning about the types of others.

$\therefore$ Bayes equilibrium exists (special case of result for extensive form games, or use fixed point proof.)

The half-Kuhn homework problem could be viewed as a Bayesian game. It began by dealing one card from a deck of three cards ($\{L, M, H\}$) to each of player 1 and player 2. The card of player i can be viewed as a type. In that example, types are dependent.

<u>Example</u> A Cournot game with incomplete information. (see problem set 5) Given $c > a > 0$, $0 < p < 1$.

Suppose production cost (per unit) for player 1 is c_1 with probability p ($\theta_1 = 1$)
0 " " $1-p$ ($\theta_1 = 0$)

and for player 2 is c (with probability one).

cost is 0. cost is c

Player 1 strategy is a pair $(q_{1,0}, q_{1,1})$

2 " " a number q_2

$$
\begin{cases}
q_{1,0} = \arg\max_{q_1} q_1(a - q_2 - q_1) \\
\quad = \frac{(a-q_2)_+}{2} \\
q_{1,1} = \arg\max_{q_1} q_1(a - c - q_2 - q_1) \\
\quad = \frac{(a-c-q_2)_+}{2} \quad \text{(player 1 produces less if his cost is larger)} \\
q_2 = \arg\max_{q_2'} (1-p) q_2'(a - q_{1,0} - c - q_2') \\
\qquad + p q_2'(a - q_{1,1} - c - q_2')
\end{cases}
$$

$$= \arg\max_{q_2'} q_2'(a - \bar{q}_1 - c - q_2') = \frac{(a-c-\bar{q}_1)_+}{2}$$

where $\bar{q}_1 = (1-p) q_{1,0} + p q_{1,1}$. Note: $\bar{q}_1 = \frac{a-pc-q_2}{2}$ if "$(\ )_+$" above are inactive.

So we have fixed point equations for $(\bar{q}_1, q_2)$. ... continued in ps 5 #4.

Fudenberg and Tirole (1988) introduced a notion of perfect Bayes equilibrium (see their book). Their PBE is a somewhat weaker property than sequential equilibrium in the equivalent extended game (with complete information, using first node controled by nature.)

(Discussed in Myerson book).

The Vikrey-Clarke-Groves (VCG) mechanism.

Let $\mathcal{C}$ represent a set of possibilities for an allocation, or social choice, affecting N players.

Examples:

(a) $\mathcal{C}$ could be a set of possible locations of a new school in a community.

(b) Auction of a single object.

$\mathcal{C} = \{1, \cdots, N\}$; i means player i gets object.

(c) Auction of objects in a finite set $\mathcal{O}$.
$\mathcal{C} = \{(A_1, \cdots, A_n) : A_i \subset \mathcal{O},\ A_i A_j = \emptyset\}$

Suppose player i has a utility function $u_i(x)$. A <u>mechanism</u> $M = (g, m_1, \cdots, m_N)$ takes a vector of bids $(s_1, \cdots, s_N)$ of the players and assigns an allocation $g(s) \in \mathcal{C}$ and a set of payments $(m_1(s), \cdots, m_N(s))$ to be made by the players.

A mechanism induces a game among the players. The payoff function for player i is

$$\pi_i(s) = u_i(g(s)) - m_i(s)$$

If $x^* = g(s^{NE})$, where s^{NE} is a Nash equilibrium for the game, and x^* satisfies some property P, we say that the mechanism is a Nash <u>implementation</u> of property P.

The VCG mechanism, described next, gives an implementation in weakly dominated strategies (so stronger than Nash equilibrium) of maximum welfare allocation.

The (social) welfare of an assignment x is defined by $W(x) = \sum_{i=1}^{N} u_i(x)$.

The set of maximum welfare assignments: $\arg\max_{x \in \mathcal{C}} W(x)$

<u>VCG mechanism</u>

Space of bids is space of utility functions. So use $\hat{u}_i$ instead of s_i for notation.

Given the bids $\hat{u}_i$, x^* is selected with

$$x^* \in \arg\max_{x \in \mathcal{C}} \sum_{i=1}^{n} \hat{u}_i(x)$$

(i.e. x^* maximizes welfare for <u>reported</u> utility functions.

Payments

$$m_i(\hat{u}) = -\sum_{j: j \neq i} \hat{u}_j(x^*) + \underbrace{t_i(\hat{u}_{-i})}_{\text{transfer not influenced by bid of player } i}$$

<u>Claim</u> Bidding u_i (i.e. truthful bidding) is a weakly dominant strategy of player i.

Notes: (a) $m_i(\hat{u})$ depends on $\hat{u}_i$ only through the fact $\hat{u}_i$ influences the selection of x^*.

(b) Above, $\hat{u}_{-i}$ stands for the vector of <u>functions</u>, $\hat{u}_{-i} = (\hat{u}_j : j \neq i)$. It is <u>not</u> the vector of functions evaluated at x^*. So $\hat{u}_{-i}$ does not depend on the bid, $\hat{u}_i$, of player i. A common choice for t_i is discussed below.

<u>Proof of claim</u> The payoff of player i is given by

$$\pi_i(\hat{u}) = \underbrace{u_i(x^*) + \sum_{j:j\neq i} \hat{u}_j(x^*)}_{\substack{\text{social welfare of } x^* \\ \text{using } u_i \text{ and reports} \\ \text{of other players}}} - \underbrace{t_i(\hat{u}_{-i})}_{\substack{\text{Not influenced} \\ \text{by } \hat{u}_i}}$$

Thus, the payoff of player i is the social welfare (computed using u_i and $\hat{u}_{-i}$) minus a term not depending on $\hat{u}_i$. No matter what $\hat{u}_{-i}$ was reported by other players, by reporting u_i truthfully (i.e. taking $\hat{u}_i = u_i$) player i directs the mechanism to select x^* to maximize $\pi_i(\hat{u})$ with respect to $\hat{u}_i$. $\square$

(Basically, the payment rule aligns the interests of each player i with the social good.)

<u>A common choice for $t_i(\hat{u}_{-i})$</u>

If player i were not present, the maximum social welfare for the other players, based on their reported bids,

$$\max_{x \in \mathcal{C}} \sum_{j: j \neq i} \hat{u}_j(x)$$

This is a function of $\hat{u}_{-i}$, and is a common choice for $t_i(\hat{u}_{-i})$. For this choice,

$$(*) \qquad m_i(\hat{u}) = \left(\max_{x \in \mathcal{C}} \sum_{j: j \neq i} \hat{u}_j(x) \right) - \sum_{j: j \neq i} \hat{u}_j(x^*)$$

Then $m_i(\hat{u}) \geq 0$ and it can be viewed as the loss in the total welfare of the other players caused by the actions/existence of player i.

Recovering second price auction from VCG

Recall the example of auctioning a single object: $\mathcal{C} = \{1, \cdots, N\}$. Suppose player i has a value v_i for the object. Then $u_i(x) = v_i I_{\{x=i\}}$. Each player i reports a value $\hat{v}_i$. Then

$$x^* = \arg\max_{1 \le x \le N} \hat{v}_x$$

For the choice $(*)$, the payments are

$$m_i(\hat{u}) = \begin{cases} \max_{j: j \ne i} \hat{v}_j & \text{if } x^* = i \\ 0 & \text{else} \end{cases}$$

That is, the payment of the winner is the highest bid of the other players. Losers make zero payment. The winner is a player with a maximum bid. Thus, VCG reduces to the Vickrey second price auction.