Juegos repetidos: 14.126

Sergei Izmalkov, Muhamet Yildiz
MIT

1 Introducción

1.1 Ideas

- interacción repetida, cooperación, normas sociales
- horizonte infinito, fecha final incierta
- efectos estratégicos de jugadas propias, amenazas
- castigo, venganza
- detección, supervisión
- reputación, colusión
- *aprendizaje

1.2 Principio de desviación de una sola etapa

Teorema: juego finito en varias etapas, acciones observadas.

El perfil s es SPE ⇔ para cada jugador i no existe desviación provechosa de una etapa.

Prueba: suponga que las desviaciones de una etapa no son rentables. Suponga que existe una desviación provechosa para un jugador i en un subjuego, esto es

$$\exists i, t, h^t, \hat{s}_i \quad (\hat{s}_i, \mathrm{s}_{-i})\,|_{h^t} \succ_i (s_i, \mathrm{s}_{-i})\,|_{h^t}.$$

Luego $t^* = \max t'$ **que** $\hat{s}_i(h^{t^*}) \neq s_i(h^{t^*})$**. Obviamente** $t^* > t$**. La desviación de una etapa implica** $(\hat{s}_i, \mathrm{s}_{-i})\,|_{h^{t*}} \approx_i (s_i, \mathrm{s}_{-i})\,|_{h^{t*}}$

Defina $\tilde{s}_i(h^\tau) = \hat{s}_i(h^\tau)$ **para todo** $\tau < t^*$, $\tilde{s}_i(h^\tau) = s_i(h^\tau)$ **para todo** $\tau \geq t^*$

Así, $(\hat{s}_i, \mathrm{s}_{-i})\,|_{h^\tau} \approx_i (\tilde{s}_i, \mathrm{s}_{-i})\,|_{h^\tau}$ **para todo** h^τ.

Repita para $\tilde{s}_i\ (t^* \downarrow)$.

Definición: el juego es *continuo al infinito* si por cualquier i,

$$\sup_{h, \tilde{h},\ \text{s.t. } h^t = \tilde{h}^t} \left| u_i(h) - u_i(\tilde{h}) \right| \to 0 \text{ as } t \to 0.$$

Ejemplo: descuento + pagos por etapas acotados.

Teorema: principio de desviación de una etapa para juegos con pagos continuos.

Prueba: sea ε el tamaño de la mejora. Recorte el extremo que tenga menor importancia que $\varepsilon/2$. El resto es juego finito. No hay lugar para mejora.

Comentario: si los pagos se definen de forma diferente, p.ej. la media por el transcurso del tiempo, no es necesario que se mantenga el principio.

1.3 Ejemplos

El dilema del prisionero

	C	D
C	$1,1$	$-1,2$
D	$2,-1$	$0,0$

	L	M	R
U	$0,0$	$3,4$	$6,0$
M	$4,3$	$0,0$	$0,0$
D	$0,6$	$0,0$	$5,5$

2 Juegos repetidos con acciones observables

2.1 El modelo

Juego de etapa G: $\mathcal{I}$**-jugadores,** A_i**-acciones,** $g_i : A = \times_{i \in \mathcal{I}} A_i \to \mathbb{R}$**-pagos,** $\mathcal{A}_i$**-distribuciones de probabilidad sobre** A_i**.**

Juego repetido: $\mathrm{a}^t \equiv (a_i^t)_{i \in \mathcal{I}}$, h^0**-historial nulo,** $h^t = (\mathrm{a}^0, \mathrm{a}^1, \ldots, \mathrm{a}^{t-1})$, $H^t = (A)^t$

Estrategia (pura) $s_i \equiv (s_i^t)$, $s_i^t : H^t \to A_i$; **(mezclado)** $\sigma_i \equiv (\sigma_i^t)$, $\sigma_i^t : H^t \to \mathcal{A}_i$.

Pagos:

- **Descuento:** $G(\delta)$,

$$u_i = E\boldsymbol{\sigma}(1-\delta)\sum_{t=0}^{\infty}\delta^t g_i(\sigma^t(h^t)) \to \max.$$

- **Criterio del promedio temporal**

$$\lim \inf_{T\to\infty} E\frac{1}{T}\sum_{t=0}^{T} g_i(\sigma^t(h^t)) \to \max.$$

- **Criterio del adelantamiento**

$$g = (g^0, g^1, \ldots) \precsim_i h = (h^0, h^1, \ldots)$$
$$\text{if } \exists T', \forall T > T', \quad \sum_{t=0}^{T} g^t \le \sum_{t=0}^{T} h^t.$$

Observación: si α^* es un equilibrio de Nash en G, entonces "jugar α^* (α_i^* por cada i) para todo t" es SPE.

2.2 Teoremas de Folk

Pagos factibles, racionales individualmente.

Utilidad de reserva (minimax):

$$\underline{v}_i = \min_{\alpha_{-i}}\left[\max_{\alpha_i} g_i(\alpha_i, \alpha_{-i})\right].$$

Defina m^i para ser un perfil minimax.

Observación: en cualquier equilibrio de Nash estático o equilibrio de Nash de juego repetido, el pago de i no es más bajo que $\underline{v}_i$.

Prueba: juegue $a_i(h^t)$ para maximizar $Eg_i(a_i, \sigma_{-i}(h^t))$, donde σ es estrategia de Nash.

Pagos factibles (con aleatorización):

V = c.h. $\{v = g(a)$, *para a* $\in A\}$.

Teorema: (Folk) para cualquier $\mathrm{v} \in V$, con $v_i > \underline{\mathrm{v}}_i$ para todo i, existe un $\delta^* < 1$, tal que para todo $\delta \in (\delta^*, 1)$ existe equilibrio de Nash con pagos v.

Prueba: castigo por minimax.

Teorema: (Friedman, amenazas Nash) α^* es un Nash estático con pagos *e*. Entonces para cualquier $\mathrm{v} \in V$, con $v_i > e_i$ para todo i, existe un $\delta^* < 1$ tal que para todo $\delta \in (\delta^*, 1)$ existe SPE de $G(\delta)$ con pagos *v*.

Prueba: castigo por Nash. SPE sigue de la observación anterior.

Teorema: (Aumann, Shapley) criterio de media temporal, entonces o cualquier $\mathrm{v} \in V$, con $v_i > \underline{\mathrm{v}}_i$ para todo i, existe un SPE de $G(\delta)$ con pagos *v*.

Prueba: castigo por minimax por un tiempo determinado. Los efectos a largo plazo son cero.

Teorema: (Fudenberg, Maskin) suponga $\dim V = \#\mathcal{I}$. Entonces por cualquier $\mathrm{v} \in V$, con $v_i > \underline{\mathrm{v}}_i$ para todo i, existe un $\delta^* < 1$, tal que para todo $\delta \in (\delta^*, 1)$ existe SPE de $G(\delta)$ con pagos *v*.

Prueba: la idea es recompensar a los que castigan. Suponga que por todo *v* tenido en cuenta, existe un $g(\mathrm{a}) = \mathrm{v}$. Ya que $\dim V = \#\mathcal{I}$, $\exists \mathrm{v}' \in V$, $\underline{\mathrm{v}}_i < v_i' < v_i$ para todo i, y $\mathrm{v}'(i) \in V$, que

$$\mathrm{v}'(i) = \left(v_1' + \varepsilon, \ldots, v_{i-1}' + \varepsilon, v_i', v_{i+1}' + \varepsilon, \ldots, v_I' + \varepsilon\right).$$

Suponga $\mathrm{a}'(i)$ existe que $g(\mathrm{a}'(i)) = \mathrm{v}'(i)$.

Fase 1. Juegue a hasta que la acción realizada sea a o difiera de a en ≥ 2 componentes. Si $a_j' \neq a_j$ cambie a fase 2_j.

Fase 2_j. Juegue m^j durante N periodos con tal que la acción realizada sea m^j o difiera de m^j en ≥ 2 componentes. Cambie a fase 3_j. Si alguna k se desvía, cambie a fase 2_k.

Fase 3_j. Juegue $\mathrm{v}'(j)$ siempre con tal que la acción realizada sea $\mathrm{a}'(j)$ o difiera de $\mathrm{a}'(j)$ en $\geq$ 2 componentes. Si el licitador k se desvía, cambie a la fase 3_k.

Utilice el principio de desviación de una sola vez.

Problema: si $\mathrm{a}'(i)$ está mezclado, hay que garantizar el mismo pago de continuación para todas las acciones de apoyo.

Teorema: (Abreu, Dutta, Smith) condición NEU en lugar de $\dim V = \#\mathcal{I}$.

Definición: ***NEU*** (utilidades no equivalentes) se satisface si por cualquier (i, j), $\nexists c, d \in \mathbb{R}_+$ que $g_i(\mathrm{a}) = c + dg_j(\mathrm{a})$ para todo $\mathrm{a} \in A$.

Prueba: NEU $\Longrightarrow \exists \left[\mathrm{v}^1, \ldots, \mathrm{v}^I\right]$ tal que $\forall i, j,\ v_i^i < v_i^j$.

Aprox.: sustituya $\mathrm{v}'(i)$ por v^i.

2.3 Juegos finitos

Teorema: (Benoit, Krishna) criterio de promedio temporal. Suponga $\forall i$ existe Nash estático $\alpha^*(i)$ con $g_i(\alpha^*(i)) > \underline{\mathrm{v}}_i$. Entonces el conjunto de pagos del equilibrio de Nash de G^T converge como $T \to \infty$ con el conjunto de pagos IR factibles de G^∞.

Prueba: fase de recompensa terminal. ciclo $R \times I$:

$([\alpha^*(1), \ldots, \alpha^*(I)])^R$—ruta equilibrio Nash.

Da estrictamente más que $\underline{\mathrm{v}}_i$ a cada i. Para una R mayor la amenaza de la estrategia minimax sobre periodos RI impide toda desviación.

Ajuste $\varepsilon > 0$. Existe T, tal que el pago sobre $T - RI$ periodos aproxima v_i para todo i en ε. ...

2.4 Oponentes distintos

2.4.1 Jugadores a corto plazo frente jugadores a largo plazo

Si los jugadores a corto plazo mueven primero, se puede lograr la "cooperación".

Principio: jugador S-R juega C, jugador L-R responde C con tal que (C, C) se jugase en el pasado. En otro caso, D.

Movimientos simultáneos: el jugador S-R siempre juega BR.

1, . . . , l - jugadores L-R ,

$l+1, \ldots, I$ - jugadores S-R,

$B : \times_{i=1}^{l} \mathcal{A}_i \to \times_{j=l+1}^{I} \mathcal{A}_j$ − BR correspondencia.

$$\underline{v}_i = \min_{\alpha \in \text{graph}(B)} \left[\max_{a_i} g_i(a_i, \alpha_{-i}) \right],$$

$V = \text{c.h.} \left\{ \text{v} = (g_i(\text{a}))_{i=1}^{l} \in \mathbb{R}^l, \right.$ **para** $\alpha \in$ **grafo**(B)**}**

La observabilidad de las acciones mezcladas es importante. Los jugadores a largo plazo han de estar indiferentes entre las acciones puras a las que asignan probabilidades positivas.

$$\bar{v}_i = \max_{\alpha \in \text{graph}(B)} \left[\min_{a_i \in \text{supp}(\alpha_i)} g_i(a_i, \alpha_{-i}) \right].$$

Teorema: (Fudenberg, Kreps, Maskin).

Suponga $\dim V = l$.

Para cualquier $\text{v} \in V$**, con** $\bar{v}_i > v_i > \underline{v}_i$ **para todo** i**, existe un** $\delta^* < 1$ **tal que para todo** $\delta \in (\delta^*, 1)$ **existe un SPE de** $G(\delta)$ **con pagos** v**.**

2.4.2 Generaciones superpuestas

Los jugadores viven T periodos. Cada generación tiene la misma masa.

Las acciones son observables: *work* o *shirk* (IR, NE estático). Todo *work* es eficiente.

Los pagos son promedios durante la vida.

Resultado (Crémer): el equilibrio de Nash existe donde todos excepto el mayor *work*.

Teoremas de Folk: Candori, Smith.

2.4.3 Correspondencia aleatoria

¿Qué es observable? ¿Qué es recordado? Información pública frente a información privada.

El dilema del prisionero: juegue C con tal que se haya jugado (C, C). D en otro caso.

Sostenible con tal que δ sea lo bastante alto y se conozca alguna información del oponente.

Si sólo son observables resultados privados pasados, con N lo bastante alto, las estrategias de "contagio" tal vez no sean un equilibrio.

Razón: responder C sobre D disminuye el contagio.

2.5 Perfección de Pareto

$$Eff(C) = \{x \in C, \nexists y \in C, y \geqq x, y \neq x\}.$$

Definición: (Bernheim, Peleg, Whinston) Considere que G^T, P^T **es el conjunto de pagos SPE de estrategia pura de** G^T. $Q^1 = P^1$, $R^1 = Eff(P^1)$.

Para $T > 1$, $Q^T \subseteq P^T$**—el conjunto de pagos de estrategia pura y equilibrio perfecto impuesto por** R^{T-1}**. Sea** $R^T = Eff(Q^T)$

SPE σ **es perfecto Pareto si,** $\forall t$ **y** $\forall h^t$**, los pagos continuos supuestos por** σ **están en** R^{T-t}.

Ejemplo: (Benoit, Krishna) $\delta = 1$.

	b_1	b_2	b_3	b_4
a_1	$0,0$	$2,4$	$0,0$	$5.5,0$
a_2	$4,2$	$0,0$	$0,0$	$0,0$
a_3	$0,0$	$0,0$	$3,3$	$0,0$
a_4	$0,5.5$	$0,0$	$0,0$	$5,5$

3 Juegos repetidos con información pública imperfecta

3.1 El modelo

$a \in A \to \Delta(y)$, $y \in Y$**—observable públicamente.**

$\pi_y(a) \in \Delta(y)$; $\pi(a)$

$r_i(a_i, y)$**—pago a** i**, (!) independiente de** a_{-i}.

$g_i(a) = \sum_y \pi_y(a) r_i(a_i, y)$.

$h^t = (y^0, y^1, \ldots, y^{t-1})$**—historia pública**.

z_i^t**—historia privada (acciones pasadas).**

Estrategia (mezclada) $\sigma_i \equiv (\sigma_i^t)$, $\sigma_i^t : H^t \times Z_i^t \to \mathcal{A}_i$.

Definición: σ_i **es una estrategia pública si** $\sigma_i(h^t, z_i^t) = \sigma_i(h^t, \tilde{z}_i^t)\ \forall t, h^t, z_i^t, \tilde{z}_i^t.$

Observación: el pago de un equilibrio de estrategia pura se puede apoyar como el pago de un equilibrio en estrategias públicas.

Definición: σ **es un *equilibrio público perfecto* si para todo** i**,** σ_i **es una estrategia pública y** $\forall t, h^t$, **estrategias** $\sigma|_{h^t}$ **forman un equilibrio Nash.**