Capítulo 11

Juegos repetidos

11.1. Introducción

Dos conceptos nuevos surgen cuando consideramos que un juego puede repetirse un número finito o infinito de veces: coordinación y reputación. La característica principal de los juegos repetidos es que cada vez que se deben tomar acciones, el juego es el mismo. Cuando las acciones o el juego que se juega van cambiando, el modelo es el que se conoce como juegos estocásticos. La repetición finita de un juego puede introducir equilibrios nuevos (equilibrios de Nash o equilibrios perfectos en subjuegos), diferentes a jugar siempre el equilibrio del juego estático. Entre más grande sea el horizonte, más equilibrios pueden surgir.

11.2. Ejemplo horizonte finito

Vamos a estudiar algunos ejemplos en horizonte finito que ilustran algunas de las ideas principales.

Ejemplo 11.1. Considere el siguiente juego repetido dos veces:

1\2	D	C
D	1,1	4,0
C	0,4	3,3

En este juego, todos los equilibrios del juego repetido implican jugar (D, D) en cada etapa. Sin embargo, la estrategia C puede utilizarse por fuera del camino de equilibrio. Por ejemplo, considere la siguiente estrategia para cada jugador. En la primera etapa jugar D y en la segunda etapa, si el adversario jugó D jugar D, de lo contrario jugar la mixta: $(\frac{7}{8}, \frac{1}{8})$, que significa que se juega D con probabilidad $\frac{7}{8}$ y C con probabilidad $\frac{1}{8}$. Para ver esto consideremos dos potenciales casos de desviaciones unilaterales del jugador 1 (como el juego es simétrico, basta con verificar el caso de un jugador). Utilizando las estrategias sugeridas el pago de cada jugador es 2, la suma de los pagos en las dos etapas.

Caso I: El jugador 1 se desvía a C en la primera etapa y en la segunda etapa sigue la estrategia de equilibro, es decir, juega D. En este caso el jugador 2 sigue la estrategia de equilibrio y jugará C en la primera etapa y $(\frac{7}{8}, \frac{1}{8})$ en la segunda etapa. Luego el pago para el jugador 1 es $0 + \frac{7}{8} * 1 + \frac{1}{8} * 4 < 2$, y por lo tanto no hay incentivos a desviarse.

Caso II: El jugador 1 juega D en la primera etapa y se desvía a C en la segunda etapa. El jugador 2 sigue su estrategia de equilibrio y juega D en ambas etapas. Claramente esta desviación no beneficia al jugador 1 que obtiene un pago total de 1.

Ahora, a diferencia del ejemplo anterior, en el que todos los equilibrios son equivalentes en términos del resultado, puede ocurrir que surjan equilibrios nuevos.

Ejemplo 11.2. Considere el siguiente juego repetido dos veces:

1\2	D	C	P
D	1,1	4,0	-1,0
C	0,4	3,3	-1,0
P	0,-1	0,-1	-2,-2

El único equilibrio del juego estático es (D, D). Sin embargo, existen nuevos equilibrios del juego dinámico. Jugar C en la primera etapa y si el oponente juega C en la primera etapa jugar D. Caso contrario jugar P. Para ver esto de nuevo consideremos dos posibles desviaciones unilaterales del jugador 1. Si ambos jugadores usan las estrategas sugeridas el pago es 4 para cada jugador.

Caso I: El jugador 1 se desvía a D en la primera etapa y en la segunda etapa sigue la estrategia de equilibro, es decir, juega D. En este caso el jugador 2 sigue la estrategia de equilibrio jugará C en la primera etapa y P en la segunda etapa, luego el pago para el jugador 1 es $4 - 1 < 4$, luego no hay incentivos a desviarse. Alternativamente, el jugador se pudo haber desviado a P en la primera etapa y después continuar la estrategia de equilibrio y jugar D. En este caso el jugador 2 juega C en la primera etapa y P en la segunda. Po rlo tanto el pago para el jugador 1 es $0 - 1 < 4$.

Caso II: El jugador 1 juega C en la primera etapa y se desvía a C en la segunda etapa. El jugador 2 sigue su estrategia de equilibrio y juega C en la primera etapa y D en la segunda etapa. En este caso el pago para el jugador 1 es $3 + 0 < 4$. Si en la segunda etapa el jugador decide mejor desviarse a P entonces su pago sería $3 + 0 < 4$.

11.3. Horizonte infinito

Ejemplo 11.3. Considere un juego repetido una infinidad de veces del juego estático:

1\2	D	C
D	-10,-10	0,-12
C	-12,0	-1,-1

y suponga que la utilidad de cada jugador es de la forma:

$$\pi_i^\delta = (1-\delta)\sum_{t=1}^{\infty} \delta^{t-1}\pi_i \tag{11.1}$$

donde π_i es el pago para i de acuerdo al juego estático y $\delta \in (0,1)$. El factor de descuento se puede interpretar como el nivel de impaciencia de los jugadores que valoran más los pagos en el presente que en el futuro. Entonces la utilidad intertemporal puede verse como un promedio ponderado de la utilidad en cada uno de los periodos (i.e., el factor $1-\delta$ hace que los pesos sumen 1).

Ahora considere la siguiente estrategia (*trigger strategy*):

1. En $t \geq 1$ jugar C si ningún jugador ha jugado D en $t-1$ o antes.
2. Caso contrario jugar D.

Si $\delta \geq \frac{2}{3}$ entonces esta estrategia es un equilibrio de Nash. Por simplicidad supongamos $\delta = \frac{2}{3}$. Para ver esto, obsérvese que el pago en el equilibrio propuesto π^* es (momentáneamente ignoremos el factor de normalización $\frac{1}{1-\delta}$):

$$\pi^* = \sum_{t=1}^{\infty} (\frac{2}{3})^{t-1}(-1) = -3 \tag{11.2}$$

Donde hemos utilizado que para todo número real $x \in (0,1)$, $\sum_{t=1}^{n} x^{t-1} = \frac{1-x^n}{1-x}$. Ahora, si alguien intenta desviarse en $t_0 = 0$ entonces su pago será:

$$\sum_{t=1}^{t_0-1} (\frac{2}{3})^{t-1}(-1) + 0 + \sum_{t=t_0+1}^{\infty} (\frac{2}{3})^{t-1}(-10) < -3. \tag{11.3}$$

Obsérvese que solamente hemos considerado desviaciones del equilibrio en una sola iteración del juego. Para una demostración de que esto es suficiente véase el Apéndice del libro. Un aspecto muy interesante, que será el objeto de estudio en el resto de este capítulo es, en un juego repetido, donde los agentes son lo suficientemente pacientes, es posible soportar como un equilibrio, resultados eficientes y colaborativos.

El modelo general que vamos a considerar es el siguiente. El juego estático lo denotamos por $(N, (A_i), (\pi_i))$. Este juego se repite en varias ocasiones y todos los jugadores pueden observar lo que todos han jugado en el pasado. Las funciones de pago que vamos a considerar son de la forma:

$$\pi^{\delta} = (1-\delta) \sum_{t=1}^{\infty} \delta^{t-1} \pi_i \tag{11.4}$$

Denotamos este juego repetido una infinidad de veces por $R^{\delta}(\pi)$.

11.3.1. Equilibrios de Nash

Sea $V = Conv\{v \in R^n : v = (\pi_1(a_1)....\pi_n(a_n)), a_i \in A_i\}$. Estos son los pagos posibles del juego estático incluyendo la posibilidad de que los jugadores utilicen un mecanismo de coordinación estocástico. Sea $\underline{v}_i$ el menor valor al que puede ser forzado i si todos los demás coordinan para castigarlo. Esto es:

$$\underline{v}_i = min_{\alpha_{-i}\in\Delta_{-i}} max_{\alpha_i\in\Delta_i} \pi_i(\alpha_i, \alpha_{-i}) \tag{11.5}$$

$\underline{v}_i$ puede interpretarse como una restricción de racionalidad individual ya que cualquier valor menor puede ser bloqueado por i.

Teorema 11.4 (Equilibrio de Nash en Horizonte Infinito)**.** *Sea*$(v_1, ...v_1) \in V, v_i > v_i$ para todo i. Entonces si δ es lo suficentemente grande, existe un equilibrio de Nash del juego $R^\delta(\pi)$ con pago v_i para cada jugador.

Prueba. Por simplicidad supongamos que existen $a_i \in A_i$ tal que $v_i = \pi_i(a_1, ..., a_n)$.[1] Sean $(\alpha_1^i, ..., \alpha_n^i)$ las estrategias que implementan el valor minmax para el jugador i. Esto es, para $j \neq i, \alpha_j^i$ son las estrategias de los demás jugadores que si coordinan fuerzan a i al mayor castigo y α_i^i es la mejor respuesta de i.

Ahora considere la estrategia para el jugador i: en t jugar a_i si antes de ningún jugador j se ha desviado de forma unilateral de jugar a_j. Caso contrario jugar α_i^j donde j es el pirmer jugador que se desvió de forma unilateral. Este es un equilibrio de Nash.

Si i intenta desviarse en t_0 a $\overline{a}_i$ el pago es:

$$(1-\delta)\sum_{t=1}^{t_0-1} \delta^{t-1}\pi(a_i, a_{-i}) + (1-\delta)\delta^{t_0-1}\pi(\overline{a}_i, a_{-i}) + (1-\delta)\sum_{t=t_0+1}^{\infty} \delta^{t-1}\pi(\alpha_i^i, \alpha_{-i}^i) \tag{11.6}$$

$$\leq (1-\delta^{t_0-1})v_i + (1-\delta)\delta^{t_0-1}\pi(\overline{a}_i, a_{-i}) + \delta^{t_0}\underline{v}_i \tag{11.7}$$

donde $\overline{a}_i$ es la mejor respuesta de i en el juego estático a a_{-i}.

Sea $f(\delta) = (1-\delta)v_i + (1-\delta)\delta^{t_0-1}\pi(\underline{a}_i, a_{-i}) + \delta^{t_0}\underline{v}_i$ y obsérvese que $\lim_{\delta\to 1} f(\delta) = \underline{v}_i < v_i$. Luego si δ es lo sufientemente cerca de 1 no existen incentivos a desviarse.

■

Ejemplo 11.5. Considere los juegos estáticos de la página 290 de Vega - Redondo. En ambos casos el valor minmax en estrategias puras para cada

[1] En este caso se puede definir $\underline{v}_i$ utilizando únicamente estrategias puras.

jugador es 1. En el primer juego es posible implementar el pago eficiente $(4,4)$. En el segundo ejemplo, obsérvese que solo hay un equilibrio de Nash en el juego estático. Ahora en el juego repetido es posible implementar el pago $(2,2)$ que es inferior a jugar el equilibrio de Nash del juego estático en todas las repeticiones del juego.

11.3.2. Equilibrios Perfectos en Subjuegos

En el caso de equilibrios perfectos en subjuegos, existe un teorma similar al anterior. Sea $\widetilde{v}_i$ el menor pago que el jugador i puede llegar a tenr en un algún equilibrio de Nash del juego estático. Debe ser claro que $\widetilde{v} \geq \underline{v}_i$.

El teorema principal dice que si tenemos un vector de pagos $(v_1, ..., v_n)$ tal que $v_i > \widetilde{v}_i$ entonces si δ es lo suficientemente grande, existe un equilibrio perfecto en subjuegos del juego repetido $R^{\delta}(\pi)$ que implementa este vector de pagos.

11.4. Principio de una única desviación es suficiente

Fijemos un jugador i y un conjunto de estrategias $(\gamma^*_{-i,t})_{t=1,...}$ para los demás jugadores. Sea $(\gamma^*_{i,t})_{t=1,...}$ una estrategia del jugador i y π^* el pago asociado cuando todos juegan $(\gamma^*_{i,t})_{t=1,...,i=1,...N}$. Supongamos que el jugador i tiene una estrategia $(\gamma_{i,t})_{t=1,...}$ tal que el pago asociado es $\pi > \pi^*$ cuando el usa esta estrategia y todos los demás usan $(\gamma^*_{-i,t})_{t=1,...}$. Entonces existe una estrategia $(\gamma'_{i,t})_{t=1,...}$ para el jugador i tal que:

1. $\gamma'_{i,t} = \gamma^*_{i,t}$ para todo t excepto un único valor de $t = T$.
2. $\pi^{'} > \pi^*$ donde $\pi^{'}$ es el pago de i cuando el usa la estrategia $(\gamma'_{i,t})_{t=1,...}$ y los demás usan $(\gamma^*_{-i,t})_{t=1,...}$.

Para ver esto obsérvese que si $\pi > \pi^*$ entonces no puede ser cierto que para todo t: $\pi_t(\gamma^*_{i,t}, \gamma^*_{-i,t}) \geq \pi_t(\gamma_{i,t}, \gamma^*_{-i,t})$. Sea T cualquier t para el cual: $\pi_T(\gamma^*_{i,T}, \gamma^*_{-i,T}) < \pi_T(\gamma_{i,T}, \gamma^*_{-i,T})$ entonces la estrategia $(\gamma'_{i,t})_{t=1,...}$ para el jugador i tal que:

1. $\gamma'_{i,t} = \gamma^*_{i,t}$ para todo $t \neq T$
2. $\gamma'_{i,t} = \gamma_{i,t}$ para $t = T$

satisface las propiedad deseada.