GIỚI THIỆU LÝ THUYẾT TRÒ CHƠI VÀ MỘT SỐ ỨNG DỤNG TRONG KINH TẾ HỌC VI MÔ

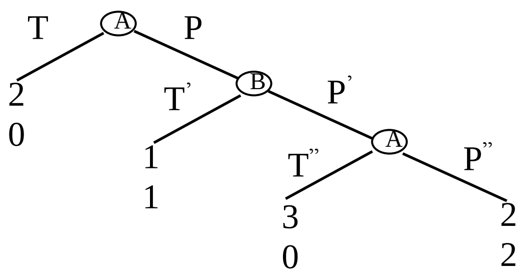
Phần 2: Trò chơi động với thông tin đầy đủ

Trò chơi động (dynamic game) diễn ra trong nhiều giai đoạn, và một số người chơi sẽ phải hành động ở mỗi một giai đoạn. Trò chơi động khác với trò chơi tĩnh ở một số khía cạnh quan trọng. ***Thứ nhất***, trong trò chơi động, thông tin mà mỗi người chơi có được về những người chơi khác rất quan trọng. Như ở Phần 1 đã phân biệt, một người có thông tin đầy đủ (complete information) khi người ấy biết hàm thỏa dụng (kết cục - payoff) của những người chơi khác. Còn một người có thông tin hoàn hảo (perfect information) nếu như tại mỗi bước phải ra quyết định (hành động), người ấy biết được toàn bộ lịch sử của các bước đi trước đó của trò chơi. ***Thứ hai***, khác với các trò chơi tĩnh, trong trò chơi động mức độ đáng tin cậy (credibility) của những lời hứa (promises) hay đe dọa (threats) là yếu tố then chốt. Và ***cuối cùng***, để tìm điểm cân bằng cho các trò động, chúng ta phải vận dụng phương pháp quy nạp ngược (backward induction).

Trò chơi động với thông tin đầy đủ và hoàn hảo

Ví dụ 1: *Một trò chơi tưởng tượng*

Thử tưởng tượng một trò chơi động với thông tin đầy đủ và hoàn hảo và có cấu trúc như hình



vẽ. Tại mỗi nút hoặc A hoặc B phải ra quyết định. Không gian hành động của họ chỉ gồm hai khả năng: hoặc chọn trái (T), hoặc chọn phải (P). Những con số ở ngọn của các nhánh trong cây quyết định chỉ kết quả thu được của hai người chơi, trong đó số ở trên là kết quả của A.

Để tìm điểm cân bằng của trò chơi này, chúng ta không thể bắt đầu từ giai đoạn đầu tiên, mà ngược lại, chúng ta sẽ dùng phương pháp quy nạp ngược, tức là bắt đầu từ giai đoạn cuối cùng của trò chơi.

Lưu ý là phương án tối ưu cho người chơi thứ nhất là kết cục T'', ở đó A được 3 và B không được gì. Còn phương án tối ưu cho B là kết cục P'', trong đó B được 2 và A được 2. Nhìn từ góc độ xã hội, dường như P'' là lựa chọn tối ưu vì nó giúp tối đa hóa tổng phúc lợi cho cả A và B (hiệu quả), đồng thời đạt được tính công bằng cho hai người chơi khi họ hợp tác một cách thiện chí. Nhưng nếu mục đích của mỗi người là tối đa hóa độ thỏa dụng của mình mà không quan tâm đến phúc lợi của người khác thì kết quả này sẽ không xảy ra. Tại sao vậy?

Nếu trò chơi kéo dài đến giai đoạn 3 thì A chắc chắn sẽ chọn T'' (vì 3 > 2). Còn nếu B được ra quyết định ở giai đoạn 2 và biết điều này chắc chắn sẽ không chọn P' mà chọn T' (vì 1 > 0). Và ở giai đoạn 1, A dự đoán trước được những hành động kế tiếp của cả hai người nên chắc chắn sẽ chọn T (vì 2 > 1).[1] Như vậy, trò chơi kết thúc ở ngay giai đoạn thứ nhất với việc A chọn T, và do vậy B không có cơ hội để hành động.

Bây giờ chúng ta quay lại thảo luận vấn đề mức độ tin cậy của lời hứa hẹn hay đe dọa. Giả sử trước khi bắt đầu chơi, B đề nghị với A như sau. Trong lần chơi đầu tiên anh nên chọn P. Nếu thế, khi đến lượt tôi thì tôi sẽ chọn P′, và rồi trong giai đoạn cuối cùng anh sẽ chọn P″ để mỗi chúng ta cùng được 2. Liệu A có nên tin vào lời đề nghị (hứa hẹn) bằng miệng này của B hay không?[2] Nếu đây là trò chơi xảy ra một lần và mục đích của mỗi người chơi đơn thuần chỉ là tối đa hóa lợi ích của mình thì câu trả lời hiển nhiên là không. Lý do là đến giai đoạn 2, B biết chắc là nếu A đổi ý và chọn T″ thì anh ta sẽ không được gì, còn A sẽ được 3 (là kết cục tốt nhất của A). Lường trước điều này, B chỉ đợi A chọn P là sẽ chọn T′ để được 1, đồng thời A cũng chỉ được 1. Đứng trước tình huống này, với những thông tin cho trước và nếu A là người duy lý thì chắc chắn A sẽ không dại gì nghe theo lời hứa hẹn ngon ngọt của B. Kết quả là A sẽ chọn T trong giai đoạn đầu tiên như chúng ta đã phân tích ở trên. Nói một cách ngắn gọn, những hứa hẹn và đe dọa trong tương lai mà không đáng tin cậy sẽ không hề có tác động gì, dù là nhỏ nhất, tới ứng xử của những người chơi trong giai đoạn hiện tại. Trong một phần khác, chúng ta sẽ nghiên cứu tình huống trong đó lời hứa/ đe dọa đáng tin cậy và do đó có ảnh hưởng đến hành vi của những người chơi ngay trong giai đoạn hiện tại.

***Ví dụ 2**: Mô hình độc quyền song phương Stackelberg (1934)*

Nhớ lại trình tự thời gian của trò chơi này như sau:

1) Hãng 1 chọn sản lượng $q_1 \geq 0$
2) Hãng 2 quan sát q_1 rồi sau đó chọn sản lượng $q_2 \geq 0$
3) Hai hãng sản xuất với sản lượng q_1, q_2 và lợi nhuận tương ứng là π_1 và π_2

$\pi_1(q_1, q_2) = q_1[P(Q) - c]$; $Q = q_1 + q_2$

$\pi_2(q_1, q_2) = q_2[P(Q) - c]$; $P(Q) = a - Q = a - (q_1 + q_2)$

trong đó hằng số c là chi phí cận biên, đồng thời là chi phí trung binh của cả 2 hãng.

Để tìm điểm cân bằng của trò chơi này, chúng ta lại áp dụng phương pháp quy nạp ngược bằng cách bắt đầu với hãng thứ 2. Đầu tiên chúng ta phải tìm hàm phản ứng tốt nhất của hãng 2 đối với quyết định sản lượng q_1^* của hãng thứ nhất trong giai đoạn 1 :

$$\max_{q_2 \geq 0} \pi_2(q_1, q_2) = q_2[a - c - q_1^* - q_2] \Rightarrow q_2 = (a - c - q_1^*)/2$$

Lưu ý rằng về mặt hình thức thì hàm phản ứng $q_2(q_1^*)$ ở đây giống như trong mô hình Cournot. Tuy nhiên, có một điểm khác biệt quan trọng là trong mô hình Cournot, q_1^* là một giá trị giả định, còn trong mô hình này, khi ra quyết định q_2 hãng 2 đã <u>quan sát được</u> và biết giá trị của q_1^*.

Vì đây là bài toán với thông tin đầy đủ và hoàn hảo nên hãng thứ nhất có thể đặt mình vào vị trí của hãng thứ hai và do vậy biết rằng nếu mình quyết định sản lượng là q_1^* thì hãng thứ hai sẽ sản xuất $q_2 = (a - c - q_1^*)/2$. Vì vậy, trong giai đoạn 1, hãng thứ nhất sẽ chọn q_1 sao cho

---

[1] Để ý rằng phương pháp quy nạp ngược được sử dụng ở đây một cách dễ dàng là nhờ cấu trúc thông tin đầy đủ và hoàn hảo của bài toán (tưởng tượng) này. Trong các bài toán thực tế, cấu trúc thông tin thường phức tạp hơn nhiều.

[2] Vì là hợp đồng miệng nên nó không thể bị chế tài nhờ trọng tài.

Max $\pi_1(q_1, q_2(q_1)) = q_1[a - c - q_1 - q_2(q_1)] = q_1 \dfrac{a-c-q_1}{2}$ $\Rightarrow q_1^* = \dfrac{a-c}{2}$

$\Rightarrow q_2^* = \dfrac{a-c}{4}$

Lợi nhuận tương ứng là :

$$\pi_{S1}^* = \frac{(a-c)^2}{8} > \pi_{c1}^* = \frac{(a-c)^2}{9}$$

$$\pi_{S2}^* = \frac{(a-c)^2}{16} > \pi_{c2}^* = \frac{(a-c)^2}{9}$$

Câu hỏi đặt ra là tại sao hãng 1 có thể đạt được mức sản lượng và lợi nhuận tương đương với mức sản lượng và lợi nhuận độc quyền trong khi hãng 2 thậm chí còn không đạt được mức lợi nhuận trong độc quyền song phương Cournot? Câu trả lời không thuần túy chỉ nằm ở trình tự thời gian mà quan trọng hơn là do thông tin. Trong ví dụ này, *cả hai hãng đều biết nhiều thông tin hơn so với trường hợp độc quyền song phương Cournot*: Hãng 2 có thể quan sát quyết định về sản lượng của hãng 1, còn hãng 1 biết là hãng 2 biết sản lượng của mình. *Tuy nhiên hãng 1 có thể sử dụng thông tin bổ sung này để làm lợi cho mình trong khi hãng 2 khi có thêm thông tin lại bị thiệt. Hay nói một cách chính xác hơn, việc hãng 2 làm cho hãng 1 biết là hãng 2 biết sản lượng của hãng 1 làm cho hãng 2 bị thiệt.* Để thấy điều này, giả sử bằng một cách nào đó, hãng 2 gây nhiễu thông tin làm cho hãng 1 không biết được là liệu hãng 2 có biết sản lượng của mình hay không. Khi ấy, bài toán trở thành tương tự như với trường hợp độc quyền Cournot trong đó 2 bên quyết định sản lượng mà không hề biết sản lượng thực tế của bên kia (thông tin không hoàn hảo)

Ví dụ 3: *Mặc cả luân phiên* (Rubinstein sequential bargaining) – xem bài đọc thêm.

Trò chơi động với thông tin đầy đủ nhưng không hoàn hảo (xem bài đọc thêm)

Trò chơi lặp lại (repeated games)

Mục đích của tiểu mục này là xem xét liệu các đe dọa hay hứa hẹn tương lai đáng tin cậy ảnh hưởng thế nào tới hành vi hiện tại của những người chơi.

Ví dụ 1: *Thế lưỡng nan trong trò chơi lặp hai giai đoạn*

Quay lại bài toán lưỡng nan của người tù được trình bày dưới dạng chuẩn tắc như trong bảng bên.

Cân bằng Nash duy nhất là (không hợp tác, không hợp tác) và kết cục là (1, 1). Bây giờ giả sử trò chơi này (gọi là trò chơi giai đoạn – stage game) được lặp lại lần thứ hai, bảng kết quả được trình bày trong bảng dưới đây.

		Người 1	
		Không hợp tác	*Hợp tác*
Người 2	*Không hợp tác*	1 , 1	5 , 0
	Hợp tác	0 , 5	4 , 4

Cân bằng Nash duy nhất vẫn là (không hợp tác, không hợp tác) và kết cục hợp tác vẫn không đạt được như là một điểm cân bằng

		Người 1	
		Không hợp tác	*Hợp tác*
Người 2	*Không hợp tác*	2 , 2	6 , 1
	Hợp tác	1 , 6	5 , 5

<u>*Nhận xét*</u>:

- Nếu trò chơi giai đoạn (stage game) chỉ có một cân bằng Nash duy nhất thì nếu trò chơi ấy được lặp lại nhiều lần thì cũng sẽ chỉ có một cân bằng Nash duy nhất, đó là sự lặp lại cân bằng Nash của trò chơi giai đoạn.
- Rõ ràng là nếu trò chơi này được lặp lại nhiều lần thì thiệt hại từ việc không hợp tác sẽ rất lớn. Câu hỏi đặt ra là liệu có cách nào để thiết lập sự hợp tác hay không? Ở đây chúng ta tạm thời không quan tâm tới khía cạnh đạo đức và lương tâm của mỗi người chơi mà chỉ xem xét thuần túy về động cơ kinh tế của họ.

Ví dụ 2: *Thế lưỡng nan trong trò chơi lặp vĩnh viễn*

Bây giờ giả sử trò chơi được lặp lại một cách vĩnh viễn. Chúng ta sẽ xem xét khả năng một đe dọa hay hứa hẹn tương lai đáng tin cậy ảnh hưởng thế nào tới hành vi hiện tại của những người chơi?

Nhớ lại công thức tính hiện giá của thu nhập, trong đó một người nhận được π_1 trong giai đoạn 1, π_2 trong giai đoạn 2 v.v. Tổng thu nhập của người đó tính theo giá hiện tại là $\Sigma PV = \pi_1 + \delta\pi_2 + \delta^2\pi_3 + \ldots$; trong đó δ là nhân tố chiết khấu (discount factor)[3].

Bây giờ chúng ta sẽ chứng minh rằng ngay cả khi trò chơi giai đoạn chỉ có một cân bằng Nash duy nhất thì vẫn có cách để buộc những người chơi duy lý hợp tác với nhau, với điều kiện δ đủ lớn. Cách thức để đạt được sự hợp tác này là thực hiện chiến lược "trừng phạt" (trigger strategy) mà thực chất là một lời đe dọa trả đũa đáng tin cậy đối với những hành vi vi phạm hợp đồng. Chiến lược trừng phạt này được thực hiện như sau:

- Trong giai đoạn 1, cả hai người chơi chọn hành động "hợp tác"
- Trong giai đoạn t, mỗi người chơi tiếp tục chọn "hợp tác" chừng nào trong (t-1) giai đoạn trước người kia cũng chọn "hợp tác"
- Chuyển sang chơi "không hợp tác" nếu trong giai đoạn (t-1), người kia phá bỏ hợp đồng chơi "hợp tác"

Giả sử trong suốt (t-1) giai đoạn đầu tiên, cả hai người chơi đều tuân thủ thỏa ước và chọn "hợp tác". Nhưng tại giai đoạn thứ t, một người toan tính việc vi phạm thỏa ước vì thấy cái lợi trước mắt. Khi ấy, người này phải so sánh 2 giá trị thu nhập kỳ vọng của hợp tác và không hợp tác.

Nếu trong giai đoạn t người ấy không hợp tác thì người ấy được 5, và từ (t+1) trở đi người kia sẽ chọn không hợp tác để trừng phạt người này, và khi ấy phản ứng tốt nhất tương ứng của người này cũng sẽ là không hợp tác. Như vậy, tổng giá trị kỳ vọng thu nhập của người ấy theo hiện giá là:

$$\begin{aligned} PV_{\overline{C}} &= \delta^{t-1}.5 + \delta^{t}.1 + \delta^{t+1}.1 + \ldots \\ PV_{\overline{C}} &= \delta^{t-1}[5 + \frac{\delta}{1-\delta}] \end{aligned} \qquad (1)$$

Còn nếu trong giai đoạn t người ấy vẫn chọn hợp tác thì khi ấy, tổng thu nhập của anh ta theo hiện giá sẽ là:

---

[3] Nhn tố chiết khấu δ = 1/(1 + r), trong đó r là suất chiết khấu (discount rate).

$$PV_C = \delta^{t-1}.4 + \delta^t.4 + \delta^{t+1}.4 + \ldots$$

$$PV_C = \delta^{t-1}.\frac{4}{1-\delta} \qquad (2)$$

So sánh (1) và (2) ta thấy $PV_C \geq PV_{\overline{C}} \Leftrightarrow \quad \frac{4}{1-\delta} \geq 5 + \frac{\delta}{1-\delta}$

<=> 4 ≥ 5(1-δ) + δ = 5 -4δ

<=> δ ≥ 1/4

Như vậy, nếu δ ≥ 1/4 thì chiến lược trừng phạt là một cân bằng Nash. Nói cách khác, với δ đủ lớn (tức là những người chơi chiết khấu tương lai đủ ít) thì khi theo đuổi mục tiêu vị kỉ là tối đa hóa lợi ích của mình thì tất cả người chơi đều có động cơ tôn trọng thỏa ước hợp tác.

Ví dụ 3: *Trở lại với độc quyền song phương Cournot*

Chúng ta đã biết rằng trong trường hợp độc quyền song phương Cournot:

$q_{c1}^* = q_{c2}^*$=(a-c)/3 và do vậy Q_C^* = 2(a-c)/3 > Q_m^* = (a-c)/2 (= mức tổng cầu khi hai doanh nghiệp cấu kết lũng đoạn thị trường độc quyền). Như vậy, hai hãng này có thể áp dụng chiến lược trừng phạt để đạt được sự hợp tác trong sản xuất. Để kiểm tra lại mức độ hiểu các nội dung trình bày ở ví dụ 2, chúng ta có thể làm một bài tập nhỏ sau. Giả sử trò chơi Cournot này được lặp lại mãi mãi, hãy tìm giá trị tối thiểu của δ để giải pháp hợp tác là một cân bằng Nash (SPNE)?

Chiến lược trừng phạt như sau:

- Bắt đầu chơi bằng việc chọn mức sản lượng $Q_{m/2}^*$ (=(a-c)/4) trong giai đoạn 1
- Nếu trong (t-1) giai đoạn đầu tiên, bên kia chọn $Q_{m/2}^*$ thì tiếp tục chọn $Q_{m/2}^*$. Bằng không thì chuyển sang $Q_{c/2}^*$ (= (a-c)/3) mãi mãi.

Giả sử ở giai đoạn t, hãng 1 toan tính chuyện phá vỡ thỏa ước ban đầu. Hãng này biết là hãng 2 sẽ chuyển sang chọn $q_2^* = q_{c2}^*$ kể từ giai đoạn thứ (t+1). Vì vậy, hãng 1 đứng trước hai lựa chọn:

- Phá vỡ thỏa ước:

$$\pi^{\overline{C}} = \delta^{t-1}.\pi_d + \delta^t \pi_C + \delta^{t+1}\pi_C + \ldots$$
$$= \delta^{t-1}(\pi_d + \delta\pi_C + \delta^2\pi_C + ..)$$

$$\pi^{\overline{C}} = \delta^{t-1}(\pi_d + \frac{\delta}{1-\delta}\pi_C)$$

Nếu hãng 2 tiếp tục chọn hợp tác trong giai đoạn t, tức là tiếp tục chọn $q_2^* = Q_{m/2}^*$ = (a - c)/4 thì q_{d1}^* sẽ max q_{d1}[a - c - q_{d1} – (a-c)/4] => q_{d1}^* = 3(a-c)/8 => $\pi_d = 9(a-c)^2/64$

- Tôn trọng thỏa ước:

$$\pi^C = \delta^{t-1}.\pi_m + \delta^t \pi_m + \delta^{t+1}\pi_m + \ldots$$

$$\pi^C = \delta^{t-1}\frac{\pi_m}{1-\delta}$$

So sánh $\pi^C \geq \pi^{\overline{C}}$:

$$\Leftrightarrow \frac{\pi_m}{1-\delta} \geq \pi_d + \frac{\delta}{1-\delta}\pi_C$$

$$\Leftrightarrow \frac{(a-c)^2}{8(1-\delta)} \geq \frac{9(a-c)^2}{64} + \frac{\delta}{1-\delta}\frac{(a-c)^2}{9}$$

Một lần nữa chúng ta lại thấy là nếu δ đủ lớn (tức là những người chơi chiết khấu tương lai đủ ít) thì khi theo đuổi mục tiêu vị kỉ là tối đa hóa lợi nhuận của mình thì hai công ty cùng có động cơ tôn trọng thỏa ước hợp tác.

Tài liệu tham khảo

Robert Gibbons, "Game Theory for Applied Economists", Princeton University Press, 1992