Equivalenza di due condizioni per ESS: dimostrazione

Appunti di
Fioravante PATRONE

http://www.diptem.unige.it/patrone/default.htm

Decisori (razionali) interagenti

versione del 12 maggio 2007

Indice

Fioravante PATRONE
Dipartimento di Ingegneria della Produzione, Termoenergetica e Modelli Matematici
P.le Kennedy - Pad D
16129 Genova - ITALY
patrone@diptem.unige.it

http://www.diptem.unige.it/patrone homepage
http://tdg.dima.unige.it web teaching
http://www.citg.unige.it/citg.htm web server "CITG"
http://www.scallywag.it web page del gruppo Scaλλywag

http://www.diptem.unige.it/patrone/DRI.htm Decisori (razionali) interagenti

1 Dimostrazione dell'equivalenza di due condizioni per ESS

Queste note sono richiamate a pag. 138 del testo "Decisori (razionali) interagenti" e riguardano l'idea che è alla base della giustificazione delle ESS. Vediamo i dettagli formali. Come prima cosa, abbiamo un gioco *simmetrico*: (X, T, f, g), con $X = Y$ e con $f(x, y) = g(y, x)$ (cosa soddisfatta nel gioco "falco/colomba": $X = Y = \{F, C\}$ e la condizione di "simmetria" per i payoff è soddisfatta, come si può agevolmente verificare[1]). Una strategia x^* è un ESS se è una strategia che soddisfa una opportuna condizione di "stabilità", che vorrebbe tradurre l'idea che è una strategia resistente rispetto ad altre strategie che "compaiano" fra quelle usate dagli individui nella popolazione, giocate da una piccola frazione di individui nella popolazione (strategie "mutanti"). L'idea è che il payoff atteso da un "mutante", presente nella popolazione in una frazione pari a ε, il quale giochi la strategia x sia espresso dalla formula:

$$(1 - \varepsilon) f(x, x^*) + \varepsilon f(x, x)$$

e la ragione è ovvia, visto che si assume che il mutante incontri un altro mutante con probabilità ε e invece un non mutante con probabilità $1 - \varepsilon$. Invece, un non mutante ha un payoff pari a:

$$(1 - \varepsilon) f(x^*, x^*) + \varepsilon f(x^*, x)$$

Quindi, (x^*, x^*) viene detto ESS se, per ogni $x \in X$ diverso da x^*, esiste $\bar{\varepsilon} > 0$ tale che la relazione seguente valga per ogni $\varepsilon > 0$ tale che $\varepsilon < \bar{\varepsilon}$:

$$(1 - \varepsilon) f(x^*, x^*) + \varepsilon f(x^*, x) > (1 - \varepsilon) f(x, x^*) + \varepsilon f(x, x) \tag{1}$$

La condizione di essere un ESS può essere riscritta in modo più facilmente leggibile. Infatti, si prova che $x^* \in X$ è un ESS se e solo se, per ogni $x \in X, x \neq x^*$, si ha:

$$f(x^*, x^*) \geq f(x, x^*) \tag{2}$$

oppure

$$f(x, x^*) = f(x^*, x^*), \text{ nel qual caso deve essere } f(x^*, x) > f(x, x) \tag{3}$$

[1] Si noti che molti dei giochi più noti sono giochi simmetrici: lo è il dilemma del prigioniero, la batteglia dei sessi, il "pari o dispari", il gioco di puro coordinamento e quello che abbiamo usato per introdurre gli equilibri correlati...

Dimostrazione

• Se x^* soddisfa la condizione (1), basta passare al limite per $\varepsilon \to 0$ per ottenere la (2). Infatti (1) ci dice che (qui x e x^* sono fissati e la disuguaglianza vale per ogni ε *sufficientemente piccolo*[2]):

$$(1-\varepsilon)f(x^*,x^*) + \varepsilon f(x^*,x) - (1-\varepsilon)f(x,x^*) + \varepsilon f(x,x) > 0$$

e, passando al limite[3]:

$$\lim_{\varepsilon \to 0} [(1-\varepsilon)f(x^*,x^*) + \varepsilon f(x^*,x) - (1-\varepsilon)f(x,x^*) + \varepsilon f(x,x)] \geq 0$$

ovvero (ricordo che x e x^* sono fissati):

$$f(x^*,x^*) - f(x,x^*) \geq 0$$

e quindi:

$$f(x^*,x^*) \geq f(x,x^*)$$

La condizione (3) è del tutto ovvia (basta dividere per ε ciò che resta di entrambi i membri dopo avere eliminato i due addendi uguali).

• Per il viceversa, supponiamo di avere (2). Se $f(x^*,x^*) > f(x,x^*)$, ancora il teorema di permanenza del segno ci dice che vale la (1).

A dire il vero, non c'è neanche bisogno di utilizzare il teorema di permanenza del segno, visto che abbiamo una funzione che è *lineare* nella variabile ε e quindi, volendo, potremmo anche fornire delle stime esplicite. Vediamolo, per curiosità. Dati x e x^*, definiamo, per comodità, $\delta = f(x^*,x^*) - f(x,x^*)$. E' sufficiente scegliere ε (possiamo anche dire esplicitamente quanto deve valere ε, volendo) in modo che siano soddisfatte le seguenti relazioni: $|\varepsilon f(x^*,x^*)| < \delta/4$, $|\varepsilon f(x^*,x)| < \delta/4$, $|\varepsilon f(x,x^*)| < \delta/4$ e $|\varepsilon f(x,x)| < \delta/4$.

Allora:

$$(1-\varepsilon)f(x^*,x^*) + \varepsilon f(x^*,x) - (1-\varepsilon)f(x,x^*) - \varepsilon f(x,x) =$$

$$= f(x^*,x^*) - f(x,x^*) - \varepsilon f(x^*,x^*) + \varepsilon f(x^*,x) + \varepsilon f(x,x^*) - \varepsilon f(x,x) >$$

$$> f(x^*,x^*) - f(x,x^*) - 4 \cdot \frac{\delta}{4} = f(x^*,x^*) - f(x,x^*) - \delta = 0$$

---

[2]Per la precisione, si è supposto che esista un $\bar{\varepsilon} > 0$ per cui la (1) valga per ogni ε t.c. $0 < \varepsilon < \bar{\varepsilon}$.

[3]Segue dal teorema di permanenza del segno il quale dice che le disuguaglianze, al limite, si conservano, pur se indebolite. Infatti, possiamo solo garantire che il limite sia maggiore o uguale a 0.

Se invece $f(x^*, x^*) = f(x, x^*)$, come sopra la validità di (1) è ovvia (basta moltiplicare per ε che è un numero positivo i due membri della disuguaglianza che abbiamo in (3) e poi aggiungere i due addendi uguali). Q.E.D.