1 Best reply dynamics

Consideriamo un gioco in forma strategica (X, Y, f, g).

Una dinamica "naturale", di tipo *discreto* è la dinamica che prevede che ogni giocatore reagisca alla strategia giocata dall'altro al tempo precedente scegliendo la sua miglior risposta a quella strategia (sperando che questa miglior risposta esista; inoltre, non è detto che la miglior risposta sia unica, per cui dovremmo comunque dire "una miglior risposta").

Se assumiamo che per ogni (x, y) sia univocamente determinata la miglior risposta di I ad y e di II ad x, la dinamica che si ottiene è la seguente:

$$(x_{n+1} = R_I(y_n), y_{n+1} = R_{II}(x_n))$$

Quindi, se immaginiamo che si parta da un punto iniziale $(\hat{x}, \hat{y})$, si ottiene la "traiettoria" (x_n, y_n) descritta ricorsivamente (ovvero, per induzione matematica) nel modo seguente:

$$\begin{cases} (x_1, y_1) = (\hat{x}, \hat{y}) & \text{BASE} \\ (x_{n+1}, y_{n+1}) = (R_I(y_n), R_{II}(x_n)) \quad \forall n \in \mathbb{N} & \text{PASSO} \end{cases}$$

Questa dinamica, ad esempio, converge all'unico equilibrio di Nash nel modello di duopolio di Cournot, nella sua versione più semplice, ovvero quella in cui gli spazi di strategie sono $X = Y = [0, +\infty[$ ed i payoff sono i seguenti:

$$f(q_1, q_2) = q_1 \cdot p(q_1 + q_2) - c \cdot q_1$$

$$g(q_1, q_2) = q_2 \cdot p(q_1 + q_2) - c \cdot q_2$$

e dove la funzione $q \mapsto p(q)$ è così definita:

$$p(q) = \begin{cases} a - q & \text{se} \quad 0 \le q \le a \\ 0 & \text{se} \quad q \ge a \end{cases}$$

Essendo a e c costanti positive assegnate, con $c < a$[1].

Questa dinamica ha delle varianti, tipo quella in cui le "reazioni" dei due giocatori avvengono a tempi alternati (I reagisce alla strategia usata da II, poi al "tempo" successivo il giocatore II reagisce alla strategia usata da I, etc.). Anche questa variante converge all'equilibrio di Nash nel duopolio.

Si noti come questa dinamica corrisponda ad un grado di intelligenza piuttosto basso dei giocatori, nel senso che loro reagiscono solo a ciò che è

[1] Sennò non c'è gusto...

avvenuto all'ultimo istante precedente e non fanno alcuna analisi della futura dinamica (sono "miopi").

Concludo con una osservazione matematica che non mi sento di evitare: dovrò assumere che su X ed Y vi sia una struttura topologica se voglio che il problema della convergenza (o meno) della "best reply dynamics" abbia un senso. Sarà a questo scopo sufficiente che sia X che Y siano spazi metrici (ad esempio, che siano sottoinsiemi di uno spazio euclideo di dimensione finita).

2 Tra la "best reply dynamics" ed il "fictitious play"

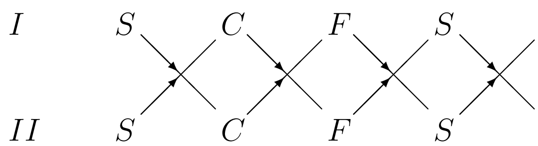
Prima di passare all'altro classico esempio di dinamica, può essere istruttivo (a mio parere) osservare cosa avviene se applichiamo la "best reply dynamics" a giochi finiti (intendo dire: rimanendo con le sole strategie pure, senza passare alle miste).

Cominciamo con la "morra cinese".

$I \backslash II$	S	C	F
S	$0, 0$	$-1, 1$	$1, -1$
C	$1, -1$	$0, 0$	$-1, 1$
F	$-1, 1$	$1, -1$	$0, 0$

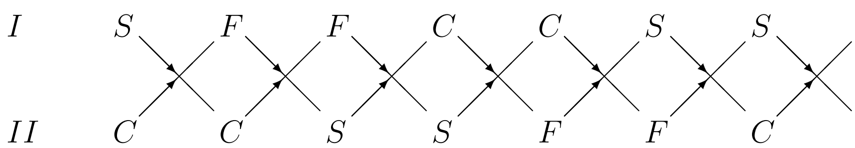
Possiamo immaginare di partire da una coppia di strategie (ad esempio, da (S, S)) e vedere come si comporta la dinamica (a tempi discreti) della "best reply dynamics".

Rappresento l'evoluzione del sistema con il seguente diagramma, che dovrebbe essere chiaro:



come si vede, si ha un ciclo di ordine 3. Cioè, dopo tre iterazioni si ritorna alla stessa coppia di strategie.

Partendo da un'altra coppia, ad esempio (S, C), si ha:



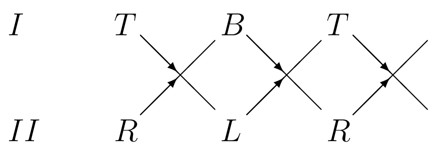
anche questa volta si ottiene un ciclo, di ordine 6.

E’ abbastanza facile verificare come questi siano gli unici due casi che si possono creare, ovvero un ciclo di ordine 3 o di ordine 6. In ogni caso, la dinamica continua a “saltellare” senza convergere da nessuna parte. Possiamo però notare come la *frequenza* con la quale le strategie S, C, F sono giocate da entrambi i giocatori tende verso 1/3 (ciò è immediata conseguenza del fatto che, sia nel ciclo di lunghezza 3 che in quello di lunghezza 6, questa è la frequenza con la quale compaiono le tre strategie).

Visto che la strategia mista che assegna 1/3 a ciascuna delle strategie pure è l’unico equilibrio del gioco, potremmo sperare di avere a disposizione una “ricetta” generale per trovare gli equilibri in strategie miste (tra l’altro, una cosa del tutto analoga avviene anche per il“pari o dispari”). Vedremo che così non è, ma prima guardiamo cosa capita in un gioco interessante (non a somma zero) come la battaglia dei sessi. L’interesse di questo gioco, rispetto a quello che stiamo facendo, è che ha equilibri in strategie pure e quindi potrebbe presentare un comportamento diverso rispetto alla “morra cinese” (essendo un gioco senza equilibri in strategie pure, era un po’ difficile sperare[2] che la dinamica potesse convergere verso un equilibrio!).

$I \backslash II$	L	R
T	$2,1$	$0,0$
B	$0,0$	$1,2$

Se partiamo da un equilibrio, ad esempio (T, L), è ovvio che la dinamica da lì non si muove. Se però partiamo da un altro punto, ad esempio da (T, R), vediamo come anche qui si abbia un ciclo:



il ciclo è di lunghezza 2. Il che è una “cattiva notizia”, perché ci porta (volendo seguire la congettura formulata sulla base della “morra cinese” e

[2]Nota “matematica”. A dire il vero, varrebbe la pena notare come su un insieme finito ogni metrica è equivalente alla metrica discreta e quindi una successione è convergente (rispetto a qualsiasi metrica) se e solo se è definitivamente costante. Pertanto, o la successione da un certo punto in poi è un equilibrio di Nash oppure non converge. Ciò non toglie che ci possano essere comportamenti significativamente diversi. Ad esempio, nel dilemma del prigioniero ogni dinamica di “best reply”, qualunque sia il punto di partenza, converge all’equilibrio (nel senso che dopo un passo al massimo i termini della successione coincidono tutti con l’equilibrio).

del "pari o dispari") verso una coppia di strategie miste che consistono nel giocare T, B con pari probabilità ed idem per L, R. Ma questa coppia di strategie miste non è affatto un equilibrio per la "battaglia dei sessi".

Come mai il metodo fallisce? Semplice. Non tiene conto dei valori dei payoff. Ciò che determina la dinamica è solo il fatto che T è l'unica miglior risposta per I ad L, e per questo serve solo che $f(T, L) > f(B, L)$. Non importa il fatto che $f(T, L)$ valga 2 oppure 1. Tanto è vero, che il gioco di puro coordinamento:

$I \backslash II$	L	R
T	$1, 1$	$0, 0$
B	$0, 0$	$1, 1$

ha lo stesso identico pattern di comportamento della "battaglia dei sessi", per quanto riguarda la dinamica della "best reply".

Quindi, sembra che occorra qualche idea migliore, se si vuole ottenere un metodo convergente. Prima di passare ad esaminare il "fictitious play" (che, se vogliamo, può essere visto come un modo per cercare di tenere conto della precedente osservazione), vediamo cosa succede per un altro gioco a somma zero, in modo da toglierci tutte le illusioni.

Consideriamo il gioco:

$I \backslash II$	L	R
T	$-2, 2$	$3, -3$
B	$3, -3$	$-4, 4$

che è anche "giocabile" in rete:
`http://www.dima.unige.it/~patrone/s0v3.htm`
e che ha un equilibrio in strategie miste il quale prevede che I giochi T con probabilità $7/12$.

Per quanto abbiamo detto dopo la "battaglia dei sessi", è chiaro che la dinamica della "best reply" è identica a quella del "pari o dispari" e quindi, se la congettura fosse corretta, ci aspetteremmo una strategia mista in cui la probabilità assegnata a T da I sia $1/2$. Quindi non troviamo la strategia di equilibrio.

E allora, dobbiamo fare un passo in avanti. Ed è proprio quanto ci apprestiamo a fare nel paragrafo seguente.

3 Fictitious play

Una variante interessante consiste nell'utilizzare sempre l'idea di "best reply", ma anziché reagire solo a quello che è avvenuto nello stadio precedente si tiene conto di tutta la "storia passata".

Un modo per fare questo consiste nel considerare la "best reply" ad una media delle azioni scelte precedentemente dall'altro giocatore. E, di queste, la più semplice è la media aritmetica.

Quindi si considera la seguente dinamica[3] (partendo da una coppia iniziale di strategie $(\hat{x}, \hat{y})$ scelta "arbitrariamente"):

$$(x^{k+1} = R_I(\frac{y^1 + \ldots + y^k}{k}), y^{k+1} = R_{II}(\frac{x^1 + \ldots + x^k}{k}))$$

Oltre alle considerazioni già fatte per la "best reply dynamics", relative alla necessità che la miglior risposta esista ed alla convenienza che essa sia unica, in questo caso si aggiunge una ulteriore considerazione. Occorre che abbia senso poter considerare $x^1 + \ldots + x^k$ e poter dividere per k (naturalmente la stessa cosa serve per le "y"). Le operazioni indicate (somma e divisione per k) potrebbero non avere alcun senso in X. Un modo per garantire che si possano fare è assumere che X ed Y siano sottoinsiemi convessi di uno spazio vettoriale. Ciò è vero, ad esempio, nel caso del duopolio ($X = Y = [0, \infty[$, e $[0, \infty[$ è un intervallo e pertanto è un sottoinsieme convesso di $\mathbb{R}$) e nel caso dell'estensione mista di un gioco finito ($\Delta(X)$ e $\Delta(Y)$ sono insiemi convessi).

Conviene esaminare un poco più in dettaglio il caso della estensione mista di un gioco finito. Supponiamo che sia $Y = \{y_1, \ldots, y_n\}$. Allora, $\Delta(Y)$ può essere identificato con il simplesso di $\mathbb{R}^n$ e le particolari strategie "miste" che rappresentano le n strategie pure saranno gli n versori degli assi (ovvero, saranno gli n elementi della base canonica di $\mathbb{R}^n$): $(1, 0, \ldots, 0), (0, 1, \ldots, 0), \ldots, (0, 0, \ldots, 1)$. Allora, l'elemento $(\frac{y^1 + \ldots + y^k}{k})$ non è altro che un elemento di $\mathbb{R}^n$ il quale può essere interpretato nel modo seguente: assegnamo alla strategia y_i la probabilità che venga giocata pari al numero di volte in cui y_i è stata usata come strategia da II nei turni precedenti, fratto il numero totale dei turni precedenti (k).

Nel caso specifico dell'estensione mista di giochi finiti questa dinamica è stato introdotta da Brown (G. W. Brown (1951): *Iterative Solution of Games by Fictitious Play*, in: "Activity Analysis of Production and Allocation" (T. C. Koopmans, ed.), Wiley, New York, 374-376) e Julia Bowman Robinson (può valere la pena vedere uno sketch della sua biografia in:
`http://www.agnesscott.edu/lriddle/women/robinson.htm`)
ha provato che questo metodo converge ad un equilibrio nel caso a somma zero (Robinson, J. (1951): *An iterative method of solving a game*, Annals of Mathematics, **54**, 296-301).

[3] L'indice in alto (lo "esponente") identifica l'iterazione. Viene messo in alto per evitare confusione tra la strategia giocata al passo k, che è x^k per I ed y^k per II, e il generico elemento dello spazio X o Y, che useremo nel seguito ed indicheremo con $x_1, \ldots, x_m$ o $y_1, \ldots, y_n$

E' opportuno precisare che, quando parliamo di convergenza del metodo, non intendiamo che si abbia la convergenza della successione (x^k, y^k). Questa successione normalmente "saltellerà" qui e là (come si può verificare agevolmente applicando, ad esempio, il metodo al "pari o dispari"). Ciò che possiamo sperare che converga è la successione $((\frac{x^1+\ldots+x^k}{k}), (\frac{y^1+\ldots+y^k}{k}))$.

Il metodo, tuttavia, nonostante questa precisazione non è detto che fornisca in generale (cioè, al di fuori del caso dei giochi a somma zero) una successione convergente ad un equilibrio di Nash. Può in effetti avvenire che la successione creata da questo processo iterativo (o sistema dinamico discreto che dir si voglia), non converga, come nell'esempio di Shapley[4]:

$I\backslash II$	L	C	R
T	$0,0$	$0,1$	$1,0$
M	$1,0$	$0,0$	$0,1$
B	$0,1$	$1,0$	$0,0$

La successione (x^k, y^k), come ci si può in generale attendere, "oscilla" fra i sei punti: (M, R), (T, R), (T, C), (B, C), (B, L) ed (M, L). Fin qui, nulla di male. Il problema è che la successione (x^k, y^k) staziona per tempi via via più lunghi in questi sei punti, in modo che la frequenza relativa (cioè, la successione $((\frac{x^1+\ldots+x^k}{k}), (\frac{y^1+\ldots+y^k}{k})))$ non converga affatto.

Un esempio semplice ed istruttivo di applicazione del metodo è fornito dalla sua utilizzazione per il "pari o dispari". Il lettore diligente potrebbe provare a ricostruirne "con carta e penna" alcuni passi o, meglio ancora, scrivere un programmino che in qualche linguaggio implementi il metodo. Non è difficile rendersi conto del fatto che, qualunque sia il punto di partenza, si ha la convergenza di $((\frac{x^1+\ldots+x^k}{k}), (\frac{y^1+\ldots+y^k}{k}))$ all'unico punto di equilibrio (cosa garantita dal teorema di Robinson). Ovvero, la frequenza con cui vengono giocate le due strategie a disposizione tende ad approssimare il valore $1/2$ per entrambi i giocatori.

Si noti però che questo metodo di gioco o di "apprendimento" prefigura un basso grado di intelligenza. Infatti, le scelte fatte da un giocatore che adotti

[4]A dire il vero, Shapley prova che il metodo non converge per una classe di giochi che include ad esempio il gioco seguente:

$I\backslash II$	L	C	R
T	$0,0$	$1,2$	$2,1$
M	$2,1$	$0,0$	$1,2$
B	$1,2$	$2,1$	$0,0$

L'esempio citato nel testo è analizzato da Shapley che, oltre ad asserire la non convergenza del metodo anche in questo caso, ne studia il "ciclo limite".

il "fictitious play" sono prevedibili: un avversario abbastanza furbo sarà in grado rapidamente di capire il metodo impiegato dall'altro e lo potrà quindi sfruttare a suo vantaggio (niente di meglio che giocare a "pari o dispari" con uno del quale siete in grado di prevedere la mossa!).

4 Dinamica del replicatore

Questo paragrafo è scopiazzato da Binmore "Fun and Games".

Si comincia con una versione discreta (cioè a tempo discreto). La dinamica avviene in un "giorno" di durata τ (diciamo che un anno è lungo 1 e quindi τ sarà espresso come frazione di un anno; se si trattasse di anni e giorni veri, sarebbe $\tau = 1/365$).

C'e' una popolazione composta di una frazione p di H e una frazione $1-p$ di D.

I payoff, esprimenti la "fitness" sono:

$I \backslash II$	D	H
D	$U+1, U+1$	$U, U+2$
H	$U+2, U$	$U-1, U-1$

I payoffs[5] indicano il numero di "offsprings" attesi "in un anno" da una madre che sia senza nido (U), che abbia un nido tutto per sé ($U+2$), che divida il nido con un'altra ($U+1$), che abbia combattuto aspramente per avere un nido ($U=1$). Assumiamo che la riproduzione sia *asessuata* e che *da una madre D nasca una figlia D* (idem per H).

Allora, per ottenere gli "offsprings" in un "giorno", le quantità indicate nella matrice vanno moltiplicate per τ.

Assumiamo che la popolazione totale sia mantenuta stabile al livello di N individui (dopo la nascita, competono per il cibo e prima di un nuovo ciclo riproduttivo solo N sopravvivono: evidentemente, assumiamo una situazione in cui le risorse disponibili sono date e fissate).

Vediamo come si evolve una popolazione in cui al tempo t siano presenti una frazione $p = p(t)$ di individui del tipo H e (ovviamente) $1-p$ individui di tipo D.

[5] Di fatto, come si vede nel fare i calcoli, la quantità U sarà irrilevante. Cioè, otterremmo lo stesso risultato se avessimo la matrice del gioco sotto indicato, che descrive (ironia della sorte) il gioco "chicken":

$I \backslash II$	D	H
D	$1, 1$	$0, 2$
H	$2, 0$	$-1, -1$

Il numero di figli attesi di tipo D, cioè quelli nati da una madre D, sono, per ciascuna di queste madri:

$$\tau f_D(p) = \tau \cdot [pU + (1-p)(U+1)] = \tau[U + (1-p)]$$

Poiché ci sono $N(1-p)$ madri di tipo D in giro, e queste madri sono vive anche dopo la riproduzione (le morti avvengono solo per effetto della competizione per il cibo), abbiamo che "il giorno dopo", prima di iniziare a competere per il cibo, vi saranno

$$N(1-p)(1 + \tau f_D(p))$$

individui di tipo D. Analogamente, avremo

$$Np(1 + \tau f_H(p))$$

individui di tipo H. Possiamo notare che si ha:

$$\tau f_H(p) = \tau \cdot [p(U-1) + (1-p)(U+2)] = \tau[U + 2(1-p) - p]$$

Alla "sera" seguente, dopo la competizione per il cibo e prima della riproduzione, la popolazione sarà stata ricondotta al livello di numerosità N. Si noti che, dal punto di vista della competizione per il cibo, non viene fatta alcuna distinzione fra individui D o H, nel senso che la probabilità di non riuscire a procurarsi il cibo è identica per i due tipi (ipotesi non da poco...). Allora, se al tempo t avevamo la frazione $p(t)$ di individui di tipo H, al tempo $t+\tau$ la loro frazione nella popolazione sarà:

$$p(t+\tau) = \frac{Np(t)(1 + \tau f_H(p(t)))}{N \cdot ((1-p(t))(1 + \tau f_D(p(t))) + p(t)(1 + \tau f_H(p(t))))} =$$

$$= \frac{Np(t)(1 + \tau f_H(p(t)))}{N \cdot (1 + \tau(1-p(t))f_D(p(t))) + \tau p(t) f_H(p(t))} = \frac{p(t)(1 + \tau f_H(p(t)))}{1 + \tau \bar{f}(p(t))} =$$

$$= p(t)\frac{1 + \tau f_H(p(t))}{1 + \tau \bar{f}(p(t)}$$

Dove abbiamo posto

$$\bar{f}(p) = (1-p) f_D(p) + p f_H(p)$$

Facendo i calcoli:

$$\frac{p(t+\tau) - p(t)}{\tau} = p(t)\frac{f_H(p(t)) - \bar{f}(p(t))}{1 + \tau \bar{f}(p(t))}$$

E, per $\tau \to 0$:

$$p'(t) = p(t)(f_H(p(t)) - \bar{f}(p(t)))$$

Questa è l'equazione della dinamica del replicatore. Che si può generalizzare al caso di n tipi diversi di individui, ottenendo:

$$p_i'(t) = p_i(t)(f_i(p(t)) - \bar{f}(p(t)))$$

Dove

$$f_i(p) = \sum_j a_{ij} p_j;$$

$$\bar{f}(p)(= \sum_{i,j} p_i a_{ij} p_j) = \sum_i p_i f_i(p)$$

(la matrice a_{ij} è naturalmente la matrice della fitness).

Tornando al nostro caso, se sostituiamo ad f_H ed a $\bar{f}$ la loro espressione (abbiamo trovato che $f_H(p) = U + 2(1-p) - p$ e $f_D(p) = U + (1-p)$, da cui ci ricaviamo anche $\bar{f}(p)$), otteniamo l'equazione:

$$p' = p(1-p)(1-2p)$$

Possiamo studiare la stabilità (asintotica) di questa equazione. E vediamo che $p = 1/2$ è un equilibrio stabile.

Calcoli analoghi fatti usando una matrice dei payoff che, invece essere quella del gioco "chicken" (vedi nota), sia quella del dilemma del prigioniero, forniscono una dinamica che converge all'unico equilibrio di Nash.