

# Aprendizagem 3

## Dinâmica do Replicador e Ajuste com ruído persistente

*Sergei Izmalkov e Muhamet Yildiz*

### Roteiro

1. Definição
2. Comportamento assintótico da dinâmica do replicador
  - a) RD x Racionabilidade
  - b) RD x ESS
  - c) RD x Equilíbrio Perfeito
3. Generalização de RD
4. Fundamentos do Aprendizado de RD
  - a) Aprendizado social
  - b) Resposta a estímulo
5. Modelos de ajuste com aleatoriedade persistente

## 1. Notação

- $G = (S, A)$  é um jogo simétrico, de dois jogadores em que
- $S$  é um espaço de estratégia;
- $A_{i,j} = u_1(s_i, s_j) = u_2(s_j, s_i)$ ;
- $x, y \in ?$  são estratégias mistas;  $u(x, y) = x^T A y$ ;
- $u(ax + (1 - a)y, z) = au(x, z) + (1 - a)u(y, z)$ .

## 2. ESS

Definição: Diz-se que uma estratégia (mista) é *evolucionária estável* se, e somente se, dados quaisquer  $y \neq x$ , e exista  $\epsilon_y > 0$ , tal que

$$u(x, (1 - \epsilon)y + \epsilon y) > u(y, (1 - \epsilon)x + \epsilon y),$$

para cada  $\epsilon$  em  $(0, \epsilon_y]$ .

Fato:  $x$  é evolucionária estável se, e somente se,  $\forall y \neq x$ ,

1.  $u(x, x) \geq u(y, x)$ , e

2.  $u(x, x) = u(y, x) \implies u(x, y) > u(y, y)$ .

### 3. Dinâmica do Replicador

- $p_i(t) = \# \text{pessoa que joga } s_i \text{ em } t$ .
- $p(t) = \text{população total em } t$ .
- $x_i(t) = \frac{p_i(t)}{p(t)}$ ;  $x(t) = (x_1(t), \dots, x_k(t))$ .
- $\dot{x}_i = [u(s_i, x) - u(x, x)] x_i = u(s_i - x, x) x_i$ .

### 4. RD no Jogo de Pedra-Papel-Tesoura

	R	S	P
R	1,1	2+a,0	0,2+a
S	0,2+a	1,1	2+a,0
P	2+a,0	0,2+a	1,1

- Único Equilíbrio de Nash  $(s^*, s^*)$ , onde  $x^* = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ .

- Defina  $h(x) = \log(x_1 x_2 x_3)$ .

- $\dot{h}(x) = \frac{a}{2} (3 \|x\|^2 - 1)$ .

- $\min_x \|x\| = 1$ ,  $\arg \min = (\frac{1}{3}, \frac{1}{3}, \frac{1}{3})$ .

## 4.1 Dinâmica

Três cenários:

1.  $a = 0$  — RSP original, todas as trajetórias são ciclos.

2.  $a < 0$  —  $x^*$  é instável.

3.  $a > 0$  —  $x^*$  é estável.

## 5. Racionalizabilidade

- $\xi(t, x_0)$  é a solução para a dinâmica do replicador iniciando em  $x_0$ .

Teorema: se uma estratégia pura  $i$  é estritamente dominada (por  $y$ ), então  $\lim_t \xi_i(t, x_0) = 0$  para qualquer  $x_0$  interior.

Prova: defina  $v_i(x) = \log(x_i) - \sum_j y_j \log(x_j)$ . Então,

$$\begin{aligned} \frac{dv_i(x(t))}{dt} &= \frac{\dot{x}_i}{x_i} - \sum_j y_j \frac{\dot{x}_j}{x_j} \\ &= u(s_i - x, x) - \sum_j y_j u(s_j - x, x) \\ &= u(s_i - y, x) \leq -\epsilon < 0. \end{aligned}$$

Por isso,  $v_i(\xi(t, x_0)) \rightarrow -\infty$ , so  $\xi_i(t, x_0) \rightarrow 0$ .

Teorema: Se  $i$  não é racionalizável, então  $\lim_t \xi_i(t, x_0) = 0$  para qualquer  $x_0$  interior.

## 6. Teoremas

Teorema: cada ESS  $x$  é um estado estável assintoticamente da dinâmica do replicador.

(Se os indivíduos puderem herdar as estratégias mistas, o inverso também será verdadeiro.)

Prova: defina  $C = \text{supp}(x)$ ,  $Q = \{y | C \subset \text{supp}(y)\}$ ,  
 $H(y) = \sum_{i \in C} x_i \log(y_i)$ .

1.  $x$  é uma máxima local de  $H$ , e
2.  $\exists$  uma vizinhança  $n(x)$  tal que  $H$  é crescente ao longo de qualquer trajetória em  $Q \cap n(x)$ .

$$\dot{H} = \sum_{i \in C} x_i \frac{\dot{y}_i}{y_i} = \sum_{i \in C} x_i u(s_i - y, y) = u(x - y, y) > 0.$$

NE  $\rightarrow$  estado estável em RD;

Estável SS em RD  $\rightarrow$  NE.

Teorema: se  $x$  é um estado estável assintoticamente da dinâmica do replicador, então  $(x, x)$  é um perfeito equilíbrio de Nash.

Prova:

1.  $(x, x)$  é um equilíbrio de Nash.

- a)  $x$  é estável  $\Rightarrow \dot{x}_i = u(s_i - x, x)x_i = 0$ .
- b) Suponha que  $(x, x) \notin NE$ .
- c)  $\exists i \notin \text{supp}(x) : u(s_i - x, x) > 0$ . [por 1 e 2]
- d)  $\exists \delta > 0, n(x) : u(s_i - y, y) > \delta \forall y \in n(x)$ .
- e)  $\xi_i(t, y^0) > y_i^0 e^{\delta t}$  if  $\xi_i(\cdot, y^0)$  restando em  $n(x)$ .

2.  $x$  não é fracamente dominado (desde ASS).

## 7. Estabilidade assintótica não-ESS

	L	M	R
L	0,0	1,-2	1,1
M	-2,1	0,0	4,1
R	1,1	1,4	0,0

- $NE = \left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$ ; mutante  $= \left(0, \frac{1}{2}, \frac{1}{2}\right)$ .
- RD é assintoticamente estável.
- Nota: se as estratégias podem ser herdadas,  $\left(\frac{1}{3}, \frac{1}{3}, \frac{1}{3}\right)$  se torna estável.

## 8. Dinâmica geral

Definição: Um processo é monótono de ganho se, e somente se, a cada interior  $x$ ,

$$u(s_i, x) > u(s_j, x) \Leftrightarrow \frac{\dot{x}_i}{x_i} > \frac{\dot{x}_j}{x_j}.$$

Teorema: sob qualquer dinâmica monótona de ganho “regular”, se a estratégia  $i$  for eliminada pelo processo de dominância estrita de estratégia pura iterada, então  $\lim_t x_i(t) = 0$ .

## 9. Aprendizado social

- Pergunte aos demais; se a outra pessoa fizer melhor, adote sua estratégia.

Dinâmica de emulação (“intermediário-melhorando”):

O jogador 2 é fictício,  $p(L) = \frac{1}{3}$ .

	L	R
U	9,0	0,0
D	2,0	2,0

- Pergunte aos demais; se o outro faz  $u'$ , e você  $u$ , então troque com probabilidade máxima  $\max\{0, b(u' - u)\}$ .
- Níveis de aspiração.

## 10. Resposta a estímulo

- $u(x, y) \in [0, 1]$
- $x_i^k(t+1) = (1 - \gamma u(s^k(t), \cdot))x_i^k(t) + F(s^k(t), i)\gamma u(s^k(t), \cdot)$ ,  
onde

$$F(s^k(t), i) = 1 \text{ se } s^k(t) = i,$$

$$F(s^k(t), i) = 0 \text{ por outro lado.}$$

- Resultado: como  $\gamma$  vai para 0, as trajetórias convergem para as trajetórias de RD.