

# Jogos Repetidos: 14.126

Sergei Izmailkov e Muhamet Yildiz  
MIT

## 1. Introdução

### 1.1 Idéias

- interação repetida, cooperação, normas sociais
- horizonte infinito, data terminal incerta
- efeitos estratégicos das próprias ações, ameaças
- punição, vingança
- detecção, monitoramento
- reputação, conluio
- \*aprendizagem

## 1.2 Princípio do Desvio de Fase Única

Teorema: jogo finito de múltiplas fases, ações observadas.

O perfil  $s$  é SPE  $\Leftrightarrow$  para cada jogador  $i$  não existem desvios lucrativos de fase única.

Prova: suponha que os desvios de fase única não sejam lucrativos. Suponha que existe um desvio lucrativo de fase única para algum jogador  $i$  em um subjogo, ou seja

$$\exists i, t, h^t, \hat{s}_i \quad (\hat{s}_i, s_{-i})|_{h^t} \succ_i (s_i, s_{-i})|_{h^t}.$$

Então  $t^* = \max t'$  de modo que  $\hat{s}_i(h^{t^*}) \neq s_i(h^{t^*})$ .

Obviamente  $t^* > t$ . O desvio de fase única envolve

$$(\hat{s}_i, s_{-i})|_{h^{t^*}} \approx_i (s_i, s_{-i})|_{h^{t^*}}.$$

Defina  $\tilde{s}_i(h^\tau) = \hat{s}_i(h^\tau)$  para todo

$\tau < t^*$ ,  $\tilde{s}_i(h^\tau) = s_i(h^\tau)$  para todo  $\tau \geq t^*$ .

Portanto,  $(\hat{s}_i, s_{-i})|_{h^\tau} \approx_i (\tilde{s}_i, s_{-i})|_{h^\tau}$  para todo  $h^\tau$ .

Repita para  $\tilde{s}_i(t^* \downarrow)$ .

Definição: o jogo é *contínuo no infinito* se para qualquer  $i$ ,

$$\sup_{h, \tilde{h}, \text{ s.t. } h^t = \tilde{h}^t} |u_i(h) - u_i(\tilde{h})| \rightarrow 0 \quad \text{como } t \rightarrow 0.$$

Exemplo: descontos + ganhos de fase limitada.

Teorema: O princípio do desvio de fase única para jogos com ganhos contínuos.

Prova: Seja  $\epsilon$  o tamanho da melhoria. Apare as extremidades que importam menos do que  $\epsilon/2$ . O restante é o jogo finito. Não há melhoria possível.

Observação: se os ganhos são definidos de maneira diferente, por exemplo, o tempo extra médio, o princípio não precisa se manter.

### 1.3 Exemplos

		C	D
Dilema do	C	1, 1	-1, 2
Prisioneiro	D	2, -1	0, 0

	L	M	R
U	0, 0	3, 4	6, 0
M	4, 3	0, 0	0, 0
D	0, 6	0, 0	5, 5

## 2. Jogos Repetidos com Ações Observáveis

### 2.1 O Modelo

Jogo de fase  $G$ :  $\mathcal{I}$ -jogadores,  $A_i$ -ações,  
 $g_i : A = \times_{i \in \mathcal{I}} A_i \rightarrow \mathbb{R}$ -ganhos,  $A_i$ -distribuições de  
 probabilidades sobre  $A_i$ .

Jogo repetido:  $\mathbf{a}^t \equiv (a_i^t)_{i \in \mathcal{I}}$ ,  $h^0$ -história nula,  $h^t =$   
 $(\mathbf{a}^0, \mathbf{a}^1, \dots, \mathbf{a}^{t-1})$ ,  $H^t = (A)^t$ .

Estratégia (pura)  $s_i \equiv (s_i^t)$ ,  $s_i^t : H^t \rightarrow A_i$ ; (mista)  
 $\sigma_i \equiv (\sigma_i^t)$ ,  $\sigma_i^t : H^t \rightarrow \mathcal{A}_i$ .

Ganhos:

- Descontando:  $G(\delta)$ ,

$$u_i = E_{\sigma}(1 - \delta) \sum_{t=0}^{\infty} \delta^t g_i(\sigma^t(h^t)) \rightarrow \max.$$

- Critério do tempo médio

$$\liminf_{T \rightarrow \infty} E \frac{1}{T} \sum_{t=0}^T g_i(\sigma^t(h^t)) \rightarrow \max.$$

- Critério alcançado

$$g = (g^0, g^1, \dots) \preceq_i h = (h^0, h^1, \dots) \left| \begin{array}{l} \text{if } \exists T', \forall T > T', \quad \sum_{t=0}^T g^t \leq \sum_{t=0}^T h^t. \end{array} \right.$$

Observação: se  $\alpha^*$  é um equilíbrio de Nash em  $G$ , então “jogada  $\alpha^*$  ( $\alpha_i^*$  para cada  $i$ ) pois todo  $t$ ” é SPE. Se existe um equilíbrio estático  $m$ , qualquer combinação deles é SPE.

## 2.2 Teoremas de Folk

*Ganhos racionais individualmente possíveis*

Utilidade de reserva (minimax):

$$\underline{v}_i = \min_{\alpha_{-i}} \left[ \max_{\alpha_i} g_i(\alpha_i, \alpha_{-i}) \right].$$

Defina  $m^i$  como perfil minimax.

Observação: em qualquer equilíbrio de Nash estático ou equilíbrio de Nash do jogo repetido, o ganho de  $i$  não é menor que  $\underline{v}_i$ .

Prova: jogue  $a_i(h^t)$  para maximizar  $E g_i(a_i, \sigma_{-i}(h^t))$ , na qual  $\sigma_{-i}$  são estratégias de Nash.

Ganhos possíveis (com aleatoriedade):

$$V = \text{c.h.} \{v = g(a), \text{ para } a \in A\}.$$

Teorema: (teorema de folk) Para qualquer  $v \in V$ , com  $v_i > \underline{v}_i$  para todo  $i$ , existe um  $\delta^* < 1$ , tal que para todo  $\delta \in (\delta^*, 1)$  existe um equilíbrio de Nash com ganhos  $v$ .

Prova: punir por minimax.

Teorema: (Friedman, ameaças de Nash)  $\alpha^*$  é um Nash estático com ganhos  $e$ . Então, para qualquer  $v \in V$ , com  $v_i > e_i$  para todo  $i$ , existe um SPE de  $G(\delta)$  com ganhos  $v$ .

Prova: punir por Nash. SPE vem das observação acima.

Teorema: (Aumann, Shapley) critério da média temporal, então qualquer  $v \in V$ , com  $v_i > \underline{v}_i$  para todo  $i$ , existe um SPE de  $G(\delta)$  com ganhos  $v$ .

Prova: punir por minimax por um tempo limitado. Efeitos de longo prazo são nulos.

Teorema: (Fudenberg, Maskin) Suponha  $\dim V = \infty$ . Então para qualquer  $v \in V$ , com  $v_i > \underline{v}_i$  para todo  $i$ , existe um  $\delta^* < 1$ , tal que para todo  $\delta \in (\delta^*, 1)$  existe um SPE de  $G(\delta)$  com ganhos  $v$ .

Prova: a idéia é recompensar os punidores. Suponha que para todo  $v$  considerado, existe  $a$ ,  $g(a) = v$ . Já que  $\dim V = \infty$ ,  $\exists v' \in V$ ,  $\underline{v}_i < v'_i < v_i$  para todo  $i$ , e  $v'(i) \in$  de modo que

$$v'(i) = (v'_1 + \varepsilon, \dots, v'_{i-1} + \varepsilon, v'_i, v'_{i+1} + \varepsilon, \dots, v'_I + \varepsilon).$$

Suponha que  $a'(i)$  existe de modo que  $g(a'(i)) = v'(i)$ .

Fase 1. Jogue  $a$  até que a ação seja  $a$  ou diferente de  $a$  por dois ou mais componentes. Se  $a'_j \neq a_j$ , passe para a fase  $2_j$ .

Fase  $2_j$ . Jogue  $m^j$  por  $N$  períodos de maneira que a ação realizada seja  $m^j$  ou seja diferente de  $m^j$  por 2 ou mais componentes. Passe para a fase  $3_j$ . Se algum  $k$  desviar-se, passe para a fase  $2_k$ .

Fase 3j. Jogue  $v^j(j)$  durante o tempo em que a ação realizada for  $a^j(j)$  ou diferente de  $a^j(j)$  por dois ou mais componentes. Se o licitante  $k$  desvia-se, passe para a fase 3k.

Use o princípio do desvio de fase única.

Problema: se  $a^j(i)$  é misto, o mesmo ganho de continuação deve ser garantido para todas as ações em suporte.

Teorema: (Abreu, Dutta, Smith) condição de NEU em vez de  $\dim V = \#I$ .

Definição: NEU (utilidades não-equivalentes) é satisfeito se para qualquer  $(i, j)$ ,  $\exists c, d \in \mathbb{R}_+$  de modo que  $g_i(a) = c + dg_j(a)$  para todo  $a \in A$ .

Prova: NEU  $\implies \exists [v^1, \dots, v^I]$ , tal que  $\forall i, j, v_i^i < v_i^j$ .

Aproximadamente: Substitua  $v^i$  por  $v^j(i)$ .

## 2.3 Jogos Finitos

Teorema: (Benoit, Krishna) critério da média temporal. Suponha que para  $\forall i$  existe um Nash estático  $\alpha^*(i)$  com  $g_i(\alpha^*(i)) > \underline{v}_i$ . Então, o conjunto dos ganhos em equilíbrio de Nash de  $G^T$  converge quando  $T \rightarrow \infty$  para o conjunto dos ganhos possíveis de  $G^\infty$ .

Prova: fase de recompensa terminal. Ciclo  $R \times I$ :

$([\alpha^*(1), \dots, \alpha^*(I)])^R$  — via de equilíbrio de Nash.

Fornece exatamente mais que  $\underline{v}_i$  para cada  $i$ . Para um  $R$  grande a ameaça do minimax durante os períodos  $RI$  evita todos os desvios.

Fixe  $\varepsilon > 0$ . Existe um  $T$ , tal que o ganho durante os períodos  $T - RI$  se aproxima de  $v_i$  para todo  $i$  dentro de  $\varepsilon$ .

...

## 2.4 Variando os oponentes

### 2.4.1. Jogadores de curto prazo vs longo prazo

Se os jogadores de curto prazo movimentam primeiro, a “cooperação” é alcançável.

Princípio: jogador(es) S-R joga(m)  $C$ , jogador(es) L-R responde(m) com  $C$  conquanto que  $(C, C)$  tenha sido jogado no passado. Caso contrário,  $D$ .

Movimentos simultâneos: jogadores S-R sempre jogam BR.

$1, \dots, l$  — jogadores L-R

$l + 1, \dots, I$  — jogadores S-R

$B : \times_{i=1}^l \mathcal{A}_i \rightarrow \times_{j=l+1}^I \mathcal{A}_j$  — correspondência com BR

$$\underline{v}_i = \min_{\alpha \in \text{graph}(B)} \left[ \max_{a_i} g_i(a_i, \alpha_{-i}) \right],$$

$$V = \text{c.h.} \left\{ v = (g_i(a))_{i=1}^l \in \mathbb{R}^l, \text{ para } \alpha \in \text{graph}(B) \right\}.$$

A observação das ações mistas é importante. Jogadores de longo prazo devem ser indiferentes em relação às ações puras às quais atribuem probabilidades positivas.

$$\bar{v}_i = \max_{\alpha \in \text{graph}(B)} \left[ \min_{a_i \in \text{supp}(\alpha_i)} g_i(a_i, \alpha_{-i}) \right].$$

Teorema: (Fudenberg, Kreps, Maskin).

Suponha que  $\dim V = l$ .

Para qualquer  $v \in V$ , com  $\bar{v}_i > v_i > \underline{v}_i$  para todo  $i$ , existe um  $\delta^* < 1$ , tal que para todo  $\delta \in (\delta^*, 1)$  existe SPE de  $G(\delta)$  com ganhos  $v$ .

#### 2.4.2 Gerações sobrepostas

Os jogadores subsistem por  $T$  períodos. Cada geração tem a mesma massa.

As ações são observáveis: *trabalho* ou *ociosidade* (IR, NE estático). Todo *trabalho* é eficiente.

Os ganhos são médios durante o tempo de permanência do jogador.

Resultado (Crémer): existe equilíbrio de Nash onde todos, exceto o mais velho, *trabalham*.

Teoremas de Folk: Candori, Smith.

#### 2.4.3 Competições aleatórias

O que é observável? O que é lembrado? Informações Públicas vs Privadas.

Dilema do Prisioneiro: jogue  $C$ , conquanto que  $(C, C)$  tenha sido jogado. Caso contrário,  $D$ .

Suportável desde que  $\delta$  seja suficientemente alto e que se conheça alguma informação sobre o oponente.

Se apenas os resultados privados passados são observáveis, com  $N$  suficientemente grande, as estratégias de “contágio” podem não ser um equilíbrio.

Razão: responder  $C$  em  $D$  atrasa o contágio.



## 2.5 Perfeição de Pareto

$$Eff(C) = \{x \in C, \nexists y \in C, y \geq x, y \neq x\}.$$

Definição: (Bernheim, Peleg, Whinston) Considere que  $G^T, P^T$  é o conjunto de ganhos SPE de estratégia pura de  $G^T$ .  $Q^1 = P^1$ ,  $R^1 = Eff(P^1)$ .

Para  $T > 1$ ,  $Q^T \subseteq P^T$  — o conjunto de ganhos de equilíbrio perfeito de estratégia pura forçado por  $R^{T-1}$ . Seja  $R^T = Eff(Q^T)$ .

SPE  $\sigma$  é o *Perfeito de Pareto* se,  $\forall t$  e  $\forall h^t$ , os ganhos de continuação determinados por  $\sigma$  estiverem em  $R^{T-t}$ .

Exemplo: (Benoit, Krishna)  $\delta=1$

	$b_1$	$b_2$	$b_3$	$b_4$
$a_1$	0, 0	2, 4	0, 0	5.5, 0
$a_2$	4, 2	0, 0	0, 0	0, 0
$a_3$	0, 0	0, 0	3, 3	0, 0
$a_4$	0, 5.5	0, 0	0, 0	5, 5

## 3. Jogos Repetidos com Informações Públicas Imperfeitas

### 3.1 O Modelo

$a \in A \rightarrow \Delta(y)$ ,  $y \in Y$  — publicamente observável.

$$\pi_y(a) \in \Delta(y); \pi(a)$$

$r_i(a_i, y)$  — ganho para  $i$ , (!) independente de  $a_{-i}$ .

$$g_i(a) = \sum_y \pi_y(a) r_i(a_i, y).$$

$h^t = (y^0, y^1, \dots, y^{t-1})$  — história pública.

$z_i^t$  — história privada (ações passadas).

Estratégia (mista)  $\sigma_i \equiv (\sigma_i^t)$ ,  $\sigma_i^t : H^t \times Z_i^t \rightarrow \mathcal{A}_i$ .

Definição:  $\sigma_i$  é uma *estratégia pública* se

$$\sigma_i(h^t, z_i^t) = \sigma_i(h^t, \tilde{z}_i^t) \quad \forall t, h^t, z_i^t, \tilde{z}_i^t.$$

Observação: o ganho de equilíbrio da Estratégia Pura pode ser sustentado como um ganho de um equilíbrio em estratégias públicas.

Definição:  $\sigma$  é um equilíbrio público perfeito se para todo  $i$ ,  $\sigma_i$  é uma estratégia pública, e para  $\forall t, h^t$ , estratégias  $\sigma|_{h^t}$  formam o equilíbrio de Nash.