GAME THEORY

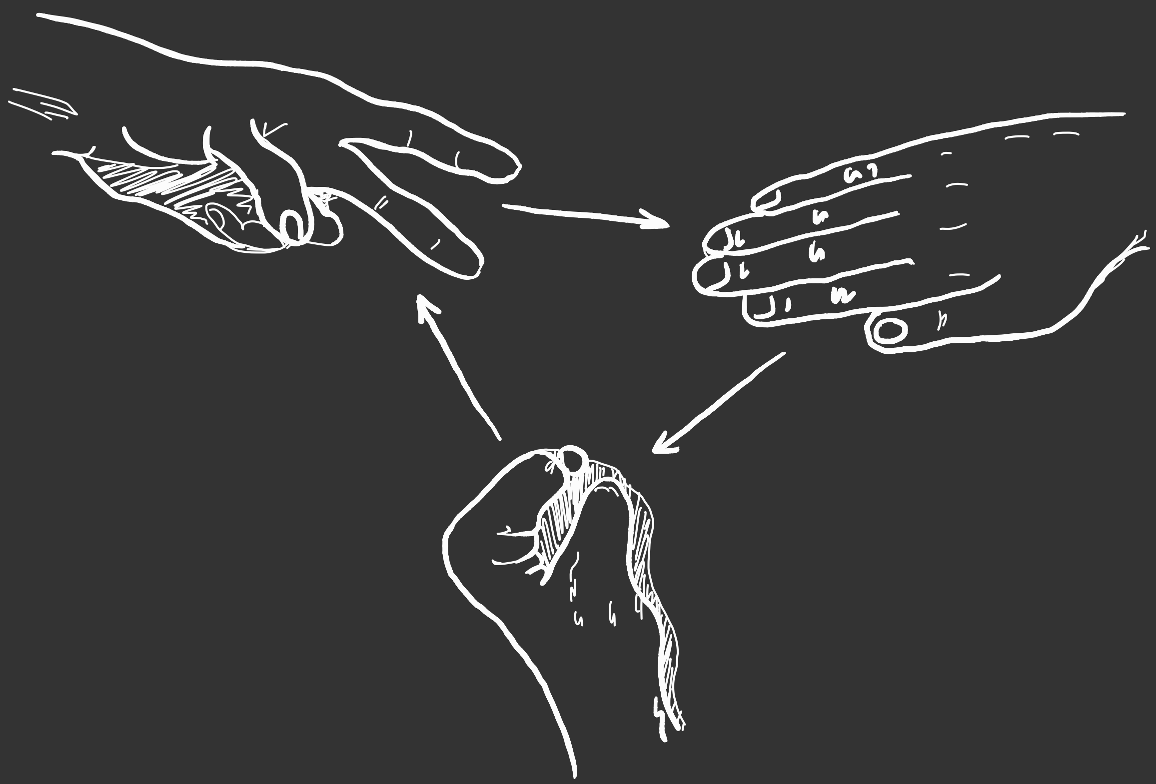
Prof. Hamed Kebriaei

Spring 1401

Contents

CHAPTER 1

~ Basic concepts
~ Simple form of games
~ Mixed strategy



Shiva Shakeri

بهینه‌سازی

سیستم تک‌عامله: $\max_{x\in X\subseteq \mathbb{R}} u(x) \to x^*$

سیستم چندعامله: $\max_{x_i\in X_i\subseteq\mathbb{R}} u_i(x_i) \to x_i^*$ ← لزوماً مسئله‌ی Game نداریم!

تاثیرپذیری از سایر بازیگران :

تعریف ۱. $u_i(x_1,x_2,\dots,x_n)$

$\to X=\begin{pmatrix}x_1\\ \vdots\\ x_n\end{pmatrix} \to u_i(X)$

تعریف 2. $u_i(x_i,\underbrace{x_{-i}}_{})$ $\to * \; x_{-i}$: همه‌ی عامل‌ها به‌غیر از عامل iام

$$\to \max_{x_i\in X_i\subseteq\mathbb{R}} \overset{\text{خروجی}}{u_i}(\underset{\text{ورودی}}{x_i},x_{-i}) = \boxed{x_i^*=f_i(x_{-i})}$$

Best Response Function

یک تابع از تصمیم سایر عوامل

- حل همکارانه :

$$\to \max_{X\in \mathbf{X}\subseteq\mathbb{R}^n} \boxed{\sum u_i(x_i,x_{-i})} \to X^* = \begin{pmatrix}x_1^*\\ \vdots\\ x_n^*\end{pmatrix}, \qquad \mathbf{X}=\prod_{i=1} X_i$$

* لزومی نداره هر عامل به بهینه‌ترین تصمیم برسه.

* متغیر تصمیم‌گیری: اسکالر ← برداری تبدیل شده

- حل غیرهمکارانه :

$\to \underbrace{x_i^*=f_i(x_{-i})}_{\text{n معادله، n مجهول}} \;;\; i=1,2,\dots,n$ تابعی از تصمیمات بقیه بازیگران

نتیجه $\to X^*=\begin{bmatrix}x_1^*\\ \vdots\\ x_n^*\end{bmatrix}$

بازی‌های استراتژیک

۱. بازی غیرهمکارانه

2. اطلاعات کامل از توابع هدف بازیگران

3. تصمیم‌گیری بازیگران بدون اطلاع از یکدیگر

4. بازی استاتیک

5. بازی تک مرحله‌ای

$$G(I,(S_i),(u_i))$$

- $I := \{1,2,\dots,n\} \to$ مجموعه‌ی بازیگران
- S_i: مجموعه‌ی استراتژی‌های بازیگر i
  - تابع توصیف نحوه‌ی عملکرد در بازی
- $u_i : (s_1,\dots,s_n) \to$ تابع سود

فرض Rationality برای agentها

→ fully rational

→ bounded rational

PLAYER 1

$S_2=\{U,M,D\}$ (PLAYER 2) — $S_1=\{L,M,R\}$ (PLAYER 1)

	L	M	R
U	4,3	5,1	6,2
M	2,1	8,4	3,6
D	3,0	9,6	2,8

Figure 1.1

PAGE 5
of the reference

استراتژی غالب

finite space

PRISONER'S DILEMMA

The police put each suspect in a different cell to prevent the two suspects from communicating with each other. The police tell each suspect that if he testifies against (doesn't cooperate with) the other, he will be released and will receive a reward for testifying, provided the other suspect does not testify against him. If neither suspect testifies, both will be released on account of insufficient evidence, and no rewards will be paid. If one testifies, the other will go to prison; if both testify, both will go to prison, but they will still collect rewards for testifying. In this game, both players simultaneously choose between two actions. If both players cooperate (C) (do not testify), they get-4each. If they both play noncooperatively (D, for defect), they obtain-2. If one cooperates and the other does not, the latter is rewarded (gets-1) and the former is punished (gets -5). Although cooperating would give each player a payoff of 1, self-interest leads to an inefficient outcome with payoffs 0.

↳ As an strategic game:

$$G(I,(S_i),(u_i))$$

- $I=\{1,2\}$ ⇝ 2 players
- $S_1=S_2=\{$ Confess , Don't Confess $\}$
- u_i :

Player 1 \ Player 2	C	D
C	-4 , -4	-1 , -5
D	-5 , -1	-2 , -2

player 1

player 2

← استراتژی غالب هر بازیگر ← C

تعریف. استراتژی غالب

استراتژی $s_i \in S_i$ یک استراتژی غالب است؛ اگر:

$$\forall s_i' \in S_i ,\ \forall s_{-i} \in S_{-i} :$$

$$u_i(s_i, s_{-i}) \geq u_i(s_i', s_{-i})$$

تعریف. تعادل استراتژی غالب

استراتژی $s^*=(s_1^*, \dots, s_n^*)$ یک تعادل استراتژی غالب است، اگر s_i^* برای هر بازیگر $i \in I$، یک استراتژی غالب باشد.

استراتژی مغلوب

استراتژی ای که فارغ از عملکرد سایر بازیگران، برای یک بازیگر سود کمتری دارد.

تعریف. استراتژی (کاملاً) مغلوب

استراتژی $s_i \in S_i$ یک استراتژی مغلوب برای بازیگر i است، اگر:

$$\exists s_i' \in S_i ,\ \forall s_{-i} \in S_{-i} :$$

$$u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$$

تعریف. استراتژی مغلوب ضعیف

استراتژی $s_i \in S_i$ یک استراتژی مغلوب برای بازیگر i است، اگر:

$$\exists s_i' \in S_i :$$

$$,\ \forall s_{-i} \in S_{-i}: \ u_i(s_i', s_{-i}) \geq u_i(s_i, s_{-i})$$

$$,\ \text{for some } s_{-i} \in S_{-i}: \ u_i(s_i', s_{-i}) > u_i(s_i, s_{-i})$$

DOMINANCE SOLVABLE

← در برخی از بازی ها، هیچ کدام از استراتژی ها غالب نیست!

پس می‌توان با حذف استراتژی های مغلوب، مسأله را حل کرد!

infinite space

COURNUT COMPETITION

The strategies are quantities. Firm 1 and firm 2 simultaneously choose their respective output levels, q_i, from feasible sets $Q_i = [0, \infty)$, say. They sell their output at the market-clearing price $p(q)$, where $q = q_1 + q_2$. Firm i's cost of production is $c_i(q_i)$, and firm i's total profit is then

$$u_i(q_1, q_2) = q_i p(q) - c_i(q_i).$$

↳ As an strategic game:

$$G(I,(S_i),(u_i))$$

- $I=\{1,2,\dots,n\}$ ⇝ n players : تولیدکننده
- $s_i \in [0,\infty) = S_i$ ⇝ میزان تولید
- $u_i(s_i, s_{-i}) = \overset{\text{مقدار تولید}}{s_i}\ \overbrace{P(s_i, s_{-i})}^{\text{قیمت فروش}} - \overbrace{C_i(s_i)}^{\text{هزینه تولید}}$

↳ $P(s_i, s_{-i}) = \alpha - \beta \sum_{i=1}^{n} s_i$

جنس تو بازار ↑ قیمت فروش ↓

NASH EQUILIBRUM POINT

$G(I,(S_i),(u_i))$

تعریف. تعادل نش

برای بازی G استراتژی $S^*=(s_1^*,s_2^*,\dots,s_n^*)\in S^*$ یک استراتژی تعادلی نش است:

$$\forall s_i\in S_i,\forall i:\quad u_i(s_i^*,s_{-i}^*)\geq u_i(s_i,s_{-i}^*)$$

آیا این لزوماً بهترین استراتژی برای هر بازیگر خواهد بود؟ خیر!

آیا اگر همه‌ی بازیگرها از استراتژی S^* به $\tilde{S}$ بروند، ممکن است سود همه‌ی بازیگرها افزایش پیدا کند؟ بله!

$\rightsquigarrow$ مثال: در معمای زندانی

نقطه‌ی تعادل نش

Player 1 \ Player 2	C	D
C	-4 , -4	-1 , -5
D	-5 , -1	-2 , -2

player 1

player 2

$S^*=(C,C)$; آیا برقرار است؟ $u_1(C,C)\overset{?}{\geq}u_1(s_1,C)$ ✓

$u_2(C,C)\overset{?}{\geq}u_2(C,s_2)$ ✓

$s_2,s_1\in\{C,D\}$

مثلاً برای بررسی نامساوی اول، استراتژی بازیگر 2 را ثابت C نگه می‌داریم و می‌پرسیم اگر استراتژی بازیگر 1 از C به D تغییر کند، سودش کم می‌شود؟ جواب بله است. پس نامساوی برقرار است.

$\rightsquigarrow$ مثال:

Player 1 \ Player 2	a'	b'	c'
a	4 , 4	3 , 3	5 , 3
b	3 , 1	5 , 1	2 , 4
c	3 , 2	4 , 3	4 , 3

$\rightsquigarrow S^*=(a,a')$

← اگر بازیگر 2 استراتژی a' داشته باشد، بازیگر 1 تمایلی به تغییر استراتژی اش از a به b یا c ندارد، چون بیش‌ترین سود را دارد.

اگر بازیگر 1 استراتژی a داشته باشد، بازیگر 2 تمایلی به تغییر استراتژی اش از a' به b' یا c' ندارد، چون بیش‌ترین سود را دارد.

ادامه ص بعد

COURNUT COMPETITION حل

فرض: تعداد بازیگر $= 2$ ← $i=1,2$

$C_i(s_i)=s_i$, $P(s_1,s_2)=\text{Max}(0,2-s_1-s_2)$

Best Strategy ← $B_i(s_{-i})\in \text{Argmax}\ u_i(s_i,s_{-i})$

بهترین پاسخ به استراتژی رقیب

$$\to B_1(s_2)\in\underset{s_1\geq0}{\text{Argmax}}\Big(\underbrace{s_1(2-s_1-s_2)-s_1}_{u_1(s_1,s_2)}\Big)$$

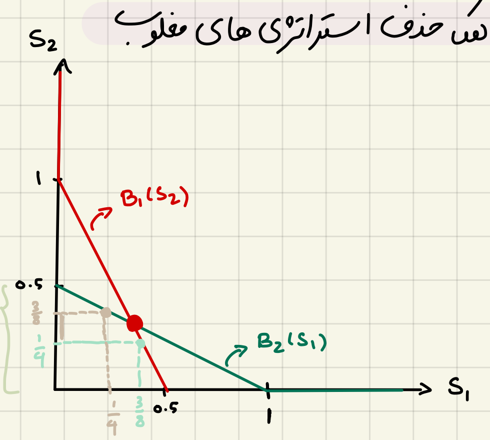
$$\frac{\partial u_1(s_1,s_2)}{\partial s_1}=2-2s_1-s_2-1=0$$

$$\to s_1^*=B_1(s_2)=\begin{cases}\frac{1-s_2}{2} & s_2\leq1\\ 0 & o.w.\end{cases}$$

به همین ترتیب داریم:

$$\to s_2^*=B_2(s_1)=\begin{cases}\frac{1-s_1}{2} & s_1\leq1\\ 0 & o.w.\end{cases}$$

حل به کمک حذف استراتژی های مغلوب



مرحله‌ی صفر: $S_1^0=[0,+\infty)$, $S_2^0=[0,+\infty)$

$S_1^1=[0,\frac{1}{2}]$, $S_2^1=[0,\frac{1}{2}]$

$S_1^2=[\frac{1}{4},\frac{1}{2}]$, $S_2^2=[\frac{1}{4},\frac{1}{2}]$

$S_1^3=[\frac{1}{4},\frac{3}{8}]$, $S_2^3=[\frac{1}{4},\frac{3}{8}]$

⋮

با ادامه‌ی این روند و حذف استراتژی‌های مغلوب، استراتژی‌ای که در نهایت باقی می‌ماند، همان نقطه‌ی برخورد توابع بهترین پاسخ بازیگران است

مثال :

		Player 2	
		a'	b'
Player 1	a	1 , 4	0 , 0
	b	0 , 0	4 , 1

در این مثال (a,a') و (b,b') نقاط تعادل نش بازی هستند.

← اینکه نقطه‌ی تعادل چند تاست، چه مفهومی دارد؟

این بازی یک نتیجه‌ی ثابت به ما نمی‌دهد. تعادل نش را طوری بدست می‌آوریم که انگار به بازی اعتماد کامل داریم. اما در عمل بازیگرهای ما bounded-rational هستند و اطلاعات ناقص دارند؛ یک تعبیر از تعادل نش نقاطی هستند که ما انتظار داریم بازیگران ما با اطلاعات ناقص در نهایت به آنها همگرا شوند. وقتی همگرا شدند، درنهایت می‌بینند که نمی‌توانند با تغییر استراتژی‌شان به سود بیش‌تری برسند ← تمایلی به تغییر ندارند. یعنی درواقع با فرآیند یادگیری learning به نقطه‌ی تعادل همگرا می‌شوند. حال اینکه این فرآیند یادگیری لزوماً به یک نقطه همگرا نمی‌شود.

مثال :

		Player 2	
		a'	b'
Player 1	a	2 , 2	-1 , 1
	b	1 , -1	0 , 0

در این مثال (a,a') و (b,b') نقاط تعادل نش بازی هستند. پس بازی ممکن است تعادل های مختلفی را به ما معرفی کند؛ به عنوان نقاطی که بازیگرها درنهایت رقابت در آن نقاط قرار بگیرند. و اینکه بین آن نقاط کدام یک valid هستند، به فرآیند learning و نوع تصمیم گیری‌شان وابسته خواهد بود همچنین به سودآوری آن نقطه نیز قطعاً وابسته است مثلاً در این مثال در حین learning بازیگرها می‌فهمند (a,a') سود بیش‌تری دارد و درعین حال مفهوم تعادلی را نیز دارد پس (a,a') نسبت به (b,b') ارجحیت دارد.

← مفهوم دیگر از تعادل نش :

بازیگر 1 عمل a را انجام می‌دهد، درصورتی که حدس بزند بازیگر 2، عمل a' را انجام می‌دهد.

بازیگر 2 عمل a' را انجام می‌دهد، درصورتی که حدس بزند بازیگر 1، عمل a را انجام می‌دهد.

پس درنهایت حدس دوطرف با عمل دوطرف منطبق شد ← تقاطع BR بازیگران

با همین تعریف (a,b') و (a',b) نیز می‌توانند تعادل نش باشند. پس با این تعبیر چرا (a',b) تعادل نش نیست؟ نمی‌تواند روی BR بازیگر 1 قرار بگیرد.

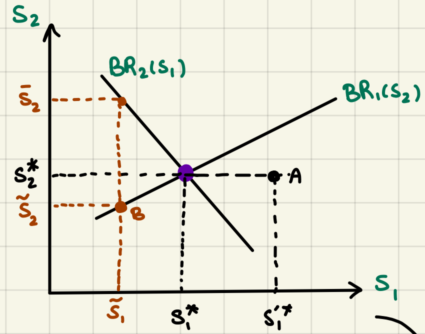
Best Response

$$s_1^* = BR_1(s_2) = \begin{cases} a & ; s_2 = a' \\ b & ; s_2 = b' \\ a & ; s_2 = c' \end{cases}$$

$$s_2^* = BR_2(s_1) = \begin{cases} a' & ; s_1 = a \\ c' & ; s_1 = b \\ b',c' & ; s_1 = c \end{cases}$$

← برای یافتن تعادل نش، نقطه‌ی برخورد Best Response

نقطه‌ی برخورد : $s^* = (a,a')$ ← نقطه‌ی تعادل نش



A

اگر فرض کنیم بازیگر 1، درصورتی‌که بازیگر 2، s_2^* را بازی کند، استراتژی $s_1'^*$ را بجای s_1^* بازی کند، از خط Best Response خودش خارج می‌شود. پس s_1^* بهترین استراتژی دراین شرایط است

B

← اگر فرض کنیم نقطه‌ی B نقطه‌ی تعادل باشد، این نقطه فقط روی BR بازیگر 1 قرار دارد. بنابراین اگر بازیگر 2، $\tilde{s}_2$ را بازی کند، بهترین کار برای بازیگر 1، استراتژی $\tilde{s}_1$ است. اما بازیگر 2، در صورتی که بازیگر 1، $\tilde{s}_1$ را بازی کند با بازی کردن $\bar{s}_2$ سود بیش‌تری دارد و ترجیح می‌دهد $\bar{s}_2$ را بازی کند. پس نامساوی تعادل نش برای بازیگر 2 صدق نمی‌کند.

* نامساوی های مربوط به تعادل نش باید برای همه بازیگرها برقرار باشد.

⇓

محل برخورد BR تمام بازیگرها = نقطه‌ی تعادل نش

↓

ممکن است در چند نقطه یا هیچ نقطه

مثال :

		Player 2	
		H	T
Player 1	H	$-1, 1$	$1, -1$
	T	$1, -1$	$-1, 1$

اگر تمام استراتژی ها را بررسی کنیم، خواهیم دید که این بازی، تعادل نش ندارد.

اگر بازیگر 2، حدس بزند که $\underbrace{\text{بازیگر 1، H را انجام دهد}}$، H را انجام می دهد.

بازیگر 1 در چه صورتی H را انجام می دهد؟

درصورتی‌که بازیگر 2، T را انجام دهد، بازیگر 1 با این فرض، H را انجام می دهد.

← پس در نهایت حدس ها و اعمال دو بازیگر بریکدیگر منطبق نمی‌شود!

با بررسی تمام حالت‌ها به همین نتیجه خواهیم رسید این بازی تعادل نش ندارد!

جمع‌بندی تعادل نش

← به ما ابزار قدرتمندتر می دهد تا نتایج بازی را تحلیل و پیش‌بینی کنیم.

← مفهوم آن در بازی‌های غیرهمکارانه معنا دارد.

← در این تعادل هیچ بازیگری به تمایلی تعاملی به تغییر استراتژی اش ندارد. چون با تغییر استراتژی اش نمی‌تواند به سود بیش‌تری برسد و چون بازی غیرهمکارانه است پس مفهوم تعادلی دارد.

← تعادل نش روی تقاطع BR همه‌ی بازیگرهاست.

← تعادل نش می‌تواند روی یک یا چند یا هیچ نقطه باشد.

← نقطه‌ی تعادل نش، نقطه‌ای ست که عکس العمل هر بازیگر با توجه به حدسی که از رقبایش دارد با حدس رقبا و عکس العمل آنها نسبت به آن حدس دارند، منطبق است.

← در فرمول نامساوی ای که داریم، نقطه‌ی تعادل نش را با وجود اطلاعات کامل بدست می آوریم؛ امّا این به این معنی نیست که از این تعادل در مسائل بدون اطلاعات کامل استفاده نمی‌شود.

← تعادل استراتژی غالب از تعادل نش یک تعریف قوی‌تر است.

تعادل استراتژی غالب ← تعادل نش

تعادل نش در بازی های پیوسته

مثال :

بازی کورنات

$$I = \{1, 2\} \quad , \quad q_i = [0, \infty)$$

$$P(Q) = \alpha - \beta Q \quad , \quad Q = q_1 + q_2$$

$$u_i(q_1, q_2) = P(Q)\, q_i - C_i(q_i) \quad , \quad C_i(q_i) = c\, q_i$$

$$\rightarrow u_i(q_1, q_2) = (\alpha - \beta(q_1 + q_2))\, q_i - c q_i$$

$$\rightarrow u_1 = [\alpha - \beta(q_1 + q_2)]\, q_1 - c q_1$$

$$BR_1(q_2) = \underset{q_1 \geq 0}{\operatorname{Arg\,max}}\ u_1(q_1, q_2)$$

$$\frac{\partial u_1}{\partial q_1} = [\alpha - \beta(q_1 + q_2)] - \beta q_1 - c = 0$$

$$\rightarrow q_1^* = BR_1(q_2) = \frac{\alpha - c - \beta q_2}{2\beta}$$

$$\frac{\partial^2 u_1}{\partial q_1^2} = -2\beta < 0 \ \checkmark$$

در جهت ماکزیمم کردن هستیم.

$$\rightarrow u_1 = [\alpha - \beta(q_1 + q_2)]\, q_2 - c q_2$$

$$\frac{\partial u_2}{\partial q_2} = [\alpha - \beta(q_1 + q_2)] - \beta q_2 - c = 0$$

$$\rightarrow q_2^* = BR_2(q_1) = \frac{\alpha - c - \beta q_1}{2\beta}$$

برای رسیدن به تعادل نش، برخورد BR هردو بازیگر را محاسبه می‌کنیم.

$$q_1^* = q_2^* = \frac{\alpha - c}{3\beta} \longrightarrow$$ تعادل نش

$$Q = q_1^* + q_2^* = \frac{2(\alpha - c)}{3\beta} \rightarrow P^* = \alpha - \beta(q_1^* + q_2^*) = \frac{\alpha + 2c}{3}$$

سود $u_i^* = \left(\frac{\alpha + 2c}{3}\right)\left(\frac{\alpha - c}{3\beta}\right) - c\,\frac{(\alpha - c)}{3\beta} = \left(\frac{\alpha - c}{3\beta}\right)\left(\frac{\alpha - c}{3}\right)$ سود هر بازیگر در تعادل نش

می‌توانیم نشان دهیم اگر دو بازیگر باهم همکاری می‌کردند، می‌توانستند به سود بیش‌تری برسند.

اگر فروش خود را کاهش بدهند:

$$\tilde{q}_1 = \tilde{q}_2 = \left(\frac{\alpha - c}{4\beta}\right)$$

$$Q = \tilde{q}_1 + \tilde{q}_2 = \frac{\alpha - c}{2\beta} \longrightarrow \tilde{P} = \alpha - \beta(\tilde{q}_1 + \tilde{q}_2) = \frac{\alpha + \beta}{2}$$

سود جدید $\tilde{u}_i = \left(\frac{\alpha + c}{2} - c\right)\left(\frac{\alpha - c}{4\beta}\right) = \left(\frac{\alpha - c}{2}\right)\left(\frac{\alpha - c}{4\beta}\right)$

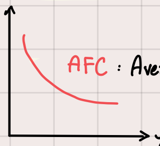
$$\Longrightarrow \tilde{u}_i > u_i^*$$

تابع هدف

← تابع هزینه (Cost func.)

• وقتی بازیگرها تولیدکننده هستند.

• هزینه: ۱- ثابت 2- متغیر (نسبت به تولید)



AFC : Average Fixed Cost = $\frac{Cost}{y}$

AVC : Average Variable Cost

مثال: $VC = C_0 + C_1 y + C_2 y^2$

$\rightarrow AVC = \frac{C_0}{y} + C_1 + C_2 y$

$\Rightarrow AC = AFC + AVC$

$\frac{C_0}{y} + C_1 + C_2 y$

مثال: $AC = \frac{C_3}{y} + C_1 + C_2 y$

$\Rightarrow Cost = C_2 y^2 + C_1 y + C_3$

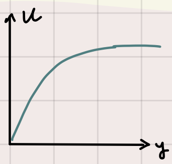
• هزینه‌ی حدی (MC)

افزایش هزینه به ازای افزایش یک واحد تولید

پیوسته: $\frac{dC(y)}{dy}$ / گسسته: $\frac{\Delta C(y)}{\Delta y}$

← تابع ارزش (utility func.)

• مدل کردن ارزش دستاورد بازیگر در بازی



افزایشی و مقعر ⇝

مثال → $\log(1+y)$

مثال → $\begin{cases} -ay^2 + by & ; y \leq y^* \\ c = -ay^{*2} + by^* & ; y > y^* \end{cases}$

مثال تعادل نش

بازی برترند: رقابت بر سر قیمت‌دهی بین تولیدکنندگان

$I = \{1,2\}$, قیمت تولید $P_i \in [0,+\infty)$

مقدار نیاز بازار به ازای قیمت (demand func.) $Q(P)$

$$\pi_i(P_i, P_{-i}) = \begin{cases} Q(P_i)P_i - \overset{MC}{C}\,Q(P_i) & P_{-i} > P_i \\ \frac{1}{2}\left(Q(P_i)P_i - C\,Q(P_i)\right) & P_{-i} = P_i \\ 0 & P_{-i} < P_i \end{cases}$$

نش در کدام حالت؟

← درحالتی که قیمت بازیگر i بیش‌تر از رقیب باشد. $P_i > P_{-i} > C$

خیر! P_i دراین حالت سود صفر بدست می‌آورد. پس انگیزه دارد با تغییر استراتژی با ثابت بودن استراتژی رقیب، قیمت را کمتر کند تا سود بیش‌تری بکند.

← درحالتی که قیمت بازیگر i برابر با رقیب باشد: $P_i = P_{-i} > C$

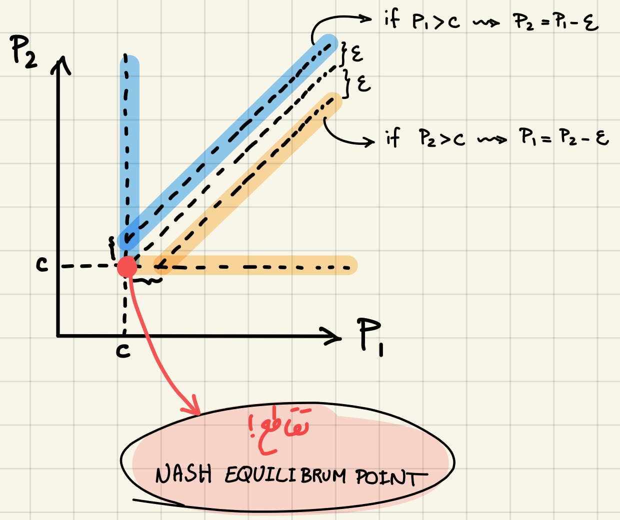
خیر! چون هردو بازیگر دراین صورت انگیزه دارند که با کاهش قیمت، بتوانند سود بیش‌تری داشته باشند.

← درحالتی که قیمت بازیگر i بیش‌تر از رقیب باشد. $P_i > P_{-i} = C$

خیر! بازیگر i دراین حالت انگیزه‌ای برای تغییر استراتژی ندارد نیازی ندارد که برای سود، قیمتش را کاهش دهد. چرا؟ چون دراین حالت که سودش صفر است. حال اگر فرض کنیم قیمتش را تا برابر شدن با C یا P_{-i} کاهش دهد، باز هم طبق ضابطه سودش صفر است. پس برایش فرقی ندارد پس استراتژی اش را تغییر نمی‌دهد. اما بازیگر دیگر تمایل دارد با افزایش قیمت بسیار کم سودش را از صفر بیش‌تر کند.

← درحالتی که قیمت بازیگر i برابر با رقیب باشد: $P_i = P_{-i} = C$

بله! هر دو بازیگر انگیزه‌ای برای تغییر قیمت ندارند هرکدام با کاهش قیمت سود منفی خواهند داشت. و با افزایش قیمت سودشان صفر می‌شود. پس این نقطه می‌تواند نقطه‌ی تعادل نش بازی باشد.



مثال تعادل نش

← بازی second price auction

رقابت بین خریداران

کسی که بیش‌ترین قیمت را بده برنده است. اما قیمتی برابر با نفر دوم باید بپردازد.

$I=\{1,\dots,n\}$ ، پیشنهاد قیمت خرید bid ~ $b_i \in [0,+\infty)$

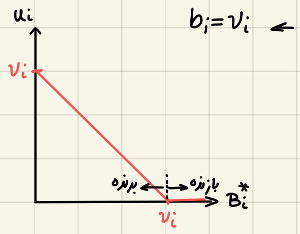
$v_1 > v_2 > \dots > v_n$ ⟿ ارزش کالای مورد نظر برای بازیگر i : v_i تابع سود

$B_i^* = \underset{j\neq i}{\mathrm{Max}}\ b_j$ ، $i^* = \underset{i}{\mathrm{Argmax}}\ b_i$ برنده‌ی بازی

پول پرداختی نفر اول → $B^* = \underset{j\neq i^*}{\mathrm{Max}}\ b_j$ دومین بیش‌ترین قیمت

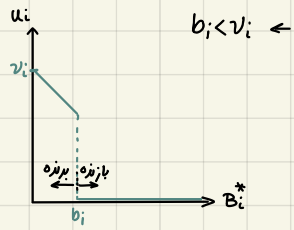
سود بازیگر i : $u_i = \begin{cases} v_i - B^* & ; \ i = i^* \\ 0 & ; \ o.w. \end{cases}$

← $b_i = v_i$

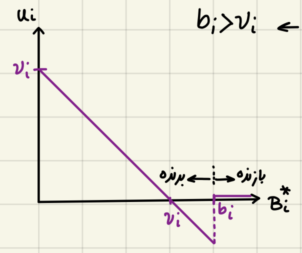


استراتژی غالب ←
⇓
نقطه‌ی تعادل نش
$(v_1, v_2, \dots, v_n)$

← $b_i < v_i$



← $b_i > v_i$



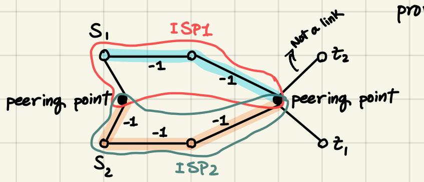
← سایر تعادل های نش $(v_1 > v_2 > \dots > v_n)$

- $(v_1, 0, 0, \dots, 0)$ → نفر اول برنده
- $(v_2, v_1, \dots, 0)$ → نفر دوم برنده

مثال تعادل نش

← بازی ISP Routing

دوتا سرویس provider



هدف SP1 : رساندن ترافیک از S_1 به t_1

هدف SP2 : رساندن ترافیک از S_2 به t_2

- هر SP برای استفاده از هر لینک هزینه‌ی 1- دارد. ← مسیر آبی هزینه‌ی 2- دارد.

- هر SP این امکان را دارد از لینک طرف مقابل نیز استفاده کند ← هزینه‌ی باقی انتقال بر عهده‌ی SP مقابل
  پس SP مقابل هم هزینه‌ی ارسال ترافیک خود را می‌دهد هم هزینه‌ی SP مقابلش ← مسیر نارنجی
  مثلاً P2 باید S3 هزینه برای ارسال ترافیک SP1 تحمل کند.

ISP1

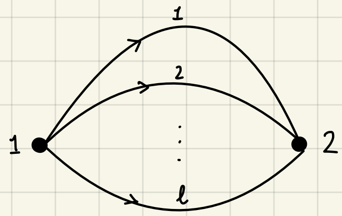
ISP2	C	N
C	-2,-2	-5,-1
N	-1,-5	-4,-4

تعادل نش ← (-4,-4)

N : همکاری کنند.

C : هرکس لینک خودش

مثال تعادل نش



بازی Routing

یک شبکه از لینک های موازی

d واحد جریان از گره 1 به 2

صاحب هر لینک، بازیگر.

← بازیگران لینک ها را قیمت دهی می کنند.

$I = \{1, \dots, l\}$,

* هزینه عبور واحد جریان بدلیل تاخیر، اشغال لینک بصورت $\overbrace{}^{l_i(x_i)}$ تابعی از کل جریان عبوری از لینک i می باشد $\underbrace{}_{x_i}$

* بازیگر i، قیمت P_i را برای هر واحد جریان عبوری از لینک i، قرار می دهد.

$\Longrightarrow$ هزینه ی موثر استفاده از لینک i $\quad C_i = P_i + l_i(x_i)$

* کاربران مقدار reservation utility (R) دارند. یعنی اگر $C_i > R$ شد، هیچ ترافیکی از لینک i عبور نمی کند

WARDROP E.q.

← فرض می کنیم d واحد جریان طوری بین l لینک تقسیم می شود که هزینه استفاده از لینک ها باهم برابر شود.

$$\begin{cases} \forall i \quad P_i + l_i(x_i) = \min_j \left[P_j + l_j(x_j)\right] \\ P_i + l_i(x_i) < R \\ \sum x_i = d \end{cases}$$

حل ←

فرض ها $\quad l = 2$, $\quad l_1(x_1) = 0,\ l_2(x_2) = \frac{3}{2} x_2$

$d = 1$ $\qquad 0 \leq P_1, P_2 \leq 1$

$R = 1$

$$u_1(P_1, P_2) = P_1 x_1 = (P_1 + l_1(x_1))\, x_1 - l_1(x_1)\, x_1$$

$$u_2(P_1, P_2) = P_2 x_2 = (P_2 + l_2(x_2))\, x_2 - l_2(x_2)\, x_2$$

$\rightarrow x_1 + x_2 = 1$

$$P_1 + \underbrace{\cancel{l_1(x_1)}}_{0} = P_2 + \underbrace{\cancel{l_2(x_2)}}_{\frac{3x_2}{2}} \Longrightarrow x_2 = \begin{cases} \frac{2}{3}(P_1 - P_2) & P_1 \geq P_2 \\ 0 & \text{o.w.} \end{cases}$$

$x_1 = 1 - x_2$

← بهترین پاسخ بازیگرها.

بازیگر 1 : $\max\ P_1 x_1(P_1, P_2) \quad ; \quad 0 \leq P_1 \leq 1$

$P_1(1 - x_2) = P_1\left(1 - \frac{2}{3}(P_1 - P_2)\right) \quad ; \quad \underline{P_1 \geq P_2}$

$\leadsto \max\ P_1\left(1 - \frac{2}{3}(P_1 - P_2)\right) \quad ; \quad \underline{P_1 \geq P_2}$

$$\frac{\partial P_1\left(1 - \frac{2}{3}(P_1 - P_2)\right)}{\partial P_1} = 1 - \frac{2}{3}(P_1 - P_2) - \frac{2}{3} P_1$$

$$\frac{\partial^2}{\partial P_1^2} = \frac{-4}{3} < 0 \rightarrow P_1 = \frac{3}{4} + \frac{P_2}{2} \ ; \ 0 \leq P_1 < 1 \ , \ P_1 \geq P_2$$

$$BR_1(P_2) = \min\left\{1, \frac{3}{4} + \frac{P_2}{2}\right\}$$

بازیگر 2 : $\max\ P_2 x_2(P_1, P_2) \quad ; \quad 0 \leq P_2 \leq 1$

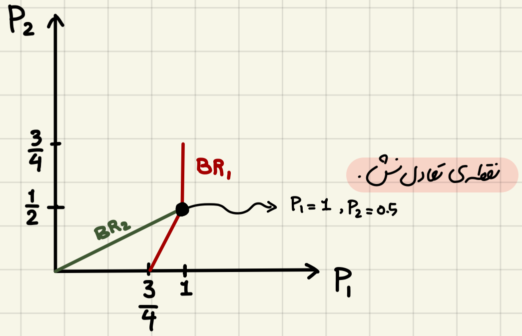
$P_2(1 - x_2) = P_2\left(1 - \frac{2}{3}(P_1 - P_2)\right) \quad ; \quad \underline{P_1 \geq P_2}$

$\leadsto \max\ P_2\left(1 - \frac{2}{3}(P_1 - P_2)\right) \quad ; \quad \underline{P_1 \geq P_2}$

$$\frac{\partial P_2\left(1 - \frac{2}{3}(P_1 - P_2)\right)}{\partial P_2} = \frac{2}{3}(P_1 - P_2) - \frac{2}{3} P_2$$

$$\frac{\partial^2}{\partial P_1^2} = \frac{-4}{3} < 0 \rightarrow P_2 = \frac{P_1}{2} \ , \ P_1 \geq P_2$$

$$BR_2(P_1) = \frac{P_1}{2}$$



استراتژی مخلوط

مثال :

Player 2

Player 1	H	T
H	−1, 1	1, −1
T	1, −1	−1, 1

Pure Strategy

اگر تمام استراتژی ها را بررسی کنیم، خواهیم دید که این بازی ، تعادل نش ندارد.

اگر بازیگر 2 ، حدس بزند که $\underbrace{\text{بازیگر 1 ، H را انجام دهد}}$ ، H را انجام می دهد .

بازیگر 1 درچه صورتی H را انجام می‌دهد؟

درصورتی‌که بازیگر 2 ، T را انجام دهد، بازیگر 1 با این فرض ، H را انجام می دهد .

← پس در نهایت حدس ها و اعمال دو بازیگر بر یکدیگر منطبق نمی‌شود!

با بررسی تمام حالت‌ها به همین نتیجه خواهیم رسید این بازی تعادل نش ندارد!

← فرض : بازیگر 1 با احتمال $\frac{1}{2}$ ، عمل 1 و 2 را انجام دهد.

بازیگر 2 با احتمال $\frac{1}{2}$ ، عمل 1 و 2 را انجام دهد.

$\left(\left(\frac{1}{2},\frac{1}{2}\right),\left(\frac{1}{2},\frac{1}{2}\right)\right)$ ⇝ فضای stochastic

؟ چه توزیعی روی فضای استراتژی بازیگران یک تعادل نش است؟

← در حالت کلی :

Player 2

Player 1	1	2
1	a, a'	b, b'
2	c, c'	d, d'

$$\delta_1 = \begin{pmatrix} P_1^1 \\ P_1^2 \end{pmatrix} \quad \Rightarrow \quad P_1^1 + P_1^2 = 1 \,/\, P_1^1, P_1^2 \geqslant 0$$

P_1^1 ⇝ احتمال عمل 1 توسط بازیگر 1

P_1^2 ⇝ احتمال عمل 2 توسط بازیگر 1

$$\delta_2 = \begin{pmatrix} P_2^1 \\ P_2^2 \end{pmatrix} \quad \Rightarrow \quad P_2^1 + P_2^2 = 1 \,/\, P_2^1, P_2^2 \geqslant 0$$

P_2^1 ⇝ احتمال عمل 1 توسط بازیگر 2

P_2^2 ⇝ احتمال عمل 2 توسط بازیگر 2

← برای توابع سود :

$$A = \begin{bmatrix} a & b \\ c & d \end{bmatrix} , \quad B = \begin{bmatrix} a' & b' \\ c' & d' \end{bmatrix}$$

برای داشتن یک مقدار قطعی برای سود :

$$E(u_1) = U_1 \quad , \quad E(u_2) = U_2$$

پس با کمک امید ریاضی تابع سود را تعریف می کنیم :

$$E(u_1) = U_1 = P_1^1 P_2^1 a + P_1^1 P_2^2 b + P_1^2 P_2^1 c + P_1^2 P_2^2 d$$

$$E(u_2) = U_2 = P_1^1 P_2^1 a' + P_1^1 P_2^2 b' + P_1^2 P_2^1 c' + P_1^2 P_2^2 d'$$

$$\Rightarrow \quad E(u_1) = U_1 = \delta_1^T A \delta_2 \quad ; \quad E(u_2) = U_2 = \delta_2^T B^T \delta_1 = \delta_1^T B \delta_2$$

بازی استراتژیک

$$G(I, (S_i), (u_i))$$

مجموعه‌ی استراتژی خالص بازیگر i ام :

$$S_i = \{ a_i^1, \dots, a_i^{m_i} \} \quad , i \in I$$

استراتژی مخلوط یا Mixed Strategy بازیگر i ام :

احتمال انجام عمل a_i^1

$$\delta_i = (P_i^1, \dots, P_i^{m_i})^T \quad , i \in I$$

$$\sum_{k=1}^{m_i} P_i^k = 1 \quad , \quad P_i^k \geqslant 0$$

$$\delta_i \in \Sigma_i \subseteq [0 \;\; 1]^{m_i}$$

مثلاً وقتی 2 عمل وجود دارد، P_i^1 , P_i^2 داریم که چون $P_i^1 + P_i^2 = 1$ ، خطی که روی نمودار تشکیل می‌دهند واضحاً 1 بُعدی است. این δ_i که یک خط است درواقع در یک فضای 2 بعدی (صفحه) قرار می‌گیرد که به آن Σ_i می‌گوییم.

پس Σ_i یک فضای $m_i - 1$ بُعدی است.

$$\rightarrow \quad \delta_i = (\overbrace{\delta_i(a_i^1)}^{P_i^1}, \dots, \delta_i(a_i^{m_i}))^T$$

$$\rightarrow \quad \boxed{\delta_i(s_i) = P_i^k} \quad \text{for} \quad s_i = a_i^k$$

$\hookrightarrow s_i \in S_i$

تابع سود :

سود بازیگر i وقتی بازیگر 1 ، s_1 ... و بازیگر n ، s_n را بازی می‌کنند.

احتمال

$$E(u_i) = U_i(\delta_1, \dots, \delta_n) = \sum u_i(s_1, \dots, s_n)\, \delta_1(s_1) \cdots \delta_n(s_n)$$

؟ بدون فرض استقلال : به‌کل s_1 تا s_n بعنوان S نگاه می‌کنیم. احتمال وقوع S ، را با $\delta(S)$ نمایش می دهیم.

$$\delta \in \Sigma = \prod_{i \in I} \Sigma_i$$

$= \delta(s_1, \dots, s_n)$

* چون استقلال نداریم نمی‌توان این جمله را بصورت ضرب احتمالات نوشت.

احتمال اینکه بازیگر 1 ، s_1 را انتخاب کند و... و بازیگر n ، s_n را انتخاب کند.

$$U_i(\delta) = \sum_{s \in S} u_i(s)\, \delta(s)$$

$\downarrow$ استراتژی مخلوط بازیگران

← در حالت پیوسته:

هر بازیکن چه تابع توزیعی را انتخاب می‌کند؟

$$u_i(\delta) = \int_S u_i(s)\, d\underbrace{\delta(s)}_{\text{تابع احتمال مشترک بازیکنان}}$$

← در حالت استقلال:

$$u_i(\delta) = \int_{S_n}\cdots\int_{S_1} u_i(s)\, \overbrace{f_1(s_1)\cdots f_n(s_n)}^{\text{تابع احتمال}}\, ds_1\cdots ds_n$$

$$d\delta(s) = d\delta_1(s_1)\ldots d\delta_n(s_n)$$

$$d\delta_i(s_i) = f_i(s_i)\, ds_i$$

هدف!

جمع‌بندی استراتژی مخلوط

← پروفایل استراتژی مخلوط بازیکن i : δ_i

← احتمال انجام عمل s_i توسط بازیکن i : $\delta_i(s_i)$

← پروفایل استراتژی مشترک مخلوط بازیکنان : $\delta = (\delta_1^T, \ldots, \delta_n^T)^T$

← احتمال وقوع پروفایل $S=(s_1,\ldots,s_n)$ توسط بازیکنان : $\delta(s) = \delta(s_1,\ldots,s_n)$

تعادل نش مخلوط ← Mixed N.E

اون پروفایل مخلوطی که در آن هیچ بازیکنی با تغییر پروفایل استراتژی شخصی‌اش نتواند به Expected utility بیش‌تری برسد.

تعریف. تعادل نش مخلوط

استراتژی δ^* یک تعادل نش مخلوط است اگر:

$$\forall \delta_i\in\Sigma_i,\ \forall i \in I,\ \delta^* = (\delta_1^{*T},\ldots,\delta_n^{*T})^T = (\delta_i^{*T}, \delta_{-i}^{*T})^T$$

$$\rightsquigarrow U_i(\delta_i^*, \delta_{-i}^*) \geq U_i(\delta_i, \delta_{-i}^*)$$

ادعا می‌کنیم این تعریف معادل معادله‌ی زیر است:

$$\boxed{U_i(\delta_i^*, \delta_{-i}^*) \geq U_i(s_i, \delta_{-i}^*)\ ;\ \forall i\in I,\ \forall s_i \in S_i}$$

مفهوم: برای اینکه نشان دهیم که δ_i^*، BR نسبت به δ_{-i}^* است، لازم نیست برای همه‌ی پروفایل‌های استراتژی مخلوط ممکن (δ_i) نامساوی موجود در تعریف را چک کنیم. این معادله ساده‌تر است زیرا می‌گوید δ_i^*، BR است نسبت به استراتژی خالص یا pure، آنگاه می‌توان گفت δ_i^* نسبت به همه‌ی پروفایل‌های δ_i، BR است.

این عمل را با احتمال 1 انجام دهد.

$$s_i = (0,0,\ldots,\underset{K\text{th}}{1},\ldots,0) \rightsquigarrow \text{pure strategy}$$

چرا؟ اثبات؟

← معادله‌ی اول یعنی تعادل نش مخلوط، معادله‌ی دوم را نتیجه می‌دهد:

* درواقع نامساوی دوم یک حالت خاص از نامساوی اول است پس می‌توان آنرا از نامساوی دوم نتیجه گرفت.

← معادله‌ی دوم معادله‌ی اول را نتیجه می‌دهد:

وقتی در نامساوی دوم، شرایط را به ازای همه‌ی استراتژی‌های s_i چک می‌کنیم. چون δ_i یک ترکیب خطی از s_i ها، می‌توان ضریب ترکیب خطی را احتمال وقوع آن استراتژی در نظر بگیریم.

$$U_i(\delta_i, \delta_{-i}^*) = \sum_{\substack{\forall s_i\in S_i\\ \forall s_{-i}\in S_{-i}}} u_i(s_i,s_{-i})\,\delta_i(s_i)\,\delta_{-i}(s_{-i})$$

$$= \sum_{\forall s_i\in S_i}\sum_{\forall s_{-i}\in S_{-i}} u_i(s_i,s_{-i})\,\delta_i(s_i)\,\delta_{-i}(s_{-i})$$

$$= \sum_{\forall s_i\in S_i}\delta_i(s_i)\underbrace{\sum_{\forall s_{-i}\in S_{-i}} u_i(s_i,s_{-i})\,\delta_{-i}(s_{-i})}$$

$U_i(s_i, \delta_{-i})$ = متوسط سود بازیکن i وقتی بازیکن i عمل s_i را انجام دهد (بطور قطعی) و بقیه‌ی بازیکن‌ها (i-) با پروفایل استراتژی δ_{-i} دارند اعمال s_{-i} را انجام می‌دهند.

* پس عبارت $U_i(\delta_i, \delta_{-i}^*)$ برابر است با $\sum_{\forall s_i\in S_i}\delta_i(s_i)U_i(s_i,\delta_{-i}^*)$. یعنی یک ترکیب خطی است از $U_i(s_i,\delta_{-i}^*)$.

← معرفی نامساوی سوم:

به ازای تمام $s_i\in S_i$ که در support بردار δ_i^* وجود دارند: (با احتمال غیرصفر)

$$\boxed{U_i(\delta_i^*, \delta_{-i}^*) = U_i(s_i, \delta_{-i}^*)}$$

مفهوم: استراتژی تعادلی نش برای هر بازیکن i با پروفایل استراتژی نش δ_i^*، متوسط سودش در مقابل استراتژی نش سایر بازیکنان برابر است با همه‌ی استراتژی‌های حالتی که بازیکن i می‌تواند بازی کند در مقابل استراتژی مخلوط نش سایر بازیکنان. فقط یک شرط دارد: s_i هایی در این تساوی صدق می‌کنند که در پروفایل δ_i^* با احتمال غیرصفر ظاهر شده باشند.

چرا؟ برهان خلف: فرض می‌کنیم $\exists s_i \in S_i$ که با احتمال غیرصفر ظاهر شده در δ_i^*، و سود ناشی از آن کمتر از سود سایر استراتژی‌ها باشد پس اگر این سود را کمتر کنیم (وزن یا احتمالش را کم کنیم)، سود بیش‌تر خواهد شد. که این تناقض است با اینکه پروفایل δ_i^* تعادل نش باشد.

مثال استراتژی مخلوط نش

← مثال تساوی سوم:

		Player 2: H	T
Player 1	H	−1 , 1	1 , −1
	T	1 , −1	−1 , 1

گفته بودیم که این بازی نقطه‌ی تعادل نش خالص ندارد.

$$\sigma^* = \Big(\underbrace{(\tfrac{1}{2},\tfrac{1}{2})}_{\sigma_1^{*T}},\underbrace{(\tfrac{1}{2},\tfrac{1}{2})}_{\sigma_2^{*T}}\Big)^T$$ ← تعادل نش مخلوط

طبق تساوی سوم → $U_1(H,\sigma_2^*) = \frac{1}{2}(-1) + \frac{1}{2}(1) = 0$

$U_1(T,\sigma_2^*) = \frac{1}{2}(1) + \frac{1}{2}(-1) = 0$

$U_1(\sigma_1^*,\sigma_2^*) = \frac{1}{2}\frac{1}{2}(-1) + \frac{1}{2}\frac{1}{2}(1) + \frac{1}{2}\frac{1}{2}(1) + \frac{1}{2}\frac{1}{2}(-1) = 0$

پس این تساوی تأیید می‌شود: $U_i(\sigma_i^*,\sigma_{-i}^*) = U_i(s_i,\sigma_{-i}^*)$

مثال استراتژی مخلوط نش

← مثال تساوی سوم:

		Player 2: B	F
Player 1	B	2 , 1	0 , 0
	F	0 , 0	1 , 2

این بازی نیز نقطه‌ی تعادل نش خالص ندارد.

$$\sigma^* = \Big(\underbrace{(\overset{B}{\tfrac{2}{3}},\overset{F}{\tfrac{1}{3}})}_{\sigma_1^{*T}},\underbrace{(\overset{B}{\tfrac{1}{3}},\overset{F}{\tfrac{2}{3}})}_{\sigma_2^{*T}}\Big)^T$$ ← تعادل نش مخلوط

طبق تساوی سوم → $U_2(\sigma_1^*,B) = \frac{2}{3}(1) + \frac{1}{3}(0) = \frac{2}{3}$

$U_2(\sigma_1^*,F) = \frac{2}{3}(0) + \frac{1}{3}(2) = \frac{2}{3}$

$U_2(\sigma_1^*,\sigma_2^*) = \frac{2}{3}\frac{1}{3}(1) + \frac{2}{3}\frac{2}{3}(0) + \frac{1}{3}\frac{1}{3}(0) + \frac{1}{3}\frac{2}{3}(2) = \frac{2}{3}$

پس این تساوی تأیید می‌شود: $U_i(\sigma_i^*,\sigma_{-i}^*) = U_i(s_i,\sigma_{-i}^*)$

نحوه‌ی محاسبه تعادل نش مخلوط

← با یک مثال نحوه‌ی محاسبه‌ی تعادل نش مخلوط را نشان می‌دهیم:

		Player 2: 1	2
Player 1	1	1 , 3	0 , 2
	2	2 , 0	−1 , 1

- ماتریس سود بازیکن 1: $A = \begin{bmatrix} 1 & 0 \\ 2 & 1 \end{bmatrix}$ ، ماتریس سود بازیکن 2: $B = \begin{bmatrix} 3 & 2 \\ 0 & 1 \end{bmatrix}$
- استراتژی مخلوط بازیکن 1: $P = \begin{pmatrix} P_1 \\ P_2 \end{pmatrix}$، $P_1 + P_2 = 1$ ، استراتژی مخلوط بازیکن 2: $q = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}$، $q_1 + q_2 = 1$
- تابع سود بازیکن 1: $U_1 = P^T A q$ ، تابع سود بازیکن 2: $U_2 = P^T B q$

← بررسی می‌کنیم که بازی تعادل نش خالص ندارد.

فرض: تعادل نش مخلوط در (P^*,q^*) باشد:

$1-\ P^{*T} A q^* \geq P^T A q^*$ ← استفاده مستقیم از نامساوی نش:

$P_1^* q_1^* \times 1 + 0 + (1-P_1^*) q_1^* \times 2 + (1-P_1^*)(1-q_1^*) \times (-1) \geq$

$P_1 q_1^* + 0 + (1-P_1) q_1^* \times 2 + (1-P_1)(1-q_1^*) \times (-1)$

$\rightarrow -2P_1^* q_1^* + P_1^* \geq -2P_1 q_1^* + P_1 \rightarrow P_1^*(1-2q_1^*) \geq P_1(1-2q_1^*)$

$2-\ P^{*T} B q^* \geq P^{*T} B q$

$P_1^* q_1^* \times 3 + P_1^*(1-q_1^*) \times 2 + 0 + (1-P_1^*)(1-q_1^*) \times 1 \geq$

$P_1^* q_1 \times 3 + P_1^*(1-q_1) \times 2 + 0 + (1-P_1^*)(1-q_1) \times 1$

$\rightarrow 2P_1^* q_1^* - q_1^* \geq 2P_1^* q_1 - q_1 \rightarrow q_1^*(2P_1^*-1) \geq q_1(2P_1^*-1)$

$P_1^*(1-2q_1^*) \geq P_1(1-2q_1^*)$

$q_1^*(2P_1^*-1) \geq q_1(2P_1^*-1)$

✓ اگر جملات زرد شده صفر باشند. هر دو نامساوی به تساوی تبدیل می‌شوند.

$\rightsquigarrow P_1^* = q_1^* = \frac{1}{2}$

✗ اگر فقط در نامساوی بالا جمله زرد شده صفر باشد: در این صورت $q_1^* = \frac{1}{2}$ ؛ در نامساوی دوم: $\frac{1}{2} \geq q_1$ یا $\frac{1}{2} \leq q_1$ ✗.

✗ اگر هیچ کدام از جملات زرد شده صفر نباشند:

$1-2q^* \neq 0$: $q_1^* < \frac{1}{2} \rightarrow P_1^* \geq P_1 \rightarrow P_1^* = 1$ ؛ $q_1^* > \frac{1}{2} \rightarrow P_1^* \leq P_1 \rightarrow P_1^* = 0$

هیچ حالتی درست و جور درنمی‌آید!

$2P^* - 1 \neq 0$: $P_1^* < \frac{1}{2} \rightarrow q_1^* \leq q_1 \rightarrow q_1^* = 0$ ؛ $P_1^* > \frac{1}{2} \rightarrow q_1^* \geq q_1 \rightarrow q_1^* = 1$

پس هیچ کدام جواب نیستند.

پس پاسخ برابر است با: $(P^*,q^*) = \Big((\frac{1}{2},\frac{1}{2}),(\frac{1}{2},\frac{1}{2})\Big)$

← پس برای اینکار باید بررسی کنیم در مسأله‌ی ما استراتژی مغلوب وجود نداشته باشد.
که پس از بررسی هر دو استراتژی مشاهده می‌کنیم هیچ‌کدام مغلوب نیستند.

فرض: $p_1^*, q_1^* \rightsquigarrow \underline{NASH}$

استراتژی‌های: $\sigma_1^* = \begin{pmatrix} p_1^* \\ 1-p_1^* \end{pmatrix} \quad , \quad \sigma_2^* = \begin{pmatrix} q_1^* \\ 1-q_1^* \end{pmatrix}$

← طبق نتیجه‌گیری که در جلسه‌ی قبل داشتیم:

$$\rightsquigarrow U_1(1, \sigma_2^*) = U_1(2, \sigma_2^*) = U_1(\sigma_1^*, \sigma_2^*)$$

$$U_1(1, \sigma_2^*) = 1 \times q_1^* + 0 \times (1-q_1^*) = q_1^*$$
$$U_1(2, \sigma_2^*) = 2 \times q_1^* + (-1)(1-q_1^*) = 3q_1^* - 1$$
$$q_1^* = 3q_1^* - 1 \to q_1^* = \frac{1}{2}$$

$$\rightsquigarrow U_2(\sigma_1^*, 1) = U_2(\sigma_1^*, 2) = U_2(\sigma_1^*, \sigma_2^*)$$

$$U_2(\sigma_1^*, 1) = 3 \times p_1^* + 0 \times (1-p_1^*) = 3p_1^*$$
$$U_2(\sigma_1^*, 2) = 2 \times p_1^* + 1 \times (1-p_1^*) = p_1^* + 1$$
$$3p_1^* = p_1^* + 1 \to p_1^* = \frac{1}{2}$$

؟ اگر در این حالت استراتژی مغلوب در مسأله وجود داشته باشد؟

← با حل یک مثال نشان می‌دهیم: (معمای زندانی)

Player 2

Player 1	1	2
1	-4 , -4	-1 , -5
2	-5 , -1	-2 , -2

نقطه‌ی تعادل نش

واضح است استراتژی 2 نسبت به 1 مغلوب است. به دنبال استراتژی مخلوط:

$$U_1(1, \sigma_2^*) = U_1(2, \sigma_2^*)$$
$$-4q_1^* - (1-q_1^*) = -5q_1^* - 2(1-q_1^*)$$
$$\Rightarrow 0 = -1 \quad ✗$$

پس حذف استراتژی مغلوب (قوی!) برای استفاده از این روش مهم است.

مثال دیگر:

Player 2

Player 1	1	2
1	-4 , -4	-1 , -5
2	-5 , -1	-1 , -1

بازی دو تعادل نش خالص دارد!

← روش دوم. تقاطع BR (بهترین پاسخ‌ها)

Player 2

Player 1	1	2
1	1 , 3	0 , 2
2	2 , 0	-1 , 1

$$\rightsquigarrow E(u_1) = p_1 q_1 + 0 + 2(1-p_1)q_1 - (1-p_1)(1-q_1)$$
$$= p_1(q_1 - 2q_1 - q_1 + 1) + 2q_1 - (1-q_1)$$
$$= (1-2q_1)p_1 + 3q_1 - 1$$

در بهینه‌سازی نقش ندارد

$$\Rightarrow BR_1(q_1) = \underset{p_1}{Argmax}\ U_1(p_1, q_1) = \underset{p_1}{Argmax}\left[(1-2q_1)p_1 + 3q_1 - 1\right]$$

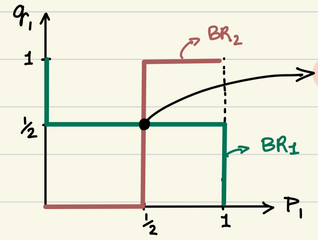
خطی

$$p_1^*(q_1) = \begin{cases} 1 & ; \ q_1 < \frac{1}{2} \\ \text{Don't Care} & ; \ q_1 = 1/2 \\ 0 & ; \ q_1 > \frac{1}{2} \end{cases}$$

$$\rightsquigarrow E(u_2) = p^T B q = 2p_1 q_1 + p_1 + 1 - q_1 = q_1(2p_1 - 1) + 1 + p_1$$

$$\Rightarrow BR_2(p_1) = \underset{q_1}{Argmax}\ U_2(p_1, q_1) = \underset{q_1}{Argmax}\left[q_1(2p_1-1) + 1 + p_1\right]$$

$$q_1^*(p_1) = \begin{cases} 1 & ; \ p_1 > \frac{1}{2} \\ \text{Don't Care} & ; \ p_1 = 1/2 \\ 0 & ; \ p_1 < \frac{1}{2} \end{cases}$$



NASH E.q

$\Rightarrow p_1^* = q_1^* = \frac{1}{2}$

← روش سوم: استفاده از تساوی حاصل از نامساوی نش

یادآوری:

$$\boxed{U_i(\sigma_i^*, \sigma_{-i}^*) = U_i(s_i, \sigma_{-i}^*)}$$

Player 2

Player 1	1	2
1	1 , 3	0 , 2
2	2 , 0	-1 , 1

- فرض اول: استراتژی‌هایی سودشان باهم برابر و برابر با استراتژی نش است که با احتمال غیرصفر در استراتژی مخلوط ظاهر می‌شوند.

زمانی با احتمال صفر ظاهر می‌شوند
که استراتژی مغلوب باشند
(نسبت به استراتژی خالص یا مخلوط)

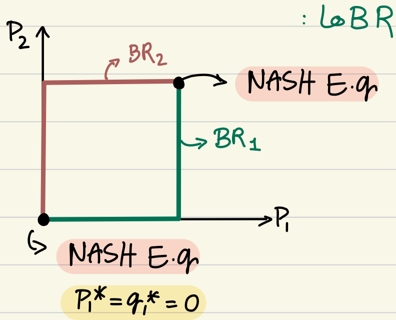
← در این مثال استراتژی 2 ، مغلوب ضعیف! است.

$$U_1(①, \sigma_2^*) = U_1(②, \sigma_2^*)$$

$$-4q_1^* - (1-q_1^*) = -5q_1^* - (1-q_1^*)$$

$$\Rightarrow q_1^* = 0 \xrightarrow[\text{متقارن}]{\text{بازی}} P_1^* = 0$$

← با استفاده از BRها :



* پس استراتژی مغلوب ضعیف! تساوی حاصل از نامساوی نش را نقض نمی‌کند.

← یک استراتژی ممکن است نسبت به یک استراتژی مخلوط، مغلوب باشد:

← مثال:

Player 2 / Player 1

	L	R
U	2,0	-1,0
M	0,0	0,0
D	-1,0	2,0

؟ آیا استراتژی M نسبت به ترکیب خطی از استراتژی های U و D مغلوب است؟

فرض: $\sigma = (\frac{1}{2}, 0, \frac{1}{2}) \rightarrow$

$$U_1(\sigma, L) = \frac{1}{2}\times 2 + 0 + (-1)\times\frac{1}{2} = \frac{1}{2}$$

$$U_1(\sigma, R) = \frac{1}{2}\times(-1) + 0 + \frac{1}{2}\times 2 = \frac{1}{2}$$

$$U_1(M, L) = 0 \;,\; U_1(M, R) = 0$$

$$U_1(\sigma, L) > U_1(M, L)$$

$$U_1(\sigma, R) > U_1(M, R)$$

استراتژی M نسبت به σ مغلوب است. ← پس برای بدست آوردن یک استراتژی مخلوط، می‌توان از ابتدا M را حذف کرد.

← مثال:

Player 2 / Player 1

	L	R	N
U	2,0	-1,0	2,-1
M	0,0	0,0	0,2
D	-1,0	2,0	-1,-1

← تعادل های نش خالص: (D,R) , (U,L)

← برای یافتن نش‌های مخلوط:

- حذف استراتژی های مغلوب:

این سطر و ستون هایی که در آن‌ها نش خالص پیدا شده، نمی‌توانند مغلوب باشند! مثلاً U و D برای بازیکن 1 و L و R برای بازیکن 2 نمی‌توانند مغلوب باشند.

سطر M و ستون N باقی می‌مانند. باید بررسی کرد آیا در تمام حالت ها M نسبت به U و D مغلوب است یا خیر؟

← در M، 0 باید بصورت ترکیبی از 2 و -1 مغلوب شود.

استراتژی مخلوط غالب نسبت به M $\leftarrow (\frac{1}{2}, 0, \frac{1}{2})$: مثلاً

$$\frac{1}{2}\times 2 + 0\times 0 + \frac{1}{2}(-1) = \frac{1}{2} > 0$$

استراتژی مخلوط غالب نسبت به M $\leftarrow (\frac{2}{5}, 0, \frac{3}{5})$: یا مثلاً

$$\frac{2}{5}\times 2 + 0\times 0 + \frac{3}{5}(-1) = \frac{1}{5} > 0$$

$\Leftarrow$ M یک استراتژی مغلوب نسبت به استراتژی مخلوط است.

← * بصورت iterative می‌توان استراتژی های مغلوب را حذف کرد.

پس از حذف M ، سود حاصل از استراتژی N فقط -1 و -1 خواهد بود. در این صورت نسبت به هر دو استراتژی L و R مغلوب است.

جدول جدید:

Player 2 / Player 1

	L	R
U	2,0	-1,0
D	-1,0	2,0

مثلاً اگر $P_1 < P_2$ ، P_1 انگیزه دارد که حوزش را افزایش دهد و برعکس.

⇐ پس حالت غیرتساوی نیز تعادل نش نخواهد بود

← آیا بازی تعادل نش مخلوط دارد؟

← روش حذف استراتژی مغلوب

با توجه به اینکه هر بازیگر مطمئن است که $\frac{1}{3}$ تقاضای بازار به آن می‌رسد. اگر این را بداند و قرار است بازنده باشد، بیش‌ترین سودی که می‌تواند کند، این است که قیمت را 1 دهد و در این صورت سودش $\frac{1}{3}$ خواهد شد.

اگر هرکدام از بازیگرها قیمت را کمتر از $\frac{1}{2}$ دهند و فرض کنیم که برنده هم می‌شود :

کمترین حالت سود

$$P_1 < \frac{1}{2} \longrightarrow \text{سود} = \frac{2}{3}\overset{1}{\cancel{D}} \times P_1 - 0 = \frac{2}{3}P_1 < \frac{1}{3}$$

اگر در این صورت استراتژی اش را تغییر می‌داد به 1 و بازنده هم باشد:

بدترین حالت

$$P_1 = 1 \longrightarrow \text{سود} = \frac{1}{3} \times 1 \times P_1 = \frac{1}{3}$$

⇐ فضای استراتژی $P_1, P_2 \in [0, \frac{1}{2})$ برای هردو بازیگر، مغلوب است نسبت به استراتژی $P_2 = 1$ و $P_1 = 1$ (مستقل از اینکه بازیگر رقیب چه کار می‌کند.)

← روش معادله‌ی سوم :

نیازی به اندیس نیست چون بازی متقارن است.

$$E(U_1(P_1, f_2^*(P_2))$$

متوسط سود بازیگر 1 وقتی بازیگر 1 استراتژی خالص P_1 را بازی کند در مقابل استراتژی مخلوط رقیب.

← بازیگر 1 بازنده باشد، $P_1 > P_2$:

$$\text{سود} = \left[\frac{1}{3}D \times P_1 - 0 \times \frac{1}{3}D\right] P\{P_2 < P_1\}$$

+

← بازیگر 1 برنده باشد، $P_2 > P_1$:

$$\text{سود} = \left[\frac{2}{3}D \times P_1 - 0 \times \frac{1}{3}D\right] P\{P_1 < P_2\}$$

جمع این‌ها = Expected Utility of Player 1

← یافتن نش مخلوط .

$$U_1(U, (\overset{\text{احتمال انجام } L}{q^*}, \overset{\text{احتمال انجام } R}{1-q^*})) = U_1(D, (q^*, 1-q^*))$$

$$\Rightarrow 2q^* - (1-q^*) = -q^* + 2(1-q^*) \Rightarrow q^* = \frac{1}{2}$$

$$U_2(L, (p^*, 1-p^*)) = U_2(R, (p^*, 1-p^*))$$

$$\Rightarrow 0 \times p^* + 0 \times (1-p^*) = 0 \times p^* + 0 \times (1-p^*)$$

$$\Rightarrow p^* \in (0, 1) \quad \text{و} \quad p^* = 0, p^* = 1$$

⇐ تعادل نش مخلوط : $\{((p^*, 0, 1-p^*), (\frac{1}{2}, \frac{1}{2}, 0))$, $p^* \in (0,1)$

* بی‌نهایت نش مخلوط

← مثال : بازی پیوسته ← برترند

NASH E.q. $P_1^* = P_2^* = C$

← ظرفیت محدود تولیدکنندگان .

فرض : هر بازیگر حداکثر $\frac{2}{3}$ تقاضای بازار را می‌تواند تأمین کند.

← تعادل نش قبلی دیگر تعادل نش در این شرایط نیست .

$$\text{تقاضای بازار} = D = 1, \quad P_1, P_2 \in [0, 1], \quad C = 0$$

اگر قیمت را صفر بدهم، سود هر جفت 0 خواهد شد. در این صورت بازیگر 1 این تمایل را خواهد داشت که مقدار قیمت را مثلاً به 1 تغییر دهد در این صورت با اینکه بازنده‌ی بازی می‌شود، چون قیمت بالاتری دارد، اما $\frac{1}{3}$ عرضه‌ی بازی را تأمین می‌کند.

$$P_1 : 0 \longrightarrow 1 \Rightarrow \frac{D}{3} \times 1 - 0 > 0$$

ادعا : حالت هایی که قبلاً نش بودند نیز دیگر نش نیستند .

اگر $P_1 = P_2 = a > 0$ ، آنگاه مثلاً P_1 این انگیزه را خواهد داشت که به اندازه‌ی ε از a قیمتش را کمتر دهد که $\frac{2}{3}$ عرضه را تأمین کند.

($\frac{1}{2}$ را به $\frac{2}{3}$ افزایش می‌یابد) پس سود می‌کند ←

⇐ به ازای هیچ a ، $P_1 = P_2 = a$ ، تعادل نش وجود ندارد.

Expected Utility.

احتمال باخت × سود باخت + احتمال برد × سود برد

$$\frac{1}{3}P_1\,F^*(P_1)+\frac{2}{3}P_1\left(1-F^*(P_1)\right) \Leftarrow$$

$$=\frac{1}{3}P_1\int_0^{P_1} f^*(P_2)\,dP_2+\frac{2}{3}P_1\int_{P_1}^{1} f^*(P_2)\,dP_2$$

$$=\frac{1}{3}P_1\,F^*(P_1)+\frac{2}{3}P_1-\frac{2}{3}F^*(P_1)=\underbrace{\frac{2}{3}P_1-\frac{1}{3}P_1F^*(P_1)}$$

مقدار این عبارت به ازای همه P_1هایی که در support تابع چگالی

احتمال نش f^* موجود در برابر است. پس به یک چیز کلی بجای P_1 می‌گذاریم p:

$$\rightarrow \frac{2}{3}p-\frac{1}{3}p\,F^*(p)=K$$

همه مقادیر P_1 که در support تابع احتمال وجود دارند: p

$$F^*(1)=1 \quad \left(\rightsquigarrow P\{P_2\leqslant 1\}=1 \checkmark\right)$$

$$\rightarrow \frac{2}{3}-\frac{1}{3}\times 1=K \rightarrow K=\frac{1}{3}$$

$$\rightsquigarrow \frac{2}{3}p-\frac{1}{3}p\,F^*(p)=\frac{1}{3} \Rightarrow F^*(p)=2-\frac{1}{p}$$

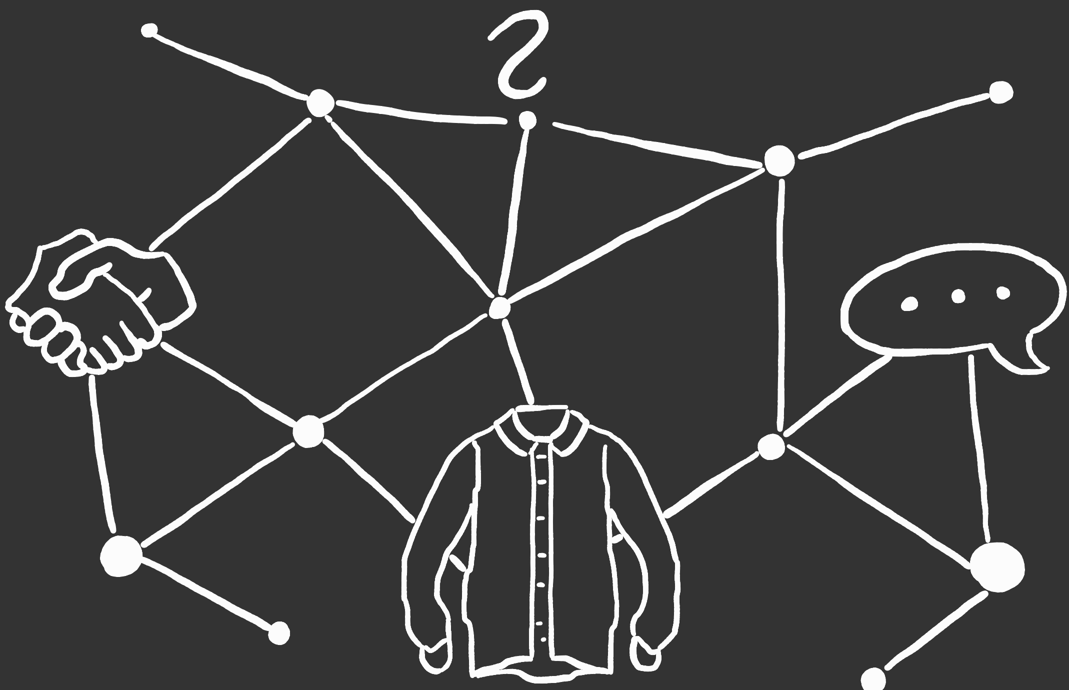
$$F^*(p)=\begin{cases}0 & 0\leqslant p<\frac{1}{2} \\ 2-\frac{1}{p} & \frac{1}{2}\leqslant p\leqslant 1 \\ 1 & p>1\end{cases}$$

استراتژی مغلوب

استراتژی تعادلی نش مخلوط

CHAPTER 2

~ Extensive - form games
~ Repeated game
~ Backward induction



فرم گسترده بازی (بازی های درختی)

بازی: یک ابزاری که بخواهد نتیجه‌ی interaction یک سری تصمیم‌گیر را که تصمیماتشان روی یکدیگر اثر دارد را نهایتاً Predict کند.

از هم اطلاعات کامل دارن؟ با هم رقیب هستن؟

همزمان تصمیم می‌گیرند؟ یا ترتیبی؟

- بازی‌های استراتژیک:
  - اطلاعات کامل
  - تصمیم‌گیری بدون ترتیب
  - بازی 1 مرحله‌ای
  - غیر همکارانه

- فرم گسترده‌ی بازی‌ها:
  - اطلاعات کامل
  - تصمیم‌گیری با امکان وجود ترتیب (بعضی از بازیگرها با ترتیب یا بعضی بدون ترتیب)
  - امکان وجود چندین مرحله تصمیم‌گیری
  - غیر همکارانه

مثال: ← نشان دادن بازی با ترتیب

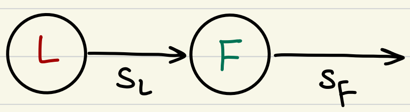
یک بازی با 2 بازیگر: رهبر (Leader) و پیرو (follower)

ابتدا L تصمیم می‌گیرد و F علق اطلاع از تصمیم L، تصمیم می‌گیرد.

$I=\{L,F\}$, $s_i \in S_i$, $\underbrace{J_F(s_F,s_L), J_L(s_F,s_L)}_{\text{توابع سود}}$

رهبر می‌داند که استراتژی پیرو در تابع سودش مؤثر است. پس با خودش میگه اگر من تصمیم s_L را بگیرم، پیرو این را خواهد دید و بر اساس این تصمیم خواهد گرفت و قاعدتاً چون پیرو rational است، بهترین تصمیمش را خواهد گرفت با توجه به تصمیمی که از من می‌بینه.

پس من چه تصمیمی بگیرم؟



L ← : پس من چه تصمیمی بگیرم؟

من باید از بین همه‌ی تصمیم‌هایی که می‌توانم بگیرم در S_L، باید تصمیمی را انتخاب کنم که با توجه به reaction پیرو نسبت به این تصمیم (که من می‌توانم تابع این reaction را حساب کنم)، در مجموع سود من را Max کند!

F ← چه کار خواهد کرد؟

پیرو بهترین کار را خواهد کرد ←

$$s_F^*(s_L) = \underset{s_F}{\text{Argmax}}\ J_F(s_F, s_L)$$

از آنجایی که بازی با اطلاعات کامل است، رهبر می‌تواند $s_F^*(s_L)$ را محاسبه کند و برای تابع هدف خودش بکار ببره:

$J_L \to J_L(s_F^*(s_L), s_L) \to s_L$: تنها متغیر

$$\boxed{\begin{aligned} s_L^* &= \underset{s_L}{\text{Argmax}}\ J_L(s_F^*(s_L), s_L) \\ s_F^* &= s_F^*(s_L^*) \end{aligned}}$$

روند تحلیل

1. $s_F^*(s_L) = \underset{s_F}{\text{Argmax}}\ J_F(s_F, s_L)$
2. $J_L \to J_L(s_F^*(s_L), s_L)$
3. $s_L^* = \underset{s_L}{\text{Argmax}}\ J_L(s_F^*(s_L), s_L)$
4. $s_F^* = s_F^*(s_L^*)$

مفهوم تعادل نش برقرار است.

← به تعادل آن Stackelberg گفته می‌شود!

مثال های leader-follower

← مثال 1 :

$P_2 \rightarrow F$

$P_1 \downarrow L$

	y_1	y_2	y_3
x_1	6,4	9,4	6,3
x_2	7,10	6,9	3,7
x_3	9,8	5,5	2,8

اعداد، هزینه هستند.

روند تفکر L : اگر من x_1 را انتخاب کنم، F ← y_3

اگر من x_2 را انتخاب کنم : F ← y_3

اگر من x_3 را انتخاب کنم : F ← y_2

← کدام برای L بهتر است؟

x_2 . چون هزینه‌ی کمتری دارد.

که در ادامه‌ی آن P_2 یا F ، y_3 را انتخاب می‌کند.

(x_2, y_3) stackelberg E.q.

← حال اگر P_1 ← F و P_2 ← L باشند :

روند تفکر L :

اگر من y_1 را انتخاب کنم، F ← x_1 ← هزینه L : 4

اگر من y_2 را انتخاب کنم : F ← x_3 ← هزینه L : 5

اگر من y_3 را انتخاب کنم : F ← x_3 ← هزینه L : 8

← کدام برای L بهتر است؟

y_1 . چون هزینه‌ی کمتری دارد.

که در ادامه‌ی آن P_1 یا F ، x_1 را انتخاب می‌کند.

(x_1, y_1) stackelberg E.q.

← در حالت بازی همزمان (استراتژیک) ؟

وقتی P_1 رهبر بود، BR برای P_2 را حساب کردیم. استراتژی (x_3, y_2) در هر دو حالت جزء گزینه‌های بررسی روند تفکر رهبر قرار گرفت. یک بار $BR_2(P_1)$، یک بار دیگر هم $BR_1(P_2)$ وقتی که P_2 رهبر بود. در واقع استراتژی (x_3, y_2) با هزینه‌ی (5,5)، تقاطع BR هاست. اگر فرم بازی را استراتژیک در نظر می‌گرفتیم، (5,5) تعادل نش بود.

* در بازی Stackelberg، سود رهبر قطعاً بزرگتر - مساوی سود همان بازیگر در تعادل نش است.

همواره گزینه‌ی نش (تقاطع BR ها) برای رهبر وجود دارد. رهبر نگاه می‌کند اگر این برایش بهترین گزینه بود، آن را انتخاب می‌کند.

← مثال 2 : بازی COURNUT

مجموعه‌ی تولیدکنندگان : $I = \{1, 2\}$

مقدار تولید یا استراتژی : $q_1, q_2 \in [0, \infty)$

تابع سود . $\pi_i = \underbrace{q_i}_{\text{درآمد}}\overbrace{(P - C_i)}^{\text{هزینه}}$,

• تابع هزینه : $C_i(q_i) = c q_i$

تابع قیمت : $P = \alpha - \beta(q_1 + q_2)$

فرض می‌کنیم P_1 ← L و P_2 ← F

رهبر، reaction func ، پیرو را محاسبه می‌کند.

$$\frac{\partial \pi_2}{\partial q_2} = -2\beta q_2 + (\alpha - \beta q_1 - C) = 0 \rightarrow q_2 = \frac{\alpha - \beta q_1 - C}{2\beta}$$

BR بازیگر 2 در پاسخ به بازیگر 1

← ادامه‌ی مثال قبل

$$P=\alpha-\beta(q_1+q_2) \xrightarrow[q_2]{\text{جایگذاری}} \alpha-\beta q_1-\beta\left(\frac{\alpha-\beta q_1-c}{2\beta}\right)$$

$$\Rightarrow \pi_1 = q_1\left(\alpha-\beta q_1-\beta\left(\frac{\alpha-\beta q_1-c}{2\beta}\right)-c\right)$$

$$\rightarrow \frac{\partial \pi_1}{\partial q_1}=0 \rightarrow \alpha-2\beta q_1+\beta q_1-c=\frac{\alpha-c}{2}$$

$$\Rightarrow \boxed{q_1^*=\frac{\alpha-c}{2\beta}}$$

با جایگذاری q_1^* در رابطه‌ی q_2 داریم:

$$q_2=\frac{\alpha-\beta q_1-c}{2\beta} \xrightarrow{q_1^*} \boxed{q_2^*=\frac{\alpha-c}{4\beta}}$$

Stackelberg E.q ← $q_2^*=\frac{\alpha-c}{4\beta}$, $q_1^*=\frac{\alpha-c}{2\beta}$

NASH E.q ← $q_1^*=q_2^*=\frac{\alpha-c}{3\beta}$

- سود بازیکن 1 در حالت رهبر: $\pi_1=\frac{5}{16}\frac{(\alpha-c)^2}{\beta}$
- سود بازیکن 2 در حالت پیرو: $\pi_2=\frac{5}{32}\frac{(\alpha-c)^2}{\beta}$
- سود هر بازیکن در تعادل نش: $u_{1,2}=\frac{2}{9}\frac{(\alpha-c)^2}{\beta}$

$$\pi_1>u_1 \quad , \quad \pi_2<u_2$$

* در بازی Stackelberg، سود رهبر "قطعاً" بزرگتر - مساوی سود همان بازیکن در تعادل نش است.

← مثال 3:

اگر سکه‌ی P_1 و P_2 متفاوت رو شود، P_1 برنده و اگر مشابه رو شود، P_2 برنده‌ی بازی است.

		P_2: H	P_2: T
P_1	H	-1 , 1	1 , -1
	T	1 , -1	-1 , 1

همانطور که در فصل 1 بررسی کردیم، این بازی تعادل نش خالص ندارد!

فرض می‌کنیم P_1 ← L و P_2 ← F

روند تفکر L یا P_1:

اگر من H را انتخاب کنم، F ← H ← سود L : -1

اگر من T را انتخاب کنم، F ← T ← سود L : -1

پس این بازی دو تعادل stackelberg دارد.

(H,H) , (T,T)

در این حالت سود رهبر، -1 است.

این مثال را با استراتژی مخلوط نیز بررسی کردیم. سود بازیکن 1 در آن حالت، صفر شده بود.

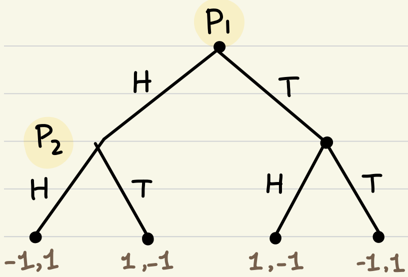
سوال: مگر ادعا نکرده بودیم سود رهبر در تعادل stackelberg بیش‌تر از سود همان بازیکن در تعادل نش است؟

نکته‌ی مهم این است که تعادل نش مخلوط در این ادعا درست نیست. زیرا در نش مخلوط ما expected utility را محاسبه می‌کنیم که نش مخلوط را غیرقابل مقایسه می‌کند.

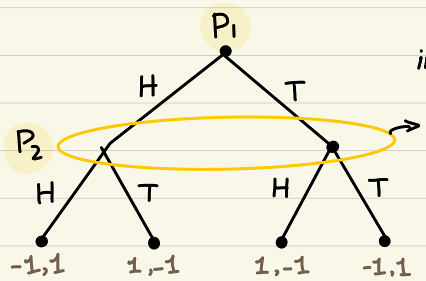
* در بازی Stackelberg، سود رهبر "قطعاً" بزرگتر - مساوی سود همان بازیکن در تعادل نش خالص است.

نمایش درختی

نمایش معتبر: ترتیب داره



سوال: آیا با این درخت می‌توان بازی استراتژیک را نمایش داد؟



information set

این خط نشان می‌دهد که P1 نمی‌داند که در کدام یک از این node ها قرار دارد.

P2 باید بدون اطلاع از تصمیم P1، تصمیم بگیرد. بازی استراتژیک

⇐ فرقی نداره P1 یا P2 بالاتر باشند.

سوال: برای نمایش درختی نیز نمایش ماتریسی وجود دارد؟

در نمایش ماتریسی leader-follower ظاهر نیست. P2 یک استراتژی دارد، P1 یک تصمیم دارد که H یا T را باید انتخاب کند اما P2 قبل از اینکه تصمیم P1 معلوم شود، استراتژی مشخص دارد. P2 در این حالت 4 تا استراتژی دارد: HH، HT، TH، TT

این گسترش کمک می‌کند که بتوانیم ماتریسی نمایش دهیم.

P1 \ P2	HH	HT	TH	TT
H	-1,1	-1,1	1,-1	1,-1
T	1,-1	-1,1	1,-1	-1,1

چون اطلاعات P1 از P2 را لحاظ کردیم می‌توان تعادل نش را محاسبه کرد که حالات کردیم.

نکات مثبت نمایش درختی

⇐ اطلاع داشتن یا نداشتن بازیکنان از یکدیگر را می‌توان نشان داد.

همانطور که در ستون قبل توضیح دادیم، بازی‌ها با فرم گسترده یا استراتژیک را می‌توان نمایش داد.

⇐ اگر تعداد بازیکنان بیش‌تر شود نیز نمایش درختی بهتر است.

در این حالت ابعاد نمایش ماتریسی بیش‌تر شده و مثلاً ماتریس 3 بازیکنه، مکعب خواهد شد.

⇐ در نمایش درختی، تکرار در تصمیم‌گیری را می‌توان لحاظ کرد.

کافی است از node های انتهایی درخت را ادامه دهیم.

تعریف. نمایش درختی

نمایش درختی یک بازی روشی برای نمایش تعاملات تعدادی بازیگر است که در تعدادی مرحله تصمیم‌گیری می‌کنند؛ بطوریکه در هر مرحله مجموعه‌ای از استراتژی‌ها برای بازیگران موجود است.

علاوه بر این، اطلاعات در دسترس برای تصمیم‌گیری در هر مرحله در مدل درختی قابل نمایش است.

تعریف. زیر بازی - sub game

زیر بازی G' از بازی گسترده‌ی درختی G شامل یک node و همه‌ی زیرشاخه‌های آن است؛ بطوریکه:

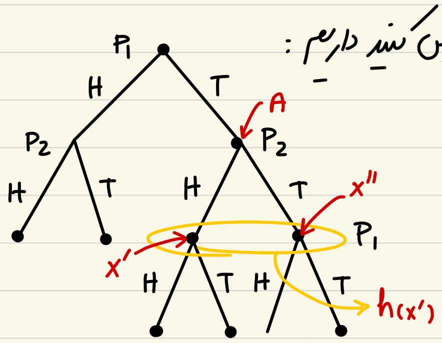
$$\text{if } x'' \in h(x') \ \& \ x' \in V_{G'} \text{ then } x'' \in V_{G'}.$$

V_G : مجموعه‌ی node های بازی درختی G

$V_{G'}$: مجموعه‌ی node های بازی درختی G'

$h(x')$: مجموعه‌ی node هایی که با node، x' در یک بسته‌ی اطلاعاتی قرار دارند.

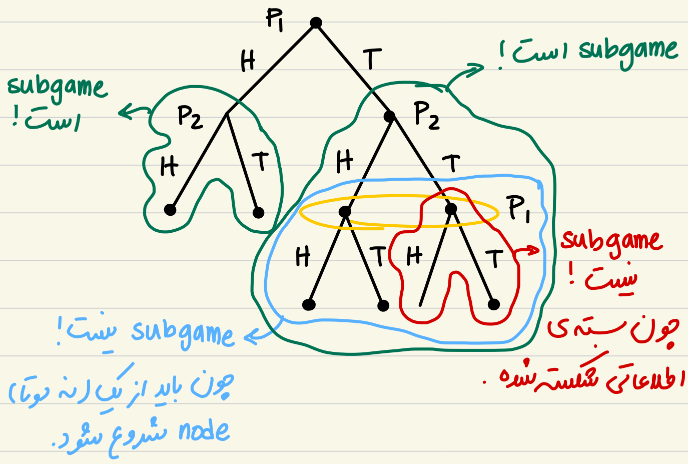
فرض کنید داریم:



← node هایی که در دسته‌ی اطلاعاتی مشابه خط زرد وجود ندارند، در واقع با خودشان در یک دسته‌ی اطلاعاتی قرار می‌گیرند.

← زیربازی شامل یک node مثل A و همه‌ی زیرشاخه‌های آن بطوریکه اگر یک node از این زیربازی در نظر گرفتیم، تمام node‌ها که در دسته‌ی اطلاعاتی آن node هستند، عضو آن باشند.

ما حق شکاندن دسته‌ی اطلاعاتی را در زیر درخت نداریم.



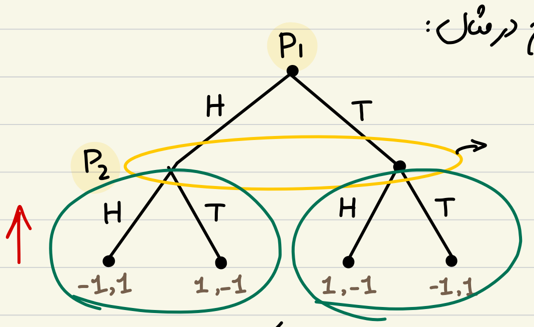
یافتن تعادل نش از طریق نمایش درختی

نحوه‌ی حل بازی stackelberg، در یک درخت، از پایین به بالا حل می‌کنیم. چون رهبر ابتدا BR پیرو را نگاه می‌کند، اول BR پیرو محاسبه می‌شود و بعد به ازای BR پیرو، بهترین تصمیم رهبر محاسبه می‌شود. پس یک اصل داریم:

اصل Bottom-up principle / Backward Induction

مسأله‌ای را که بصورت زنجیره‌ای وجود دارد و چندمرحله‌ای است را باید از انتها به ابتدا حل کنیم تا به جواب بهینه برسیم.

← توضیح در مثال:



این مثال را از پایین به بالا حل کردیم: اگر P1، H را انتخاب کند، اگر P2، H را انتخاب کند سود P2، 1 و اگر P2، T را انتخاب کند سودش -1 خواهد شد. پس H را انتخاب می‌کند. همین تحلیل از پایین به بالا است.

برای این تحلیل کوچکترین جزئی که احتیاج دارم از پایین به بالا، از درخت جدا کنم ← subgame است.

وقتی پایین‌تر اطلاعاتی وجود دارد که بالا وجود ندارد، پس باید دید که مرحله‌ی آخر چه می‌شود؟ وقتی مرحله‌ی آخر بهینه باشد، مرحله‌ی قبل آن با توجه به بهینگی مرحله‌ی آخر تعیین می‌شود و مرحله قبل از یکی مانده به آخر نیز با توجه به بهینگی مرحله‌ی یکی مانده به آخر تعیین می‌شود و تا آخر ... ← Dynamic Programming

پس از تعریف چند مفهوم می‌توانیم تعادل نش برای درخت را بیان کنیم:

تعریف. subgame perfect equilibrium (SPE)

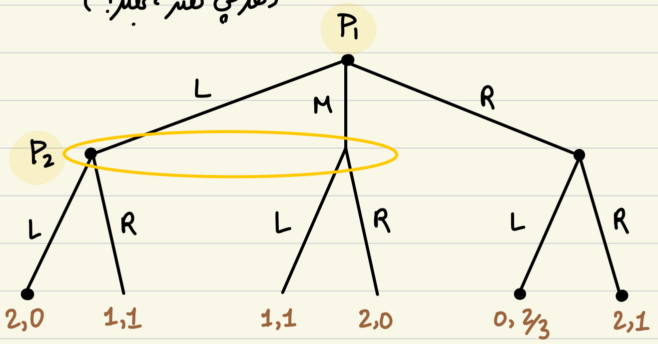
استراتژی s^* یک SPE برای بازی G است، اگر برای هر زیربازی G' از G، $s^*|_{G'}$ یک تعادل نش برای بازی G' باشد.

$s^*|_{G'} \to$ تصویر استراتژی s^* متعلق به بازی G بر بازی G'

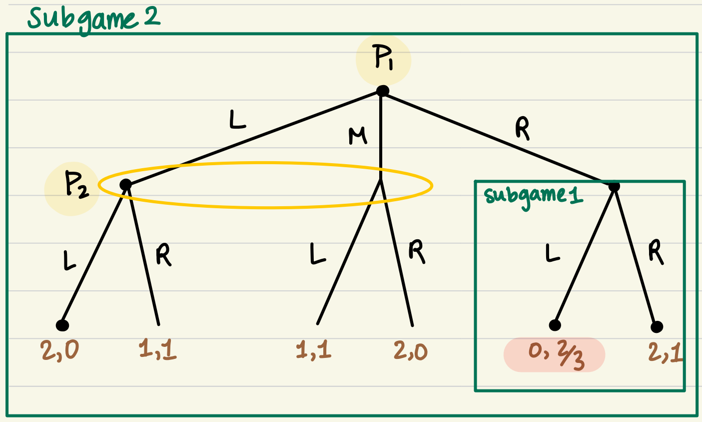
مثال های بازی درختی

← مثال 1: یافتن SPE

* اعداد، هزینه هستند (هرچی کمتر، بهتر!)



← در گام اول subgame هارا مشخص می‌کنیم.



← subgame 1: P_1 تصمیم R را گرفته است. P_2 اگر تصمیم L رابگیرد Cost $\frac{2}{3}$ و اگر تصمیم R رابگیرد Cost 1 می‌شود.

پس P_2 تصمیم L را خواهد گرفت. Cost P_1 ← 0

← subgame 2: اگر P_1 تصمیم R را نگیرد، وارد یک بازی ماتریسی با P_2 خواهد شد:

		P_2: L	P_2: R
P_1	L	2,0	1,1
	M	1,1	2,0

→ تعادل نش خالص ندارد!

← پس به دنبال نش مخلوط می‌گردیم:

$$u_1(L, q^*) = u_1(M, q^*)$$
$$2q^* + (1-q^*) = q^* + 2(1-q^*)$$
$$\Rightarrow q^* = 0.5$$

$$u_2(L, p^*) = u_2(R, p^*)$$
$$0 \times p^* + 1(1-p^*) = 1 \cdot p^* + 0$$
$$\Rightarrow p^* = 0.5$$

$$\Rightarrow \delta^{*T} = \left(\left(\tfrac{1}{2}, \tfrac{1}{2}\right), \left(\tfrac{1}{2}, \tfrac{1}{2}\right)\right)$$

Mixed Nash E.q.

$$\begin{cases} \text{Cost } P_1 = u_1(L \text{ یا } M, q^*) = q^* + 1 = 1.5 \\ \text{Cost } P_2 = u_2(p^*, L \text{ یا } R) = p^* = 0.5 \end{cases}$$

← پس P_1، R را انتخاب می‌کند یا بین L و M تصمیم می‌گیرد؟

اگر R: Cost $P_1 = 0$

اگر بازی ماتریسی M و L: Cost $P_1 = 1.5$

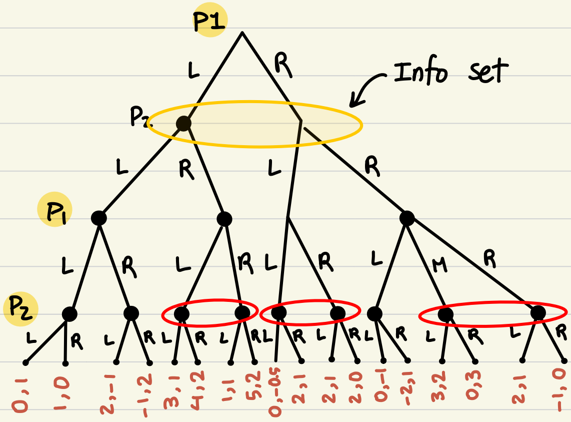
⟸ پس R را انتخاب می‌کند.

SPE: (R, L)

مثال های بازی های درختی

← مثال 2 .

*اعداد، هزینه هستند
(هرچی کمتر، بهتر!)



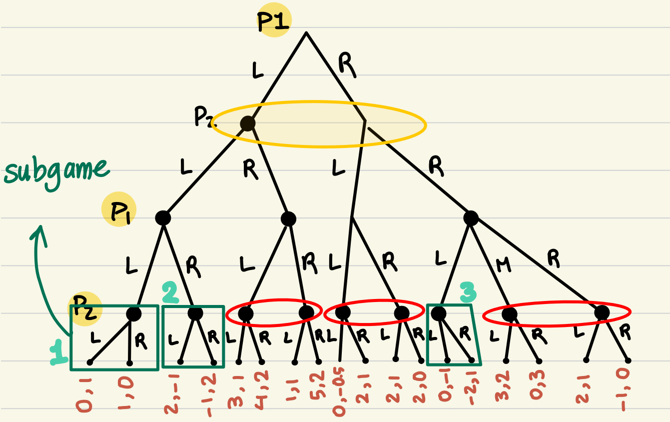
↲ مثال شامل 2 بازیگر ← هر بازیگر در 2 مرحله تصمیم می‌گیرد.
در مرحله‌ی اول (first stage)، بازیگران بدون اطلاع از
تصمیم یکدیگر (همزمان) تصمیم گیری می‌کنند. (actions: L,R)
در مرحله‌ی دوم (second stage)، $P_{1,2}$ کاملاً آگاهی خواهند داشت

* یادآوری. دو مفهوم subgame و backward induction

ما باید اول subgame ها را تشخیص دهیم و
از پایین به بالا براساس اصل backward induction،
تعادل های ش subgame ها را بدست آورده و کم کم
کل درخت را تحلیل می‌کنیم.

پیاده سازی در این مثال:

• subgame ها را پیدا می‌کنیم (در مرحله‌ی آخر تصمیم گیری)

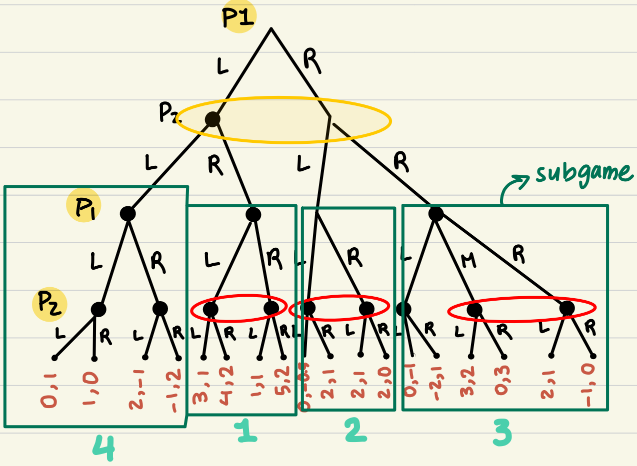


← در subgame 1 . P_2 اگر L را انتخاب کند: هزینه‌اش = 1
P_2 اگر R را انتخاب کند: هزینه‌اش = 0
⇐ پس R را انتخاب می‌کند
NASH

← در subgame 2 . P_2 اگر L را انتخاب کند: هزینه‌اش = -1
P_2 اگر R را انتخاب کند: هزینه‌اش = 2
⇐ پس L را انتخاب می‌کند
NASH

← در subgame 3 . P_2 اگر L را انتخاب کند: هزینه‌اش = -1
P_2 اگر R را انتخاب کند: هزینه‌اش = 1
⇐ پس L را انتخاب می‌کند
NASH

← سراغ subgame های مرحله‌ی آخر اما بزرگتر می‌رویم:



← بازی ماتریسی subgame 1 :

P_1 \ P_2	L	R
L	3,1	4,2
R	1,1	5,2

M L
(1,1) : یک تعادل ش خالص

← بازی ماتریسی subgame 2 :

P_1 \ P_2	L	R
L	$0, -\frac{1}{2}$	2,1
R	2,1	2,0

دو تعادل ش . (L,L), (R,R)

در یک استراتژی مغلوب ضعیف
برای P1 رخ می‌دهد. استراتژی R
نسبت به L برای P1 مغلوب ضعیف است.

* در بازی های 2×2 اگر دو تعادل نش خالص باشد، معمولاً نش مخلوط نیز داریم.

پس در بازی ماتریسی subgame **2**

		P_2: $L \to q$	$R \to 1-q$
P_1	$L \to p$	$0, -\frac{1}{2}$	$2, 1$
	$R \to 1-p$	$2, 1$	$2, 0$

← از روش سوم (تساوی حاصل از نامساوی نش) استفاده کنیم،

فرض کنیم استراتژی ها با احتمال غیرصفر ظاهر می‌شوند،

پس استراتژی L و R برای P1 باید سود یکسانی دهد.

$$EU_1(L,(q,1-q)) = 0\times q + 2\times(1-q) = 2(1-q)$$

$$EU_1(R,(q,1-q)) = 2\times q + 2(1-q) = 2q + 2(1-q)$$

طبق روش سوم: $\rightsquigarrow EU_1(L) = EU_1(R)$

$\boxed{q=0} \to$ تناقض با فرض اولیه

$$EU_2((p,1-p),L) = -\frac{1}{2}\times p + 1(1-p)$$

$$EU_2((p,1-p),R) = 1\times p + 0$$

طبق روش سوم. $\rightsquigarrow EU_2(L) = EU_2(R)$

$\boxed{p = \frac{2}{5}}$

نتیجه قابل اعتماد نیست! پس از تقاطع BR ها استفاده می‌کنیم که خیالمون راحت باشه!

ادامه در ستون بعدی

$$\to EU_1 = 0\cdot pq + 2\,p(1-q) + 2(1-p)q + 2(1-p)(1-q)$$

$= p(-2q) + \sim \to$ در بهینه سازی نقش ندارد! ← چون یا عدد ثابته یا تابعی از q.

Minimize (چون تابع هزینه)

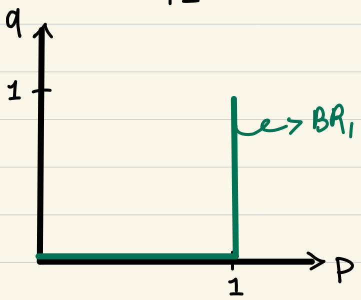
• اگر $q>0 \leftarrow -2q<0 \leftarrow$ پس P باید بیشترین مقدار خودش یعنی 1 باشد تا هزینه کمینه شود.

$\Rightarrow \forall q>0 \rightsquigarrow p=1$

• اگر $q=0 \leftarrow$ P فرقی نمی‌کند چه باشد:

$\Rightarrow q=0 \rightsquigarrow p =$ Don't care

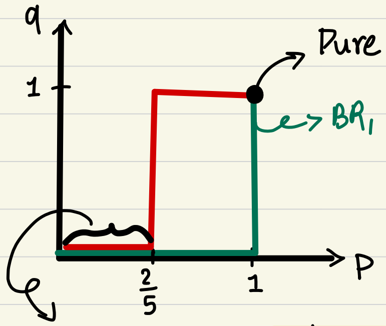
پس BR_1 را رسم می‌کنیم:



$$\to EU_2 = -\frac{1}{2}\cdot pq + 1\cdot p(1-q) + 1(1-p)q + 0(1-p)(1-q)$$

$$= q\left(-\frac{5}{2}p + 1\right) + \sim$$

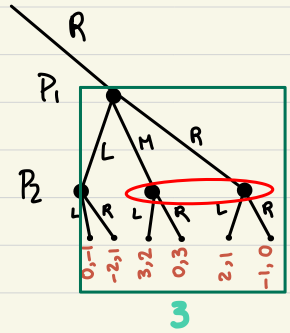
- if $-\frac{5}{2}p + 1 < 0 \rightsquigarrow q = 1$
- if $-\frac{5}{2}p + 1 = 0 \rightsquigarrow q =$ Don't care
- if $-\frac{5}{2}p + 1 > 0 \rightsquigarrow q = 0$



Pure Nash: $p_1^* = q_1^* = 1$

Mixed Nash: $p_1^* \in (0, \frac{2}{5}]$, $q_1^* = 0$

← بازی ماتریسی 3 subgame :



بخش سمت چپ تحلیل شد. (خودش subgame بود.)

بخش سمت راست :

		P_2 L	R
P_1	M	3,2	0,3
	R	2,1	-1,0

یک تعادل نش خالص

← اگر P1، تصمیم L را نگیرد، وارد بازی ماتریسی بالا با P2 خواهد شد که تعادل نش خالص آن (-1,0) است و هزینه‌ی P1 در این صورت (-1) خواهد شد؛ اگر P1 تصمیم L را بگیرد، همانطور که قبلاً گفتیم برای برقراری نش، P2، L را انتخاب کرده و هزینه‌ی P1، 0 خواهد شد. هزینه‌ی P1 در حالت -1 کمتر است که (R,R) این را نتیجه می‌دهد.

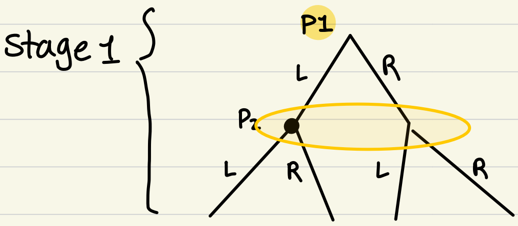
Subgame Perfect Equilibrium = $(\overset{R}{-1},\overset{R}{0})$

← بازی ماتریسی 4 subgame :

subgame های زیرین این زیربازی را بررسی کردیم. گفتیم اگر P1، L را انتخاب کند، P2، R را انتخاب و اگر P1، R را انتخاب کند، P2، L را انتخاب خواهد کرد. که هزینه‌ی آنها، به ترتیب برای P1، 1 و 2 خواهد شد. پس P1، L را انتخاب می‌کند که هزینه‌ی کمتری دارد.

Subgame Perfect Equilibrium = $(\overset{L}{1},\overset{R}{0})$

← حال اطلاعاتی که از stage 2 بدست آوردیم را در بررسی stage 1 بکار می‌بریم.



طبق 4 subgame، بررسی شده در نهایت، payoff ها را می‌نویسیم:

- اگر P1 ← L و P2 ← L را بازی کنند، وارد subgame 4 می‌شویم که گفتیم نش نهایی آن (1,0) خواهد بود.
- اگر P1 ← L و P2 ← R را بازی کنند، وارد subgame 1 می‌شویم که گفتیم نش نهایی آن (1,1) خواهد بود.
- اگر P1 ← L و P2 ← R را بازی کنند، وارد subgame 2 می‌شویم که گفتیم 3 نش پایینی از بازی داشت. که یک نش قوی‌تر داشت، $(0,-\frac{1}{2})$ تعادل نش خالص داشت که ناشی از استراتژی غالب بازیگر 1 بود. نش بعدی، یک استراتژی مغلوب ضعیف بازیگر 1 حاصل شد، (2,0)، یک تعادل نش مخلوط داشت که با تقاطع BR ها بدست آمد. در این 3 نش باید payoff هر بازیگر را بدست آوریم.

نش های خالص: $(0,-\frac{1}{2})$ و $(2,0)$

نش مخلوط: payoff مربوط را محاسبه کردیم.

$\Rightarrow P_1^* \in (0,\frac{2}{5}]$ $q_1^*=0$ R ← P2

در این صورت payoff، P1، 2 خواهد شد.

برای محاسبه‌ی payoff بازیگر 2 :

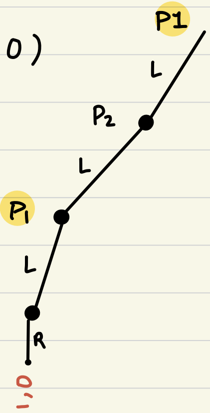
		P_2 L	R
P_1	L	$0,-\frac{1}{2}$	2,1
	R	2,1	2,0

$U_2 = 1\times P_1 + 0$

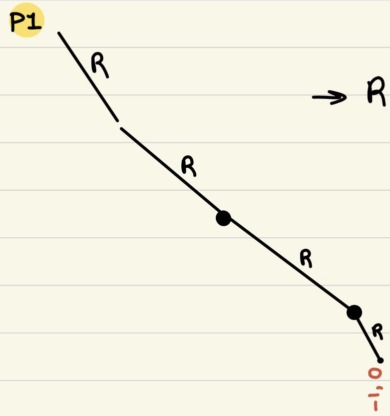
$\rightarrow U_2 \in (0,\frac{2}{5}]$

پس در این حالت 2 تعادل نش خالص داریم.

→ LLLR : (1,0)



→ RRRR : (-1,0)



← چون دو تعادل نش خالص داریم، می‌توان تعادل نش مخلوط داشت:

		P_2: L (q)	R ($1-q$)
P_1	L (p)	1,0	1,1
	R ($1-p$)	2,K	-1,0

* با استفاده از تساوی حاصل از نامساوی نش:

$$\begin{cases} EU_1(L,(q,1-q)) = 1\times q + 1\times(1-q) = 1 \\ EU_1(R,(q,1-q)) = 2\times q + (-1)\times(1-q) = 3q-1 \end{cases}$$

$$\Rightarrow EU_1(L) = EU_1(R) \rightarrow 1 = 3q-1$$

$$\Rightarrow q^* = 2/3$$

$$\begin{cases} EU_2((p,1-p),L) = 0\times p + K\times(1-p) = K-Kp \\ EU_2((p,1-p),R) = 1\times p + 0(1-p) = p \end{cases}$$

$$\Rightarrow p^* = \frac{K}{K+1} \xrightarrow{K\neq 0} p^* = \frac{K}{K+1}\,,\ q^* = \frac{2}{3}$$

(subgame2) ← Mixed N.E.

پس payoff بازیکن 2، بین 0 تا $\frac{2}{5}$ خواهد شد ← عدد K

$\in [0, \frac{2}{5}]$

(2,0) Pure NE

		P_2: L	R
P_1	L	1,0	1,1
	R	2,K	-1,0

→ if $K \in (0, \frac{2}{5})$. نش مخلوط خواهیم داشت

• اگر P1 ← L و P2 ← R را بازی کنند، وارد subgame 3 می‌شویم که گفتیم نش نهایی آن (-1,0) خواهد شد.

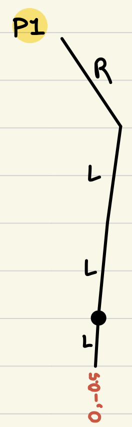
← پس برای تحلیل stage 1 وارد 2 حالت می‌شویم.

1- تعادل نش خالص بازی وقتی پی‌آف ما از حل subgame 2، برابر $(0,-\frac{1}{2})$ باشد:

Pure NE

		P_2: L	R
P_1	L	1,0	1,1
	R	$0,-\frac{1}{2}$	-1,0

⇒ RLLL



2- تعادل نش خالص بازی وقتی پی‌آف ما از حل subgame 2، برابر (2,K) باشد:

		P_2: L	R
P_1	L	1,0	1,1
	R	2,K	-1,0

→ Pure NE

if $k=0 \rightsquigarrow$ Pure NE (in subgame 2)

← از روش تقاطع BR ها:

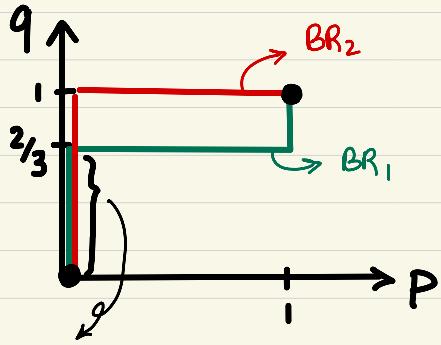
$$E(u_1) = 1(pq) + 1(p)(1-q) + 2(1-p)q + (-1)(1-p)(1-q)$$

$$= (-3q+2)p + \sim$$

$$\begin{cases} q<\frac{2}{3} \rightarrow p^*=0 \\ q=\frac{2}{3} \rightarrow \text{Don't care} \\ q>\frac{2}{3} \rightarrow p^*=1 \end{cases}$$

$$E(u_2) = p(1-q) = (-p)q + p$$

$$\begin{cases} p<0 \rightarrow \text{Not Possible} \\ p=0 \rightarrow \text{Don't care} \\ p>0 \rightarrow q^*=1 \end{cases}$$



Mixed NE : $\begin{cases} p^*=0 \\ q^* \in (0, \frac{2}{3}] \end{cases}$

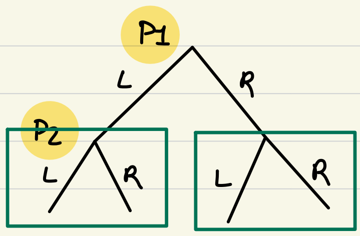
قضیه روش استقرای معکوس یا Backward Induction، همه‌ی SPE های (Subgame Perfect Equilibrium) بازی گسترده را به دست می‌دهد.

قضیه هر بازی محدود (finite) با اطلاعات کامل، یک استراتژی SPE خالص دارد.

* همه‌ی information set ها، تک node باشند. (یک عضو داشته باشند.)

* منظور از اطلاعات کامل، اطلاع از تصمیم بازیگرانی‌ست که در مرتبه‌ی بالاتر تصمیم می‌گیرند.

مثال ←



← مسأله‌ی بهینه‌سازی ست و بحث تصمیم‌گیری همزمان نیست؛ چون P1 قبلاً تصمیمش رو گرفته. P2 باید بین دو گزینه انتخاب کند، یا هردو برایش بهینه هستند یا یکی. پس هر دو subgame مشخص شده، حل بهینه دارد.
P1 هم با توجه به حل این دو subgame عمل می‌کند یا L و R برایش مساوی است یا یک کدام ممتر است.

← وقتی بازی بصورت سلسله مراتبی حل می‌شود، حتماً تعادل نش خالص دارد چون در آخرین level، نش برای آخرین subgame وجود دارد و این به level های بالاتر می‌رود. یعنی چند بهینه‌سازی زنجیره وار انجام می‌شود.

قضیه هر بازی درختی محدود، یک تعادل SPE دارد.
↳ Pure/Mixed

اصل. One Stage Deviation Principle

در یک بازی با مراحل محدود، استراتژی s^* برای زیربازی G یک SPE است، اگر هیچ بازیگری نتواند با تغییر استراتژی خود در یک stage، و ادامه آن استراتژی پس از آن به سود بیش‌تری دست یابد. ← چک کردن همه تغییرات ممکن برای SPE، لازم نیست!

⟸ یعنی چی؟

امکان دارد یک بازیگر در 2 stage، در درخت تصمیم‌گیری کند. نیاز نیست تغییر استراتژی آنرا در هر 2 stage چک کنیم همانکه در stage 1، استراتژی‌اش را تغییر دهید و به سود بیش‌تری نرسد، و در stage 2 هم استراتژی‌اش را تغییر دهید و به سود بیش‌تری نرسد، در این صورت این تعادل، تعادل نش است.

«یک stage کفایت می‌کند»

شرط لازم است.

اگر s^* یک SPE برای بازی G باشد، اگر هر بازیگر استراتژی‌اش را عوض کند (با ثابت ماندن استراتژی سایر بازیگران)، این بازیگر نباید به سود بیش‌تری برسد. که این همان مفهوم تعادل نش است.

شرط کافی است.

اگر در یک درخت در آخرین لایه بررسی کنیم، در این لایه یک سری SPE پیدا می‌شود. اگر در stage آخر، کسی از استراتژی‌اش deviate کند، چون BR است، ضرر می‌کند. حل مراحل پایین‌تر به عنوان ورودی حل مراحل بالاتر قرار می‌گیرد. در stage های بالایی هم اگر یک بازیگر، استراتژی‌اش را تغییر دهد، به این معناست که BR بازی نکرده نسبت به استراتژی مرحله‌ی پایین‌ترش، بنابراین این باعث ضررش خواهد شد. اگر اول تعادل،

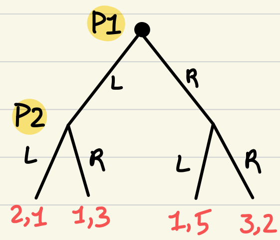
← SPE باشد، اگر از استراتژی اش deviate کند و به سود بیش‌تری نرسد، پس این BR ش بوده نسبت به مراحل پایین‌تر. بنابراین اگر از پایین‌ترین لایه‌ی درخت در نظر بگیریم و به بالا برویم، نشان می‌دهد که این one stage ها دارند تغییرات ممکن را نشان می‌دهند و لزومی ندارد که تغییر 2 stage مختلف را در نظر بگیریم.

از اصل Backward Induction یا بهینگی Dynamic Programming حاصل می‌شود.

* نکته: finite game با finite horizon game ؟
finite game: استراتژی ها محدود هستند.
finite horizon game: در این game چند stage محدود باشد.
کاربرد این اصل

مثال های OSDP

← مثال 1.



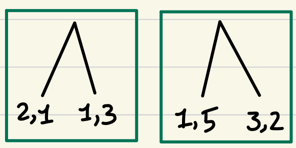
* اعداد هزینه هستند

← اطلاعات کامل P2 از تقسیم P1.

* نکته. یعنی مواقع اطلاعات کامل منظور اطلاع از تابع هدف بازیگران است که payoff ها معلوم است وقتی داریم « اطلاع از تقسیم بازیگر 1 » ، یعنی همین درخت بالایی»

← درست آوردن SPE :

2 تا subgame مشخص شده را جدا می‌کنیم.



1 2

← 1. (2,1), (1,3)

بین این دو هزینه‌ی عدد 1 برای P2 کمتر است پس L را انتخاب کرده و (2,1) انتخاب می‌شود.

هزینه‌ی P1 = 2

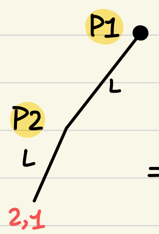
← 2. (1,5), (3,2)

بین این دو هزینه‌ی عدد 2 برای P2 کمتر است پس R را انتخاب کرده و (3,2) انتخاب می‌شود.

هزینه‌ی P1 = 3

⇐ پس P1 اگر وارد زیربازی 1 شود، هزینه‌ی کمتری می‌دهد.

P1، L را انتخاب می‌کند.



تعادل بازی : => $(\overset{L}{2}, \overset{L}{1})$

⇐ OSDP Verification :

← برای P2، اگر از L به R تغییر استراتژی دهد، هزینه‌اش 3 خواهد شد ← ضرر خواهد کرد

← برای P1، با ثابت ماندن P2، اگر P1 استراتژی اش را از L به R تغییر دهد، آیا payoff، P1 را در حالت (R,L) که 1 می‌شود بررسی می‌کنیم؟ NO. این طرز بررسی اشتباه است چون استراتژی بازیگر 2 را لحاظ نکردیم. در (R,L)، L، action است. ما باید استراتژی P2 را ثابت نگه داریم و استراتژی P1 را تغییر دهیم. استراتژی P2 چیست؟ که اگر P1، L بازی کند، P2، L بازی می‌کند. اگر P1، R بازی کند، P2، R بازی می‌کند. پس استراتژی P2، LR است.

پس اگر P1 استراتژی اش را از L به R تغییر دهد و P2 استراتژی اش که LR است را ثابت نگه دارد، payoff بازیگران درخت (3,2) قرار گرفته و P1 در این صورت ضرر می‌کند.
action (R,R)

One stage deviation principle for infinite horizon game

← برای بازی‌های finite horizon توضیح دادیم. چرا finite، چرا infinite را نکردیم؟

اصل Backward Induction برای حالتی است که تعداد مراحل بازی محدود است. چون باید از آخرین مرحله، به مراحل بالاتر پیش رویم. وقتی یک بازی انتها نداشته باشد نمی‌توان این اصل را استفاده کرد و آن‌را در اثبات OSDP بکار برد.

از طرف دیگر بازی‌های infinite کجا بکار برده می‌شوند؟

کاربرد این بازی‌ها در بازی‌های repeated یا تکراری است؛ که نمی‌دانیم تعداد مراحل چندتاست یا خیلی زیاد است.

در چه صورت می‌توان از این اصل در این بازی‌ها استفاده کرد؟

شرط پیوستگی در بی‌نهایت ← Continuous at infinity

فرض کنید بازی فرم گسترده که در بی‌نهایت مرحله انجام می‌شود را با G^∞ نمایش دهیم. اگر تاریخچه‌ی تصمیم بازیگران تا مرحله t، را به صورت مقابل نشان دهیم: $h^t=(a^0,a^1,\dots,a^{t-1})$، و تاریخچه‌ی کل تصمیمات بازیگران: $h=(a^0,a^1,\dots)$، بطوریکه $a^i=(a_1^i,\dots,a_N^i)$ و N تعداد بازیگران باشد. آنگاه:

بازی G^∞ پیوسته در بی‌نهایت است اگر $\forall i\in I$ با تابع سود u_i داشته باشیم:

$$\lim_{t\to\infty}\ \sup_{\forall h,\tilde h \text{ s.t. } h^t=\tilde h^t} |u_i(h)-u_i(\tilde h)| \to 0$$

مفهوم: اگر این بازی در مراحل زیاد انجام شود، اگر ما تاریخچه‌ی بازی را تا تعداد مراحل زیاد به اندازه کافی نشان دهیم، از آن‌جا به بعد h، $\tilde h$ تغییر می‌کنند (اما تا لحظه‌ی t یکی‌اند) این عوض شدن history از مرحله‌ای به بعد تأثیری در payoff نداشته باشد.

← انگار از یک جایی به بعد تصمیم بازیگران اثری در سودشان ندارد.

انگار بازی را با حالت finite تقریب می‌زند.

دسته‌ای از بازی‌های infinite horizon با شرط پیوستگی در بی‌نهایت.

Repeated Games

← این دسته بصورت زیر تعریف می‌شوند:

$$u_i=\sum_{t=0}^{\infty}(\delta_i)^t\, g_i^t(a^t)$$

- $g_i^t(a^t)$: stage payoff
  $\forall t: \quad |g_i^t(a^t)|\le B$

فرق بازی repeated با درختی

در بازی درختی پس از اینکه تصمیمات گرفته شدند، payoff حساب می‌شد. در آخرین node، payoff معلوم می‌شد. در این نوع بازی‌ها، پس از هر stage (که ممکن است هر stage خود یک بازی درختی باشد)، payoff ها مشخص می‌شود. پس به stage بعدی می‌رویم.

- a^t : $(a_1^t,\dots,a_N^t)$ پروفایل تصمیم بازیگران در مرحله t
- δ_i : <1 . Discount factor

$\delta_i<1$، به توان t می‌رسد:

در یک بازی با تعداد مراحل زیاد، agent به مراحل نزدیک‌تر و پیش رو اهمیت بیش‌تری نسبت به مرحله‌های خیلی دور، قائل است.

δ_i، ارزش reward های مراحل نزدیک را بیش‌تر از دورتر می‌کند.

با به توان رسیدن، ارزش مراحل خیلی دور را نزدیک صفر می‌کند.

$$t\to\infty \Rightarrow (\delta_i)^t\to 0 \Rightarrow u_i\to 0$$

مثال بازی repeated

← مثال 1:

		P2	
		C	D
P1	C	1,1	-1,2
	D	2,-1	0,0

فرض کنید این بازی را در stageهای مختلف انجام می‌دهند و در هر stage، payoff دریافتی‌شان با تابع g_i حساب می‌شود.

در واقع اگر تابع g معرفی شده

* که در این‌جا در تمام مراحل g ثابت است.

$$\rightsquigarrow u_i(a_i, a_{-i}) = \sum_{t=0}^{T} (\delta_i)^t g_i(a_i^t, a_{-i}^t)$$

← این بازی شبیه معمای زندانی است و استراتژی D برای P1، استراتژی غالب است. هم‌چنین برای P2 نیز، D غالب است.

• اگر بازی را استاتیک فرض کنیم و یک بار انجام شود.

(D, D) → تعادل نش غالب

• حال اگر فرض این بازی بخواهد تکرار شود و در چندین مرحله انجام شود:

• تعداد مراحل محدود: (T محدود باشد)

اگر هر دو بازیکن بدانند که این بازی مثلاً 3 بار انجام می‌شود

چون تعداد مراحل محدود است، bottom-up برقرار است.

← بار آخر (سوم): استراتژی غالب هر دو D است و باید D را انتخاب کنند.

← مرحله دوم. payoff ها در دو مرحله مستقل هستند. دوباره (D,D) نش غالب خواهد بود.

← مرحله اول: متعاقباً این مرحله نیز بازیکنان (D,D) برمی‌گزینند

* اگر یک تعادل نش در بازی وجود داشته باشد و بازی به تعداد محدود انجام شود و افراد هم به این محدودیت آگاهی داشته باشند، آنگاه این تعادل نش برای بازی تکراری همان تعادل نش در یک مرحله است که به تعداد مراحل تکرار خواهد شد. ← براساس استقرای معکوس

⇐ پس (D,D)، SPE یکتای بازی در T مرحله است.

قضیه. اگر بازی G دارای تعادل نش با پروفایل a^* باشد، آنگاه $G^T(\delta)$، یعنی همان بازی با T بار تکرار و ضریب DF (Discount Factor) برابر δ، دارای SPE بصورت زیر می‌باشد:

$G^T(\delta)$: SPE $\qquad a^t = a^* \;;\; t = 0, \dots, T$

$$\rightarrow \tilde{a} = (a^1, a^2, \dots, a^T) = (a^*, \dots, a^*)$$

• تعداد مراحل نامحدود یا نامعلوم.

* در این صورت نمی‌توان از روش استقرای معکوس استفاده کرد.

* برای بازی‌های infinite horizon repeated games یک دسته استراتژی معرفی کردند که قابلیت تعادل نش بودن را دارند.

Non-Forgiving Trigger Strategy

یا Grim Trigger Strategy

		P2	
		C	D
P1	C	1,1	-1,2
	D	2,-1	0,0

← بیا را بر دوستی می‌گذاریم اما اگر رقیب اشتباهی کرد باهاش cut! :))

استراتژی s^* را برای هر دو بازیکن بصورت زیر درنظر بگیرید:

s^*
1. C بازی کن اگر تا به حال D بازی نشده است.
2. اگر D مشاهده کردی تا آخر D بازی کن.

← اگر بخواهیم از اصل OSDP استفاده کنیم باید اولاً سود بازیکنان اگر هر دو s^* بازی کنند چه خواهد بود؟ دوماً اگر یک بازیکن از s^* deviate کند و دوباره به این استراتژی برگردد، سود چه خواهد بود؟

• اگر هر دو بازیکن از s^* تبعیت کنند:

در مرحله‌ی اول هر دو C بازی می‌کنند و تا آخر چون D دیده نمی‌شود، C بازی خواهند کرد.

$$u_1 = u_2 = \underbrace{1 + \delta + \delta^2 + \cdots}_{\text{تصاعد هندسی}} = \frac{1}{1-\delta}$$

• فرض کنیم یکی از بازیگران از s^* سرپیچی کند:

فرض کنیم بازیگر 1 در مرحله‌ی 1 از s^*، deviate کند، P2 دارد از s^* تبعیت می‌کند پس P2 دارد C بازی می‌کند، P1 چون از s^* سرپیچی کرده، D بازی می‌کند. پس P1 در مرحله 1، سود 2 بدست می‌آورد. چون One Stage Deviation، داریم. در مرحله‌ی بعد P1 می‌خواهد s^* را بازی کند. s^* می‌گوید اگر D مشاهده کردی تا آخر D بازی کن پس P1 تا آخر D بازی خواهد کرد. P2 نیز طبق s^* چون D مشاهده کرده تا آخر D را بازی می‌کند پس از مرحله 1 هردو D بازی کرده و سودشان صفر خواهد شد.

$\to u_1 = 2+0+0+\dots = 2$

⇐ حال بررسی می‌کنیم، P1 با deviationی که انجام داد، سود بیش‌تری گرفت یا نه؟

~ مقدار δ وابسته می‌شود. اگر بخواهیم s^* نش شود، سود P1 باید در حالت بازی s^* بزرگتر از سودش در حالت سرپیچی از s^* شود:

$$\frac{1}{1-\delta} \geq 2 \to \delta \geq \frac{1}{2}$$

⇐ حال deviation، P1 از s^* را در مرحله‌ی kام می‌نویسیم:

$$u'_1 = 1+\delta+\delta^2+\dots+\delta^{k-1}+2\delta^k+0+0+\dots$$

تفاوت این سود با حالت تبعیت از s^*.

$$u_1=\frac{1}{1-\delta} \Rightarrow u_1-u'_1 = \delta^k(1+\delta+\delta^2+\dots-2)$$
$$= \delta^k\left(\frac{1}{1-\delta}-2\right) \geq 0$$

$\to \delta \geq \frac{1}{2}$

⇐ پس بررسی در مرحله‌ی اول یا kام فرقی نداشت و یک نتیجه را داد.

⇐ پس اگر $\delta \geq \frac{1}{2}$ باشد، s^* می‌تواند تعادل نش باشد.

نکات:

نقش نامحدود بودن بازی را کجا دیدیم؟ چرا اگر بازی محدود بود، s^*، نش نفی شد؟

اگر تعداد بازی محدود بود و P1 در مرحله‌ی آخر تخطی یا نافرمانی می‌کرد، به سود بیش‌تری می‌رسید. پس s^* نمی‌توانست نش باشد.

آیا این s^* که به عنوان Grim Strategy معرفی شد، تنها تعادل نش بازی است؟ یا تعادل نش قبلی که هردو بازیگر همواره D بازی کنند، آیا در این حالت نیز نش است؟

اگر هردو بازیگر D بازی کنند، هرکس در یک مرحله deviate کند و دوباره ~ D بازی کردن برگردد، در همان مرحله سودش کمتر می‌شود و پس در مجموع نیز سودش کمتر خواهد شد. بنابراین (D,D) که تعادل نش غالب بازی یک مرحله ای بود، همچنان برای این بازی نیز نش است. نکته این است که تعادل های نش متنوع تری مثل Trigger Strategy وجود دارد.

ما برای این مثال خاص $\delta \geq 1/2$ بدست آوردیم. آیا تعمیمی برای آن وجود دارد؟

قضیه Friedman.

اگر a^* تعادل نش خالص بازی G با پروفایل سود e^* باشد، برای هر پروفایل سود شدنی v که در آن $\forall i \in I$ $v_i > e_i^*$ باشد، ثابت $\underline{\delta}<1$ وجود دارد بطوریکه $\forall \delta \geq \underline{\delta}$، استراتژی Trigger یک SPE برای بازی $G^\infty(\delta)$ با سود v می‌باشد.

G:
- $I=\{1,\dots,N\} \to$ بازیگران
- $a^*=(a_1^*,\dots,a_N^*) \to$ پروفایل نش
- $e^*=(e_1^*,\dots,e_N^*) \to$ پروفایل سود نش

- $v_i > e_i^* \to$ پروفایل استراتژی دیگری مانند b وجود دارد که در آن سود همه‌ی بازیگرها از سود حاصل از a^* بیش‌تر است.

← مثال 2 : (MIT Lectures 15)

Repetition Can Lead to Bad Outcomes

- The following example shows that repeated play can lead to *worse* outcomes than in the one shot game:

P2

P1

	A	B	C
A	2, 2	2, 1	0, 0
B	1, 2	1, 1	−1, 0
C	0, 0	0, −1	−1, −1

- For the game defined above, the action A strictly dominates B, C for both players, therefore the unique Nash equilibrium of the stage game is (A, A).
- If $\delta \geq 1/2$, this game has an SPE in which (B, B) is played in every period.
- It is supported by a slightly more complicated strategy than grim trigger:

s^* {
I. Play B in every period unless someone deviates, then go to II.
II. Play C. If no one deviates go to I. If someone deviates stay in II.

استراتژی غالب هر دو بازیکن، A است و (2,2) نش خالص بازی.

تعادل دیگر نش بر اساس s^* تعریف شده معرفی می‌شود. پیروی از s^* به این معناست که در تمام مراحل هردو B را بازی کنند که سودشان در این صورت (1,1) خواهد بود ← پس سود کمتری از نش خالص بازی خواهند داشت. می‌خواهیم نشان دهیم چرا این بازی تحت استراتژی s^* یک تعادل نش است؟

برای نشان دادن تعادل نش، سود P1, P2 را تحت s^* حساب می‌کنیم اگر هردو از s^* پیروی کنند، در تمام مراحل مقدار سودشان 1 خواهد شد پس.

$$v_1 = v_2 = 1+\delta+\delta^2+\dots = \frac{1}{1-\delta}$$

حال فرض کنید که P1 در یک مرحله از s^*، deviate کند و دوباره به s^* باز گردد. اگر در یک مرحله deviate کند، پس در یک مرحله به جای B، A یا C را بازی می‌کند. حال بهترین حالت را برایش درنظر می‌گیریم: که deviate کند به A (چون A نسبت به C نفع دارد. پس C، payoff کمتری برایش خواهد داشت).

پس فرض می‌کنیم در مرحله‌ی kام از B به A، deviate کند:

سود P1 قبل از deviate : $u_1 = 1+\delta+\dots+\delta^{k-1}$

← وقتی P1، deviate کند، استراتژی P2 همان s^* است

پس P1 ← A، P2 ← B را بازی می‌کنند در این مرحله داریم:

(2,1)

سود موقع deviate : $u_1 = 1+\delta+\dots+\delta^{k-1}+2\delta^k$

← بلافاصله این، P1 به s^* باز می‌گردد طبق s^* باید خط دوم s^* را اجرا کند یعنی C بازی کند. P2 هم طبق s^*، C بازی می‌کند.

(−1,−1)

سود پس از تخطی : $u_1 = 1+\delta+\dots+\delta^{k-1}+2\delta^k+(-1)\delta^{k+1}$

← هردو C بازی کردند. طبق خط دوم s^* می‌گوید اگر هیچ کدام deviate نکرد به خط اول استراتژی برو. چون الان هردو از s^* پیروی کردند و کسی مجدداً deviate نکرده بنابراین هردو به استراتژی خط اول رفته و دوباره B بازی خواهند کرد.

(2,1)

$$u_1' = 1+\delta+\dots+\delta^{k-1}+2\delta^k-\delta^{k+1}+\delta^{k+2}+\delta^{k+3}+\dots$$

⇐ حال $u_1 = \frac{1}{1-\delta}$ و u_1' که بدست آوردیم را مقایسه می‌کنیم:

برای اینکه s^* نش باشد باید: $u_1 - u_1' = -\delta^k + 2\delta^{k+1} \geq 0$

$$\to \delta^k(2\delta-1) \geq 0 \to \delta \geq 1/2$$

← آیا برگشت به s^* پس از یک بار deviate کردن، BR برای P1 است؟ قاعدتاً اگر s^* نش باشد، جمله‌ی بالا درست است.

حالت اول

حالتی را در نظر می‌گیریم که P1 پس از deviate به s^* باز نگردد. مثلاً P1 به A، deviate کند و همواره A را بازی کند. فرض کنیم که k−1 مرحله‌ی اول را که از s^* پیروی کرده و در مرحله k، A بازی کند:

$$u_1 = 1+\delta+\dots+\delta^{k-1}+2\delta^k$$

از این جا به بعد P2 به خط دوم s^* رفته و C بازی می‌کند. وقتی P1، A بازی کند سودشان طبق ماتریس : (0,0) می‌شود.

⟵ ادامه: پس در نهایت سود P1 را با u'' در این حالت نشان می‌دهیم:

$$u_1'' = 1+\delta+\cdots+\delta^{k-1}+2\delta^{k}+0$$

حال u_1' و u_1'' را مقایسه می‌کنیم: (با شرط $\delta \geq \frac{1}{2}$) ⟵

$$u_1' - u_1'' = -\delta^{k+1} + \delta^{k+2}\left(\frac{1}{1-\delta}\right) \geq 0$$

چون می‌خواهیم بگیم اینکه P1 پس از deviate دوباره به s^* بازگردد، Best Response برای P1 خواهد بود. پس سود بیشتری از حالتی که به s^* باز نگردد کسب می‌کند.

$$\Rightarrow -\delta^{k+1} + \delta^{k+2}\left(\frac{1}{1-\delta}\right) \geq 0$$

$$\Rightarrow \frac{\delta}{1-\delta} - 1 \geq 0 \to \delta \geq 1-\delta \to \delta \geq \frac{1}{2}$$ ✓

مطابق فرض اولیه‌مان بدست آمد. ⟸

پس بازگشت به s^* پس از deviate، بهتر از بازگشت به A است.

حالت دوم

⟸ فرض کنیم P1 به‌جای بازی کردن B، A بازی کند در مرحله‌ی K ام. بعد به مرحله‌ی 2 می‌رود. این جا برای اینکه P2 را تنبیه دهد، در این مرحله C را انجام دهد که طبق خط 2 در s^*، شرط No one deviate برقرار نشود و هر دو بازیکن به خط اول s^* بروند. حال P1 در خط اول s^* دوباره deviate کند و A انجام دهد پس از deviate دوباره به خط دوم s^* می‌روند، P2، C را بازی کند و از این به بعد همواره C بازی خواهد کرد. (به عنوان punishment برای P1). در این صورت بهترین تصمیم برای P1 طبق ماتریس همان A بازی کردن است.

$$u_1''' = \underbrace{1+\delta+\cdots+\delta^{k-1}}_{(B,B)} + \underbrace{2\delta^{k}}_{(A,B)} \underbrace{-\,\delta^{k+1}}_{(C,C)} + \underbrace{2\delta^{k+2}}_{(A,B)} + \underbrace{0}_{(A,C)}$$

حال u_1' و u_1''' را مقایسه می‌کنیم:

$$u_1' - u_1''' = \delta^{k+2}\cdot\frac{2\delta-1}{1-\delta} \geq 0 \to \delta \geq \frac{1}{2}$$ ✓

⟸ در حالت $\delta \geq \frac{1}{2}$.

1- OSDP را انجام داده‌ایم و مطمئنیم به سود بیشتری نمی‌انجامد.

2- گفتیم کاری که به عنوان OSDP انجام دادیم آیا درست بود؟

چه وقتی درست است؟ وقتی مراحل بعدی BR باشند. بررسی کردیم برگشتن به s^* پس از deviate، BR است.

Bayesian Game

← در بازی‌های قبلی بازیگران از تابع هدف سایر بازیگران اطلاع کامل
← ممکن بود از تصمیم یکدیگر خبر نداشته باشند.

در بازی‌های بیزی اطلاعات از تابع هدف یکدیگر بصورت تابع توزیع احتمالاتی
← imperfect information

← منشأ این عدم قطعیت:

- عدم قطعیت در محیط بازی ← محیط در state های مختلف payoff های متفاوت دارد.

هر دو P1, P2 نسبت به محیط عدم قطعیت دارند.
به جای payoff از expected payoff استفاده می‌کنیم.

- عدم قطعیت خود بازیگران ← اصطلاحاً بازیگران mode های متفاوت دارند و مطابق mode هاش payoff متفاوت.
این mode برای یک بازیگر احتمال دارد.

P1 از احوال خودش آگاه است اما از احوال P2 آگاه نیست.

مثال بازی Bayesian

P_1 \ P_2	T	F
T	2,1	0,0
F	0,0	1,2

← مثال 1:
استراتژی ها: Tennis, Football

← اگر بازی استراتژیک ساده باشد. (1,2),(2,1) ← نش خالص
+ 1 تعادل نش مخلوط دارد.

← فرض کنیم P1 نسبت به اینکه اصولاً P2 علاقه مند است با او بازی کند یا خیر، عدم قطعیت دارد. فرض می‌کنیم P2، 50% mode بازی کردن با او را دارد. ⇐ P2 از دید P1 دو نوع (Type) دارد و هر Type با احتمال $\frac{1}{2}$ رخ می‌دهد.

ادامه ستون بعد

P_1 \ P_2	T	F
T	2,1	0,0
F	0,0	1,2

Type 1

↓

P2 را دیدن P1 حوش حل می‌شود.

P_1 \ P_2	T	F
T	2,0	0,2
F	0,1	1,0

Type 2

↓

P2 را دیدن P1 حوش حل می‌شود.

← P2 دقیقاً می‌داند در کدام Type قرار دارد پس باید استراتژی مناسب بچیند.

⇐ به کمک روش تقاطع BR ها تعادل نش را می‌یابیم:

$$P1\begin{cases} T \to BR_2(T)=\begin{cases}\text{Type 1}: T\\ \text{Type 2}: F\end{cases}\\ F \to BR_2(F)=\begin{cases}\text{Type 1}: F\\ \text{Type 2}: T\end{cases}\end{cases}$$

← P2 در کل 4 استراتژی دارد: TT, TF, FT, FF
در BR_2 فقط دو حالت آن اتفاق افتاد: TF, FT

پس BR_1 را براساس استراتژی ها TF, FT، P2 حساب می‌کنیم:

T: (Type1) P2   F: (Type2) P2

$$P2\begin{cases}\begin{cases}\text{Type 1}: T\\ \text{Type 2}: F\end{cases} \to BR_1(TF)\begin{cases}T: 2\times\frac{1}{2}+0\times\frac{1}{2}=1\\ F: 0\times\frac{1}{2}+1\times\frac{1}{2}=\frac{1}{2}\times\end{cases}\\ \begin{cases}\text{Type 1}: F\\ \text{Type 2}: T\end{cases} \to BR_1(FT)\begin{cases}T: 0\times\frac{1}{2}+2\times\frac{1}{2}=1\\ F: 1\times\frac{1}{2}+0\times\frac{1}{2}=\frac{1}{2}\times\end{cases}\end{cases}$$

P2: T   P2: F

تعادل نش بیزی → (T, (TF))

وجود Type در بازی و اطلاع داشتن P2 باعث شد استراتژی داشته باشد و برای هر Type ش برنامه داشته باشد.

اگر P2 نیز نسبت به Type های خودش عدم آگاهی داشت ← عدم آگاهی از محیط بازی

Bayesian Game

• اجزای بازی های بیزی:

1- Players:

a finite set → $I=\{1,2,\dots,N\}$

2- State:

متغیر تصادفی

a set Ω , $\omega\in\Omega$

شرایط مختلف بازی

اطلاعات بازیگران از این شرایط: متفاوت

بازیگران یک توزیع از state دارند و براین اساس expected payoff حساب می‌کنند.

3- Action:

set of player i's actions → A_i

↳ $i\in I$

4- Type:

set of player i's types → T_i

$t_i\in T_i$ ← یک نوع مشخص از انواع ممکن برای P_i

یک Random Variable است.

P_i از T_i اطلاع دارد.

5- Blief:

تابع احتمال $P_i : T_i \longrightarrow \Delta(\Omega\times\prod_{j\neq i} T_j)$

$P_i(\omega, t_{-i}\,|\,t_i)$

ورودی تابع: t_i ← باتوجه به type ای که دارد

یک تخمین احتمالاتی نسبت به type سایر

بازیگران و محیط می‌دهد.

فرض اولیه بازی های بیزی

$p(\omega,t)\in\Delta(\Omega,\prod T_i)$ فرض common prior :

فرض می‌کنیم تابع احتمال مشترک Type ها و State های بازیگران،

همه‌ی بازیگران ازش اطلاع دارند. با این فرض:

$$P(\omega,t_{-i}\,|\,t_i)=\frac{P(\omega,t)}{P(t_i)}$$

6- Payoff:

$u_i : A\times T\times\Omega\longrightarrow\mathbb{R}$

$u_i(a,\omega,t)$

↳ $a=(a_1,\dots,a_N)\in A=\prod_{i\in I}A_i$

→ $\omega\in\Omega$

→ $t=(t_1,\dots,t_N)\in T=\prod_{i\in I}T_i$

← سود بازیگر i، وقتی مشخص باشد هر بازیگر چه تصمیمی گرفته، دقیقاً شرایط محیطی الان چیست و دقیقاً هر بازیگر درکدام Type است، می‌توان سود بازیگر i را محاسبه کرد. چون نمی‌دانیم دقیقاً ω, t چه هستند بجای payoff از expected payoff استفاده می‌کنیم.

7- Strategy:

تابع $s_i . T_i\longrightarrow A_i$

Action خروجی تابع (به نام Strategy) از Type است.

8- Expected Payoff:

متوسط سود بازیگر i درحالتی که نوع T_i را داشته باشد و

استراتژی S_i را دنبال کند.

$$E\Big(u_i\big(\underbrace{(\underbrace{s_i(t_i)}_{a_i},\underbrace{s_{-i}(\tilde{t}_{-i})}_{a_{-i}})}_{a}\big),(t_i,\tilde{t}_{-i}),\tilde{\omega})\,\Big|\,t_i\Big)$$

* متغیرهایی که روی آنها علامت "~" دارد، متغیرهای تصادفی اند.

$$\sum_{\omega\in\Omega}\sum_{t_{-i}\in T_{-i}} u_i\big((s_i(t_i),s_{-i}(t_{-i})),(t_i,t_{-i}),\omega\big)\cdot P_i(\omega,t_{-i}\,|\,t_i)$$

تعریف • تعادل نش بیزی (BNE)

استراتژی $s^*=(s_1^*,\dots,s_N^*)$ یک استراتژی BNE خالص است، اگر به‌ازای هر $i\in I$، s_i^*، پاسخ بهینه (BR) نسبت به s_{-i}^* باشد، به‌ازای هر t_i مشخص.

$$E(u_i(s_i^*(t_i), s_{-i}^*(t_{-i}), t_i, \tilde{t}_{-i}, \tilde{\omega})\mid t_i) \geqslant$$

$$E(u_i(a_i, s_{-i}^*(t_{-i}), t_i, \tilde{t}_{-i}, \tilde{\omega})\mid t_i), \forall a_i\in A_i, t_i\in T_i, i\in I$$

Session 11-1

← یادآوری Expected Utility :

$$E(u_i(s_i(t_i), s_{-i}(\tilde{t}_{-i}), t_i, \tilde{t}_{-i}, \tilde{\omega})\mid t_i)$$

تابع توزیع تجمعی

پیوسته → $= \int u_i(s_i(t_i), s_{-i}(t_{-i}), t_i, t_{-i}, \omega)\, dF(\omega, t_{-i}\mid t_i)$

تابع چگالی احتمال

$$= \int_{\omega\in\Omega}\int_{t_{-i}\in\prod_{j\neq i}T_j} u_i(s_i(t_i), s_{-i}(t_{-i}), t_i, t_{-i}, \omega)\, f(\omega, t_{-i}\mid t_i)\, d\omega\, dt_{-i}$$

← Mixed Strategy :

در حالت Pure، استراتژی ما یک تابع بود با ورودی تابع Type، و خروجی action است. در حالت Mixed به جای action، احتمال انتخاب action را داریم. پس تابع Mixed Strategy را تعریف می‌کنیم.

احتمال اینکه تحت استراتژی مخلوط δ_i، عمل a_i بازی شود هنگامی که نوع بازیگر i، t_i باشد. ← $\delta_i(a_i\mid t_i)$

Random Variable

$$\to E(u_i(\delta_i(a_i\mid t_i), \delta_{-i}(a_{-i}\mid t_i), t_i, t_{-i}, \omega)\mid t_i)$$

Belief

$$= \sum_{\omega\in\Omega}\sum_{t_i\in T_{-i}} P(\omega, t_{-i}\mid t_i)\left(\sum_{a\in A}\left(\prod_{j\in I}\delta_j(a_j\mid t_j)\right)u_i(a, t_i, t_{-i}, \omega)\right)$$

Session 11-2

مثال بازی Bayesian

← مثال 1 : بازی بین کلانتر و مضنون که هردو مسلح هستند.

P_2 → sherrif

Suspect P_1

Type 1. innocent $(1-p)$

	Shoot	Not
sh	-3,-1	-1,-2
Not	-2,-1	0,0

Type 2. criminal (p)

	Shoot	Not
sh	0,0	2,-2
Not	-2,-1	-1,1

← آیا بازی تعادل دارد؟ بر این اساس بهترین تصمیم برای کلانتر؟

در Type1: برای P1 استراتژی غالب ← Not shoot

در Type2: برای P1 استراتژی غالب ← Shoot

حال برای P2 یا کلانتر بررسی می‌کنیم.

باتوجه به اینکه BR_1 مشخص است، پس استراتژی P1 را در Type1، Not shoot و در Type2، Shoot، fix. کرده، Expected payoff بازیکن 2 را طبق این می‌سنجیم:

$$E_2(\text{Shoot}) = (-1)(1-p)+(0)(p) = p-1$$

$$E_2(\text{Not}) = (0)(1-p)+(-2)p = -2p$$

فرض : $E(\text{shoot}) > E(\text{Not})$ ←

$$p-1>-2p \to 3p>1 \to p>\frac{1}{3}$$

اگر کلانتر با احتمال بزرگتر از $\frac{1}{3}$ تشخیص دهد که مضنون، مجرم است، بهتر است که shoot کند.

$\Rightarrow P \geqslant \frac{1}{3}$: BNE((NS,S))

$P \leqslant \frac{1}{3}$: BNE((NS,N))

← مثال 2 .

		I_{21} (P2)			I_{22} (P2)	
		U	D		U	D
I_{11} (P1)	U	2,0	0,2	U	2,2	0,3
	D	0,2	2,0	D	3,0	1,1
		p= 0.3			p= 0.1	
I_{12} (P1)	U	2,2	0,0	U	2,1	0,0
	D	0,0	1,1	D	0,0	1,2
		p=0.2			p=0.4	

احتمال اینکه Type هردو بازیگر 2 باشد.

اگر این احتمالاتی که نوشته شده، صرفاً یک عدم قطعیت محیطی بود، یعنی صرفاً بازیگران عدم قطعیت داشتند که در کدام یک از 4 نوع قرار می‌گیرد، در این صورت حل بازی با حالت عادی Baysian فرقی نمی‌کرد و باید احتمال هر Type که در جدول نوشته شده را در payoff های بازی مربوط به آن ضرب می‌کردیم و با هم جمع می‌کردیم در نهایت یک بازی، به وجود می‌آید که Expected Utility بازیگران در آن وجود داشت، باید آن‌را ماکزیمم می‌کردیم.

← این جا فرض می‌کنیم در این بازی، احتمالات نوشته شده، مربوط به عدم قطعیت از Type هاست.

• اگر P1 بخواهد BR اش را نسبت به P2 محاسبه کند: بازیگر 2، در هر Type می‌تواند U یا D را انتخاب کند. پس در مجموع 4 استراتژی دارد. به ازای هر 4 تا، باید BR_1 را محاسبه کنیم و برعکس. سپس تقاطع‌شان را بررسی کنیم.

$$P2: \begin{cases} I_{21}: U \\ I_{22}: U \end{cases} \qquad \frac{P(I_{21}, I_{11})}{P(I_{11})} = \frac{0.3}{0.3+0.1} \qquad \frac{0.1}{0.4}$$

$$BR_1(UU) = \begin{cases} I_{11} \begin{cases} U: 2 \times P(I_{21}|I_{11}) + 2 \times P(I_{22}|I_{11}) = 1.75 \checkmark \\ D: 0 \times P(I_{21}|I_{11}) + 3 \times P(I_{22}|I_{11}) = 0.75 \end{cases} \\ I_{12} \begin{cases} U: 2 \times P(I_{21}|I_{12}) + 2 \times P(I_{22}|I_{11}) = 2 \checkmark \\ D: 0 \times P(I_{21}|I_{12}) + 0 \times P(I_{22}|I_{11}) = 0 \end{cases} \end{cases}$$

$$\frac{P(I_{21}, I_{12})}{P(I_{12})} = \frac{0.2}{0.2+0.4} \qquad \frac{0.4}{0.6}$$

در حالت I_{11}، payoff ، 1.75 بیشتر است پس BR_1، وقتی بازیگر یک، Type1 دارد ← U است.

در حالت I_{21}، payoff ، 2 بیشتر است پس BR_1، وقتی بازیگر یک، Type2 دارد ← U است.

→ $BR_1(UU) = UU$

حال BR_2 را محاسبه می‌کنیم، فرض:

$$P1: \begin{cases} I_{11}: U \\ I_{12}: U \end{cases}$$

$$BR_2(UU) = \begin{cases} I_{21} \begin{cases} U: 0 \times \frac{0.3}{0.5} + 2 \times \frac{0.2}{0.5} = 0.8 \\ D: 2 \times \frac{0.3}{0.5} + 0 \times \frac{0.2}{0.5} = 1.2 \checkmark \end{cases} \\ I_{22} \begin{cases} U: 2 \times \frac{0.1}{0.5} + 1 \times \frac{0.4}{0.5} = 1.2 \checkmark \\ D: 3 \times \frac{0.1}{0.5} + 0 \times \frac{0.4}{0.5} = 0.6 \end{cases} \end{cases}$$

→ $BR_2(UU) = DU$

چون در حالت (UU ,UU) تقاطع نشد پس استراتژی تعادلی نش نمی‌تواند باشد.

* در خانه 6 حالت دیگر را بدست آورید >_<

← مثال 3 : بازی کورنات با اطلاعات ناقص :

• تابع هزینه‌ی خطی ← برای هر دو

$$\begin{cases} P_1 : C \\ P_2 : \begin{cases} C_H : 1-\theta \ \text{احتمال} \\ C_L : \theta \ \text{احتمال} \end{cases} \end{cases}$$

• تابع سود بازیگران.

$$\Pi_1(q_1,q_2) = q_1 P(q_1+q_2) - c q_1$$

$$\Pi_2(q_1,q_2) = q_2 P(q_1+q_2) - C_t q_2 \qquad t \in \{L, H\}$$

• تابع قیمت :

$$P(q_1+q_2) = \alpha - (q_1+q_2)$$

← تعادل نش بازی براساس تقاطع BR ها :

$$(q_1^*, q_2^*(t)) \rightarrow \text{NASH}$$

نمایش را عوض می‌کنیم : $q_2(H) = q_H, \ q_2(L) = q_L$

$$q_1^* = BR_1(q_L, q_H) = \underset{q_1 \geq 0}{\operatorname{Arg\,max}} \Big[\theta(P(q_1+q_L)-c)q_1 + (1-\theta)(P(q_1+q_H)-c)q_1\Big]$$

$$q_L^* = BR_L(q_1) = \underset{q_L \geq 0}{\operatorname{Arg\,max}} \big[(P(q_1+q_L)-C_L)q_L\big]$$

$$q_H^* = BR_H(q_1) = \underset{q_H \geq 0}{\operatorname{Arg\,max}} \big[(P(q_1+q_H)-C_H)q_H\big]$$

هرسه quadratic هستند. با حل سه معادله سه مجهول :

$$q_1^* = \frac{1}{3}(\alpha - 2c + \theta c_L + (1-\theta)c_H)$$

$$q_L^* = \frac{1}{3}(\alpha - 2c_L + c) - \frac{1}{6}(1-\theta)(c_H - c_L)$$

$$q_H^* = \frac{1}{3}(\alpha - 2c_H + c) - \frac{1}{6}\theta(c_H - c_L)$$

← وجود اطلاعات ناقص چه تأثیری در نتایج این بازی دارد؟

می‌توان نشان داد که $q_L^* > q_H^*$. وقتی P2 کنار Type خودش آگاهی دارد، وقتی ضریب تابع هزینه‌اش کمتر است پس هزینه‌اش کمتر است، تمایل به تولید بیش‌تری دارد.

$$q_L^* = \frac{1}{3}(\alpha - 2c_H + c) + \frac{2}{3}(c_H - c_H) - \frac{1}{6}(c_H - c_L) + \frac{1}{6}\theta(c_H - c_L)$$

$$= q_H^* + \frac{1}{2}(c_H - c_L) > q_H^*$$

← یادآوری نش در حالت بدون عدم قطعیت :

$$q_1^* = \frac{\alpha - 2c_1 + c_2}{3}, \quad q_2^* = \frac{\alpha - 2c_2 + c_1}{3}$$

? آیا اطلاعات ناقص P1 از P2 در تصمیم‌گیری P2 تأثیر دارد؟

تأثیر داشته.

$$q_H^* = \frac{1}{3}(\alpha - 2c_H + c) \underbrace{- \frac{1}{6}\theta(c_H - c_L)}_{\text{تأثیر}}$$

این اثر متقابل در کجا به نفع و ضرر بازیگر 2 است؟

وقتی که Type بازیگر 2، L است. اثر این اطلاعات با علامت منفی ظاهر شده. چون Expected Cost بازیگر 2 از دید P1، از C_L بیش‌تر است. چون یک ترکیب خطی از C_L, C_H است.

انگار بازیگر 1، Over estimate از P2 داشته، و بازیگر 2 به‌خاطر این Over estimation، مجبور شده کمتر تولید کند و اثر این به‌صورت $-\frac{1}{6}(1-\theta)(c_H - c_L)$ در q_L^* ظاهر شده و قاعدتاً در قیمت مشابه، تولید کمتر باعث ضرر خواهد شد و در آن سو، q_1^* مقدار تولیدش بیش‌تر شده $(C_L > C_H)$

وقتی Type، P2، H باشد. باعث شده تولیدش زیاد شود. و بازیگر 1، under estimate کرده (چون $C_H > C_L$) و تولیدش کاهش می‌یابد.

کاربرد بازی های بیزی
در Auction ها

- Players : $I=\{1,\dots,n\}$

- Types . $v_i \in [0\ \ \bar{v}]$ $\quad i=1,\dots,n$

فرض می‌کنیم $v_1,\dots,v_n$ مستقل هستند و از توزیع یکسانی می‌آیند.

i i d: independent & identically distributed

← v_1 ~ تابع چگالی احتمال f با تابع توزیع تجمعی F

Common prior : همه‌ی بازیکنان مطلع‌اند

- Strategy: مثلاً $\beta_1 : [0\ \ \bar{v}] \to \mathbb{R}^+$

$b_1 = \beta_1(v_i)$

* هم Type، هم استراتژی بازیکنان پیوسته است.

- Payoff :

→ First Price Auction (FPA)

$$(\pi_1)_{FPA} = \begin{cases} v_i - b_i \ ; & b_i > \max_{j\neq i} b_j \ . \text{ winner} \\ 0 \ ; & o.w. \end{cases}$$

→ Second Price Auction (SPA)

$$(\pi_1)_{FPA} = \begin{cases} v_i - \max_{j\neq i} b_j \ ; & b_i > \max_{j\neq i} b_j \\ 0 & o.w. \end{cases}$$

* در حالت کلی پیدا کردن استراتژی BNE در بازی‌های بیزی که هم Type و هم محیط بازی پیوسته است، کمی مشکل است.

چرا؟ استراتژی، تابعی از Type است و این جا دنبال یک تابع هستیم و بدست آوردن این تابع از تقاطع دادن BR بازیکنان مشکل است

$$u_i = E\left(u_i(\beta_i(v_i), \beta_{-i}(v_{-i}), v_i, v_{-i}) \mid v_i\right)$$
$$= \int u_i(\beta_i(v_i), \beta_{-i}(v_{-i}), v_i, v_{-i})\, dF(v_{-i} \mid v_i)$$

محاسبات پیچیده!

← پس برای ساده‌سازی محاسبات چه می‌کنیم؟

از بین استراتژی‌های نش ممکن که می‌توان برای بازی بدست آورد، استراتژی نش را برای بازیکنان بصورت متقارن بدست می‌آوریم.

فرض می‌کنیم که $\beta_1^*=\beta_2^*=\dots=\beta_n^*$ ← تعادل نش متقارن

وقتی استراتژی تعادلی نش برای همه یکسان باشند، قاعدتاً ما با حل یکی از معادلات (نوشته شده در انتهای ستون قبل)، مواجه هستیم.

* تعادل نش متقارن به معنای action متقارن نیست!

$b_1=\beta_1^*(v_1)\ , b_2=\beta_2^*(v_2)\ , \dots , b_n=\beta_n^*(v_n)$

β_i ها را مساوی فرض کردیم. اما b_i ها به T_i هر بازیکن وابسته‌اند.

← حل بازی.

مثلا SPA.

در حالت بازی با اطلاعات کامل این بازی را تحلیل کردیم و استراتژی غالب هر بازیکن این بود که به اندازه‌ی ارزش یا v_i خود، قیمت پیشنهاد دهد ← $b_i=v_i$

با توجه به اینکه در حالت قبل استراتژی غالب داشتیم، حل مسأله در حالت بیزی بدیهی خواهد بود چون تصمیم بازیکن i، مستقل از نوع بازیکن‌های دیگر و تصمیماتشان بود. با توجه به اینکه در بازی بیزی نیز هر بازیکن Type خودش را می‌داند، همچنان در این حالت نیز $b_i=v_i$ استراتژی غالب خواهد بود و عدم قطعیت تأثیری ندارد.

پس خودبه‌خود استراتژی متقارن برای بازیکنان بدست می‌آید.

$b_i=\beta(v_i)=v_i$ ← استراتژی تعادلی نش غالب بیزی

FPA :

درحالت اطلاعات کامل، استراتژی غالب نداشت ۱

فرضیات :

(I) v_i $(i=1,\dots,n)$ دارای توزیع یکنواخت در بازه‌ی $[0,1]$ باشند و i.i.d باشند.

(II) در جست‌وجوی تعادل نش بیزی متقارن هستیم.

$$BNE = (\beta^*,\dots,\beta^*) \quad , \quad b_1^* = \beta^*(v_1)$$

(III) فرض می‌کنیم β یک تابع صعودی و مشتق‌پذیر نسبت به v_i است و داریم $\beta(0)=0$.

• صعودی : هرچه value بازیگر بالاتر برود، حاضر است bid بیش‌تری دهد.

(IV) فرض می‌کنیم : $\beta(v) = \theta v$. ما به دنبال یک استراتژی نش متقارن خطی هستیم.

لزوماً وجود ندارد. باید بررسی کنیم.

Expected Payoff.

$$E(u_i(b_i, b_{-i}, v_i, v_{-i}) \mid v_i) =$$

$$P(win \mid b_i)(v_i - b_i) + P(lose \mid b_i)\times(0) =$$

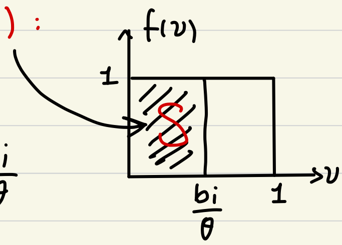
$$P(\max_{j\neq i} b_j < b_i)(v_1 - b_1) + 0 \xrightarrow[\beta(v_j)=\theta v_j]{b_j = \beta(v_j)}$$

$$= P(\max_{j\neq 1} V_j < \frac{b_i}{\theta})(v_i - b_i)$$

$$= P(V_1,\dots,V_n < \frac{b_i}{\theta})(v_i - b_i) \xrightarrow{i.i.d}$$

$$= P(V_1 < \frac{b_i}{\theta})\, P(V_2 < \frac{b_i}{\theta}) \dots P(V_n < \frac{b_i}{\theta})(v_i - b_i)$$

$* \; P(X < x)$:



$$\rightarrow S = \frac{b_i}{\theta}$$

$$\Rightarrow P(v_j < \frac{b_i}{\theta})_{j\neq 1} = \frac{b_i}{\theta}$$

$$\Rightarrow P(V_1 < \frac{b_i}{\theta})\, P(V_2 < \frac{b_i}{\theta}) \dots P(V_n < \frac{b_i}{\theta})(v_i - b_i)$$

$$= \left(\frac{b_i}{\theta}\right)^{n-1}(v_i - b_i)$$

$$\rightarrow E(u_1) = \left(\frac{b_i}{\theta}\right)^{n-1}(v_i - b_i)$$

* هدف بدست آوردن تابع و رابطه‌ی v_i, b_i است. آیا b_1 بصورت یک تبدیل خطی از θ بدست می‌آید یا خیر؟ اگر بدست بیاید یعنی هر فرضی که تا این جا داشتیم درست است

$$\frac{\partial E(u_i)}{\partial b_1} = \left(\frac{n-1}{\theta}\right)\left(\frac{b_i}{\theta}\right)^{n-2}(v_1 - b_i) - \left(\frac{b_i}{\theta}\right)^{n-1} = 0$$

$$\rightarrow \left(\frac{n-1}{\theta}\right)(v_i - b_i) - \frac{b_i}{\theta} = 0 \rightarrow b_i = \underbrace{\left(\frac{n-1}{n}\right)}_{\theta^*} v_1$$

$$\frac{\partial^2 E(u_i)}{\partial b_i^2} < 0$$ یعنی داریم ماکزیمم می‌کنیم.

یک Conjecture داشتیم که ما یک نش متقارن خطی داریم و با حل این مسأله، حدسمان تأیید شد.

$$\beta^*(x) = \frac{n-1}{n}x$$

$$E(u) = F^{n-1}(\beta^{-1}(b))(v-b)$$

← حال دنبال رابطه‌ای بین b و v که تابع β را به ما می‌دهد.

$$\frac{\partial E(u)}{\partial b} = (n-1)F^{n-2}(\beta^{-1}(b))f(\beta^{-1}(b))[\beta^{-1}(b)]' \times (v-b) - F^{n-1}(\beta^{-1}(b))$$

$$b = \beta(\beta^{-1}(b)) \xrightarrow{\partial/\partial b}$$

$$1 = [\beta^{-1}(b)]' \times \beta'(\beta^{-1}(b))$$

$$\Rightarrow [\beta^{-1}(b)]' = \frac{1}{\beta'(\beta^{-1}(b))}$$

$$\Rightarrow \frac{\partial E u}{\partial b} = \frac{(n-1)f(\beta^{-1}(b))F^{n-2}(\beta^{-1}(b))}{\beta'(\beta^{-1}(b))}(v-b) - F^{n-1}(\beta^{-1}(b)) = 0 \xrightarrow{\beta^{-1}(b)=v}$$

$$\frac{[F^{n-1}(v)]'}{\beta'(v)}(v-\beta(v)) - F^{n-1}(v) = 0$$

$$\rightarrow \underset{\Delta}{[F^{n-1}(v)]'v} = \beta'(v)F^{n-1}(v) + [F^{n-1}(v)]'\beta(v)$$

$u v' + u' v$

$$\overset{\square}{=} \frac{d}{dv}[F^{n-1}(v)\beta(v)]$$

$$\int \rightarrow \square = F^{n-1}(v)\beta(v) - F^{n-1}(0)\cancel{\beta(0)}^{\,0}$$

$$\int \rightarrow \Delta = \int_0^v x[F^{n-1}(x)]'dx$$

$$\rightarrow F^{n-1}(v)\beta(v) = \int_0^v x[F^{n-1}(x)]'dx$$

$$\beta(v) = \frac{1}{F^{n-1}(v)}\int_0^v x[F^{n-1}(x)]'dx$$

$$\Rightarrow (G(v) = F^{n-1}(v)\ ,\ g(v) = [F^{n-1}(v)]')$$

$$\rightarrow \beta(v) = \frac{1}{G(v)}\int_0^v y g(y)\,dy$$ فرم امید ریاضی را دارد.

$$= E\{Y_i \mid Y_i < v\}$$

FPA در حالت کلی‌تر:

فرضیات جدید:

(I) v_i ها، i.i.d و دارای تابع توزیع تجمعی F و تابع چگالی f می‌باشند.

← درحالت قبل توزیع یکنواخت بود

(II) تابع $\beta(\cdot)$، تابعی صعودی و مشتق‌پذیر و $\beta(0)=0$ است.

← درحالت قبل تابع β را خطی فرض کردیم.

Expected Payoff.

$$E(u_i(b_i, b_{-i}, v_i, v_{-i}) \mid v_i) =$$

$$P(win \mid b_i)(v_i - b_i) + P(lose \mid b_i) \times (0) =$$

$$P(\underset{j\neq i}{Max}\, b_j < b_i)(v_i - b_i) + 0 \xrightarrow{\text{تابع } \beta(\cdot) \text{ صعودی}}$$

$$\Rightarrow \underset{i\neq j}{Max}\, b_j = \underset{i\neq j}{Max}(\beta(v_j)) = \beta(\underset{i\neq j}{Max}(v_j))$$

$$= \beta(Y_i)\ ,\ Y_i = \underset{j\neq i}{Max}\, v_j$$

$$\rightarrow E(u_i) = P(\beta(Y_i) < b_i)(v_i - b_i) \xrightarrow[\text{و یکبه‌یک}]{\text{تابع } \beta(\cdot) \text{ صعودی}}$$

$$= P(Y_i < \beta^{-1}(b_i))(v_i - b_i)$$

$$= P(Y_1 < \beta^{-1}(b_1), \ldots, Y_{i-1} < \beta^{-1}(b_{i-1}), Y_{i+1} < \beta^{-1}(b_{i+1}), \ldots, Y_n < \beta^{-1}(b_n))(v_i - b_i)$$

$$= P(V_1 < \beta^{-1}(b_1)), \ldots, P(V_{i-1} < \beta^{-1}(b_{i-1}))P(V_{i+1} < \beta^{-1}(b_{i+1})), \ldots, P(V_n < \beta^{-1}(b_n))(v_i - b_i)$$

(برای همه یکسان است. → $\Rightarrow f_X(x) = P(X < x)$)

$$\rightarrow = F^{n-1}(\beta^{-1}(b_i))(v_i - b_i)$$

← برای سادگی نمایش از این پس. $v_i \rightarrow v\ ,\ b_i \rightarrow b$

$\beta(v) = E\{Y_i \mid Y_i < v\}$ ←

$$\begin{cases} P(Y_i < v) = G(v) \\ Y_1 = \underset{i \neq j}{\operatorname{Max}} \{v_j\} \end{cases}$$

$F^{n-1}(v)$: تابع توزیع تجمعی ماکزیمم n-1 تا متغیر تصادفی

* چک کردن مشتق دوم کمتر از صفر را نیز فراموش نکنید.

$\frac{\partial E(u_i)}{\partial b} < 0 \longrightarrow$ نشان می‌دهد که به دنبال Max هستیم

* هم چنین بررسی کنید که تابع $\beta(v)$ بدست آمده، صعودی است.

Session 13

قضایای بازی‌های بیزی

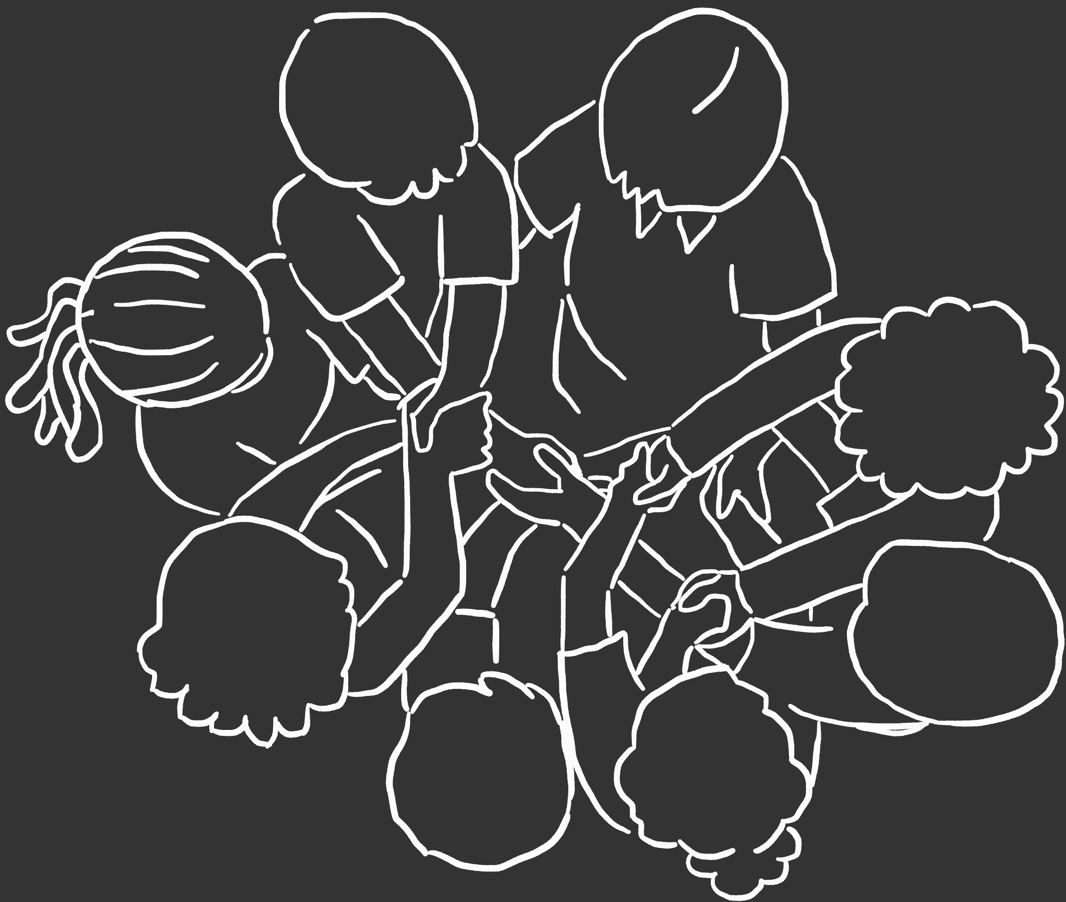
قضیه 1. یک بازی بیزی محدود (با تعداد استراتژی‌ها و Typeهای محدود) را در نظر بگیرید، در این صورت این بازی، یک تعادل نش بیزی مخلوط دارد.

قضیه 2. یک بازی بیزی با فضای استراتژی پیوسته، Type پیوسته برای بازیگران در نظر بگیرید؛ در این صورت اگر مجموعه Type و استراتژی بازیگران فشرده (compact) باشند و توابع سود بازیگران پیوسته و مقعر نسبت به استراتژی خود بازیگر باشند، در این صورت استراتژی نش خالص بیزی وجود دارد.

* Compact closed & bounded

CHAPTER 3

~ Cooperative Games
~ Coalitional Games



Shiva Shakeri

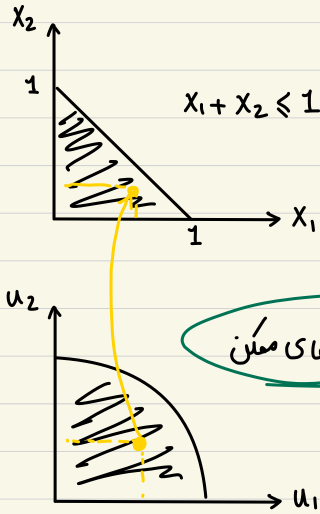
بازی های همکارانه

1. چانه‌زنی نش ، Nash Bargaining Solution
2. بازی های ائتلافی ، Coalitional Games

1 چانه‌زنی نش

رویکرد حل این مسائل این است که ، من و شمایی که با هم همکاری داریم و هرکدام تابع هدف خود را داریم ، بر سر یک سری اصول با هم توافق کنیم . اگر روی این اصول توافق کنیم ، حل مسأله بر اساس این اصول خواهد شد . این نگاه اصولی نش به این نوع مسائل است . این اصول که "Axiom" به آن‌ها گفته می‌شود ، عاقلانه و یا منطقی هستند . چه اصولی برای بازی بگذاریم ؟

← استراتژی x_1, x_2 را برای P1, P2 با تابع سودهای $u_1(x_1,x_2), u_2(x_1,x_2)$ داریم :



فضای استراتژی . $x_1+x_2\leq 1$

فضای همه‌ی payoff های ممکن

آقای نش اصول را روی این فضا بحث می‌کند .
می‌گه کلاً سودها قرار هست چگونه تقسیم شود ؟
بعد می‌بینه چه استراتژی ای متناظر این payoff هست؟

* چیزی که نیاز به توافق دارد ، سودها هستند .

← مثال 1 : تقسیم یک کیک بین دو نفر

فضای استراتژی : $X=\{(x_1,x_2) \mid x_1+x_2\leq 1,\ x_1\geq 0\}$

سهم P1 از کیک ، سهم P2

$D=(0,0)$: توافق نکنند و به هیچ‌کس کیک نرسد

توابع سود : $u_i(x)$, $i=1,2$, $x=(x_1,x_2)$

میزان رضایت $U=\{(v_1,v_2) \mid v_1=u_1(x),\ v_2=u_2(x)\}$

disagreement utility . $d=(u_1(D),u_2(D))$

← حال دنبال یک راه حل مناسب روی U هستیم :

Bargaining Problem (BP)

با (U,d) که فضای utility و disagreement هستند .

- $I=\{1,2\}$ → notation برای سادگی
- $U\subset \mathbb{R}^2$, $d\in U$

* U is convex & compact

← Convex : در یک مجموعه یعنی اگر هر دو نقطه‌ی درون آن را به هم وصل کنیم ، خط اتصال آنها ، تماماً درون مجموعه است .

← Compact : Closed & Bounded : بسته و مرزدار

$\to \exists v=(v_1,v_2)\in U : v>d$ or $(v_1>d,\ v_2>d)$

BP Solution :

B : مجموعه‌ی همه‌ی BP های ممکن

$f: B\to U$ — عضوی از فضای utility

$\to f=(f_1,f_2)=(f_1(U,d),f_2(U,d))$

هدف ما $=(v_1^*,v_2^*)\in U$

← پس برای حل یک BP به دنبال همچین f ای هستیم:

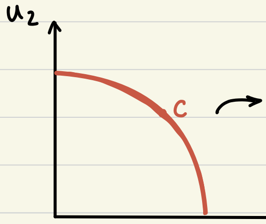
این f در درجه‌ی اول یک سری اصول است که به BP برای (U,d) وارد خواهد شد و یک نقطه در فضای استراتژی بازیکن‌ها به عنوان راه حل می‌دهد.

← f چه اصولی را باید برقرار کند؟ Axiomatic Solution

آقای نش، Nash Axioms را ارائه داده است.

(I) اصل 1. **Pareto Optimality**

ما از فضای استراتژی ممکن، به دنبال پاسخ‌هایی می‌گردیم که در آن امکان افزایش سود هردوی ما وجود نداشته باشد.



u_1

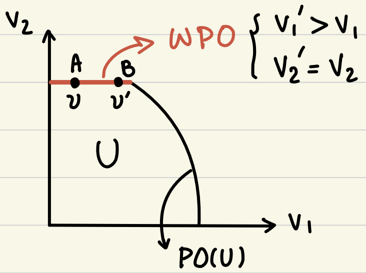
مثلاً همه نقاط روی مرز قرمز این ویژگی را دارند.

$$f((U,d)) \in \underbrace{PO(U)}_{=\{v\in U \mid \nexists v'\in U \text{ with } v' \geq v\}}$$

$$v' \geq v \Rightarrow v'_1 \geq v_1,\ v'_2 \geq v_2$$

⇝ weak PO:

$$f((U,d)) \in WPO(U) = \{v\in U \mid \nexists v'\in U \text{ with } v' > v\}$$

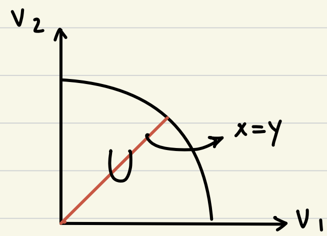


(III) اصل 2. تقارن یا **Symmetry**

اگر مجموعه‌ی U به صورتی باشد که داشته باشیم. $(v_1,v_2)\in U$، اگر و تنها اگر $(v_2,v_1)\in U$، در این صورت

$$f_1((U,d)) = f_2((U,d))$$

این اصل یک نوع شرط trigger دارد یعنی این اصل درباره‌ی join payoff هایی بحث می‌کند که کلاً متقارن هستند. یعنی اگر دو بازیکن از جهت payoff کاملاً متقارن باشند، solution ما باید طوری باشد که سود این دو باهم برابر شود.



$(d_1,d_2)\in U$
$(d_2,d_1)\in U$
$\rightarrow d_1 = d_2$

(III) اصل 3. **Scale invariance**

یعنی اگر ما فضای joint payoff بازیکن‌ها را به صورت خطی scale کنیم، solution هم به همان نسبت، scale شود.

$$\begin{cases} U' = \{(\alpha_1 v_1+\beta_1, \alpha_2 v_2+\beta_2) \mid (v_1,v_2)\in U\} \\ d' = (\alpha_1 d_1+\beta_1, \alpha_2 d_2+\beta_2) \\ f = (f_1(U,d), f_2(U,d)) \end{cases}$$

$$\Longrightarrow f_i(U',d') = \alpha_i f_i(U,d) + \beta_i \quad , i=1,2$$

(IV) اصل 4. **independence of irrelevant alternatives**

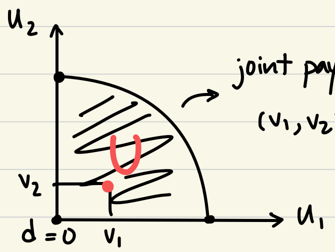
← مستقل از جواب‌های نامربوط

$$BP1: (U,d)\ , BP2: (U',d),\ U'\subseteq U,\ f(U,d)\in U'$$

$$\Longrightarrow f((U,d)) = f((U',d))$$

شکل ص بعد

← مفهوم Nash Solution



فضای joint payoff

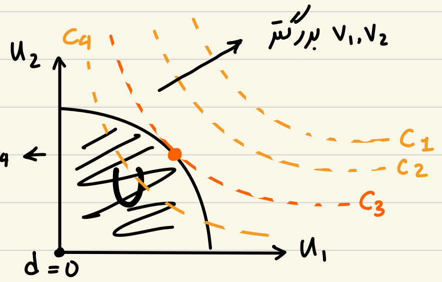
$(v_1,v_2)\in U$

$\begin{cases} u_2 \geq 0 \\ u_1\ ? \end{cases}$ $\begin{cases} u_1 \geq 0 \\ u_2\ ? \end{cases}$

$\underset{(v_1,v_2)}{Max}(v_1 v_2) = ?$ ←

به این منظور، منحنی های $v_1v_2 = C$ را رسم می‌کنیم.

به ازای C های مختلف منحنی را رسم می‌کنیم: $v_2 = \frac{v_1}{C}$

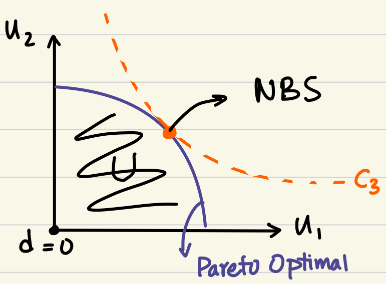


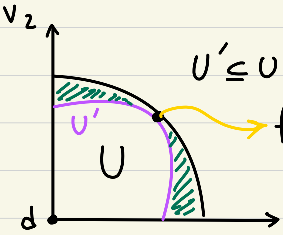
$C_1 > C_2 > C_3 > C_4$ ←

به ازای یکی از C ها، منحنی بر فضای joint payoff مماس می‌شود. بزرگترین (v_1,v_2) ای که در مجموعه‌ی feasible قرار می‌گیرد، C_3 می‌باشد. پس انگار حل این بهینه سازی نقطه‌ی تماس منحنی $v_1 v_2 = C_3$ با مجموعه‌ی feasible، joint payoff بازیگران است.

پس Max کردن v_1v_2 یعنی پیدا کردن بزرگترین C بطوریکه منحنی v_1v_2 ثابت بر منحنی Pareto بهینه سود مشترک بازیگران مماس شود.

اصل 1





$U' \subseteq U$

$f((U,d)) \in U, U'$

در این صورت جوابی که برای U' می‌دهید باید همان جواب برای U باشد.

چرا؟ به قسمت هاشور خورده‌ی سبز ← irrelevant می‌گوییم.

انگار مسأله را برای U حل کردیم و f بدست آمد، که نشان می‌دهد قاعدتاً قسمت سبز رنگ عضو جواب بهینه‌ی مسأله نبوده پس حذفشان و رسیدن به U' جواب را تغییر نخواهد داد و پاسخ U' نیز مشابه U خواهد شد.

← مثلاً فرض کنید در مجموعه‌ی $\{x_1, x_2, x_3\}$، x_1 جواب بهینه باشد حال اگر بپرسند از $\{x_1, x_2\}$ پاسخ بهینه چیست؟ شما به راحتی پاسخ می‌دهید: همان x_1!

Nash Solution

پاسخ نش که 4 اصل فوق را برقرار می‌کند، به شکل زیر خواهد بود:

(v_1^*, v_2^*) is a Nash bargaining solution if

$$(v_1^*, v_2^*) \in \underset{(v_1,v_2)}{\operatorname{ArgMax}} (v_1 - d_1)(v_2 - d_2)$$

$$\text{s.t. } (v_1,v_2)\in U \ ,\ (v_1,v_2) \geq (d_1,d_2)$$

$$\rightarrow (v_1^*, v_2^*) = f^N(U,d)$$

قضیه نش: پاسخ بهینه سازی فوق، 4 اصل مذکور را برآورده می‌کند و هر پاسخی که 4 اصل مذکور را برآورده کند، پاسخی از مسأله‌ی بهینه سازی فوق می‌باشد.

← در Nash Solution ، می‌گوییم جواب بهینه سازی وجود دارد و جوابش یکتاست.

حال این یکتایی را به قضیه‌ش اضافه می‌کنیم.

قضیه نش: پاسخ یکتای بهینه سازی فوق، 4 اصل مذکور را برآورده می‌کند و هر پاسخی که 4 اصل مذکور را برآورده کند، پاسخ یکتای مسأله‌ی بهینه سازی فوق می‌باشد.

1. وجود جواب بهینه سازی:

از آن‌جا که مجموعه‌ی U فشرده (Compact) و تابع هدف پیوسته می‌باشد، همواره جواب وجود دارد.

پیوسته : $(v_1-d_1)(v_2-d_2)$

وقتی ما یک مجموعه‌ی فشرده داریم، مجموعه solution ها از یک مجموعه‌ی محدود با مرز مشخص می‌آید. تابع هدفی هم که روی این مجموعه مشخص شده، پیوسته است. پس اگر همه‌ی نقاط موجود در U را، که فشرده است، را درنظر بگیریم و f آنها را حساب کنیم، (v_1,v_2)، از بین این ها، یک یا چندتا مقدار f بیش‌تری دارند یعنی جواب وجود دارد

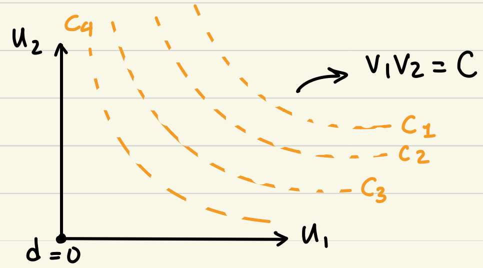
2. یکتایی جواب بهینه سازی:

از آن‌جا که تابع هدف مسأله، اکیداً شبه مقعر است، می‌توان نتیجه گرفت که جواب مسأله‌ی بهینه سازی یکتاست.

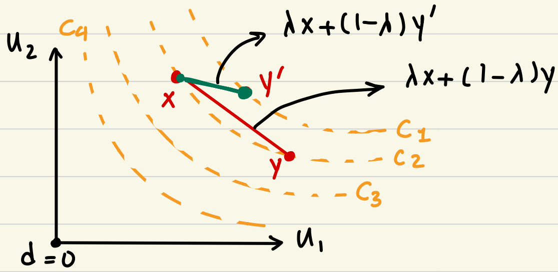
Strictly quasi-concave :

$\lambda\in(0,1)$, $f(\lambda x+(1-\lambda)y) > \min(f(x),f(y))$

$f(v_1,v_2)=v_1v_2$ ← چرا v_1v_2، اکیداً شبه مقعر است؟



برای بررسی شرط اکیداً شبه مقعر، دو نقطه x,y روی منحنی های $v_1v_2=C$ درنظر می‌گیریم. یا x، y روی یک منحنی قرار می‌گیرند یا روی منحنی های مختلف

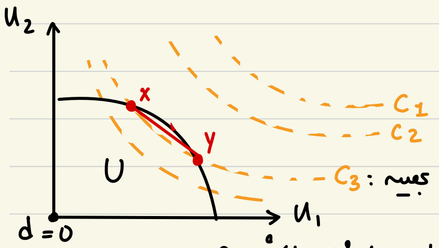


درحالت اول فرض می‌کنیم x و y روی یک منحنی باشند. دقت کنید $\lambda x+(1-\lambda)y$ در واقع همه‌ی نقاط متصل کننده‌ی x,y را می‌دهد. نقاط این خط قرمز متصل کننده‌ی x و y روی منحنی‌های C_k هایی قرار می‌گیرند که $C_k>C_2$ است. پس بنابراین عبارت زیر برقرار می‌شود.

$f(\lambda x+(1-\lambda)y) > \min(f(x),f(y))$ ✓

درحالت دوم فرض می‌کنیم x و y روی منحنی‌های متفاوت باشند. نقاط متصل کننده‌شان را با سبز نشان داده‌ایم. همه‌ی نقاط خط سبز، روی یک منحنی ثابت قرار می‌گیرند که f آنها بزرگ‌تر از C_2 و کمتر از C_1. اما شرط اکیداً شبه مقعر بودن این بود که $f(\lambda x+(1-\lambda)y)$ از $\min(f(x),f(y))$، بزرگتر باشد که دراین جا از f(x) بزرگتر است.

← چرا اکیداً شبه مقعر بودن باعث یکتایی جواب می‌شود؟



فرض کنیم $v_1v_2=C$ بر PO مماس نشود و آنرا در دو نقطه قطع کند. آیا امکان دارد؟ دو جواب داشته باشیم؟

x، y روی یک منحنی ← $f(x)=f(y)$ ← تابع اکیداً شبه مقعر

$f(\lambda x+(1-\lambda)y) > \underline{\min(f(x),f(y))} = f(x)$

داره میگه f تمام نقاطی که x، y رو به هم وصل می‌کنن (خط قرمز)، از f(x) بزرگتر است ← متناقض با بهینگی C_3 ✗.

پس جواب حتماً باید یکتا باشد!

مثال بازی همکارانه

← مثال 1: تقسیم یک کیک بین دو نفر

فضای استراتژی : $X=\{(x_1,x_2)\mid x_1+x_2\leq 1,\ x_1\geq 0\}$

سهم P1 از کیک ← ┘ └→ سهم P2

$D=(0,0)$: توافق نکنند و به هیچ‌کس کیک نرسد

توابع سود : $u_i(x)$, $i=1,2$, $x=(x_1,x_2)$

میزان رضایت $U=\{(v_1,v_2)\mid v_1=u_1(x),\ v_2=u_2(x)\}$

disagreement utility. $d=(u_1(0),u_2(0))$

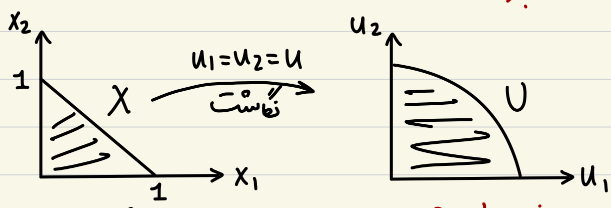
← فرض کنیم. $u(0)=0$, $u_1=u_2=u$,

u is increasing and concave

← حل: می‌توان 1- با استفاده مستقیم از 4 اصل،

2- استفاده از بهینه سازی نش، این را حل کرد.

• آیا مجموعه‌ی U، فشرده و محدب است؟



این فضا فشرده و محدب است.

این چطور؟

باید برای اثبات محدب بودن، از تعریف آن استفاده کنیم:

$(v_1,v_2)=(u(x_1),u(x_2))\in U$

$x_1+x_2\leq 1$, $(x_1,x_2)\in X$

$(v_1',v_2')=(u(x_1'),u(x_2'))\in U$

$x_1'+x_2'\leq 1$, $(x_1',x_2')\in X$

تعریف محدب $\rightarrow \alpha(v_1,v_2)+(1-\alpha)(v_1',v_2')\overset{?}{\in} U$

$\rightarrow \alpha(u(x_1),u(x_2))+(1-\alpha)(u(x_1'),u(x_2'))$

اگر این عبارت یک payoff ی باشد که از یک $x_1,x_2\in X$ تولید شده باشد، در این صورت می‌توان گفت این payoff نیز عضو U است.

← آیا joint payoff داده شده، از یک $x\in X$ تولید شده است؟

$=(\alpha u(x_1)+(1-\alpha)u(x_1'),\ \alpha u(x_2)+(1-\alpha)u(x_2'))$

$\xrightarrow{u^{-1}} (u^{-1}(\alpha u(x_1)+(1-\alpha)u(x_1')),\ u^{-1}(\alpha u(x_2)+(1-\alpha)u(x_2')))$ A

* u، concave بود. پس u^{-1} میشه convex.

$f(\alpha x+(1-\alpha)y)\leq \alpha f(x)+(1-\alpha)f(y)$

$\rightarrow A\leq(\underbrace{\alpha x_1+(1-\alpha)x_1'}_{y_1},\ \underbrace{\alpha x_2+(1-\alpha)x_2'}_{y_2})$

$(y_1,y_2)\in X$ if $y_1+y_2\leq 1$

$\rightarrow y_1+y_2=\alpha(\underbrace{x_1+x_2}_{\leq 1})+(1-\alpha)(\underbrace{x_1'+x_2'}_{\leq 1})\leq 1$

$\Longrightarrow y_1+y_2\leq 1$

$\Longrightarrow A\leq 1$

پس x مورد نظر ما عضو X بود پس payoff حاصل از آن عضو U می‌شود ⟸ پس U محدب است.

- - - - - - - - - - -

← Symmetry: اصل 2. $u_1=u_2=u$ ✓

← Pareto Op.: چون تابع صعودی است، نگاشت خط $x_1+x_2=1$، دقیقاً به مرز PO مجموعه U ها خواهد شد.

$\Longrightarrow u(x)=u(1-x)\rightarrow x=\frac{1}{2}$

$\rightarrow (v_1^*,v_2^*)=(u(\frac{1}{2}),u(\frac{1}{2}))$

• استفاده از بهینه سازی :

$$\max_{0\le z\le 1} u_1(z)u_2(1-z) \qquad PO$$

چون $x_1+x_2=1$

$$\xrightarrow{u_1=u_2=u} \max_{0\le z\le 1} u(z)u(1-z)$$

$$\xrightarrow{\text{مشتق}} u'(z)u(1-z)-u(z)u'(1-z)=0$$

$$\longrightarrow \frac{u'(z)}{u(z)}=\frac{u'(1-z)}{u(1-z)} \quad *$$

← اگر فرض کنیم پاسخ این معادله z باشد بطوریکه $z<\frac{1}{2}$.

چون تابع صعودی : $u(z)<u(z-1)$

چون تابع concave $\xrightarrow{\text{مشتق نزولی}} u'(z)>u'(1-z)$

← اگر فرض کنیم پاسخ این معادله z باشد بطوریکه $z>\frac{1}{2}$:

چون تابع صعودی : $u(z)>u(z-1)$

چون تابع concave $\xrightarrow{\text{مشتق نزولی}} u'(z)<u'(1-z)$

⇐ پس تنها در $z^*=\frac{1}{2}$ است که تساوی معادله * برقرار است.

← فرض کنیم P2، محافظه کار است! یعنی :

$u_1=u$, $u_2=h\circ u$ $\quad \begin{cases} h:\mathbb{R}\to\mathbb{R},\ h(0)=0 \\ h \text{ is increasing \& concave} \end{cases}$

؟ چرا اگر تابع بصورت $h\circ u$ باشد، می‌گوییم محافظه‌کارتر؟

فرض: $u(x)=\sqrt{x}$. صعودی و concave

$h(x)=\sqrt{x}$ $\qquad \to h\circ u=\sqrt{\sqrt{x}}$

اگر $x=\frac{1}{16}$ ← $u(x)=\frac{1}{4}$, $h\circ u(x)=\frac{1}{2}$

به هر دو بازیگر مقدار $\frac{1}{16}$ را دادیم اما میزان رضایت P2 بیش‌تر شد ← محافظه‌کارتر است و زودتر راضی می‌شود.

← استفاده از بهینه سازی :

$$\max_{0\le z\le 1} u(z)\,h(u(1-z))$$

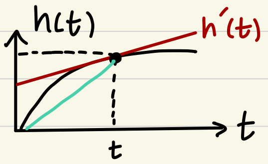
$$\xrightarrow{\text{مشتق}} u'(z)h(u(1-z))-u(z)h'(u(1-z))u'(1-z)=0$$

$$\to \frac{u'(z)}{u(z)}=\frac{h'(u(1-z))u'(1-z)}{h(u(1-z))}$$

برای حل این معادله باید توابع u و h را داشته باشیم.

؟ جوابی که از این معادله بدست می‌آید چه فرقی با حالت قبل دارد؟

از خاصیت توابع concave استفاده می‌کنیم. چون h، concave است:



$$\Rightarrow h'(t)\le\frac{h(t)}{t},\ t>0$$

میگه شیب مشتق (قرمز) کمتر از شیب خط آبی ($\frac{h(t)}{t}$) است.

$$\to \frac{u'(z)}{u(z)}=\frac{h'(u(1-z))u'(1-z)}{h(u(1-z))}$$

$$\rightsquigarrow h'(u(1-z))\le\frac{h(u(1-z))}{u(1-z)}$$

$$\Rightarrow \frac{u'(z)}{u(z)}\le\frac{h(u(1-z))}{u(1-z)}\cdot\frac{u'(1-z)}{h(u(1-z))}$$

$$\to \boxed{\frac{u'(z)}{u(z)}\le\frac{u'(1-z)}{u(1-z)}}$$

$z=\frac{1}{2}$ تنها معادله‌ی تساوی را برقرار می‌کند. برای برقراری نا‌تساوی این معادله:

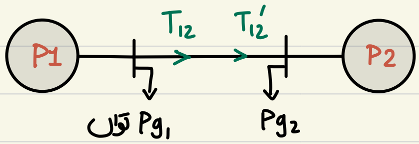
$$u(z)>u(1-z)\ ,\ u'(z)<u'(1-z)$$

$$z^{**} \Leftarrow z>1-z$$

$$z^{**}\ge z^*$$

مثال بازی همکارانه

← مثال 2 : تعادل بین تولیدکنندگان انرژی



فرض می‌کنیم که P1 امکان افزایش تولید را از Pg_1 را داشته باشد و بتواند علاوه بر Pg_1، توان اضافه‌ی T_{12} را تولید کرده و به P2 بفروشد. که با درنظر گرفتن تلفات مقدار T_{12}' به P2 خواهد رسید. این موضوع به نفع P2 می‌تواند باشد چون می‌تواند کمتر از Pg_2 تولید کند.

← مسأله چانه زنی: که ویژگی این نوع بازی‌ها این است که متغیر تصمیم‌گیری لزوماً برای هر بازیگر بصورت جداگانه تعریف نمی‌شود و گاهی متغیر تصمیم‌گیری بازیگرها، یک متغیر است که باید بر سر آن باهم توافق کنند. در این جا T_{12}، T_{12}' این چنین هستند.

• تابع هدف → $R_1 = P_T T_{12}' - (C_1(Pg_1+T_{12}) - C_1(Pg_1))$

قیمت فروش — میزان فروش یا افزایش تولید — هزینه. تابع سود جهت افزایش تولید از Pg_1 به Pg_1+T_{12} نوشته شد.

→ $R_2 = (C_2(Pg_2) - C_2(Pg_2 - T_{12}')) - P_T T_{12}'$

← 2 متغیر P_T و T_{12} یا T_{12}' وجود دارند که باید تعیین شوند.

• Disagreement Cost: $d_1 = d_2 = 0$ ←

if $T_{12}' = 0$ ($T_{12} = 0$) then

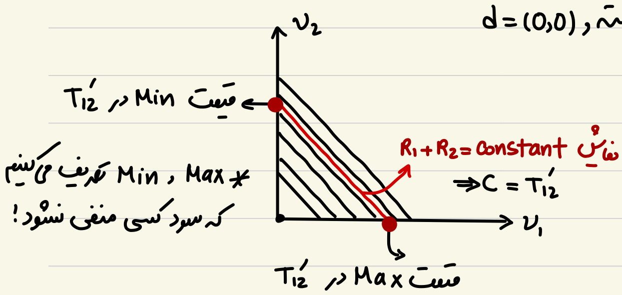
$R_1 = R_2 = 0$ ✓

← قبل از اینکه از Nash Bargaining Solution استفاده کنیم، بررسی می‌کنیم که آیا شکل joint payoff ما convex, compact را دارد یا خیر؟

باید ثابت باشد که تعادل صورت گیرد ⟶ $R_1 + R_2 = f(T_{12}')$

← به‌ازای هر T_{12}'، R_1+R_2 مقداری ثابت می‌شود و تبدیل به یک خط روی نمودار خواهد شد.

← f پیوسته، $d=(0,0)$



با قیود تعیین شده می‌توان فضای joint payoff را داشت که بصورت بالا، فضای compact و convex باشد. ✓

← حل می‌توان از روش Nash Bargaining Solution استفاده کرد.

تابع $L = R_1 \times R_2$ را باید Max کنیم: $L_{max} = \underset{P_T, T_{12}'}{\text{Max}}\ L$

با مشتق گرفتن می‌توان حل کرد: $\frac{\partial L}{\partial P_T} = 0 \ , \ \frac{\partial L}{\partial T_{12}'} = 0$

$$\bullet \ \frac{\partial L}{\partial T_{12}'} = \left(P_T - \frac{\partial C_1(Pg_1+T_{12})}{\partial T_{12}} \cdot \frac{\partial T_{12}}{\partial T_{12}'}\right) R_2 + \left(-P_T - \frac{\partial C_2(Pg_2 - T_{12}')}{\partial T_{12}'}\right) R_1 = 0$$

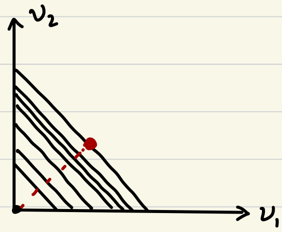
$$\bullet \ \frac{\partial L}{\partial P_T} = T_{12}'\overbrace{\left(-P_T T_{12} + (-C_2(Pg_2 - T_{12}') + C_2(Pg_2))\right)}^{R_2} - T_{12}'\overbrace{\left(P_T T_{12}' + C_1(Pg_1) - C_1(Pg_1 + T_{12})\right)}^{R_1} = 0$$

← با حل دو معادله - دو مجهول بالا به جواب می‌رسیم. فقط دقت کنید که باید نقطه‌ی بدست آمده، Max بودنش را چک کنیم: مشتق مرتبه دوم

چرا؟ مگه جواب بهینه سازی یکتا نیست؟

اثبات وجود یکتایی را در فضای joint payoff کردیم. اما این جا خود payoff ها تابعی از متغیر P_T و T_{12}' هستند و روی $R_1 R_2$ صرفاً بحث نمی‌کنیم. باید concave بودن تابع L را نسبت به P_T، T_{12}' چک کنیم.

$$J = \begin{vmatrix} \frac{\partial^2 L}{\partial P_T^2} & \frac{\partial^2 L}{\partial P_T \partial T_{12}'} \\ \frac{\partial^2 L}{\partial T_{12}' \partial P_T} & \frac{\partial^2 L}{\partial T_{12}'^2} \end{vmatrix} < 0 \rightsquigarrow$$ شرط کافی



اگر می‌خواستیم روی خود فضای joint payoff بحث کنیم نقطه • جواب تعیین می‌شد. نکته این است که آیا یک P_T و T'_{12} وجود دارد که به ما این پاسخ را بدهد؟

← این نکته را از معادلات نیز می‌توان استنتاج کرد.

① منافع تبادل به صورت مساوی تقسیم می‌شود:

از معادله‌ی $\frac{\partial L}{\partial P_T}$، می‌توان فهمید که با حذف T'_{12} ها داریم $R_1=R_2$. واضح بود چون فضای ما متقارن است.

② از معادله‌ی $\frac{\partial L}{\partial P_T}$ بدست می‌آید:

$$P_T=\frac{C_2(Pg_2)-C_2(Pg_2-T'_{12})+C_1(Pg_1+T_{12})-C_1(Pg_1)}{2T'_{12}}$$

$$=\frac{1}{2}\left[\underbrace{\frac{C_2(Pg_2)-C_2(Pg_2-T'_{12})}{T'_{12}}}_{\bar{P}_2 = P2\ \text{هزینه افزایشی}}+\underbrace{\frac{C_1(Pg_1+T_{12})-C_1(Pg_1)}{T'_{12}}}_{\bar{P}_1 = P1\ \text{هزینه افزایشی}}\right]$$

$\to P_T=\frac{1}{2}(\bar{P}_1+\bar{P}_2)$

طبق تعریف، قیمتی است که اگر فروشنده از این قیمت کمتر بفروشد، ضرر خواهد کرد. ← $\bar{P}_1$ را در R_1 جاگذاری می‌کنیم:

$\to R_1=0$ (با فرض $T_{12}=T'_{12}$)

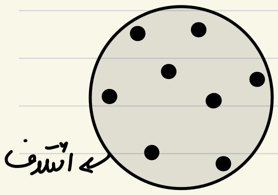
طبق تعریف، $\bar{P}_2$ نیز حداکثر قیمتی است که خریدار به‌ازای آن ضرر نخواهد کرد ← اگر $\bar{P}_2$ را در R_2 جاگذاری کنیم: $R_2=0$

$\to \bar{P}_1 < P_T < \bar{P}_2$ ⇝ Max, Min بین $\bar{P}_2$ و $\bar{P}_1$ قیمتی است که سود بازیگران منفی نشود تا توافق انجام شود.

Session 15-2

بازی‌های ائتلافی

← در این بازی‌ها، به بازیگرها نگاه کلی‌تری داریم.
کاری نداریم که افراد در این ائتلاف
هرکدام چه action ی دارند.
مثل. انتخابات، بازار برق و...



← به‌جای اینکه برای هر فرد تابع سود درنظر بگیریم، یک ارزش برای کل ائتلاف درنظر می‌گیریم و کاری نداریم هر فرد چه استراتژی‌ای دارد.

← مثل ارزش تیم‌های فوتبال

- چگونه ائتلاف تشکیل می‌شود؟
- در چه صورت یک ائتلاف پایدار است؟
- چگونه سود حاصل از یک ائتلاف را بین اعضا تقسیم کنیم؟

← تعریف بازی ائتلافی: $G(I,V)$

بازیگران: $I=\{1,\dots,N\}$

ارزش ائتلاف: $v: 2^N \to \mathbb{R}$ تابع

$S\subseteq I$: ائتلاف

2^N: تمام زیرمجموعه‌های ممکن یک مجموعه‌ی N عضوی

=> ارزش ائتلاف S $= V(S) \to V(\phi)=0$

=> سهم P_i از ائتلاف S $= x_i(S)$ یا $x_i(V(S))$, $i\in S$

$\longrightarrow X(V)=(x_1(V),\dots,x_{|S|}(V))^T \in \mathbb{R}^{|S|}$

$|S|$: تعداد اعضای مجموعه‌ی S

→ مجموع سهم بازیگران ائتلاف S برابر ارزش ائتلاف S

$$\sum_{i=1}^{|S|} x_i(V(S)) = V(S)$$

← 2 فرض داریم: 1- ارزش کل ائتلاف را با یک عدد مثل V(S) مشخص کردیم. 2- ارزش یک ائتلاف قابل تقسیم بین اعضاست، به هر شکل دلخواه.

⇓

Transferable Utility Assumption

Session 16

• سود یا ارزش ائتلاف چگونه بین اعضا تقسیم شود؟

قبلاً سود هر بازیگر به استراتژی خودش و بقیه بستگی داشت. اما این جا چیزی به نام استراتژی نداریم. پس معیار سود هر بازیگر چیست؟ قاعدتاً چون در یک ائتلاف ترکیب اعضا هستند که یک ارزش برای ائتلاف درست کردند، باید بررسی کنیم که بودن یا نبودن یک بازیگر خاص چقدر در این ارزش ائتلاف مهم است. این می‌تواند معیار باشد که چقدر سود به آن بازیگر اختصاص دهیم. این معیار یک نوع معیار fairness و عدل است.

این قانون تقسیم سود باید طوری باشد که این ائتلاف پایدار بماند. به این معنی که دوتا از اعضای این ائتلاف تمایل نداشته باشند که از ائتلاف خارج شوند و یک ائتلاف جداگانه بسازند تا سود بیشتری داشته باشند.

تعریف. Superadditive Game

اگر $S,T\subseteq I$ باشند و $S\cap T=\emptyset$ باشد:

$$V(S\cup T)\geq V(S)+V(T)$$

⇐ نتیجه. اضافه شدن افراد به ائتلاف هرگز سبب کم شدن ارزش ائتلاف نخواهد شد.

* در این نوع بازی ائتلاف سراسری بیشترین ارزش را خواهد داشت ⇐ افراد به تشکیل ائتلاف‌های بزرگتر تشویق می‌شوند.

پس در جواب سوال چگونه ائتلاف تشکیل می‌شود؟ ائتلاف سراسری را در نظر می‌گیریم که سودش در بازی superadditive بیشتر خواهد بود.

• سود یا ارزش چگونه بین اعضا تقسیم شود؟

معیارهای fair و stable

fair ← عادلانه

stable ← کسی انگیزه خارج شدن نداشته باشد.

⇐ fair. تخصیص سود (ارزش) به هر بازیگر باید متناسب با میزان تأثیر آن عضو در ائتلاف

ارزش اینکه بازیگر 1 به تنهایی ائتلاف تشکیل دهد، در یک موضوع خاص → $V(\{1\})=0$

⇐ فرض کنیم بازیگر 2 نیز چنین باشد → $V(\{2\})=0$

⇐ اما اگر بازیگر 1,2 با یکدیگر ائتلاف تشکیل دهند، ارزش این ائتلاف، 1 خواهد شد:

$$V(\{1,2\})=1$$

• سوال 1: آیا می‌توان به میزان اثرگذاری بازیگران، به آنها ارزش دهیم؟

⇐ بازیگر 1 یا 2 می‌گویند: "اگر من در ائتلاف نبودم، ارزش ائتلاف 0 می‌شد. پس اضافه شدن من به ائتلاف به اندازه‌ی 1 ارزش دارد." پس هر دو این استدلال را می‌کنند و اگر بخواهیم به میزان اثرگذاری هرکدام به آنها سود 1 دهیم، قابل انجام نیست. پس به طریقی باید لغت "متناسب بودن" را به "وزن دار کردن" تبدیل کنیم.

پس در این مثال از روش وزن دار کردن استفاده می‌کنیم.

تعریف. اصول موضوعه shapely برای تقسیم ارزش و سود:

1. فرض Transferable Utility:

$$\sum_i x_i(V(I)) = V(I)$$

2. اگر $V(S\cup\{i\})=V(S\cup\{j\})$ برای هر $S\subseteq I$ که شامل بازیگرهای i و j نباشند، در این صورت

$$x_i(V)=x_j(V)$$

اگر ارزش اضافه شدن P_i به هر ائتلاف از I بپیوندد، برابر با ارزش اضافه شدن P_j به ائتلاف S باشد، آنگاه سهم بازیگران i و j از ائتلاف سراسری باهم برابر است.

3. بازیگر خنثی : اگر $V(S)=V(S\cup\{i\})$ ، به ازای هر $S\subseteq I$ که شامل بازیگر i نباشد، آنگاه ، i بازیگر خنثی است :

$$x_i(v)=0$$

4. اگر u و v دوتابع ارزش برای یک ائتلاف باشند، آنگاه داریم :

$$x(u+v)=x(u)+x(v)$$

← یعنی تابع تقسیم سود ، خطی می باشد .

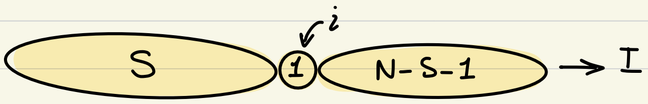
⇐ نحوه تقسیم سود براساس Shapely value :

$\forall i\in I$.

$$x_i(v(I)) = \sum_{S\subseteq I\setminus\{i\}} \frac{|S|!(N-|S|-1)!}{N!}\left(v(S\cup\{i\})-v(S)\right)$$

← مفهوم این رابطه :

بازیگر i به چه نحوی در ترکیب با اعضای دیگر ائتلاف اثر گذار است. یا اینکه اگر این ائتلاف به نحوهای مختلف می خواست شکل بگیرد، اثرگذاری بازیگر i چقدر در این شکل گیری ارزش داشت. پس تمام حالت های ممکن که یک ائتلاف N عضوی بتواند تشکیل شود. به این شکل که ما ابتدا یک ائتلاف S عضوی داشته باشیم و بعد بازیگر i به این ائتلاف اضافه شود و سپس باقی اعضا که N-S-1 عضو دیگر هستند اضافه شده تا مجموعه ی I ، که مجموعه ی سراسری است ، تشکیل شود .



$\frac{|S|!(N-|S|-1)!}{N!}$. این جایگشت نمایانگر تمام حالاتی است که این سه گروه کنار هم چیده شوند .

$(v(S\cup\{i\})-v(S))$: میزان اثرگذاری بازیگر i در پیوستن اش به ائتلاف S : ارزش افزوده

قضیه . این رابطه یک تقسیم سود یکتا $x(v(I))$ بدست می دهد که اصل 4 مذکور را برآورده می کند.

مثال بازی ائتلافی

مثال 1. یک مجلس با 4 حزب A، B، C، D با تعداد اعضای به ترتیب 45، 25، 15، 15 موجود است آنها می خواهند در مورد $100M تصمیم گیری کنند اگر هر 4 حزب را در یک ائتلاف سراسری در نظر بگیریم، سهم هر حزب از این پول چقدر است؟ با توجه به اینکه 51٪ درصد آرا، برای تصویب بودجه لازم است

← با توجه به قانون Shapley value : $N=4$

$$x_A = \frac{1!2!}{4!}\times 3\times(100-0) + \frac{2!1!}{4!}\times 3\times(100-0) = \frac{12}{24}\times 100 = 50$$

$$x_B = \frac{1!2!}{4!}\times(100-0)+\frac{2!1!}{4!}\times(100-0) = \frac{4}{24}\times100 \simeq 16.67$$

$$x_C = \frac{1!2!}{4!}(100-0)+\frac{2!1!}{4!}(100-0)\simeq 16.67$$

$$x_D = x_C$$

$$\Rightarrow \left(50, \frac{100}{6}, \frac{100}{6}, \frac{100}{6}\right)$$

← Stabality : در همین مثال 1، این مفهوم را توضیح می دهیم . بازیگر A, B مجموع سودشان $50+\frac{100}{6}$ می شود. آیا انگیزه ندارد که از ائتلاف سراسری خارج شده و دوتایی یک ائتلاف تشکیل دهند؟ در تقسیم سود به روش shapely ، A, B این انگیزه را دارند که مثلاً بصورت (75،25) ارزش ائتلاف را تقسیم کنند به این ترتیب سود بیشتری بدست آورند .

← در چه صورت تقسیم سود بازیگران به شکلی است که آنها تمایلی به ائتلاف سراسری را نداشته باشند؟

← اگر وتنها اگر نحوه تقسیم سود به گونه ای باشد که پروفایل سود مربوط در مجموعه ای به نام Core (هسته) قرار گیرد.

تعریف. Core یا هسته

بردار x متعلق به مجموعه‌ی core بازی ائتلافی G(I,V) است، اگر و تنها اگر:

$$\forall S\subseteq I \sum_{i\in S} x_i(V(I)) \geq V(S)$$

مفهوم. در صورتی تقسیم سود stable است که اگر هر مجموعه‌ی S را در نظر بگیریم به عنوان یک امکان که اعضای I مجموعه‌ی S را به عنوان ائتلاف جدید تشکیل دهند. پس منظور از S، مجموعه‌ای است که در نظر می‌گیریم که چند بازیگر بخواهند از ائتلاف سراسری خارج شوند و ائتلاف S را باهم تشکیل دهند در صورتی بازیگران انگیزه‌ی تشکیل S را خواهند داشت که ارزش ائتلافی که می‌خواهند تشکیل دهند (V(S)) کمتر مساوی مجموع ارزش هایی باشد که بازیگران الان با تقسیم سود به دست می‌آورند.

* چون ارزش ائتلاف S در هسته، کمتر یا مساوی مجموع ارزش ها در ائتلاف سراسری است، اگر یک نفر بخواهد در این ائتلاف با ارزش مجموع کمتر سود بکند، عده‌ای ضرر خواهند کرد. در این صورت دیگران اصلا حاضر نیستند همچین ائتلافی را تشکیل دهند.
این مفهوم Pareto Optimality است.

← حالت خاص: کسی به تنهایی نخواهد ائتلاف تشکیل دهد:

$$\forall i: x_i(V(I)) \geq V(\{i\})$$

← آیا core همواره وجود دارد؟ آیا core یکتاست؟
خیر! خیر!

مثال 1 را بررسی می‌کنیم. در این مثال core وجود ندارد!

مجموع سود یا ارزش B، C، D: برابر 100 / کمتر از 100

• اگر کمتر از 100 باشد: چون مجموع اعضای B، C، D به 51% می‌رسد، آنها انگیزه دارند که ائتلاف کنند و ارزش {B,C,D} برابر 100 می‌شود. پس می‌توانند یک ائتلاف {B,C,D} تشکیل دهند که بزرگتر از مجموع ارزش هایشان در ائتلاف سراسری (که فرض کردیم کمتر از 100) باشد.

• اگر ارزش هایشان را برابر 100 فرض کنیم، حزب A چون در این حالت ارزش 0 دارد، انگیزه دارد که با حداقل یکی از سایر احزاب ائتلاف کند و مثلاً ارزش {A,D} برابر 100 خواهد شد.

① $x_A(V(I))=0,\quad x_D(V(I))<100$

$\Rightarrow x_A+x_D<100$

② $V(\{A,D\})=100$

مثال 2. همان مثال 1، اما بجای 51%، حد نصاب 80% است.

ادعا می‌کنیم که $(\overset{A}{x},\overset{B}{100-x},\overset{C}{0},\overset{D}{0})$ پایدار است.

تنها ائتلاف هایی که کافی‌ست چک کنیم آیا ممکن است به سود بیشتری برای اعضا منجر شود یا خیر، {A,B,C} یا {A,B,D} است که به 80% می‌رسد.

$S=\{A,B,C\} \rightarrow V(S)=100$

$\rightarrow 100=100$ ✓

$x_A+x_B+x_C=100$

پس انگیزه تشکیل ائتلاف جدید ندارد.

پس اگر همه‌ی پول را بین A، B تقسیم کنیم چون وجودشان برای 80% آرا نیاز است، core است.

تعریف. بازی های convex

بازی ائتلافی $G(I,V)$، محدب است اگر:

$$\forall S,T \subseteq I : V(S\cup T) \geq V(S)+V(T)-V(S\cap T)$$

در super additive، این مقدار، صفر بود.

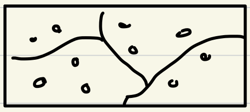
∗ super additive ⟶ convex

قضیه 1. هر بازی convex، core غیرتهی دارد.

قضیه 2. در هر بازی convex، $\underbrace{\text{shapely value}}_{\text{fair}}$ در $\underbrace{\text{core}}_{\text{stable}}$ قرار دارد.

Coalition Formation

ائتلاف سراسری لزوماً core ندارد.



⟵ این دسته بندی چه ویژگی هایی دارد؟

• Pareto Optimality. با نگاه جمعی و دیدگاه همکارانه

یک افراز از این مجموعه داشته باشیم که هیچ افراز دیگری را نتوان تشکیل داد به شکلی که سود همه ی بازیگران بیش تر شود.

افراز: $\Pi=\{S_k\}_{k=1}^{K}$, $I=\{1,\dots,N\}$

ویژگی افراز: اجتماع مجموعه ها، مجموعه ی اصلی شود. اشتراک مجموعه ها، دو به دو، تهی باشد.

$$S_k\subseteq I : \bigcup_{k=1}^{K} S_k = I \ \ \& \ \ S_i\cap S_j=\emptyset,\ i\neq j$$

Pareto Optimality:

افراز Π بهینه ی پارتو است اگر افراز $\Pi'\neq\Pi$ وجود نداشته باشد بطوریکه به بازیگری سود بیش تری برسد بدون اینکه بازیگر دیگری سودش کمتر شود.

• Movement of one player: دیدگاه رقابتی

تعریف 1. Nash stable (NS)

افراز Π، NS است اگر هیچ بازیگری نتواند با جدا شدن از ائتلاف فعلی خود و پیوستن به ائتلاف جدید، سود ببرد.

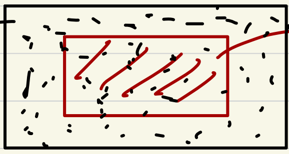
تعریف 2. Individualy stable (IS)

افراز Π، IS است اگر هیچ بازیگری نتواند با جدا شدن از ائتلاف فعلی خود و پیوستن به ائتلاف جدید، سود ببرد درحالیکه به ائتلاف مقصد ضرری نزند.

← کدام یک از IS , NS زیر مجموعه ی دیگری است؟

NS زیر مجموعه ی IS است: NS $\rightarrow$ IS

① وجود دارد فردی که با تغییر ائتلاف خود سود می کند: نقیض NS: ~NS



② وجود دارد فردی که با تغییر ائتلاف سود می کند و به ائتلاف مقصد ضرر نمی زند: ~IS

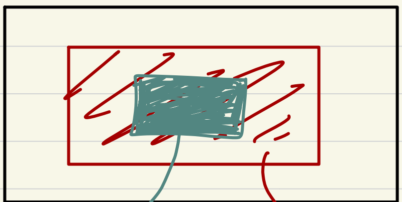
⟸ طبق این تعاریف ② ⊆ ①

~IS → ~NS ⟹ NS → IS

تعریف 3. Contractually IS (CIS)

افراز Π، CIS است اگر هیچ بازیگری نتواند با جدا شدن از ائتلاف فعلی خود و پیوستن به ائتلاف مقصد، سود ببرد؛ درحالیکه نه به ائتلاف مبدأ و نه مقصد ضرر بزند.

① با تغییر ائتلاف خود سود می کند: نقیض NS: ~NS

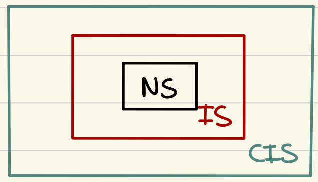


③ فردی با تغییر ائتلاف خود سود می کند و نه به مبدأ و نه به مقصد ضرر می رساند. ~CIS

② وجود دارد فردی که با تغییر ائتلاف سود می کند و به ائتلاف مقصد ضرر نمی زند: ~IS

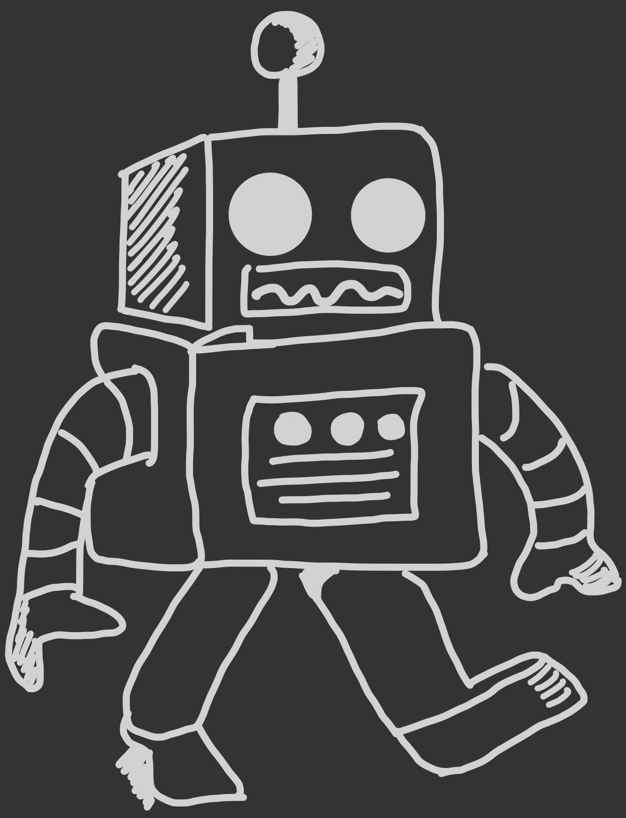
$\Longrightarrow$ ③ $\subseteq$ ② $\subseteq$ ① $\Longrightarrow \sim CIS \rightarrow \sim IS \rightarrow \sim NS$

$\Longrightarrow NS \rightarrow IS \rightarrow CIS$



CHAPTER4

~ Learning Non-cooperative Games



Shiva Shakeri

Learning in Games

• یادگیری در بازی های رقابتی یا غیرهمکارانه

• اطلاعات ناقص: ناآگاهی از توابع هدف رقبا

تاریخچه‌ی تصمیمات رقبا معلوم است!

• تکرار بازی (بازی های تکرار شونده): برای یادگیری نیاز به تکرار!

• بازیگران : $I=\{1,..,N\}$

• توابع هدف: $f_i(x_1,..,x_n)$; $i=1,..,n$

• استراتژی: $x_i \in X_i$; $i=1,..,n$

G

← این بازی در iteration های $t=1,2,..$ انجام می‌شود: G_t

← تاریخچه تصمیمات رقبا : $\{x_{-i}(t)\}_{t=0}^{k-1}$

تصمیم‌گیری در مرحله‌ی $t=K$.

بازیگر تابع هدف خود را می‌داند و می‌خواهد این تابع را در $t=k$، Max کند یا بهترین تصمیم را در مرحله‌ی k بگیرد بصورتی که تابع سودش Max شود. این تابع متأثر از تصمیم سایر بازیگران نیز است. پس اگر بتوان تخمین بزند از اینکه بازیگران دیگر باتوجه به اینکه کارهایی که قبلاً کردند و مرحله‌ی بعد آن چه می‌کنند. پس مسأله در نهایت بهینه سازی خواهد شد.

* فرآیند یادگیری ← تخمین (expectation) + تصمیم‌گیری

$$x_i^*(x_{-i}) = \underset{x_i}{\mathrm{Argmax}}\, f_i(x_i, x_{-i})$$

* معلاً "بهینه سازی" را Best Response درنظر می‌گیریم.

Iteration داریم.

$$x_i^*(t+1) = \underset{x_i}{\mathrm{Argmax}}\, f_i(x_1(t+1), x_2(t+1), \dots, x_{i-1}(t+1), x_i, x_{i+1}(t+1), \dots, x_n(t+1))$$

تخمین بازیگر i از تصمیم بازیگر j در t+1: $x_{ji}(t+1)\,;\ j\neq i$

تخمین Naive یا ساده انگارانه: $x_{ji}(t+1) = x_j(t)$

• بازیگر i فرض می‌کند بازیگر j تصمیم خود در مرحله t را در مرحله t+1 انجام می‌دهد.

• بازیگر i درمرحله t+1 خود را نسبت به تصمیم سایر بازیگران در مرحله t بهینه می‌کند.

$$x_i^*(t+1) = \underset{x_i\in X_i}{\mathrm{Argmax}}\, f_i(x_1(t), x_2(t), \dots, x_{i-1}(t), x_i, x_{i+1}(t), \dots, x_n(t))$$

تخمین Naive + BR ⇐ فرآیند learning

Session18-1

تخمین Naive + BR ⇐ فرآیند learning و بازیگران همگن هستند.

$$\rightarrow x_i(t+1) = \underset{x_i\in X_i}{\mathrm{Argmax}}\, f_i(x_i, x_{-i}(t))$$

← فرآیند یادگیری به کجا ختم می‌شود؟

$$x_i(t+1) = \overset{BR_i}{g_i}(x_1(t), \dots, x_{i-1}(t), x_{i+1}(t), \dots, x_n(t))$$

$$\rightarrow \left.\begin{array}{l} x_1(t+1) = g_1(x_2(t), \dots, x_n(t)) \\ x_2(t+1) = g_2(x_1(t), x_3(t), \dots, x_n(t)) \\ \vdots \\ x_n(t+1) = g_n(x_1(t), \dots, x_{n-1}(t)) \end{array}\right\}$$

سیستم دینامیکی و غیرخطی

سیستمی که مقدارش در لحظه‌ی بعدی وابسته به مقدارش در لحظات قبلی ست، دینامیکی است.

$\Rightarrow x = (x_1, \dots, x_n)^T$, $g = (g_1, \dots, g_n)^T$

$\Rightarrow x(t+1) = g(x(t))$

← نقطه‌ی تعادل معادله دینامیکی یادگیری بالا کجاست؟

نقطه‌ی تعادل معادله دینامیکی: بردار تصمیم‌گیری بازیگران در t و t+1 باهم برابر شود یعنی تصمیمات به یک جا ختم شده و تغییرات دیگری رخ نفی دهد:

$x(t+1) = x(t) \rightarrow \boxed{x = g(x)}$

← نقطه‌ تعادل (نقاط) تعادل دستگاه فوق چه ویژگی ای دارد؟

پس در تعادل داریم:

$$\begin{cases} x_1(\cancel{t}) = g_1(x_2(\cancel{t}),\dots,x_n(\cancel{t})) & BR_1 \\ x_2(\cancel{t}) = g_2(x_1(\cancel{t}),\dots,x_n(\cancel{t})) \\ \vdots \\ x_n(\cancel{t}) = g_n(x_1(\cancel{t}),\dots,x_{n-1}(\cancel{t})) & BR_n \end{cases}$$

→ n معادله، n مجهول

⇒ نقطه‌ی تقاطع BR بازیکنان

نقطه‌ی تعادل معادله دینامیکی یادگیری ≡ نقطه تقاطع BR

≡ نقطه‌ی تعادل نش

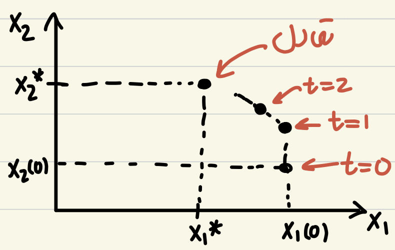
* فرآیند یادگیری که بصورت همگن که با تخمین ساده انگارانه و تصمیم گیری BR انجام می‌گیرد، منجر به معادله‌ی دینامیکی یادگیری می‌شود که نقاط تعادل آن دینامیک یادگیری، همان تعادل نش بازی غیرهمکارانه می‌باشد.

← آیا فرآیند یادگیری که ما یک معادله دینامیکی مدل شد به نقاط تعادل خود (نقاط تعادل نش) همگرا خواهد شد؟

← مثال: $I=\{1,2\}$ ،

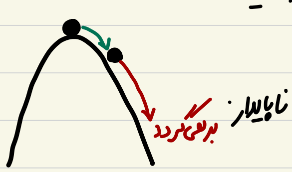
تعادل دینامیکی، $\left.\begin{array}{l} x_2 = g_2(x_1) \\ x_1 = g_1(x_2) \end{array}\right\} \Rightarrow (x_1^*, x_2^*)$: NASH

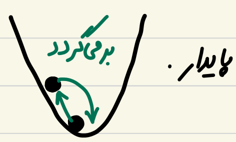
$$\left.\begin{array}{l} x(t+1) = g(x(t)) \\ x(t) = (x_1(t), x_2(t))^T \\ g = (g_1, g_2)^T \end{array}\right\}$$

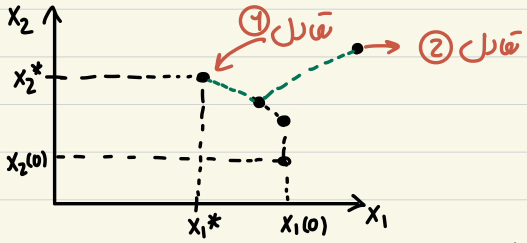


آیا به نقطه‌ی تعادل می‌رسند؟

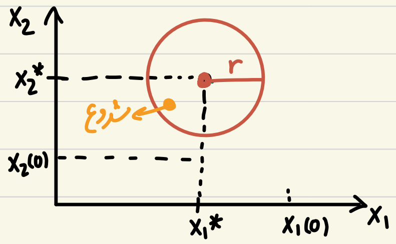
فرآیند یادگیری تضمین نمی‌کند که به نقطه تعادل می‌رسد. در سیستم‌ها نقطه تعادل پایدار و ناپایدار داریم.







← به کدام تعادل همگرا می‌شود؟ پایدار است؟



← آیا یک دایره به شعاع r حول یک تعادل نش وجود دارد که اگر شرایط اولیه از نقطه‌ی • شروع شود، بتوان تضمین کرد که تصمیم گیری شان به نقطه‌ی تعادل نش ختم شود؟

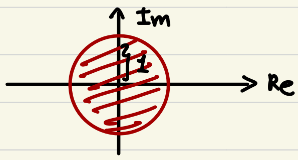
در کدام محدوده پایدار است؟

$$\begin{cases} x(t+1) = g(x(t)) \xrightarrow{\text{خطی}} x(t+1) = Ax(t) \\ x^* = (x_1^*, x_2^*)^T \quad \Rightarrow x(t)=0 \rightarrow \text{تعادل} \end{cases}$$

⇒ $x = Ax$. نقطه‌ی تعادل صفر

eigen values of A : $(\lambda_1, \dots, \lambda_n)$

← نقطه تعادل سیستم دینامیکی $x(t+1)=Ax(t)$ درصورتی پایدار (مجانبی) است که همه‌ی مقادیر ویژه‌ی ماتریس A داخل دایره‌ی واحد قرار گیرند؛



مثال یادگیری بازی

← مثال ۱ . بازی کورنات تکرارشونده با اطلاعات ناقص

• تابع هدف $\Pi_i = \lambda q_i - C_i(q_i)$, $i = 1,2$

قیمت . $\lambda = \lambda_0 - \alpha(q_1+q_2)$

cost : $C_i(q_i) = \alpha q_i^2 + b_i q_i + C_i$

← بدست آوردن مدل یادگیری (تخمین Naive و تصمیم‌گیری BR)

$$\Pi_1 = [\lambda_0 - \alpha(q_1+q_2)]q_1 - a_1 q_1^2 - b_1 q - C_1$$

$$BR_1 = \frac{\partial \Pi_1}{\partial q_1} = [\lambda_0 - \alpha(q_1+q_2)] - \alpha q_1 - 2a_1 q_1 - b_1 = 0$$

$$\rightarrow BR_1(q_2) = \frac{\lambda_0 - b_1 - \alpha q_2}{2(\alpha+a_1)} \ , \ \frac{\partial^2 \Pi_1}{\partial q_1^2} < 0 \rightarrow \text{Max}$$

symmetric $\Rightarrow BR_2 = \dfrac{\lambda_0 - b_2 - \alpha q_1}{2(\alpha+a_2)}$

$$\rightarrow \begin{cases} q_1(t+1) = \dfrac{\lambda_0 - b_1 - \alpha q_2(t)}{2(\alpha+a_1)} \\ q_2(t+1) = \dfrac{\lambda_0 - b_2 - \alpha q_1(t)}{2(\alpha+a_2)} \end{cases}$$ مدل یادگیری

← بدست آوردن نقاط تعادل (Nash) : تقاطع BR ها

تغییر متغیر

$A_1 = \dfrac{\lambda_0 - b_1}{2(\alpha+a_1)}$, $B_1 = \dfrac{\alpha}{2(\alpha+a_1)}$

$A_2 = \dfrac{\lambda_0 - b_2}{2(\alpha+a_2)}$, $B_2 = \dfrac{\alpha}{2(\alpha+a_2)}$

$$\rightarrow \begin{aligned} q_1(t+1) &= A_1 - B_1 q_2(t) \\ q_2(t+1) &= A_2 - B_2 q_1(t) \end{aligned} \quad *$$

تعادل $\rightarrow q_1(t+1) = q_1(t)$ & $q_2(t+1) = q_2(t)$

$$\left.\begin{aligned} BR_1 &\rightarrow q_1 = A_1 - B_1 q_2 \\ BR_2 &\rightarrow q_2 = A_2 - B_1 q_1 \end{aligned}\right\} \xrightarrow{\text{تقاطع}} (q_1^*, q_2^*)$$: تعادل نش و معادله دینامیکی یادگیری

تقاطع =>

$$(q_1^*, q_2^*) = \left(\frac{A_1 - A_2 B_1}{1 - B_1 B_2} \ , \ \frac{A_2 - A_1 B_2}{1 - B_1 B_2}\right)$$

← بررسی همگرایی تصمیم بازیکنان به تعادل نش دینامیک یادگیری (نش)

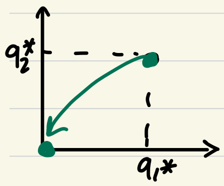
تغییر متغیر $\rightarrow q_1' = q_1 - q_1^*$, $q_2' = q_2 - q_2^*$

این تغییر متغیر را به معادله‌ی * اعمال می‌کنیم.

$q_1'(t+1) + q_1^* = A_1 - B_1(q_2'(t) + q_2^*)$ — $A_1 - B_1 q_1^*$

$q_2'(t+1) + q_2^* = A_2 - B_2(q_1'(t) + q_1^*)$ — $A_2 - B_2 q_2^*$

$$\Rightarrow \boxed{\begin{aligned} q_1'(t+1) &= -B_1 q_2'(t) \\ q_2'(t+1) &= -B_2 q_1'(t) \end{aligned}} \quad **$$



$$\rightsquigarrow \begin{pmatrix} q_1'(t+1) \\ q_2'(t+1) \end{pmatrix} = \underbrace{\begin{pmatrix} 0 & -B_1 \\ -B_2 & 0 \end{pmatrix}}_{A} \begin{pmatrix} q_1'(t) \\ q_2'(t) \end{pmatrix}$$

$\rightsquigarrow q'(t+1) = A q'(t) \longrightarrow$ فرم استاندارد

صفر نقطه‌ی تعادل است.

می‌توان مقادیر ویژه‌ی A را بررسی کرد که آیا این معادله دینامیکی پایدار است یا خیر:

$$\det(\lambda I - A) = 0 \rightarrow \det\left(\begin{bmatrix} \lambda & B_1 \\ B_2 & \lambda \end{bmatrix}\right) = 0$$

$$\rightarrow \lambda^2 - B_1 B_2 = 0 \rightarrow \lambda = \pm\sqrt{B_1 B_2}$$

داشتیم: $B_2 = \dfrac{\alpha_{>0}}{2(\alpha+a_2)_{>0}}$, $B_1 = \dfrac{\alpha_{>0}}{2(\alpha+a_1)_{>0}}$

$\Rightarrow 0 < B_2 < 0.5 \quad \Rightarrow 0 < B_1 < 0.5$

$$\Rightarrow 0 < \sqrt{B_1 B_2} < \frac{1}{2} \Rightarrow 0 < |\lambda| < 0.5$$

* هر دو مقدار ویژه‌ی A، درون دایره‌ی واحد قرار دارند.

پس مبدأ $(0,0)$ برای ** پایدار مجانبی است و بنابراین (q_1^*, q_2^*) برای معادله‌ی دینامیکی * پایدار مجانبی است $\Leftarrow$

اگر بازیکنان از هر شرایط اولیه‌ای شروع به فرآیند یادگیری کنند در نهایت به نقطه‌ی تعادل نش بازی غیرهمکارانه همگرا می‌شوند.

مرور یادگیری در بازی

• بازی های تکرار شونده با اطلاعات ناقص (غیر همکارانه)

• اجزای بازی:

- بازیگران ← $I=\{1,\dots,n\}$
- تابع هدف بازیگر i که تابعی از تصمیم خود و سایر بازیگران است ←

$$f_i(x_1,\dots,x_n)$$

- مجموعه‌ی استراتژی هر بازیگر ← $x_i \in X_i$
- مراحل متوالی بازی ← $t=1,2,\dots$
- اطلاعات موجود ← تاریخچه‌ی تصمیمات بازیگران
- اطلاعات ناموجود ← توابع هدف رقبا و تصمیم آنها در مرحله‌ی جاری بازی

• راه حل: تخمین تصمیم رقبا + تصمیم گیری بهینه

Naive. $x_j(t+1)=x_j(t)$

BR. $x_i(t+1)=\underbrace{g_i(x_{-i})}_{BR_i}$

• مدل دینامیک یادگیری ← $X(t+1)=g(X(t))$

* تعادل مدل دینامیک یادگیری ≡ تعادل نش

← تابع g یا BR لزوماً خطی نیست. در این صورت از تقریب خطی استفاده می‌کنیم

تقریب خطی: حول یک نقطه براساس بسط تیلور

$$f(x)=f(x_0)+\left(\frac{\partial f(x)}{\partial x}\right)^T(x-x_0)+\frac{1}{2}(x-x_0)^T\frac{\partial^2 f(x)}{\partial x^2}(x-x_0)$$

$$+\dots(\text{Higher order terms})$$

$X=(x_1,\dots,x_n)^T \rightarrow$ شبیه بردار تصمیم گیری بازیگران

$$\rightarrow \frac{\partial f}{\partial x}=\left(\frac{\partial f}{\partial x_1},\dots,\frac{\partial f}{\partial x_n}\right)^T, \quad X-X_0=\begin{pmatrix}x_1-x_0\\x_2-x_0\\\vdots\\x_n-x_0\end{pmatrix}$$

$$\rightarrow \frac{\partial^2 f}{\partial x^2}=\begin{pmatrix}\frac{\partial^2 f}{\partial x_1^2} & \cdots & \frac{\partial^2 f}{\partial x_n \partial x_1}\\ \vdots & \ddots & \vdots\\ \frac{\partial^2 f}{\partial x_1 \partial x_n} & \cdots & \frac{\partial^2 f}{\partial x_n^2}\end{pmatrix}$$

← استفاده از تقریب خطی:

یک دینامیک یادگیری داریم که در حالت کلی غیرخطی است

• $X(t+1)=g(X(t))$, $g=[\underset{BR_1}{g_1}\ \cdots\ \underset{BR_n}{g_n}]$

• $X=(x_1,\dots,x_n)^T$: بردار تصمیم گیری بازیگران

حال تابع g را که BR است، برای بازیگران خطی می‌کنیم:

⇐ حول x_0، بسط تیلور را می‌نویسیم (تا مرتبه 1)

$$\begin{cases}x_1(t+1)=g_1(X(t))\\ \vdots\\ x_n(t+1)=g_n(X(t))\end{cases}$$

نقطه‌ی x_0 در واقع نقطه‌ی تعادل نش باید باشد؛ چون در نهایت می‌خواهیم بررسی کنیم آیا حول این نقطه همگرا می‌شود یا خیر

$x_0=x_e \rightsquigarrow NE$

$$\rightsquigarrow x_1(t+1)=g_1(x_e)+\left(\frac{\partial g_1(x)}{\partial x}\Big|_{x=x_e}\right)^T(X(t)-x_e)$$

$$\vdots$$

$$x_n(t+1)=g_n(x_e)+\left(\frac{\partial g_n(x)}{\partial x}\Big|_{x=x_e}\right)^T(X(t)-x_e)$$

$$\Rightarrow \begin{pmatrix}x_1(t+1)\\\vdots\\x_n(t+1)\end{pmatrix}=\begin{pmatrix}g_1(x_e)\\\vdots\\g_n(x_e)\end{pmatrix}+\begin{pmatrix}\left(\frac{\partial g_1}{\partial x}\right)^T\\\vdots\\\left(\frac{\partial g_n}{\partial x}\right)^T\end{pmatrix}_{x=x_e}(X(t)-x_e)$$

$$\Rightarrow X(t+1)=g(x_e)+\left(\frac{\partial g}{\partial x}\right)^T_{x=x_e}(X(t)-x_e)$$

← در نقطه‌ی تعادل: $X(t+1)=g(X(t))=X(t)$

حل: $g(x)=x \rightarrow x_e$

پس x_e در این معادله صدق می‌کند: $g(x_e)=x_e$

$$\rightarrow X(t+1)-x_e=\left(\frac{\partial g}{\partial x}\right)^T\Big|_{x=x_e}(X(t)-x_e)$$

← خطی سازی شد.

* گفتیم که تقابل داریم. تعادل غیرصفر سیستم را، به صفر شیفت دهیم. چون برای سیستم استاندارد خطی، که نقطه‌ی تعادلش صفر است، روشی داشتیم که همگرایی را به کمک مقادیر ویژه‌ی ماتریس سیستم، بررسی کنیم.

← عمل شیفت به صفر:

$$X(t+1) - X_e = \left(\frac{\partial g}{\partial x}\right)^T\Big|_{x=x_e} (X(t) - X_e)$$

تغییر متغیر
$$\begin{cases} Y(t) = X(t) - X_e \\ Y(t+1) = X(t+1) - X_e \end{cases}$$

$$\Rightarrow Y(t+1) = \underbrace{\left(\frac{\partial g}{\partial x}\right)^T\Big|_{x=x_e}}_{} Y(t)$$

$$\begin{pmatrix} \frac{\partial g_1}{\partial x_1} & \cdots & \frac{\partial g_1}{\partial x_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_n}{\partial x_1} & \cdots & \frac{\partial g_n}{\partial x_n} \end{pmatrix}\Bigg|_{x=x_e} = J$$

ژاکوبین

$$\Rightarrow Y(t+1) = J\, Y(t)$$ → سیستم خطی استاندارد

از روی مقادیر ویژه‌ی ماتریس J همگرایی را بررسی می‌کنیم.

↲ آیا Y به صفر همگرا می‌شود؟ در این صورت $X(t) \leftarrow X_e$

بستگی دارد به مقادیر ویژه‌ی ماتریس J. اگر همه‌ی مقادیر ویژه، داخل دایره‌ی واحد بودند، آنگاه $Y(t)$ به 0 همگرا می‌شود. که معادل همگرایی تصمیم بازیگران ($X(t)$) به تعادل نش (X_e) است.

مدل تصمیم‌گیری بازیگران "لزوماً" BR نیست. چرا؟

بعضی مواقع بازیگران اطلاعات ناقص دارند، حتی نسبت به تابع هدف خودشان. یا خیلی دقیق فکر نمی‌کنند و "کاملاً" rational عمل نمی‌کنند.

Bounded Rational

در این حالت بازیگر نمی‌تواند دقیق تابع هدفش را تشکیل داده و بر اساس تصمیم‌گیری کند. "مثلاً" "صرفاً" دقت می‌کنند چقدر تصمیمشان سودآور است و اگر سودآور باشد، در جهت سودآوری آن حرکت می‌کنند و برعکس در جهت کاهش ضرر حرکت می‌کنند. ← شبیه gradient

← مدل یادگیری = تخمین + تصمیم‌گیری

gradient

← چرا روش گرادیان، bounded rational است؟

طبق توضیحات بالا، بازیگر i تصمیم خود را این طوری update می‌کند:

$$x_i(t+1) = x_i(t) + \alpha \frac{\Delta f_i}{\Delta x_i}$$

بازیگر i میگه من دفعه‌ی پیش تصمیم x_i را گرفتم. حال در چه جهتی تصمیمش را تغییر دهد؟ "اگر قبلاً" x_i را زیاد کرده و f_1 نیز زیاد شده ($\frac{\Delta f_i}{\Delta x_i} > 0$)، در این صورت x_i را در مرحله‌ی بعد ($t+1$) افزایش می‌دهد. اگر $\frac{\Delta f_i}{\Delta x_i} < 0$ باشد، یعنی x_i را کمتر کن.

⇐ مدل دقیق: $x_i(t+1) = x_i(t) + \alpha \frac{\partial f_i(x_i, x_{-i}(t))}{\partial x_i}$ ← Naive

← تعادل این سیستم کجاست؟ $i = \{1, \dots, n\}$

مدل یادگیری: تخمین Naive + تصمیم‌گیری گرادیان سود

* تعادل: $\forall i \in I : x_i(t+1) = x_i(t)$

$$x_i(t+1) = x_i(t) \longrightarrow \frac{\partial f_i(x_i, x_{-i})}{\partial x_i} = 0$$

↲ تقاطع BR به ما می‌دهد (به شرط تقعر و مشتق‌پذیری توابع).

↲ تعادل نش!!!

• Gradient Based Dynamics:

$$x_i(t+1) = x_i(t) + \alpha \frac{\Delta f_i}{\Delta x_i} \rightarrow \frac{\partial f_i}{\partial x_i}$$

• Modified Gradient Method:

$$x_i(t+1) = x_i(t) + \alpha\, x_i(t) \frac{\partial f_i(x_i, x_{-i}(t))}{\partial x_i}$$

α → learning rate

معمولا: ممکن است برخی از agentها گام‌های ثابتی برای update نداشته باشند. یعنی گام update، براساس اینکه x_i بزرگتر است یا کوچک، بزرگ یا کوچک شود. پس اگر x_i در مقادیر بزرگ ثبت شود، معنی ندارد که با گام‌های کوچک آن را update کنیم، با اینکار گام‌ها نیز متناسب با x_i می‌باشند.

* در مدل‌های تصمیم‌گیری، به دنبال مدل‌های شبیه به رفتار آدم‌ها هستیم

← تعادل Modified Gradient کجاست؟

* تعادل: $\forall i \in I: \; x_i(t+1) = x_i(t)$

$$\underbrace{\alpha x_i(t)}_{=0} \; \underbrace{\frac{\partial f_i(x_i, x_{-i}(t))}{\partial x_i}}_{=0} = 0$$

$=0$ یا $=0$ ⇝ برای $i \in I$، هرکدام از این‌ها ممکن است تعادل شود.

← 2^n حالت برای n بازیگر تعادل ممکن است باشد.

← فقط یکی از این‌ها تعادل نش است:

$$\forall i \in I: \; \frac{\partial f_i(x_i, x_{-i}(t))}{\partial x_i} = 0$$ → برای توابع مقعر و مشتق پذیر

مثال یادگیری در بازی

← مثال 1: بازی کورنات تکرار شونده

⇐ درحالتی که یک بازیگر ($P2$) با روش BR و تخمین Naive،
بازیگر دیگر ($P1$) با روش گرادیان اصلاح شده (Modified) و تخمین Naive یادگیری دارند

← در بازی cournot:

• قیمت: $P = a - bQ$

• تولید: $Q = q_1 + q_2$

• تابع سود: $\Pi_1 = q_i(a - bQ) - c_1 q_i$ = درآمد − هزینه تولید

• Cost تابع: $C_1(q_i) = c_1 q_i$

← 1- مدل یادگیری دو بازیگر:

P1 Modified Gradient + Naive →

$$q_1(t+1) = q_1(t) + \alpha q_1(t) \frac{\partial \Pi_1(t)}{\partial q_1(t)}$$

P2: BR + Naive →

$$q_2(t+1) = BR_2(q_1(t))$$

$$\leadsto \frac{\partial \Pi_i}{\partial q_i} = a - c_i - 2bq_i - bq_j \qquad i,j = 1,2,\; i \neq j$$

$$\Rightarrow \frac{\partial \Pi_2}{\partial q_2} = a - c_2 - 2bq_2 - bq_1 = 0$$

$$\Rightarrow BR_2(q_1) = \frac{a - c_2 - bq_1}{2b}$$

← مدل دینامیکی یادگیری بازیگران.

$$\begin{cases} q_1(t+1) = q_1(t) + \alpha q_1(t)\left(a - c_1 - 2bq_1(t) - bq_2(t)\right) \\ q_2(t+1) = \dfrac{a - c_2 - bq_1(t)}{2b} \end{cases}$$

← 2- بررسی نقاط تعادل:

* تعادل: $q_1(t+1) = q_1(t)$, $q_2(t+1) = q_2(t)$

$$\begin{cases} \cancel{q_1} = \cancel{q_1} + \alpha q_1 (a - c_1 - 2bq_1 - bq_2) \\ q_2 = \frac{1}{2b}(a - c_2 - bq_1) \end{cases} \Longrightarrow$$

$$\begin{cases} \alpha q_1 \underbrace{(a - c_1 - 2bq_1 - bq_2)}_{\frac{\partial \Pi_1}{\partial q_1}} = 0 \\ \underbrace{(a - c_2 - bq_1 - 2bq_2)}_{\frac{\partial \Pi_2}{\partial q_2}} = 0 \end{cases}$$ → مقعر و مشتق پذیر

← صفر بودن این دو جمله تعادل معادله دینامیکی نش است.

← اگر در جمله‌ی اول $q_1 = 0$ باشد نیز تعادل برقرار است و با جاگذاری در جمله‌ی دوم:

$$q_1 = 0 \rightarrow q_2 = \frac{a - c_2}{2b}$$

$$\Rightarrow \left(q_1 = 0,\; q_2 = \frac{1}{2b}(a - c_2)\right) \rightarrow$$ تعادل ①

تعادل ② → $\left(q_1^*=\frac{a+c_2-2c_1}{3b}\,,\,q_2^*=\frac{a+c_1-2c_2}{3b}\right)$

↳ تعادل نش بازی کورنات که تعادل دینامیکی نیز است.

قیمت باید مثبت باشد. ⇜ $\left|\begin{array}{l}2c_1-c_2<a\\2c_2-c_1<a\end{array}\right.$. شرط feasible

Session 20-2

← 3- همگرایی تعادل ها.

مدل دینامیکی یادگیری بازیگران.

$$\begin{cases} q_1(t+1)=q_1(t)+\underbrace{\alpha q_1(t)\left(a-c_1-2bq_1(t)-bq_2(t)\right)}_{f_1(q_1,q_2)} \\[2ex] q_2(t+1)=\underbrace{\dfrac{a-c_2-bq_1(t)}{2b}}_{f_2(q_1,q_2)} \end{cases}$$

⇝ $q=(q_1,q_2)^T$ ، $f=(f_1,f_2)^T$

$$\begin{cases} q_1(t+1)=f_1(q_1(t),q_2(t)) \\ q_2(t+1)=f_2(q_1(t),q_2(t)) \end{cases} \Rightarrow q(t+1)=f(q(t))$$

تعادل اول: $E_0=\left(0\,,\,\frac{a-c_2}{2b}\right)$

تعادل نش: $E_*=(q_1^*,q_2^*)=\left(\frac{a+c_2-2c_1}{3b}\,,\,\frac{a+c_1-2c_2}{3b}\right)$

خطی سازی ← $\tilde{q}(t+1)=J\,\tilde{q}(t)$

↳ ژاکوبین ⇜ $J=\begin{bmatrix}\frac{\partial f_1}{\partial q_1} & \frac{\partial f_1}{\partial q_2}\\ \frac{\partial f_2}{\partial q_1} & \frac{\partial f_2}{\partial q_2}\end{bmatrix}$

$$J=\begin{bmatrix}\frac{\partial f_1}{\partial q_1} & \frac{\partial f_1}{\partial q_2}\\ \frac{\partial f_2}{\partial q_1} & \frac{\partial f_2}{\partial q_2}\end{bmatrix}=\begin{bmatrix}1+\alpha(a-4bq_1-bq_2-c_2) & -\alpha bq_1\\ -\frac{1}{2} & 0\end{bmatrix}$$

← تعادل E_0:

$$J(E_0)=\begin{bmatrix}1+\frac{\alpha}{2}(a-2c_1+c_2) & 0\\ -\frac{1}{2} & 0\end{bmatrix}$$

مقادیر ویژه . $\det(\lambda I-J(E_0))=0$ $\quad a>2c_1-c_2$

$\Rightarrow \lambda_1=0\,,\ \lambda_2=\underbrace{1+\frac{\alpha}{2}\underbrace{(a-2c_1+c_2)}_{>0}}_{>1}$

↳ داخل دایره واحد ✓ (λ_1)

↳ خارج دایره ی واحد (λ_2)

← پس تعادل E_0 ، ناپایدار خواهد شد.

← تعادل E_*:

$$J(E_*)=\begin{bmatrix}1-2\alpha bq_1^* & -\alpha bq_1^*\\ -\frac{1}{2} & 0\end{bmatrix}$$

مقادیر ویژه: $\det(\lambda I-J(E_*))=0$

↳ The Jury Test: شرایطی که همه ی مقادیر ویژه، درون دایره واحد

$P(\lambda)=\lambda^2-Tr(J)\lambda+Det(J)$

$$\left.\begin{array}{l} ①\ Det(J)<1 \\ ②\ 1-Tr(J)+Det(J)>0 \\ ③\ 1+Tr(J)+Det(J)>0 \end{array}\right\}$$ The Jury Test برای مرتبه 2

$\det(J(E_*))=-\frac{1}{2}\alpha bq_1^*\,,\ Tr(J(E_*))=1-2\alpha bq_1^*$

① $\frac{1}{2}\alpha bq_1^*<1$

② $1-(1-2\alpha bq_1^*)+(-\frac{1}{2}\alpha bq_1^*)>0$

$\rightarrow \frac{3}{2}\alpha bq_1^*>0 \rightarrow \alpha>0, b>0, q_1^*>0$ ✓

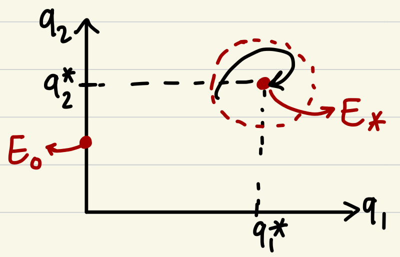
③ $1+(1-2\alpha bq_1^*)+(-\frac{1}{2}\alpha bq_1^*)>0$

$\rightarrow 2-\frac{3}{2}\alpha bq_1^*>0 \rightarrow \alpha bq_1^*<\frac{4}{5}$

$①\cap③ \rightarrow (\alpha bq_1^*<2)\cap(\alpha bq_1^*<\frac{4}{5}) \rightarrow \alpha bq_1^*<\frac{4}{5}$

$q_1^*=\frac{a+c_2-2c_1}{b} \rightarrow \alpha<\frac{12}{5}\left(\frac{1}{a+c_2-2c_1}\right)$ ✓

↳ با این شرط learning rate، به تعادل نش همگرا می شویم!



مدل یادگیری ← مدل تخمین + مدل تصمیم گیری

- مدل تصمیم گیری: BR / Gradient / Modified Gradient
- مدل تخمین: Naive

مدل تخمین تطبیقی (Adaptive expectation).

$$x_{ji}(t+1) = x_{ji}(t) + \alpha\left(x_j(t) - x_{ji}(t)\right)$$

مفهوم: عامل (agent) i، تخمینش از عامل j، در مرحله‌ی $t+1$، برحسب تخمین قبلی $(x_{ji}(t))$، به این شکل update می‌کند: با گام های $\alpha(x_j(t) - x_{ji}(t))$،

خطای تخمین بازیگر i از j ← $e_{ji}(t) = x_j(t) - x_{ji}(t)$

در مرحله‌ی t. اگر $e_{ji}(t)>0$ یعنی تصمیم واقعی بازیگر j در لحظه‌ی t، بزرگتر از تخمین بازیگر i از آن تصمیم باشد، که نشان می‌دهد بازیگر i دارد، کمتر از حالت واقعی تخمین می‌زند. بنابراین تخمین قبلی‌اش را به اندازه‌ی گام تعیین شده افزایش می‌دهد؛ و برعکس برای $e_{ji}<0$

از تجربه‌ی خطای تخمین در مرحله‌ی قبل، برای تخمین در مرحله‌ی بعد استفاده می‌کنند.

← ضریب $0 \leq \alpha \leq 1$:

- $\alpha = 0$ ← $x_{ji}(t+1) = x_{ji}(t)$: تخمین ثابت
- $\alpha = 1$ ← $x_{ji}(t+1) = x_j(t)$. تخمین Naive

تخمین Naive حالت خاصی از این تخمین است.

مدل یادگیری براساس روش تصمیم‌گیری BR، تخمین تطبیقی:

$$BR: \quad x_i(t+1) = \underbrace{g_i}_{BR_i}\left(x_{1i}(t+1), \dots, x_{ni}(t+1)\right)$$

$$\underbrace{x_{ji}(t+1),\ j \neq i}_{\to\, x_j}$$

← 1. تشکیل معادلات دینامیکی

$$x_i(t+1) = g_i\left(x_{1i}(t+1), \dots, x_{ni}(t+1)\right) \quad \} \text{ معادله } 1$$

$$\left.\begin{aligned} \Rightarrow\ & x_{1i}(t+1) = x_{1i}(t) + \alpha_{1i}\left(x_1(t) - x_{1i}(t)\right) \\ & \vdots \\ & x_{ni}(t+1) = x_{ni}(t) + \alpha_{ni}\left(x_n(t) - x_{in}(t)\right) \end{aligned}\right\} \ n-1 \text{ معادله تخمین}$$

مجموعاً n معادله‌ی دینامیکی برای هر بازیگر ← کلاً n^2 معادله‌ی دینامیکی

← 2. بررسی تعادل مدل دینامیک یادگیری.

- در تعادل تخمین ← $x_{ji}(t+1) = x_{ji}(t)$

$$\Rightarrow \alpha\left(x_j(t) - x_{ji}(t)\right) = 0 \rightarrow x_j(t) = x_{ji}(t)$$

تخمین بازیگر i از بازیگر j در لحظه‌ی t، برابر با مقدار واقعی تصمیم بازیگر j در لحظه‌ی t می‌شود.

- در تعادل معادله BR → $x_i(t+1) = x_i(t)$

از کنار هم قرار دادن این دو تعادل:

$$x_i(t) = \underbrace{g_i}_{BR_i}\left(x_1(t), \dots, x_n(t)\right); \quad i = 1, \dots, n$$

بازهم به تقاطع BR ها می‌رسیم ⇐ تعادل نش می‌دهد.

$$x^* = (x_1^*, \dots, x_n^*)$$

* در تعادل سیستم دینامیکی خواهیم داشت:

$$\text{تعادل}.\ (x^*, x^*, \dots, x^*) \Rightarrow n^2 \text{ معادله}$$

$$\begin{array}{l|ccc} \text{تعادل تصمیم} \Leftarrow & x_1^* & x_2^* & x_n^* \\ \text{تعادل تخمین} & \begin{Bmatrix} x_2^* \\ \vdots \\ x_n^* \end{Bmatrix} & \begin{Bmatrix} x_1^* \\ \vdots \\ x_n^* \end{Bmatrix} & \begin{Bmatrix} x_1^* \\ \vdots \\ x_{n-1}^* \end{Bmatrix} \end{array}$$

از کنار هم قرار دادن n^2 معادله:

$$\Rightarrow X(t+1) = F(X(t))$$

برای بررسی همگرایی. با خطی سازی معادلات حول نقطه‌ی تعادل (X^*)، ژاکوبین F را تشکیل می‌دهیم و آنرا بررسی می‌کنیم.

Session 21-2

مدل یادگیری گسسته

برای بازی های تکرار شونده با اطلاعات ناقص

Fictitious Play

اساس ساختار این مدل نیز براساس تخمین و تصمیم گیری است.

بازی استراتژیک ← $G(I, (S_i)_{i\in I}, (u_i)_{i\in I})$
که تکرار می‌شود $t=1,2,\dots$

← تخمین. نوعی تخمین مبتنی بر هیستوگرام؛ یعنی هر بازیکن فرکانس تصمیمات بقیه را در مراحل قبل درنظرگرفته و براساس آن، یک تابع توزیع را به آن fit می‌کند

- $\eta_i^t(s_{-i}) \rightarrow$ تعداد مشاهده استراتژی s_{-i} توسط عامل i تا زمان t

  $\hookrightarrow \eta_i^t: S_{-i} \to \mathbb{N}$

- $\mu_i^t(s_{-i}) = \dfrac{\eta_i^t(s_{-i})}{\sum_{s'_{-i}\in S_{-i}} \eta_i^t(s'_{-i})}$ → فرکانس مشاهده یک استراتژی مشخص: احتمال عمل s_{-i}

← تصمیم گیری براساس BR:

$$s_i^* = \underset{s_i\in S_i}{\mathrm{ArgMax}}\,(u_i(s_i, \mu_i^t(s_{-i})))$$

خروجی، استراتژی pure است.

شبیه بحث استراتژی Mixed، بازیکن i باید سود expected را نسبت به بردار احتمالی که از بقیه تخمین زده، حساب کند. پس بازیکن i باید برای تمام S_i های ممکن، سود expected خود را محاسبه کند.

← چرا تخمین بصورت استراتژی مخلوط است؟

چون عدم اطلاع کامل وجود دارد و نسبت به تصمیم بقیه، عدم قطعیت داریم ← بصورت یک بردار احتمال، احتمال انجام اعمال مختلف را تخمین می‌زنیم

← چرا استراتژی خودمان را مخلوط درنظر نمی‌گیریم؟

چون در بحث learning "واقعاً" می‌خواهیم آن عمل را انجام دهیم و "صرفاً" نمی‌خواهیم یک تعادل بدست آوریم

← فرض تخمین در Fictitious play:

استراتژی رقبا از یک توزیع پایا (stationary) پیروی می‌کند. یعنی اگر "واقعاً" رقبای من یک استراتژی مخلوط را درنظر گرفته باشند که احتمال انتخاب هر استراتژی مشخص است و براساس آن تصمیم می‌گیرند. اگر من بی نهایت بار، تصمیمات آنها را مشاهده کنم. احتمالات مربوط به هر استراتژی رقیب را می‌توانم بدست آورم. پس فرکانس مشاهدات، وقتی مشاهدات به بی نهایت میل می‌کند، تخمین درستی را از احتمال می‌دهد.

مثال مدل یادگیری گسسته

← مثال 1.

P1 \ P2	L	R
U	3,3	0,0
D	4,0	1,1

↲ باید آنقدر تخمین و تصمیم گیری را انجام دهیم که همگرا شویم.

(D,R) → Nash Equilibrium → یکتا، خالص

← پیاده سازی الگوریتم.

باید از یک initialization از تخمین از رقبا شروع شود:

$$\eta_1^t = (\eta_1^t(L), \eta_1^t(R)) \;,\; \eta_2^t = (\eta_2^t(U), \eta_2^t(D))$$

$\rightsquigarrow \eta_1^0=(3,0) \Rightarrow$ عمل L، 3 بار و R، 0 بار دیده شد.

$\eta_2^0=(1,2.5) \Rightarrow$ عمل U، 1 بار و D، 2.5 بار

بازیکن 2، مشکوک بوده که بین 2 و 3 کدام را انتخاب کند.

$$\Rightarrow \mu_1^0=\left(\frac{3}{3},\frac{0}{3}\right)=(\overset{L}{1},\overset{R}{0})$$

$$\mu_2^0=\left(\frac{1^{U}}{3.5},\frac{2.5^{D}}{3.5}\right)$$

اگر P2 با احتمال 1، L را بازی کند ← P1، D بازی می‌کند.

$\Rightarrow S_1^0=D$

برای محاسبه‌ی S_2^0، باید expected payoff برای P2 را به‌ازای اینکه L یا R بازی می‌کند، محاسبه کنیم:

$$EU_2(L)=3\times\underbrace{\frac{1}{3.5}}_{P1\to U}+0\times\underbrace{\frac{2.5}{3.5}}_{P1\to D}=\frac{3}{3.5} \checkmark\checkmark$$

$$EU_2(R)=0\times\frac{1}{3.5}+1\times\frac{2.5}{3.5}=\frac{2.5}{3.5}$$

$$EU_2(L)>EU_2(R) \Rightarrow S_2^0=L$$

حال برای η_1^1، بازیکن 1 در مرحله‌ی قبل دیده که P2، L بازی کرده، پس به تعداد مشاهداتش از L، یکی اضافه می‌شود:

$$\eta_1^1=(4,0)$$

و مشابه آن برای بازیکن 2: $\eta_2^1=(1,3.5)$

$$\Rightarrow \mu_1^1=(1,0)\ ,\ \mu_2^1=\left(\frac{1}{4.5},\frac{3.5}{4.5}\right)$$

← حال باید براساس تخمین جدیدشان، BR عمل کنند:

$S_1^1=D,\ S_2^1=?$

$$EU_2(L)=3\times\frac{1}{4.5}+0\times\frac{3.5}{4.5}=\frac{3}{4.5}$$

$$EU_2(R)=0\times\frac{1}{4.5}+1\times\frac{3.5}{4.5}=\frac{3.5}{4.5} \checkmark\checkmark$$

$\Rightarrow S_2^1=R$

$\eta_1^2=(4,1)\ ,\ \eta_2^2=(1,4.5)$

$$\rightarrow \mu_1^2=\left(\frac{4}{5},\frac{1}{5}\right),\ \mu_2^2=\left(\frac{1}{5.5},\frac{4.5}{5.5}\right)$$

$\rightarrow S_1^2=D\ ,\ S_2^2=?$

$$EU_2(L)=3\times\frac{1}{5.5}+0\times\frac{4.5}{5.5}=\frac{3}{5.5}$$

$$EU_2(R)=0\times\frac{1}{5.5}+1\times\frac{4.5}{5.5}=\frac{4.5}{5.5} \checkmark\checkmark$$

$\Rightarrow S_2^2=R$

← بازیکن 1 که کلاً استراتژی غالب دارد (D) پس تا انتهای بازی $S_1^n=D$ خواهد شد. در مورد بازیکن 2، برای η_2، چون فقط D را از بازیکن 1 مشاهده می‌کند، عدد مربوط به مشاهده‌ی D در آن افزایش می‌یابد. تأثیری که در بردار احتمال دارد این است که احتمال مربوط به U، به صفر و احتمال مربوط به D به یک میل می‌کند ← $\mu_2=(0,1)$

اثر این موضوع در expected payoff به این صورت است که مقدارش برای عمل R برای P2 غلبه خواهد داشت و $S_2=R$ می‌شود.

از این به بعد بازیکن 1، همواره D و بازیکن 2، همواره R بازی می‌کنند.
پس مدل یادگیری Fictitious Play به تعادل نش همگرا می‌شود.

همگرایی Fictitious Play

تعریف. می‌گوییم دنباله‌ی استراتژی $\{s^t\}$ به استراتژی s همگرا شده است. اگر وجود داشته باشد زمان T، بطوریکه:

$$\forall t \geq T : \quad s^t = s \quad , (s^t = (s_1^t, \dots, s_n^t)) \quad , (s = (s_1, \dots, s_n))$$

تعداد بازیکن

قضیه. اگر دنباله‌ی $\{s^t\}$ یک دنباله‌ی استراتژی تولید شده توسط FP باشد:

1- اگر دنباله‌ی $\{s^t\}$ به $\bar{s}$ همگرا شود، آنگاه $\bar{s}$ یک استراتژی تعادلی نش خالص است

2- اگر $s^t = s^*$ برای یک زمان t باشد که در آن s^* یک تعادل نش اکید است، آنگاه: $s^\tau = s^*, \tau > t$

شهود 1: وقتی یک استراتژی همگرا شود، (چون نحوه تصمیم‌گیری بازیکنان BR بود)، یعنی همه BR خود را بازی می‌کنند. بنابراین این نقطه تقاطع BR هاست یعنی تعادل نش است.

شهود 2: چرا تخمین ها باعث شدند بازی به تعادل نش برسند؟ قاعدتاً "تاریخچه (history)" که بازیکنان در t-1 داشتند، طبق آن یک تخمین از استراتژی بقیه داشتند و این تخمین بازیکنان را طوری راهنمایی کرد که BR خود را طوری بازی کنند که در s^* قرار بگیرند. $t-1 \to BR \to s^*$

فرض کنیم که استراتژی نش s^* یک بار در لحظه‌ی t اجرا شد. حال بازیکنان می‌خواهند در لحظه‌ی t+1 تصمیم‌گیری کنند. تاریخچه‌ای که از t دارند این است که همگی نش بازی کردند. پس در مرحله‌ی t+1، BR آنها همگی نش است چون BR نسبت به نش، خود نش است. پس همگی از مرحله‌ی t به بعد استراتژی نش را بازی خواهند کرد.

تعریف. همگرایی FP به استراتژی مخلوط

دنباله‌ی $\{s^t\}$ به استراتژی $\delta \in \Sigma$ به مفهوم متوسط زمانی همگرا شده است اگر:

$$\forall i, \forall s_i \in S_i \; . \; \lim \frac{\sum_{t=0} I\{s_i^t = s_i\}}{T} = \delta_i(s_i)$$

مفهوم: I در این جا indicator func است، یعنی تعداد را می‌شمارد و مقدارش 1 می‌شود زمانی که $s_i^t = s_i$ شود و در غیر این صورت مقدارش صفر است. اگر تعداد دفعاتی که استراتژی بازیکن i در لحظه‌ی t مساوی s_i شود را شمرده و بر T تقسیم کنیم، مفهوم متوسط زمانی یا فرکانس s_i را می‌دهد و این متوسط زمانی به یک مقدار مشخص همگرا خواهد شد.

تخمین سایر بازیکنان غیر از i، در زمان t، از استراتژی بازیکن i وقتی $T \to \infty$، همان فرکانس $(\delta_i(s_i))$ خواهد شد.

$$\mu_{-i}^T(s_i) \xrightarrow{T \to \infty} \delta_i(s_i)$$

قضیه فرض کنیم دنباله‌ی $\{s^t\}$ که توسط FP تولید شده، به مفهوم متوسط زمانی به استراتژی δ همگرا شود، آنگاه δ تعادل نش است.

هر بازیکن BR خود را از تخمینی که از سایر بازیکنان دارد، $\delta_i(s_i)$، انجام می‌دهد و این BR عمل کردن او باعث می‌شود خودش با یک فرکانس مشخص بازی کند. چون این فرکانس توسط BR تولید می‌شود بنابراین چون فرکانس همه fix است، یعنی تخمین ها همگرا شده، پس هرکس فرکانس تصمیم‌گیری اش را تغییر دهد، ضرر خواهد کرد پس مفهوم تعادل نش وجود دارد.

قضیه. FP از نظر متوسط زمانی در شرایط زیر همگرا می‌شود:

1- بازی‌های zero-sum دو نفره
2- بازی‌های غیر zero-sum دو نفره با دو استراتژی
:

مباحث تحقیقاتی این فصل:

- Mean Field یا Aggregative Games:

P_1 تابع هدف: $f_1(x_1, \sum_{j\neq 1} x_j)$

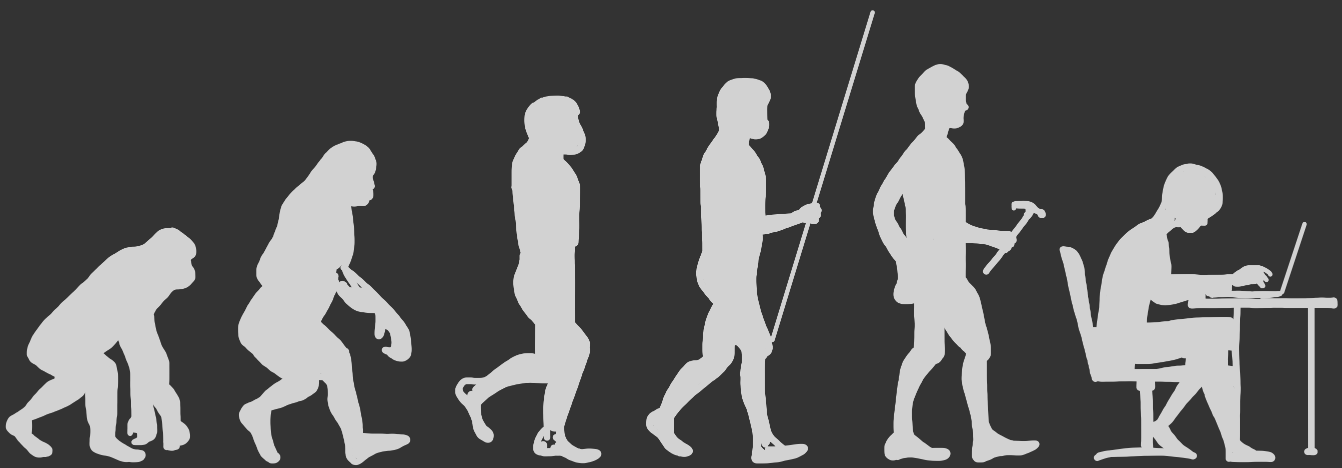
- Multi Agent Reinforcement Learning:

ممکن است تابع هدف agent $f_i(x_i, x_{-i})$
~ محیط بازیگر وابسته باشد.

ترکیب RL و تعادل نش. Nash-Q Learning ⟵

CHAPTER 5

~ Evolutionary Games



Shiva Shakeri

بازی های تکاملی

بحث تکامل: آقای داروین ←

موجودی که به دنیا می‌آد استراتژی مشخصی را دنبال می‌کند.

قابل استفاده برای مدل سازی سطح بالا از انسان ها

Natural Selection :

انتخاب گونه ها براساس یک استراتژی از پیش تعیین شده

تفاوت اصلی با مدل های قبلی.

همیشه می‌گفتیم که تصمیم گیرنده مجموعه ای از استراتژی دارد و می خواهد بین آنها انتخاب کند. اما در بحث تکاملی، در جمعیتی که داریم، چند نوع استراتژی مشاهده می‌کنیم و جمعیت را به آن چند نوع تقسیم می‌کنیم. مثل رفتار سلول ها یا رفتار حیوانات.

← پس بحث game در این نوع بازی ها، کجاست؟

در یک مجموعه ممکن است گونه های مختلف باهم، هم‌زیستی داشته باشند؛ مثل سلول های سالم و سرطانی.

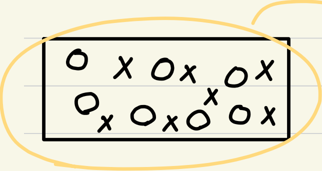
این سلول ها از دیدگاه استراتژی یک استراتژی واحد دارند.

game در جایی ورود می‌کند که این استراتژی ها با یکدیگر، تقابل داشته باشند. پس ما هر گونه ی خاصی را با استراتژی مشخص می‌بینیم؛ مثلاً خرگوش یک استراتژی مشخص دارد.

زمانی که این گونه ها تقابل داشته باشند، game پدید می‌آید. کدام استراتژی قابلیت باز تولید بیش تری خواهد داشت؟

در نهایت این تقابل به چه سمتی پیش خواهد رفت؟

← پس بحث تکامل را چگونه می‌توان مدل کرد؟



x ها و o ها هر کدام یک نوع موجودند.

x ها استراتژی ذاتی x و o ها استراتژی ذاتی o را دارند.

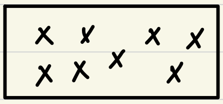
در یک محیط این دو گونه، هم زیستی می‌کنند. استراتژی = گونه

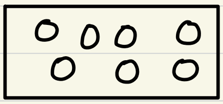
← در اثر تقابل این دو گونه به چه می‌رسیم و تکامل محیط چگونه مدل می‌شود؟

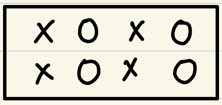
← مفهوم تعادل در بازی های تکاملی چگونه است؟

• آیا نهایتاً گونه های مختلف با استراتژی ها مشخص، به تعادل می‌رسند؟

• مدل سازی تکاملی به این ربط دارد که کدام استراتژی قابلیت بازتولید بیش تری دارد؟


← اگر x قابلیت بازتولید بیش تری داشته باشد.


← اگر o قابلیت بازتولید بیش تری داشته باشد.


← برقراری هم زیستی بین دو گونه

← روند یک نقطه ی شروع و رسیدن به نقطه ی پایانی که یکی از سه حالت بالا است، چگونه است؟

این همان مدل سازی تکاملی است که اساس آن براین است:

* هر چقدر گونه ای قابلیت fitness بیش تری داشته باشد، یعنی سود آوری بیش تری داشته باشد، آن گونه قابلیت تکثیر یا بازتولید بیش تری خواهد داشت و در محیط بیش تر می‌شود.

مدل مسأله.

1- فرض می‌کنیم یک مجموعه از جمعیت زیادی از گونه‌های مختلف داریم که بطور یکنواخت در محیط پراکنده شده اند.

2- هر عامل از هر گونه برنامه ریزی شده است که استراتژی مشخص را بازی کند.

3- فرض می‌کنیم در هر لحظه هر عامل بطور تصادفی با عامل دیگر در جمعیت برخورد دارد و در نتیجه‌ی آن یک بازی متقارن به فرم ماتریسی شکل خواهد گرفت.

~~P1~~ \ ~~P2~~	O	X	
O	a,a	a',b'	...
X	b',a'	b,b	

4- استراتژی هایی که موجب سود بیشتری می‌شوند، قابلیت بازتولید بیشتری در محیط پیدا می‌کنند و استراتژی هایی که سود کمتری در تقابل کسب می‌کنند، درصدشان در محیط کاهش می‌یابد ← قانون جنگل یا نظریه داروین :))

⇐ میزان کم و زیاد شدن نسبت یک استراتژی در محیط متناسب با میزان سودآوری آن استراتژی ست.

مثال بازی های تکاملی

← مثال 1. Hawk-Dove game (قمری و شاهین)

aggressive / defencive

سود / cost

	Hawk	Dove
Hawk	$\frac{1}{2}(v-c), \frac{1}{2}(v-c)$	$v, 0$
Dove	$0, v$	$\frac{1}{2}v, \frac{1}{2}v$

باهم مهربونن و cost ندارن :)

① $v>c$: در این صورت استراتژی Hawk به Dove غالب است. چون $\frac{1}{2}(v-c)>0$ & $v>\frac{1}{2}v$

← تعادل استراتژی غالب : (H,H)

(یعنی کل محیط درنهایت به Hawk تبدیل خواهد شد.)

② $v=c$: تعادل نش خواهیم داشت. (H,H)

(H,D)
(D,H) } → تعادل نش هستند اما در بحث تکامل مفهوم ندارند. چون یعنی همه‌ی محیط را H فرا می‌گیرد و همه‌ی محیط را D فرا می‌گیرد ✗

③ $v<c$:

$\frac{1}{2}(v-c)<0 \rightarrow$ تعادل نش (H,D),(D,H)

← تعادل نش مخلوط خواهیم داشت. مفهوم آن این است که در یک محیط دو استراتژی باهم هم‌زیستی خواهند کرد. درصدی از محیط شاهین و باقی درصد را قمری ها تشکیل خواهند داد.

← تعادل در تقابل استراتژی ها چگونه قابل تعریف است؟ ← پایداری تکاملی

← تکامل استراتژی ها در نتیجه‌ی تقابل آنها چگونه قابل مدل سازی ست؟

استراتژی پایداری تکاملی :

استراتژی (در حالت کلی مخلوط) δ^* را در نظر بگیرید. استراتژی مخلوط این جا نمایانگر درصد گونه های مختلف در محیط است.

فرض کنید تعدادی جمعیت Mutant که در آن استراتژی غیر از δ^* وجود دارد، وارد جمعیتی که استراتژی δ^* بازی می‌کنند، شوند.

ε : درصد جمعیت Mutant، δ. استراتژی جمعیت Mutant

Expected payoff of mutants:

$$=(1-\varepsilon)U(\delta,\delta^*)+\varepsilon U(\delta,\delta)$$

احتمال برخورد یک mutant با عامل های جمعیت اصلی

احتمال برخورد یک mutant با عامل های mutant

سود حاصل از برخورد دو عامل mutant

سود حاصل از برخورد دو استراتژی δ^* و δ برای عامل mutant

* $U(\delta,\delta^*)$: یعنی سود را برای جمعیت با استراتژی δ محاسبه می‌کنیم.

Expected payoff of δ^*:

$$= (1-\varepsilon)U(\delta^*,\delta^*) + \varepsilon U(\delta^*,\delta)$$

← در چه صورت استراتژی δ^*، پایدار تکاملی است؟

در صورتی که متوسط سود δ^* از متوسط سود استراتژی mutant ها ($\delta\neq\delta^*$)، بزرگ‌تر باشد. جمعیتی با δ^* تقسیم شده است. حال یک جمعیت جدید با δ اضافه شده. چون δ^* نسبت سود بیش‌تری از δ دارد پس کم‌کم نسبت خود را ترمیم می‌کند.

انگار هر عاملی که در جمعیت اصلی δ^* را بازی می‌کرده، حال هر عاملی در جمعیت mutant استراتژی δ را بازی می‌کند. اگر δ^*، سودآوری بیش‌تری داشته باشد و تعداد کمی استراتژی δ وارد محیط شده باشد، کم‌کم طبق قانون بقا و بازتولید بیش‌تر، δ^*، دوباره برگشته و کل محیط را تسخیر خواهد کرد.

$$(1-\varepsilon)U(\delta^*,\delta^*)+\varepsilon U(\delta^*,\delta) > (1-\varepsilon)U(\delta,\delta^*)+\varepsilon U(\delta,\delta)$$

تعریف 1. استراتژی $\delta^*\in\Sigma$ پایدار تکاملی است اگر $\exists\,\bar{\varepsilon}>0$ بطوریکه به ازای هر $\delta\neq\delta^*$ و $\bar{\varepsilon}>\varepsilon$ داشته باشیم.

$$(1-\varepsilon)U(\delta^*,\delta^*)+\varepsilon U(\delta^*,\delta)>(1-\varepsilon)U(\delta,\delta^*)+\varepsilon U(\delta,\delta)$$

← چون ε دقیقاً مشخص نیست، استفاده از این تعریف از نظر محاسباتی پیچیده است.

تعریف 2. استراتژی $\delta^*\in\Sigma$، استراتژی پایداری تکاملی (ESS) است، اگر به ازای هر $\delta\neq\delta^*$:

$U(\delta^*,\delta^*)\geq U(\delta,\delta^*)$ → شبیه نامساوی نش

if for some $\delta\in\Sigma$ $U(\delta^*,\delta^*)=U(\delta,\delta^*)$

then $U(\delta^*,\delta)>U(\delta,\delta)$

← مفهوم تعریف 2: تعادل نش با این تعریف بسیار مشابه است. نامساوی اول مشابه نامساوی نش است بطوریکه δ^* تعادل نش است. انگار اگر بازیگری استراتژی‌اش را از δ^* به δ تغییر دهد، سود نخواهد کرد. که اگر نش بصورت اکید برقرار نبود، یک شرط اضافه نیز می‌خواهد. سود استراتژی δ^* در مقابل جمعیت mutant با δ، باید از سود mutant ها از تقابلشان با یکدیگر بیش‌تر باشد. انگار از اینکه mutant ها در تقابل باهم، سودشان بیش‌تر شود، جلوگیری می‌کند.

Session 24-1

← این دو تعریف چه ارتباطی باهم دارند؟

قضیه. تعاریف ① و ② معادلند

← از تعریف 1 به 2

$$1:\ (1-\varepsilon)U(\delta^*,\delta^*)+\varepsilon U(\delta^*,\delta)>(1-\varepsilon)U(\delta,\delta^*)+\varepsilon U(\delta,\delta)$$

در این تعریف $\bar{\varepsilon}>\varepsilon$، یعنی ε می‌تواند به اندازه‌ی کافی کوچک باشد و به صفر میل کند. درنظر نگیریم که مساوی صفر شود چون (در این صورت به این معناست که جمعیت mutant اصلاً وجود ندارد.) پس اگر ε خیلی کوچک باشد $\varepsilon U(\delta^*,\delta)$ و $\varepsilon U(\delta,\delta)$ را می‌توان از دو طرف نامساوی حذف کرد و خواهیم داشت:

$$(1-\varepsilon)U(\delta^*,\delta^*)>(1-\varepsilon)U(\delta,\delta^*)$$

از آنجایی که $1-\varepsilon>0$ است می‌توان آنرا نیز از دو طرف ساده کرد:

$U(\delta^*,\delta^*)\geq U(\delta,\delta^*)$ → تعریف 2

اگر نامساوی برعکس باشد، به ازای $\varepsilon=0$ داریم:

$$U(\delta^*,\delta^*)<U(\delta,\delta^*)$$

حال اگر عبارت های $\varepsilon x(\sim)$ را به دو طرف نامساوی اضافه کنیم جهت نامساوی همین باقی می‌ماند و خواهیم داشت:

$$\varepsilon(\quad)+(1-\varepsilon)U(\delta^*,\delta^*)<(1-\varepsilon)U(\delta,\delta^*)+\varepsilon(\quad)$$

پیوستگی را به هم می‌زند پس علامت نامساوی همین می‌ماند.
که می‌دانیم جهتش اشتباه است (طبق تعریف 1) $\Leftarrow$
پس جهت نامساوی نمی‌تواند "$<$" باشد.

در علامت نامساوی $\geq$، امکان تساوی نیز وجود دارد،
گفتیم $\varepsilon=0$ نمی‌تواند برقرار باشد و هر مقدار مثبتی می‌تواند باشد و حالت تساوی امکان دارد که رخ دهد (؟)

اگر $U(\delta^*,\delta^*)=U(\delta,\delta^*)$ باشد:

$$(1-\varepsilon)\cancel{U(\delta^*,\delta^*)}+\varepsilon U(\delta^*,\delta)>(1-\varepsilon)\cancel{U(\delta,\delta^*)}+\varepsilon U(\delta,\delta)$$

$$\Rightarrow \cancel{\varepsilon}\,U(\delta^*,\delta)>\cancel{\varepsilon}\,U(\delta,\delta)$$

$$\Rightarrow U(\delta^*,\delta)>U(\delta,\delta) \rightarrow$$ تعریف 2

← پس از تعریف 1، تعریف 2 را نتیجه می‌گیریم.

← از تعریف 2 به 1.

تعریف 2: $U(\delta^*,\delta^*)\geq U(\delta,\delta^*)$

if for some $\delta\in\Sigma$ $U(\delta^*,\delta^*)=U(\delta,\delta^*)$

then $U(\delta^*,\delta)>U(\delta,\delta)$

• حالت اول: $U(\delta^*,\delta^*)>U(\delta,\delta^*)$

$$\rightarrow(1-\varepsilon)U(\delta^*,\delta^*)>(1-\varepsilon)U(\delta,\delta^*)$$

(بدلیل پیوستگی) $\exists\bar{\varepsilon}.\ \forall\varepsilon<\bar{\varepsilon} \xRightarrow{\text{داریم}}$

$$(1-\varepsilon)U(\delta^*,\delta^*)+\varepsilon U(\delta^*,\delta)>(1-\varepsilon)U(\delta,\delta^*)+\varepsilon U(\delta,\delta)$$

اگر یک عبارت خیلی کوچک به دو طرف اضافه شود باز هم جهت نامساوی عوض نمی‌شود.

• حالت دوم. if $U(\delta^*,\delta^*)=U(\delta,\delta^*)$

then $U(\delta^*,\delta)>U(\delta,\delta)$

$$\varepsilon U(\delta^*,\delta)>\varepsilon U(\delta,\delta) \quad (\mathrm{I})$$

$$(1-\varepsilon)U(\delta^*,\delta^*)=(1-\varepsilon)U(\delta,\delta^*)$$

$\Longrightarrow$

$$(1-\varepsilon)U(\delta^*,\delta^*)+\varepsilon U(\delta^*,\delta)>(1-\varepsilon)U(\delta,\delta^*)+\varepsilon U(\delta,\delta)$$

← پس از تعریف 2 به 1 رسیدیم.

* تعریف اول و دوم معادلند. تعریف 1، مفهومی‌ست اما از تعریف 2 در محاسبات استفاده می‌کنیم.

← ارتباط تعادل نش و استراتژی پایداری تکاملی چیست؟
(از جهت محاسباتی)

در تعریف 2: $U(\delta^*,\delta^*)\geq U(\delta,\delta^*)$ شرط نش حتماً باید برقرار باشد و می‌گویند اگر نش اکید ($U(\delta^*,\delta^*)>U(\delta,\delta^*)$) برقرار باشد حتماً δ^*، ESS است و اگر نش اکید نباشد نیز با یک شرط تعیین می‌شود که δ^*، ESS است $\Leftarrow$

نش اکید بودن شرط قوی‌تری نسبت به ESS است.

Strict NE $\subseteq$ ESS $\subseteq$ NE

حال اگر نش اکید، نش اکید غالب باشد:

* نش اکید: $U(\delta^*,\delta^*)>U(\delta,\delta^*)$

غالب بودن: این استراتژی نه تنها در تعادل نش، بلکه در مقابل هر استراتژی دیگری هم سود بیش‌تری دارد:

$$U(\delta^*,\delta)>U(\delta,\delta) \quad **$$

با استفاده از ترکیب جمع * و ** به تعریف 1 خواهیم رسید

به ازای هر ε = $\forall\varepsilon$

ESS بصورت فراگیر، global برقرار خواهد بود!

$$ESS\begin{cases}\text{Pure (Monomorphic)} : \text{جمعیت یک گونه}\\ \text{Mixed (Polymorphic)} : \text{جمعیت با چند گونه}\end{cases}$$

مثال ESS

← مثال 1: Hawk-Dove game (شاهین و قمری)

	Hawk ← q	Dove ← $1-q$
Hawk	$\frac{1}{2}(v-c),\frac{1}{2}(v-c)$	$v,0$
Dove	$0,v$	$\frac{1}{2}v,\frac{1}{2}v$

← استراتژی تکاملی؟

1- $v>c$: ن‌ش اکید و غالب (H,H)

بصورت global ← ESS

یعنی حتی اگر در جمعیت یک شاهین باشد و چندین قمری، باز هم قمری‌ها نابود خواهند شد.

2- $v=c$: سه تا ن‌ش دارد: $(\overset{\delta^*}{H},H),\cancel{(D,H)},\cancel{(H,D)}$

طبق تعریف 2:

$\underbrace{U(H,H)}_{0} \geq \underbrace{U(D,H)}_{0} \rightarrow U(H,H)=U(D,H)$

$\underbrace{U(H,D)}_{v} \overset{?}{>} \underbrace{U(D,D)}_{\frac{1}{2}v} \rightarrow U(H,D)>U(D,D)$

$\rightarrow (H,H) \leftarrow$ ESS ✓

3- $v<c$: سه تعادل ن‌ش. $\cancel{(D,H)},\cancel{(H,D)}$, Mixed

$\rightarrow$ Mixed: $\frac{1}{2}q(v-c)+(1-q)v=0\times q+\frac{1}{2}(1-q)v$

$\rightarrow \boxed{q=\frac{v}{c}} \Rightarrow \left(\underbrace{(\frac{v}{c},1-\frac{v}{c})}_{\delta^*},\overbrace{(\frac{v}{c},1-\frac{v}{c})}^{\delta^*}\right)$

ن‌ش مخلوط این بازی

• بررسی می‌کنیم که آیا این تعادل ن‌ش مخلوط (δ^*,δ^*) که بدست آوردیم ESS است؟

داشتیم: $\rightarrow U(\delta^*,\delta^*)=U(\delta,\delta^*)$

پس طبق تعریف 2 باید شرط دوم را بررسی کنیم.

$\delta=(p,1-p) \Rightarrow$

$U(\delta,\delta)=p^2\cdot\frac{1}{2}(v-c)+p(1-p)v+p(1-p)\times 0 + (1-p)^2(\frac{1}{2}v)$

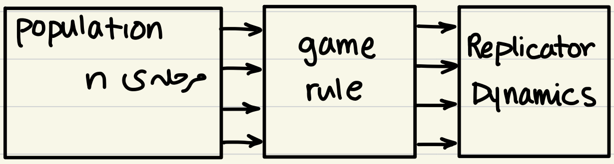
$U(\delta^*,\delta)=(\frac{v}{c})\times p\times\frac{1}{2}(v-c)+\frac{v}{c}(1-p)v+(1-\frac{v}{c})p\times 0+(1-\frac{v}{c})(1-p)(\frac{1}{2}v)$

$\rightarrow U(\delta^*,\delta)-U(\delta,\delta)=\frac{1}{2}c(\frac{v}{c}-p)^2 \geq 0$

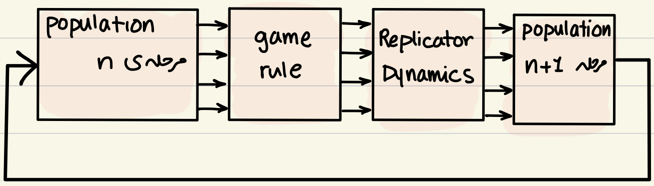
پس $U(\delta^*,\delta)>U(\delta,\delta)$ ← تعریف 2 برقرار می‌باشد

$\Leftarrow \delta^*=(\frac{v}{c},1-\frac{v}{c})$ ، ESS است!

Replicator Dynamics : مدل سازی دینامیک تکامل



یک population یا جمعیت داریم فرض می‌کنیم این population در یک مقطع زمانی مشخص، مثلاً "مرحله‌ی n" قرار دارد و یک ترکیب جمعیتی مشخص از گونه های مختلف داریم و هرگونه استراتژی خود را دنبال می‌کند. اتفاقی که می‌افتد این است که این گونه های مختلف باهم تقابل دارند و نتیجه‌ی این تعاملات براساس یک game rule مشخص می‌شود که این game rule، payoff‌ها را مشخص کرده و ماتریس payoff هاست و شبیه یک بازی ماتریسی متقارن می‌شود. حال سوالی که این جا مطرح می‌شود، این است که نتیجه‌ی این تقابل چه خواهد شد. چگونه یک تقابل باعث افزایش یا کاهش یک گونه در جمعیت می‌شود. در این جا بحث Replicator Dynamics مطرح می‌شود و می‌خواهیم مدل کنیم که در اثر تقابل گونه های مختلف، چگونه می‌شود که نسبت یک گونه بیش‌تر می‌شود یعنی بازی کردن آن استراتژی در جمعیت بیش‌تر می‌شود. آنکه سود آورتر باشد در جمعیت بیش‌تر می‌شود. باید این را بطور دقیق مدل کنیم. منطق Replicator Dynamics هم مطابق "قانون جنگل" است. قوی‌ها بیش‌تر، ضعیف‌ها کمتر!



Notation

- **S:** مجموعه‌ی استراتژی $S=\{e_1,\dots,e_n\}$ را در نظر بگیرید که n استراتژی داریم؛ یعنی n گونه مختلف در جمعیت داریم.
- **payoff:** $a_{ij}=u(e_i,e_j)$، این سود مربوط به e_1 است. a_{ij} یعنی سود حاصل از تقابل استراتژی e_i و e_j برای e_i.

← سود حاصل از استراتژی ها از تمام تقابل ها: $A=[a_{ij}]_{n\times n}$

← بازی ماتریسی هم از ماتریس A و A^T شکل خواهد گرفت

	e_1	e_2
e_1	a,a	b,c
e_2	c,b	d,d

$$\equiv A=\begin{bmatrix} a & b \\ c & d \end{bmatrix}$$

- **Mixed Strategy:** در جمعیت یک نسبت جمعیتی از گونه ها داریم.

$$p\in \underbrace{\Delta^n}_{\text{strategy simplex}}=\left\{(p_1,\dots,p_n)^T \mid p_i\geq 0, \sum_{i=1}^{n} p_i=1\right\}$$

pure : $e_i=(0,0,\dots,\overset{i\text{-th}}{1},0,\dots,0)$

- سود حاصل از تقابل دو استراتژی مخلوط:

جمعیت mutant با استراتژی $\hat{p}$ به جمعیت اصلی با استراتژی p وارد می‌شود:

$$\underbrace{U(p,\hat{p})}_{\text{سود برای } p}=\sum_{i,j=1}^{n} p_i \underbrace{U(e_i,e_j)}_{a_{ij}}\cdot\hat{p}_j$$

$$=\sum_{i,j=1}^{n} p_i a_{ij}\hat{p}_j = p^T A \hat{p}$$

$$=[p_1 \cdots p_n]\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn}\end{bmatrix}\begin{bmatrix}\hat{p}_1 \\ \vdots \\ \hat{p}_n\end{bmatrix}=[\quad]_{1\times 1}$$

Replicator Dynamics.

* تکامل تدریجی صورت می‌گیرد.

n_1 . تعداد اعضایی که در مرحله‌ی t، استراتژی e_1 را بازی می‌کنند. (تعداد اعضای گونه‌ی i-ام)

n_1' تعداد اعضایی که انتظار می‌رود در مرحله‌ی $t+1$، استراتژی e_i را بازی کنند. (تعداد اعضای گونه‌ی i-ام در $t+1$)

$$n_i' = n_i(e_i^TAP) \quad *$$

e_i^TAP . متوسط سود حاصل از تقابل استراتژی e_i و استراتژی P برای گونه‌ی i-ام

← منطق معادله‌ی * : جمعیت مرحله‌ی بعد، متناسب است با متوسط سودی که آن استراتژی داشته برای آن جمعیت.

⇐ همان نظریه تکامل داروین

← چرا ضریب این تناسب "دقیقاً" n_i است و "مثلاً" αn_i نیست؟

هر ضریب تناسبی را می‌توان در مدل‌سازی در ماتریس A، در نظر بگیریم طراح باید به خوبی ماتریس A را طراحی کند.

Mixed Strategy: $P=(P_1,\dots,P_n)^T$

$$P_i = \frac{n_i}{\sum_{j=1}^n n_j} \quad *$$

درصد گونه در جمعیت.

← نسبت جمعیتی گونه‌ی i از مرحله‌ی t به $t+1$ چگونه تغییر می‌کند؟

$$P_i' = \frac{n_i'}{\sum_{j=1}^n n_j'}$$

نسبت جمعیتی گونه i در $t+1$.

$$\Rightarrow P_i' = \frac{n_i e_i^TAP}{\sum_{j=1}^n n_j e_j^TAP} = \frac{(\sum_{k=1}^n n_k)P_i e_i^TAP}{\sum_{j=1}^n(\sum_{k=1}^n n_k)P_j e_j^TAP}$$

$$= \frac{P_i e_i^TAP}{\sum_{j=1}^n P_j e_j^TAP} =$$

ستون بعد

$$\Rightarrow \sum_{j=1}^n P_j e_j^T = (P_1,0,\dots,0)+(0,P_2,0\dots,0)+\dots+(0,0,\dots,0,P_n)$$

$$=(P_1,P_2,\dots,P_n)=P^T$$

$$\Rightarrow P_i' = P_i\frac{e_i^TAP}{P^TAP}$$

→ معادله‌ی Replicator

نسبت جمعیتی گونه‌ی i در مرحله‌ی $t+1$ را برحسب نسبت جمعیتی گونه‌ی i در مرحله‌ی t نشان می‌دهد

e_i^TAP → متوسط سود حاصل از استراتژی i

P^TAP → متوسط سود جمعیت

← آیا معادله‌ی بدیهی $\sum P_i'=1$ هم‌چنان باقی می‌ماند؟

$$\sum P_i' = \frac{\sum P_i e_i^TAP}{P^TAP} = \frac{P^TAP}{P^TAP} = 1 \checkmark$$

مدل RD در محیط پیوسته:

مقدار رشد n_i

$$\dot n_i = n_i(e_i^TAP) \quad *$$

← مثبت یا منفی بودن مقدار رشد اهمیت دارد. اگر $\dot n_i$ مثبت باشد، جمعیت افزایش می‌یابد و اگر منفی باشد، کاهش.

* ماتریس A در محیط پیوسته، متفاوت از حالت گسسته است.

$$P_i = \frac{n_i}{\sum_{j=1}^n n_j} \xrightarrow[\partial t]{\text{مشتق}}$$

$$\Rightarrow \dot P_i = \frac{\dot n_i(\sum_{j=1}^n n_j) - (\sum_{j=1}^n \dot n_j) n_i}{(\sum_{j=1}^n n_j)^2}$$

$P_i\sum n_j$, $P_j \sum n_k$

$$= \frac{n_i e_i^TAP(\sum_{j=1}^n n_j) - (\sum_{j=1}^n n_j e_j^TAP)n_i}{(\sum_{j=1}^n n_j)^2} =$$

ادامه‌ی ص بعد

$$= P_i e_i^T AP - P_i \sum_{j=1}^{n} P_j e_j^T AP$$

$$= P_i \Big(e_i^T AP - \underbrace{\sum_{j=1}^{n} P_j e_j^T}_{P^T} AP\Big)$$

$$\Rightarrow \dot P_i = P_i (e_i^T AP - P^T AP)$$

نرخ تغییرات نسبت جمعیتی گونه‌ی i برابر است با نسبت جمعیت اولیه‌اش، ضربدر اختلاف متوسط سود حاصل از استراتژی e_i برای گونه‌ی i ام و متوسط سود کل جمعیت.

← معادله‌ی $\sum_{i=1}^{n} \dot P_i = 0$ برقرار است. چرا؟

یعنی اگر چند گونه نرخ رشد مثبت داشته باشند، بقیه گونه‌ها نرخ رشد منفی و به همان اندازه دارند و در مجموع نسبت جمعیت برابر 1 باقی خواهد ماند.

$$\sum \dot P_i = \overbrace{\sum P_i e_i^T}^{P^T} AP - \overbrace{\sum P_i}^{1} P^T AP$$

$$= P^T AP - P^T AP = 0 \ \checkmark$$

Session 25-2

مثال RD

← مثال 1:

	e_1	e_2
e_1	a, a	b, c
e_2	c, b	d, d

$\rightarrow A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$

Mixed Strategy: $(P_1, P_2)^T$, $P_2 = 1 - P_1$

جمعیت شامل دوگونه با استراتژی‌های e_1, e_2

← معادله RD در حالت پیوسته:

$$\dot P_1 = P_1 (e_1^T AP - P^T AP)$$

$$= P_1 (e_1^T AP - (P_1 \ \ P_2) AP)$$

$$= P_1 (e_1^T AP - P_1 e_1^T AP - P_2 e_2^T AP) =$$

$$= P_1 \big((1-P_1) e_1^T AP - \overset{1-P_1}{P_2} e_2^T AP\big) =$$

$$= P_1 (1-P_1)(e_1^T AP - e_2^T AP)$$

جاگذاری

که $e_1^T AP = (1 \ \ 0)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} P_1 \\ P_2 \end{pmatrix} = aP_1 + bP_2$

که $e_2^T AP = (0 \ \ 1)\begin{pmatrix} a & b \\ c & d \end{pmatrix}\begin{pmatrix} P_1 \\ P_2 \end{pmatrix} = cP_1 + dP_2$

$$\Rightarrow \dot P_1 = P_1 (1-P_1)\big((a-c)P_1 + (b-d)P_2\big)$$

$$\Rightarrow \dot P_1 = P_1 (1-P_1)\big((a-c+d-b)P_1 + b - d\big)$$

← تحلیل همگرایی این دینامیک:

• نقاط تعادل RD:

در حالت دینامیک گسسته که در یادگیری در بازی‌ها داشتیم، برای مثال نقطه‌ی تعادل جایی بود که $q(t) = q(t+1)$ داشتیم.

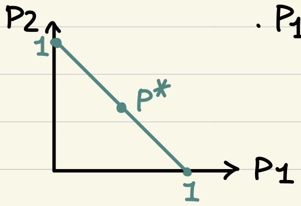
حال در حالت پیوسته شرط برقراری تعادل زمانی‌ست که:

تعادل یا rest point که $\dot P_1 = 0$

$$\dot P_1 = P_1 (1-P_1)\big((a-c+d-b)P_1 + b - d\big) = 0$$

$P_1 = 0$ یا $P_1 = 1$ یا $\boxed{0 \leq P_1 = \frac{d-b}{a-c+d-b} \leq 1}$ *

چون $P_1 + P_2 = 1$.



← دسته اول: بازی‌های Prisoner's Dilemma

$(a-c)(d-b) \leq 0$

① $(a-c) \leq 0 \ \& \ (d-b) \geq 0$

② $(a-c) \geq 0 \ \& \ (d-b) \leq 0$

← حالت 1-1:

if $|a-c| < |d-b| \rightarrow \frac{d-b \geq 0}{\underbrace{(a-c)}_{\leq 0} + \underbrace{(d-b)}_{\geq 0}} \geq 1$

if $|a-c| > |d-b| \rightarrow \frac{d-b \geq 0}{\underbrace{(a-c)+(d-b)}_{<0}} < 0$

با * در تناقض هستند.

Coordination Class ← دسته دوم :

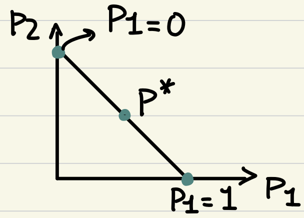
→ $a>c$ & $d>b$:

$\dot{P}_1 = P_1(1-P_1)((a-c)P_1+(d-b)(P_1-1))$

$\Rightarrow P_1^* = \frac{d-b \geq 0}{a-c+d-b \geq 0} \Rightarrow 0 \leq P_1^* \leq 1 \rightarrow$ * مطابق ✓

$\dot{P}_1 = P_1(1-P_1)(P_1-P_1^*)(a-c+d-b)$

⇐ در حالت 2، $P_1=0$, $P_1=1$ ، $P_1=\frac{d-b}{a-c+d-b}$ نقاط تعادل هستند.

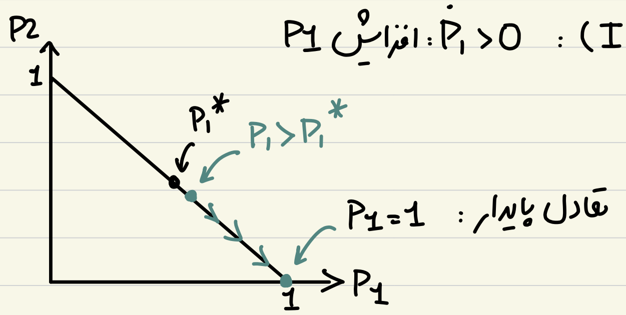


$\dot{P}_1 = \underbrace{P_1}_{\geq 0}(\underbrace{1-P_1}_{\geq 0})(P_1-P_1^*)(\underbrace{a-c+d-b}_{\geq 0})$

$\Rightarrow$ if $P_1>P_1^* \rightarrow \dot{P}_1>0$ (I)

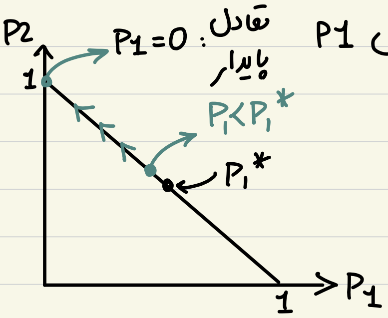
if $P_1<P_1^* \rightarrow \dot{P}_1<0$ (II)

(I) : $\dot{P}_1>0$: افزایش P1



$\Rightarrow$ if $P_1(0)>P_1^* \Rightarrow P_1 \rightarrow 1$

(II) . $\dot{P}_1<0$: کاهش P1



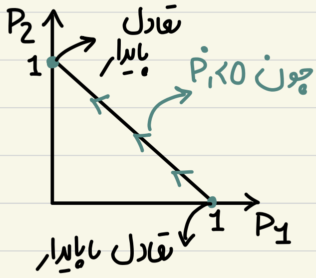
$\Rightarrow$ if $P_1(0)<P_1^* \Rightarrow P_1 \rightarrow 0$

$\Rightarrow$ if $P_1(0)=P_1^* \Rightarrow$ تعادل پایدار P_1^*

⇐ در حالت 1-1، $P_1=0$, $P_1=1$ نقاط تعادل هستند

$\dot{P}_1 = \underbrace{\underbrace{P_1}_{\geq 0}(\underbrace{1-P_1}_{\geq 0})(\underbrace{(a-c)}_{\leq 0}P_1+\underbrace{(d-b)}_{\geq 0}\underbrace{(P_1-1)}_{\leq 0})}_{\leq 0} <0$

← پس در تمام نقاط جز نقاط تعادل : $\dot{P}_1<0$



پس از جهت کاهش P_1 منحنی‌هم RD به کجا ختم می‌شود

$P_2=1, P_1=0$

← پس استراتژی 2 می‌تواند کل جمعیت را در انتها پُر کند.

← حالت 1-2 :

if $|a-c|<|b-d| \rightarrow \frac{b-d \leq 0}{\underbrace{(a-c)+(b-d)}_{\leq 0}} \geq 1$

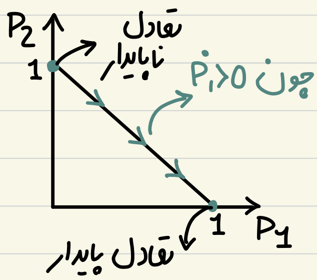
if $|a-c|>|b-d| \rightarrow \frac{b-d \geq 0}{\underbrace{(a-c)+(b-d)}_{\leq 0}} \leq 0$

* با در تناقض هستند.

⇐ در حالت 1-2، $P_1=0$, $P_1=1$ نقاط تعادل هستند

$\dot{P}_1 = \underbrace{P_1}_{\geq 0}(\underbrace{1-P_1}_{\geq 0})(\underbrace{(a-c)}_{\geq 0}\underbrace{P_1}_{\geq 0}+\underbrace{(d-b)}_{\leq 0}\underbrace{(P_1-1)}_{\leq 0})>0$

← پس در تمام نقاط جز نقاط تعادل : $\dot{P}_1>0$



پس از جهت افزایش P_1 منحنی‌هم RD به کجا ختم می‌شود

$P_1=1, P_2=0$

← پس استراتژی 1 می‌تواند کل جمعیت را در انتها پُر کند.

مثال RD

مثال 2. ← Rock-Scissors-Paper

$$A = \begin{array}{c|ccc} & R & S & P \\ R & \varepsilon & 1 & -1 \\ S & -1 & \varepsilon & 1 \\ P & 1 & -1 & \varepsilon \end{array} \quad , \ |\varepsilon|<1$$

← استراتژی پایدار تکاملی: ESS

می‌توان نقاط تعادل نش بازی را بدست آورد (شرط لازم ESS)

← می‌توان به این رسید که بازی تعادل نش خالص ندارد!

تعادل نش Mix: $P^*=(\frac{1}{3},\frac{1}{3},\frac{1}{3})$

آیا این تعادل ESS است؟ از تعریف دوم ESS:

$$P^{*T}AP^* = P^TAP^*$$

شرط ESS: $P^{*T}AP > P^TAP$

$$P^{*T}AP - P^TAP = \left(\tfrac{1}{3}\ \tfrac{1}{3}\ \tfrac{1}{3}\right)\begin{bmatrix}\varepsilon & 1 & -1\\ -1 & \varepsilon & 1\\ 1 & -1 & \varepsilon\end{bmatrix}\begin{bmatrix}P_1\\P_2\\P_3\end{bmatrix} - (P_1\ P_2\ P_3)\begin{bmatrix}\varepsilon & 1 & -1\\ -1 & \varepsilon & 1\\ 1 & -1 & \varepsilon\end{bmatrix}\begin{bmatrix}P_1\\P_2\\P_3\end{bmatrix} =$$ عملیات جبری

$$= -\varepsilon\left(P_1^2+P_2^2+P_3^2-\frac{1}{3}\right) =$$

می‌دانیم: $P_1+P_2+P_3=1 \Leftarrow$

$$= -\varepsilon\left(\left(P_1-\frac{1}{3}\right)^2+\left(P_2-\frac{1}{3}\right)^2+\left(P_3-\frac{1}{3}\right)^2\right)$$

$\rightarrow$ if $P\neq P^* \Rightarrow P^{*T}AP-P^TAP>0$ if $\varepsilon<0$

شرط ESS بودن تعادل نش Mix

← Replicator Dynamics. ص بعد

← دسته سوم: Hawk-Dove Class

$\rightarrow a<c \ \& \ b<d$

$$\dot{P}_1 = P_1(1-P_1)\big((a-c)P_1+(d-b)(P_1-1)\big)$$

$$\Rightarrow P_1^* = \frac{d-b \ (<0)}{a-c+d-b \ (<0)} \Rightarrow 0\leqslant P_1^* \leqslant 1 \rightarrow$$ مطابق

$$\dot{P}_1 = P_1(1-P_1)(P_1-P_1^*)(a-c+d-b)$$

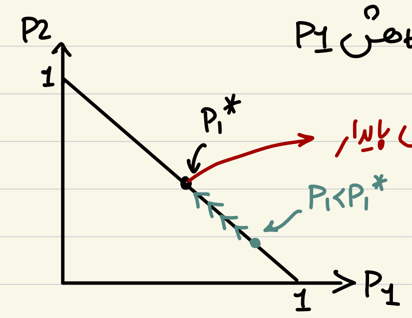
⇐ در حالت 3، $P_1=0$ ، $P_1=1$ ، $P_1=\frac{d-b}{a-c+d-b}$ نقاط تعادل هستند.

$$\dot{P}_1 = \underbrace{P_1}_{\geqslant 0}\underbrace{(1-P_1)}_{\geqslant 0}(P_1-P_1^*)\underbrace{(a-c+d-b)}_{<0}$$

$\Rightarrow$ if $P_1>P_1^* \rightarrow \dot{P}_1<0$ (I)

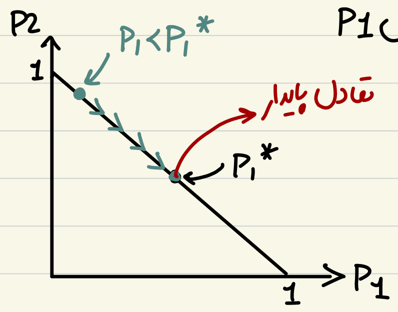
if $P_1<P_1^* \rightarrow \dot{P}_1>0$ (II)

(I): $\dot{P}_1<0$: کاهش P_1



$\Rightarrow$ if $P_1(0)>P_1^* \Rightarrow P_1\rightarrow P_1^*$

(II). $\dot{P}_1>0$: افزایش P_1



$\Rightarrow$ if $P_1(0)<P_1^* \Rightarrow P_1\rightarrow P_1^*$

$\Rightarrow$ if $P_1(0)=P_1^* \Rightarrow P_1^*$ تعادل پایدار

← فرآیند تکامل در هر صورت به P_1^*، همگرا خواهد شد.

$$\dot{p}_1 = p_1(\underline{e_1^T A p} - p^T A p), \quad e_1^T = (1\ 0\ 0)$$
$$\dot{p}_2 = p_2(\underline{e_2^T A p} - p^T A p) \quad *, \quad e_2^T = (0\ 1\ 0)$$
$$\dot{p}_3 = p_3(\underline{e_3^T A p} - p^T A p), \quad e_3^T = (0\ 0\ 1)$$
$$p^T = (p_1\ p_2\ p_3)$$

$$\Rightarrow \begin{cases} e_1^T A p = \varepsilon p_1 + p_2 - p_3 \\ e_2^T A p = -p_1 + \varepsilon p_2 + p_3 \\ e_3^T A p = p_1 - p_2 + \varepsilon p_3 \end{cases}$$

$$\Rightarrow \begin{cases} \dot{p}_1 = p_1(\varepsilon p_1 + p_2 - p_3 - p^T A p) \\ \dot{p}_2 = p_2(-p_1 + \varepsilon p_2 + p_3 - p^T A p) \\ \dot{p}_3 = p_3(p_1 - p_2 + \varepsilon p_3 - p^T A p) \end{cases}$$

← آیا تعادل نش Mix که به دست آوردیم، تعادل RD است یا خیر؟ اگر هست، پایدار است؟

در معادله ی $*$ به جای pها، P^* را قرار می دهیم، عبارت داخل پرانتز را برابر صفر قرار می دهیم. (چون اگر بخواهیم P^* نقطه ی تعادل RD باشد، باید این شرایط را داشته باشیم):

$$\begin{cases} e_1^T A P^* = P^{*T} A P^* \\ e_2^T A P^* = P^{*T} A P^* \\ e_3^T A P^* = P^{*T} A P^* \end{cases}$$

چون P^* تعادل نش مخلوط است، این سه تساوی برقرار هستند. چون سود متوسط سود حاصل از استراتژی خالص (e_1, e_2, e_3)، هرسه باهم برابر و برابر سود حاصل از استراتژی مخلوط P^* است ← طبق تعریف تعادل نش مخلوط

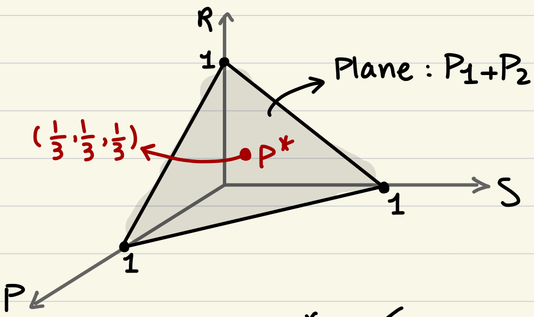
← پس در نقطه ی P^*: $\dot{p}_1 = \dot{p}_2 = \dot{p}_3 = 0$

⇐ پس P^* یک تعادل معادله ی RD می باشد

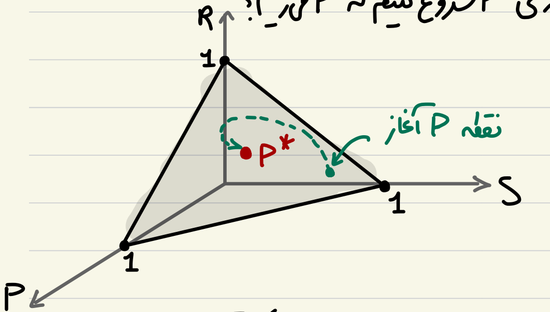
← همگرایی و پایداری نقطه ی تعادل:

در مثال 1، مقادیر P1، P2 روی یک خط تغییر می کرد و با داشتن جهت مشتق می توانستیم همگرایی را تست کنیم.

اما در این مثال فضایی بصورت صفحه ی مثلثی بوجود می آید:



← اگر مثلاً از نقطه ی P شروع کنیم به P^* می رسیم؟



این جا، جهت مشتق به ما نشان نمی دهد که آیا به سمت تعادل می رویم یا خیر. باید کل مسیر را داشته باشیم که خیلی سخت است.

دو راه حل برای تحلیل همگرایی وجود دارد:

1- خطی سازی

$$\left.\begin{aligned} \dot{p}_1 &= p_1(\varepsilon p_1 + p_2 - p_3 - p^T A p) \\ \dot{p}_2 &= p_2(-p_1 + \varepsilon p_2 + p_3 - p^T A p) \\ \dot{p}_3 &= p_3(p_1 - p_2 + \varepsilon p_3 - p^T A p) \end{aligned}\right\} \dot{p} = f(p)$$

با استفاده از بسط تیلور خطی سازی را روی f انجام می دهیم. و از روی ماتریس ژاکوبین مشابه بحث یادگیری ماری ها می توان خطی سازی کرد ← به معادلات خطی شده مثل: $\dot{\tilde{p}} = A\tilde{p}$ می رسیم.

A: ماتریس ژاکوبین در نقطه تعادل P^*

← در حالت گسسته از روی مقادیر ویژه، که باید در دایره‌ی واحد قرار می‌گرفتند، بود؛ اما چون این جا پیوسته است، شرط همگرایی محلی. همه‌ی مقادیر ویژه‌ی A باید دارای مقادیر حقیقی منفی باشند.

2_ استفاده از روش تابع پتانسیل:

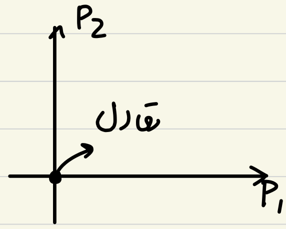
در فضای برداری نمی‌توان با جهت مشتق همگرایی را تحلیل کرد. پس در این روش فضای برداری را به اسکالر می‌بریم تا این مشکل را نداشته باشیم و از جهت مشتق بتوان استفاده کرد:

$$\begin{pmatrix}P_1\\P_2\\P_3\end{pmatrix} \xrightarrow{\text{اسکالر}} \underbrace{V(P_1,P_2,P_3)}_{\text{تابع پتانسیل}}$$

← ویژگی های تابع پتانسیل:

با یک مثال: فرض:

$$\text{سیستم}\begin{cases}\dot P_1 = f_1(P_1,P_2)\\ \dot P_2 = f_2(P_1,P_2)\end{cases}$$

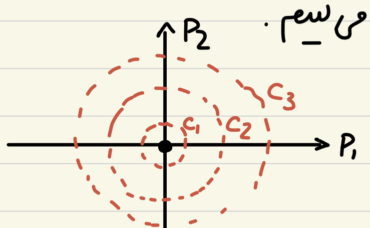


و فرض. $f_1(0,0)=f_2(0,0)=0$

همگرایی این معادله را بررسی می‌کنیم. در این جا نیز جهت مشتق به ما کمک نمی‌کند.

پیشنهادی $\begin{pmatrix}P_1\\P_2\end{pmatrix} \to V(P_1,P_2)=P_1^2+P_2^2$

پس $P_1^2+P_2^2=c$ را رسم می‌کنیم.



یک سری سطوح هم پتانسیل رسم کردیم. این باید یک سری ویژگی داشته باشد که بتوانیم همگرایی را تحلیل کنیم:

توضیح روش: تابع پتانسیل کمترین مقدار خودش را در نقطه‌ی تعادل داشته و در سایر جاها مقدار مثبت دارد و هر چقدر مقدار c بزرگ‌تر شود، سطوح هم پتانسیل بزرگ‌تر می‌شوند.

هر نقطه‌ی در فضای $\mathbb{R}^2$ روی یکی از این دایره‌ها قرار می‌گیرد.

بدون کاستن از کلیت مسأله فرض کنیم دایره‌ی c_3 را داشته باشیم.

اگر فرض کنیم $\dot V(P_1,P_2)$ را داشته باشیم؛ در این مثال:

$$\dot V(P_1,P_2)=2P_1\dot P_1+2P_2\dot P_2 = 2P_1f_1(P_1,P_2)+2P_2f_2(P_1,P_2)$$

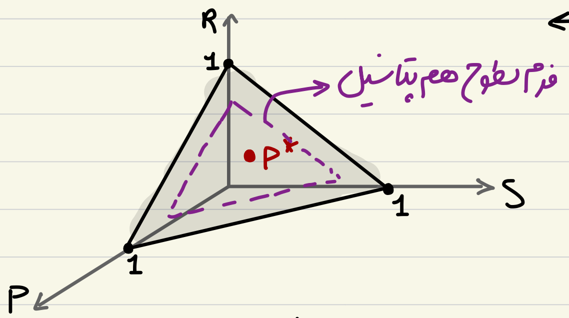
فرض کنید $\dot V(P_1,P_2)<0$ شود. یعنی چه؟

مثلاً اگر نقطه شروع از $(P_1(0),P_2(0))$ واقع روی پتانسیل c_3 باشد، منفی بودن $\dot V(P_1,P_2)$ نشان می‌دهد که در حال حرکت به سمت پتانسیل کمتر، مثل c_1, c_2 است. پس آنقدر سطح هم پتانسیل کاهش می‌یابد تا به کمترین سطح پتانسیل یعنی $(0,0)$ برسد که همان تعادلی ست که محاسبه کردیم. پس با این روش و تحلیل می‌توان همگرایی یک تعادل را بررسی کرد.

به حل مثال 2 برمی‌گردیم. فرض:

$$\begin{cases}V(P_1,P_2,P_3)=P_1P_2P_3\\ P^*=(\frac13,\frac13,\frac13)\end{cases}$$

←



بیش‌ترین مقدار $P_1P_2P_3$ اگر $P_1+P_2+P_3=1$ زمانی ست که $P_1=P_2=P_3$ باشد. پس نقطه‌ی تعادل بیش‌ترین پتانسیل را دارد. کمترین سطح پتانسیل روی اضلاع این مثلث است.

با توجه به بیش‌ترین و کمترین سطح هم‌پتانسیل، مشتق باید طوری باشد که از کمترین مقدارش به بیش‌ترین برویم تا به نقطه‌ی تعادل p^* برسیم تا ثابت کنیم که به آن همگرا می‌شویم. در این صورت مشتق باید مثبت شود. پس اگر ثابت کنیم $\dot{v}(p)>0$ است می‌توانیم همگرایی به تعادل را اثبات کنیم ⇐

$$\dot{v}(p)=\frac{d}{dt}(p_1p_2p_3)\xrightarrow{\text{عملیات جبری}}$$

$$=-\varepsilon p_1p_2p_3\Big[\underbrace{(p_1-\tfrac{1}{3})^2+(p_2-\tfrac{1}{3})^2+(p_3-\tfrac{1}{3})^2}_{>0}\Big]$$

$\to$ if $\varepsilon<0 \Rightarrow \dot{v}(p)>0$

← در این صورت p^* پایدار مجانبی فراگیر GAS

مزیت روش استفاده از سطوح هم‌پتانسیل این است همگرایی محلی (local) نیست و فراگیر است. برعکس خطی‌سازی که محلی بود.

← اگر $\varepsilon>0$ می‌بود در این صورت $\dot{v}(p)<0$ می‌شود و از بیش‌ترین سطوح هم‌پتانسیل به سمت کمترین می‌رویم یعنی اضلاع مثلث واقع در شکل. نمی‌توان دقیقاً گفت که به کدام ضلع و کجا هم‌گرا می‌شویم.

← اگر $\varepsilon=0$ باشد، $\dot{v}(p)=0$ می‌شود. این یعنی اگر از هر سطح هم‌پتانسیل که آغاز کنیم در همان سطح می‌مانیم.

فرم کلی از توابع پتانسیل. تابع لیاپانوف در حالت کلی‌تر

برای تحلیل معادلات RD ⇐

$$\begin{cases} v(p)=\prod_{i=1}^{n}(p_i)^{p_i^*} \\ \to \dot{v}(p)=v(p)(p^*-p)^T A p \end{cases}$$

در مثال 2 که $p_i^*=\frac{1}{3}$, $i=1,2,3$ بود:

$$v(p)=p_1^{\frac{1}{3}}p_2^{\frac{1}{3}}p_3^{\frac{1}{3}} \to (p_1^{\frac{1}{3}}p_2^{\frac{1}{3}}p_3^{\frac{1}{3}})^3$$

برای حل، همین تابع پیش نمادی را به توان 3 رساندیم

چون تغییری در نتیجه ایجاد نمی‌کند.

ارتباط تعادل نش، ESS و تعادل ها و تعادل های پایدار RD

قضیه 1. اگر p^* یک تعادل نش (NE) باشد، در این صورت یک حالت پایا، (تعادل یا rest point) برای معادله RD پیوسته می باشد.

← مفهوم. هر تعادل نش یک تعادل معادله ی RD سر است.

* در مورد پایداری بحث نمی کند.

$$\begin{cases} \text{if } p_i^*=0 \to \dot{p}_1\big|_{p_i^*}=0 \\ \text{if } p_i^*\neq 0 \to e_1^TAp^* = p^{*T}Ap^* \to \dot{p}_1\big|_{p_i^*}=0 \end{cases}$$

قضیه 2. اگر p^* یک تعادل پایدار مجانبی برای معادله ی RD باشد، در این صورت p^* تعادل نش است.

← مفهوم : با برهان خلف اگر فرض کنیم :

$$u(\delta^*,\delta^*) < u(\delta,\delta^*)$$

فرض خلف. یعنی اگر فرض کنیم δ^* تعادل نش نباشد و در عین حال تعادل معادله ی RD باشد.

مثال : برای استراتژی خالص ←

$\exists j\neq i.\ e_i^TAe_i < e_j^TAe_j$ ← e_i تعادل نش نباشد (1)

$P_i\ .\ e_i^TAP = P^TAP,\ P_{k \atop k\neq i} : P_k=0$ ← e_i تعادل RD، $P_i=1$ (2)

← برای اثبات ناپایداری تعادل باید اثبات کنیم اگر کمی از آن تعادل جابه‌جا شویم، به آن بار نخواهیم گشت.

$$\tilde{P} = (0,0,\dots,\underset{i\text{-th}}{1-\varepsilon},\dots,\underset{j\text{-th}}{\varepsilon},\dots,0)$$

کافی است $e_i^TA\tilde{P} - \tilde{P}^TA\tilde{P}$ را محاسبه کنیم. در معادله (2)،

فرض کردیم $P_1=1$ است در این صورت برای P_i داشتیم :

$$P_i : e_i^TAP = P^TAP \xrightarrow[P^T=e_1^T]{P_1=1} e_1^TAP - P^TAP = 0$$

اگر $e_i^TA\tilde{P}-\tilde{P}^TA\tilde{P}$ را محاسبه کنیم، خواهیم دید برای εهای به اندازه ی کافی کوچک، این عبارت کوچکتر از صفر خواهد شد. که به این معناست که P_i کاهش پیدا می کند و پایداری اش نقض خواهد شد پس فرض خلف باطل است و قضیه درست است.

← چرا عکس قضیه ی 2 صادق نیست؟

عکس قضیه. اگر p^* تعادل نش باشد آنگاه پایدار مجانبی است.

مثال نقض.

	A	B
A	1,1	0,0
B	0,0	0,0

NE

NE (مغلوب ضعیف)

$$\dot{P}_1 = P_1(1-P_1)((a-c+d-b)P_1+(b-d)) > 0$$

$$a=1,\ b=c=d=0$$

← از هر $P_1(0)$ شروع کنیم درنهایت P_1 به 1 میل می کند. یعنی به a همگرا می شویم و a پایدار کلی ست.

پس بازیکنان به نش مغلوب ضعیف همگرا نمی شوند.

قضیه 3. اگر P^*، ESS باشد در این صورت یک تعادل پایدار معادله ی RD نیز می باشد.

← مفهوم.

مثلاً داشتیم : $\nu(p)=\prod_{i=1}^{n}(p_i)^{p_i^*}$ ←

$$\dot{\nu}(p) = \nu(p)(p^* - p)^T A p$$

ادامه جلسه بعد

در صورتی P_1^* پایداری تکاملی خواهد بود که ث>0 باشد:

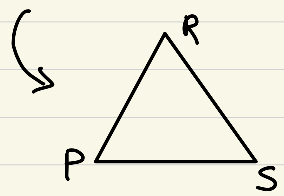
$$P^{*T}AP^* = PAP^*\ ,\ P^*AP > P^TAP \leftarrow \text{ESS}$$

← ث>0 خواهد شد پس P_1^* پایدار مجانبی است.

* عکس این "لزوما" برقرار نیست.

مثال نقض:

$$A = \begin{array}{c|ccc} & R & S & P \\ \hline R & 1 & 5 & 0 \\ S & 0 & 1 & 5 \\ P & 5 & 0 & 4 \end{array}$$



← نشان دهید نقطه‌ی X، A.S (پایدار مجانبی) برای

معادله‌ی RD می‌باشد: $X = \left(\frac{3}{18}, \frac{8}{18}, \frac{7}{18}\right)^T$

← می‌توان نشان داد این یک تعادل نش مخلوط است

و پس با 2 روش گفته شده (خطی سازی یا پتانسیل)،

نشان می‌دهیم که این تعادل پایدار مجانبی ست.

← برای اینکه نشان دهیم ESS نیست:

یک بردار دلخواه مثل $P = (0\ 0\ 1)^T$ در نظر بگیرید

شرط اول ESS. $U(X,X) = U(P,X)$

$$X^TAX = P^TAX$$

شرط دوم ESS: $\rightarrow X^TAP = \left(\frac{3}{18}, \frac{8}{18}, \frac{7}{18}\right)AP$

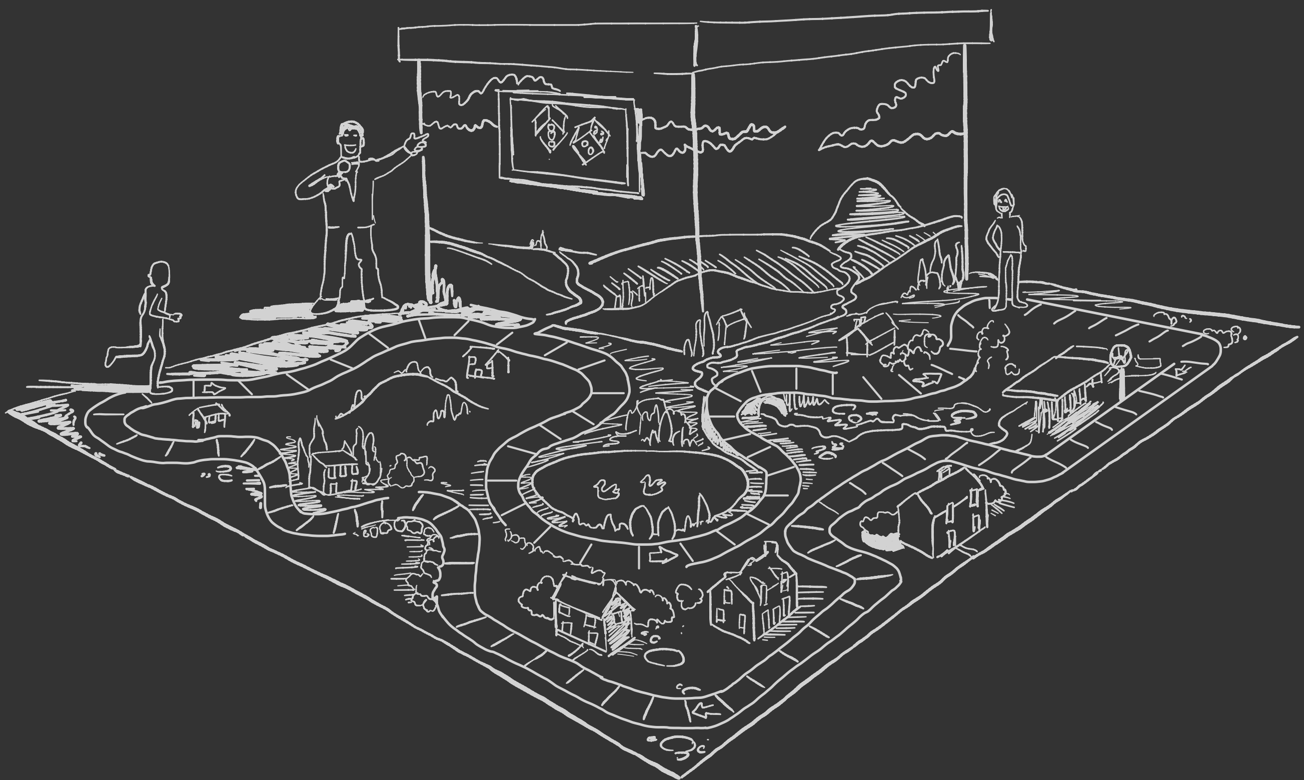
$$AP = \begin{pmatrix} 0 \\ 5 \\ 4 \end{pmatrix} \quad = \frac{68}{18}$$

$$\rightarrow P^TAP = 4 \implies X^TAP < P^TAP \implies$$

X یک ESS نیست!

CHAPTER 6

~ Mechanism Design



Shiva Shakeri

طراحی مکانیزم

← ما چگونه یک بازی را طراحی کنیم؟

← قوانین بازی طراحی می‌شوند تا هدف بازی را تعیین کند. این هدف همان تعادل بازی مطلوب است

← پس طراح بازی قوانین بازی را طوری طراحی می‌کند که بازی دارای تعادلی باشد که اهداف او را تأمین کند.

⇐ خودمان را محدود به دسته‌ی خاصی می‌کنیم:

Resource Allocation

← مانند Auction ها: هدف: تخصیص کالا - قیمت به خریداران

از طرفی در مزایده‌ها، agent بالادستی از ارزش کالا برای خریداران اطلاع ندارد

- مجموعه‌ی بازیگران $I = \{1, \dots, N\}$ ، چگالی احتمال ، توزیع احتمال
- نوع (Type) بازیگر (ارزش کالا نزد بازیگر i): $v_i \sim F_i - f_i$

← یکی از دلایلی که agent بالادستی، می‌خواهد از طریق غیرمستقیم یا طراحی بازی به اهداف خود برسد، نداشتن اطلاعات از Type یا اطلاعات private بازیگران است.

$\Rightarrow v_i \in [0 \ \ \bar{v}_i] = V_i$ ، $V = V_1 \times \dots \times V_N$

- **اجزای مکانیزم**: Mechanism (B, π, ρ)

→ B_i : مجموعه‌ی پیام‌ها (استراتژی یا تصمیم) بازیگران

$$B = B_1 \times \dots \times B_N$$

← مفهوم B: agent بالادستی، نسبت به agent های پایین، اطلاعات ناقص دارد و برای تخصیص یا allocationی که می‌خواهد انجام دهد، نیاز دارد از آنها پیام دریافت کند. این پیام‌ها به اشکال مختلف می‌توانند باشند. مثلاً می‌تواند از بازیگران بخواهد که مستقیماً Type خود را اعلام کنند و در مقابل این امکان هم وجود دارد که بازیگران صادقانه رفتار نکنند.

→ π : Allocation function

$$\pi : B \xrightarrow{\text{قانون تخصیص}} \Delta$$

مجموعه‌ی توزیع‌های احتمال روی I

← مفهوم π. تعیین می‌کند که کدام کالا به کدام بازیگر می‌رسد. مثلاً اگر کالا تقسیم‌پذیر باشد، می‌توان گفت 20% کالا به P1 و 80% آن به P2 می‌رسد. یا مثلاً 10% احتمال دارد که کالا به P1 و 90% احتمال دارد کالا به P2 برسد.

→ ρ : Payment function

$$\rho : B \to \mathbb{R}^N$$

← مفهوم ρ. ρ یک تابع قطعی یا deterministic است. مشخص می‌شود هر شخص بابت bidی که داده (و براساس آن یک کالا یا بخشی از کالا به آن تخصیص داده شده)، باید چقدر بپردازد.

* $b \in B \to (\pi(b), \rho(b))$: Outcome of the Mechanism

$b = (b_1, \dots, b_N)$

⇐ بدین ترتیب هر مکانیزم یک بازی (بیزی) بین خریداران القا می‌کند:

$$G \begin{cases} \text{Players} : I = \{1, \dots, N\} \\ \text{Strategies} : \beta_i : V_i \to B_i \\ \text{Payoff} : \underbrace{\text{Utility}}_{\pi_i(b)} - \underbrace{\text{Cost}}_{\rho_i(b)} \end{cases}$$

⇐ به این ترتیب یک تعادل نش (بیزی) برای بازی G. قابل تعریف است:

تعریف. یک تعادل نش (بیزی) $\beta(\cdot)$ برای بازی القا شده توسط مکانیزم طراحی شده است، اگر به ازای هر $i \in I$ و v_i مربوط به هر بازیگر، بهترین پاسخ به استراتژی β_{-i} باشد بطوریکه متوسط سود بازیگر i را بیشینه کند.

مثال طراحی مکانیزم

← مثال 1 : second price auction

• B : مجموعه‌ی مقدار قیمت پیشنهادی بازیگران

• π : قانون تخصیص: با احتمال 1 ، کالا به بیش‌ترین پیشنهاد تخصیص داده می‌شود.

• μ : قانون پرداخت . برنده به میزان پیشنهاد نفر دوم ، و بقیه 0 پرداخت می‌کنند.

Direct Mechanism

مکانیزم مستقیم روی نقش B در مکانیزم تمرکز دارد.

دراین نوع مکانیزم ، $B=V$ یا $B_i=V_i$ در نظر می‌گیریم.

← مفهوم: یعنی message space یا B ، دقیقاً" همان ارزش کالا یا Type بازیگران می‌باشد.

در واقع agent بالا دستی یا مدیر سیستم از بازیگران می‌خواهد که دقیقاً" اطلاعاتی که نزد (یعنی Typeشان) را برای او ارسال کنند.

← چالش :

• بازیگران ممکن است مقدار واقعی ارزش خود را بیان نکنند. پس چالش این است که π, μ چگونه طراحی شوند که بازیگران مجبور باشند صادقانه بازی کنند.

* تعادل صادقانه: اگر تحت مکانیزم مستقیم یک تعادل نش وجود داشته باشد که در آن خریداران ارزش کالای خود (تایپ خود) را درست ارسال کنند. آنگاه مکانیزم دارای تعادل صادقانه یا truthful می‌باشد.

← نتیجه‌ی مکانیزم مستقیم. $(\underbrace{\omega(x)}_{\text{تابع تخصیص مقدار}}, \underbrace{p(x)}_{\text{تابع تخصیص قیمت}})$

بردار مقدار اعلام شده از ارزش کالا : $x = (x_1, \dots, x_N)$ ←
نزد بازیگران ≡ بردار پیام

← مقدار اعلام شده در تعادل نش ≡ مقدار حقیقی ارزش کالا

⇓

تعادل نش صادقانه

• تعریف Revelation Principle

← به ازای هر مکانیزم (B, π, μ) و تعادل نش β برای این مکانیزم ، یک مکانیزم مستقیم وجود دارد که:

1- تعادل نش (بیزی) برای بازیگران این است که مقدار ارزش خود را بطور صادقانه ارسال کنند. (truthful E.q.)

2- نتیجه‌ی مکانیزم مستقیم در تعادل، مانند تعادل β در مکانیزم اصلی خواهد بود.

← ویژگی ها :

• به ازای هر مکانیزم داریم ولزوماً" مستقیم نیست.

• یک مکانیزم مستقیم با ویژگی های گفته شده ، همواره وجود دارد.

داریم: $\omega(x) = \pi(\beta(x))$. که $\omega(x)$ ، تابع تخصیص مکانیزم مستقیم ، π ، تابع تخصیص مکانیزم اصلی و β ، استراتژی تعادلی مکانیزم اصلی هستند. ← $\omega(\cdot)=\pi(\beta(\cdot))$

و داریم $p(x)=\mu(\beta(x))$ که $p(x)$ ، تابع پرداخت مکانیزم مستقیم ، μ ، تابع پرداخت مکانیزم اصلی و β ، استراتژی تعادلی مکانیزم اصلی هستند. ← $p(\cdot)=\mu(\beta(\cdot))$

← چرا دو ویژگی بیان شده ، برقرار هستند؟

در مکانیزم اصلی ، هر بازیگر یک bid مثل b_i اعلام می‌کند که $b_i=\beta(v_i)$ (با فرض متقارن بودن β). در این صورت b_i که بازیگر می‌دهد ، تابعی از ارزش واقعی اوست ← ادامه ص بعد

در مکانیزم مستقیم، $b_i = x_1$ است x_i چیزی ست که بازیگر به عنوان Type یا ارزش خود به agent بالا دستی اعلام می کند.

در مکانیزم اصلی مفهوم تعادل نش این است که هر بازیگر اگر از استراتژی تعادلی β تخطی کند، ضرر می کند. بدیهی ست این جا هر بازیگر ارزش واقعی خود را در نظر می گیرد و bid ش را براساس آن، اعلام می کند. دراین صورت اگر از β تخطی کند، ضرر می کند.

در مکانیزم مستقیم، ادعا می کنیم که تعادل نش زمانی برقرار می شود که $x_i = v_i$ باشد یعنی هر بازیگر مقدار ارزش واقعی اش را اعلام کند. چون اگر بازیگر به جای v_i مقدار دیگری را اعلام کند، معادل این است که در مکانیزم اصلی از β تخطی کرده. چون β هم در مکانیزم اصلی و هم مستقیم اثر خود را بصورت یکسان می گذارد. چگونه؟ با $\omega(\cdot) = \pi(\beta(\cdot))$ و با $p(\cdot) = \mu(\beta(\cdot))$ که گفتیم و β به کار برده شده در آنها، مربوط به مکانیزم اصلی ست و همینطور آن براین فرض استوار است که هرکس مقدار واقعی خودش را می داند و براساس Type واقعی اش باید تصمیم گیری کند. پس اگر بازیگری β را به β متفاوتی تغییر دهد ضرر کند مثل اینکه β را fix بگیریم و بگوییم اگر بازیگر از v_i اش تخطی کند، ضرر می کند و این دو معادل یکدیگرند.

دومین ویژگی به outcome مکانیزم اشاره می کند:

$$(\pi(\beta(v)), \mu(\beta(v))) = \underbrace{(\omega(v), p(v))}_{\text{مکانیزم مستقیم}}$$

مثال: second price auction

$\beta(v) = v \Rightarrow \omega = \pi ,\ P = \mu$

The VCG Mechanism

• یک مکانیزم مستقیم است.

• فرض: توابع سود بازیگران quasi linear می باشند

تابع سود $U_i = u_i(\omega, v_i) - P_i$

بردار تخصیص — نوع بازیگر — پرداخت بازیگر

همه‌ی تخصیص های ممکن $\omega \in \Omega$ ، $\omega = (\omega_1, \ldots, \omega_N)$ و

← quasi linear: تابع نسبت به یکی از پارامترهایش خطی باشد.
که در این جا U_i نسبت به P_i همواره خطی ست.

در این نوع مکانیزم:

1- B: $B = V$

2- قانون تخصیص

$$\sum_{i=1}^{N} u_i(\omega^*(\hat{v}), \hat{v}_i) \geq \sum_{i=1}^{N} u_i(\omega, \hat{v}) \quad \forall \omega \in \Omega$$

$\hat{v}$: مقدار ارزش (نوع) اعلام شده از سوی بازیگران است

$$\hat{v} = (\hat{v}_1, \ldots, \hat{v}_N) = (\hat{v}_i, \hat{v}_{-i}) \in V = B$$

← مفهوم فرمول:

agent بالا دستی، $\hat{v}$ را به رسمیت می‌شناسد و براساس آن allocate می‌کند و هدف را Max کردن سود بازیگران می‌گذارد. طبق فرمول تخصیص ω^* را به گونه ای انجام می‌دهد که مجموع utility بازیگران (که از ω^* بدست می‌آورند) نسبت به هر نوع تخصیص دیگری، Max شود. به بیان دیگر:

$$\omega^*(\hat{v}) \in \operatorname*{Argmax}_{\omega \in \Omega} \sum_{i=1}^{N} u_i(\omega, \hat{v}_i)$$

3- قانون پرداخت:

$$P_i(\hat{v}) = \overbrace{\max_{\omega} \sum_{j \neq i} u_j(\omega, \hat{v}_j)}^{\text{term 1}} - \underbrace{\sum_{j \neq i} u_j(\omega^*, \hat{v}_j)}_{\text{term 2}}$$

← مفهوم: این تابع بسیار هوشمندانه طراحی شده و باعث بوجود آمدن تعادل نش صادقانه می‌شود. term 1 همان قانون پرداخت است، وقتی که بازیگر i نباشد. یعنی اگر بازیگر i نبود، مجموع utility بازیگران (به جز i) Max شود. در term 2، همان ω^* که در حضور بازیگر i بود را در نظر می‌گیریم، اما مجموع utility همه‌ی بازیگران غیراز i را در حالتی که بازیگر i در بازی باشد، محاسبه می‌شود. بطور خلاصه داریم:

$P_i(\hat{v}) =$ مجموع utility بازیگران غیراز i، اگر i در بازی نباشد − مجموع utility بازیگران غیراز i، اگر i در بازی باشد

term 1 — term2 — social cost

← منطق جالب این تابع پرداخت:

term2: علامت پشت این عبارت منفی ست. چون اگر بودن تو باعث می‌شود بقیه utility بیش‌تری کسب کنند، payment باید منفی باشد. چون انگار جامعه به تو بدهکار میشه. پس تو بابت بودنت باید پول دریافت کنی!

term1: اگر بودنت باعث شده مجموع utility بازیگران کم بشه، وجودت برای جامعه هزینه زا بوده و تو باید پرداخت بکنی.

← ویژگی های VCG :

1- مجموع utility بازیگران را Max می‌کند:

ذاتاً" در تعریف allocation func خودش را بروز می‌دهد.
اگر یک pareto optimality در utility بازیگران
ایجاد می‌کند

2- بصورت قضیه‌ی زیر بیان می‌شود.

قضیه . مکانیزم VCG دارای تعادل صادقانه غالب است.

اثبات . برهان خلف ←

$x_i = v_i$: تایپ حقیقی بازیگر i را v_i فرض می‌کنیم.

$x_{-i} = \hat{v}_{-i}$: ارزش اعلام شده سایر بازیگران که لزوماً" صادقانه نیست.

$\hat{v}_i \neq v_i$ ← مقداری که بازیگر i اعلام می‌کند.

بازیگر i صادقانه عمل نکند.

فرض خلف : فرض کنیم بازیگر i، utility بیش‌تری کسب کرد:

$$u_i(w^*(\hat{v}_i, x_{-i}), x_i) - P_i(\hat{v}_i, x_{-i}) > u_i(w^*(x_i, x_{-i}), x_i) - P_i(x_i, x_{-i})$$

تایپ بازیگر i در فرمول P_i، در term2 تاثیر دارد. چراکه

در term2، از w^* استفاده شده که تابعی از تایپ بازیگر i

است و term1 به این پارامتر وابستگی ندارد و در دو طرف نامساوی

یکسان است.

$$u_i(w^*(\hat{v}_i, x_{-i}), x_i) + \sum_{j\neq i} u_j(w^*(\hat{v}_i, x_{-i}), x_j) > u_i(w^*(x_i, x_{-i}), x_i) + \sum_{j\neq i} u_j(w^*(x_i, x_{-i}), x_j)$$

$$\Rightarrow \underbrace{\sum_{j=1}^{N} u_j(w^*(\hat{v}_i, x_{-i}), x_j)}_{1} > \underbrace{\sum_{j=1}^{N} u_j(w^*(x_i, x_{-i}), x_j)}_{2}$$

طبق تعریف برای w^*، این عبارت بزرگتر است از :

$$\underbrace{\sum_{j=1}^{N} u_j(w, x_j)}_{3}, \ \forall w \in \Omega$$

ادامه ستون بعد

پس انگار داریم : $1 > 2 > 3$

اما در 1، خود $w^*(\hat{v}_i, x_{-i})$ هم یکی از همان $w \in \Omega$

هایی ست که نامساوی دوم را بیان کردیم پس 1 نوع

خاصی از عبارت 3 است و باید کمتر از عبارت 2 شود.

که این فرض خلف را نقض کرده و قضیه را اثبات می‌کند!

مثال VCG

← مثال 1 : SPA یک VCG است، اگر از بازیگران مقدار
ارزش آنها بعنوان استراتژی ایشان) خواسته شود.

خودتون بنویسین :))

معیارهای بهینگی مکانیزم

1) Truthfulness: در مکانیزم مستقیم اعلام صادقانه مقدار ارزش $v_i = \hat{v}_i$، یک تعادل نش باشد.

2) Efficiency: یا همان بهینگی Pareto می‌باشد:

$$\sum u_i(\omega^*, v_i) \geq \sum u_i(\omega, v_i), \quad \forall \omega \in \Omega, \ \forall v \in V$$

3) Budjet Balance: $\forall v \in V: \sum_i P_i(v) = 0$

مجموع payment بازیگران صفر می‌شود

"پول ایجاد نمیشه و از بین نمیره، بلکه بین بازیگران جابجا میشه".

مکانیزم VCG این ویژگی را ندارد.

← شرط ضعیف‌تر: Weak Budjet Balance

$$\forall v \in V: \sum_i P_i(v) \geq 0$$

پول از بین نرود و بازی موجب ضرر طراح بازی نشود.

4) Individual Rationality:

$$u_i(\omega(v), v_i) - P_i(v) \geq 0$$

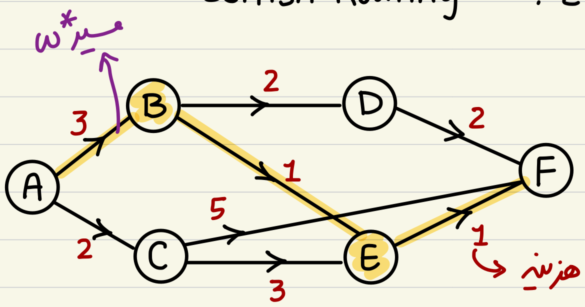
بازی موجب ضرر بازیگران نشود و بازیگران امکان شرکت در این بازی را در هر شرایطی داشته باشند.

5) Revenue Maximization: سود طراح Max شود:

$$\max_{\omega, p} E_v\left[\sum_{i=1}^{N} P_i(v)\right] \ \text{s.t.} \ldots$$

مثال VCG

← مثال 2: Selfish Routing



هدف. رفتن از A به F، با کمترین هزینه

بازیگران: صاحبان لینک‌ها: {AB, BD, EF, ...}

تابع هدف: پرداخت به agent بالادستی - هزینه لینک

$$= -u(\omega(\hat{v}), v_i) - P_i(\hat{v})$$

بردار مقداری که بازیگرها به‌عنوان هزینه خط به طراح اعلام می‌کنند.

← مکانیزم VCG:

$$\omega^* = \max_{\omega \in \Omega} -\sum_i u_i(\omega(\hat{v}), \hat{v}_i)$$

همه‌ی مسیرهای ممکن از A به F

← مثلاً برای محاسبه‌ی P_{AB}:

• اگر AB نباشه بقیه در مجموع چقدر پرداخت می‌کنند؟ در نبود AB، مسیر کم‌هزینه، AC-CE-EF بدست می‌آید و مجموع هزینه‌شان -6 می‌شود.

• اگر AB باشد، بقیه در مجموع چقدر هزینه می‌کنند؟ مسیر کم‌هزینه AB-BE-EF است و هزینه‌ی آنها به‌جز AB، مجموعاً "-2" است.

$$\Rightarrow P_{AB} = -6-(-2) = -4 \rightarrow U_{AB} = (-3)-(-4) = 1$$

← P بازیگرانی که در ω^* نباشند، صفر است. مثل AC، BD، CE، CF، DE.

مثلاً $P_{AC} = (-5)-(-5) = 0$

ادامه ص بعد

← مقدار پرداخت BE , EF .

$$P_{BE} = (-6) - (-4) = -2$$

$$P_{EF} = (-7) - (-4) = -3$$

← فرض کنید اگر AB صادقانه بازی نکند و هزینه واقعی خود (3) را اشتباه اعلام کند :

← اگر 2 اعلام کند : دوباره همان مسیر ، انتخاب می‌شود.

$$P_{AB} = (-6) - (-2) = -4$$

$U_{AB} = \underbrace{(-3)}_{\text{هزینه واقعی}} - (-4) = 1 \Rightarrow$ برابر با مقدار سود حالت صادقانه شد .

← اگر 4 اعلام کند : دو مسیر بهینه می‌شدند.

P_{AB} → مسیر قبلی w^* → $P_{AB} = -4 \rightarrow U_{AB} = 1$: "مشابه قبلا"

P_{AB} → $(AC-CE-EF) \rightarrow P_{AB} = 0 \rightarrow U_{AB} = 0$: کمتر شد

← اگر بیش‌تر از 4 اعلام کند : مسیر بهینه شامل AB نمی‌شود

$\Rightarrow P_{AB} = 0 , U_{AB} = 0$

⇐ پس در این حالت دروغ گفتن به ضرر بازیگر می‌شود .

که همان چیزی ست که از مکانیزم VCG انتظار داشتیم .