THE: Opakované hry; Korelované ekvilibrium
Repeated Games; Correlated Equilibrium

Martin Hrubý

Brno University of Technology
Brno
Czech Republic

November 9, 2010

Úvod

Čerpáno z:

- Myerson, R. B.: Game Theory, Analysis of conflict, Cambridge, Harvard University Press (1991)
- Mailath, G. J., Samuelson, L.: Repeated Games and Reputation, Oxford University Press, 2006
- Osborne, M. J.: An Introduction to Game Theory, Oxford University Press, 2004
- Osborne, M., Rubinstein, A.: A Course in Game Theory, The MIT Press, 1994
- Hrubý, M.: Algorithmic Approaches in Game-Theoretical Modeling and Simulation, AUCO Czech Economic Review, 2(3), on-line: auco.cuni.cz
- Asu Ozdaglar: přednášky "Repeated Games"
- Ratliff, J.: A Folk Theorem Sampler

War and peace (R. Aumann, Nobel prize lecture)

- Teorie her není schopna dodat magickou formuli, která vyřeší všechny konflikty. Neexistuje akademická disciplína, která by toto dokázala.
- Cílem je porozumět konfliktu. Když konfliktu porozumíme, jsme schopni hledat jeho řešení.
  - Je to jako s rakovinou: A) konvenční léčba (řešit konkrétní projevy/následky), B) snaha pochopit vnitřní principy (DNA, procesy v těle).
  - Na válku se pohlíží jako na iracionální chování a to je chyba. Válka není iracionální.
  - Jedinec je racionální, pokud koná to nejlepší dle jeho informací.
  - Hraje zřejmě nekooperativní ekvilibrium.
- Kooperace – předpokládáme, že ve hře existuje efektivnější profil než je NE.
- Profil je kooperativní, pokud žádný hráč není schopen pro sebe garantovat lepší výsledek.

Opakované hry, úvod

Zatím známe hry:

- Nekooperativní s nulovým/nenulovým součtem v normální (maticové) formě – hráči chápou situaci (jednotahovost) a strategicky analyzují množinu profilů.
- Nekooperativní hry v rozšířené formě – hráči se v tazích střídají, existuje víc tahů. Velmi záleží na stavu informace ve hře (dokonalá/nedokonalá informace o historii hry).
- Kooperativní hry definované charakteristickou funkcí – hráči hledají optimální rozložení koalic a optimální dělbu zisku přisouzené koalici (imputace).

Opakujeme tzv. "base game" (stage game) – základní hru.

Příhoda s Hodžou Nasredinem

Hodža Nasredin byl známý muslimský filosof a humorista. Jednou přišel do lázní, ale lazebník ho odbyl špatnými službami. Při odchodu z lázní Hodža zaplatil velkou mincí. Při příští návštěvě se lazebník překonával, ale Hodža při odchodu zaplatil malou mincí. Lazebník se divil: "Hodžo, posledně jsi zaplatil hodně za ... služby a za dnešní (lepší) služby málo. Proč?". Hodža odpověděl: "Posledně jsem platil za dnes a dnes jsem platil za posledně.".

Pokud lazebník pochopil Hodžovu strategii v opakované hře "návštěva lázní", pak budou od toho okamžiku oba hrát kooperativní profil.

Opakované hry nám umožňují modelovat sociální normy, vliv "trestu" na další opakování hry a přirozeně vnucenou spolupráci.

Opakované hry, úvod

Dále na úrovni nekooperativních her:

- Konfliktní situace se může opakovat. Pokud si hráči fakt opakování uvědomují, mluvíme o opakovaných hrách.
- Hráči se mohou poučit o chování protihráče, tzn. mají paměť.
- Je prostor pro vyjednávání. Je možno "trestat".
- Rozlišujeme dle opakování:
  - Konečné opakování – hra se opakuje T-krát, kde T je konkrétní předem známé číslo.
  - Pravděpodobnostní opakování – s pravděpodobností p se hra znovu zopakuje.
  - Nekonečné opakování – hra se opakuje bez ukončení. Je jasné, že všechno má konec. Spíše tím modelujeme situaci, kdy se o ukončení neuvažuje (hraje se "doživotně") nebo T je neznámé.
- Hráč na počátku hry ví o opakování a přizpůsobí tomu svůj první tah. Navazujeme tímto na sekvenční hry.

Hráč bere v úvahu, že jeho tah ovlivní bezprostřední výsledek hry, ale i budoucnost hry (vnímá svoji "reputaci").

Opakované vězňovo dilema

	C	D
C	1,1	-1,2
D	2,-1	0,0

Předpokládejme nyní strategii (nazvanou *grim trigger strategy*):

- Hraju C tak dlouho, dokud druhý hraje C.
- Pokud druhý zahraje D, od této doby hraju navždy D (trest).

Pokud někdo zahraje D, krátkodobě získá zisk 2, ale v dalších hrách už jenom 0 (namísto kooperativního 1). To znamená, že pro racionálně (citlivě k budoucnosti) uvažujícího hráče je zisk $(1, 1, 1, 1, ...)$ dominující nad $(..., 2, 0, 0, ...)$.

Strategie C je tedy best-response na *grim trigger strategy*. Tato strategie je NE v opakovaném Vězňově dilematu. Strategie (D, D) ovšem taky.

Strategie v opakované hře

Je nutno odlišovat dva pohledy na opakovou hru:

- Bázová (základní) hra má strategie dle $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$.
- Opakovaná hra má strategie definované nad celou posloupností opakování bázové hry.
- Situaci opakované hry modelujeme jako rozhodnutí hráče, jak se bude v rámci opakování her konzistentně chovat.

Příklady strategií v opakované hře:

- Vždy budu hrát ...
- Budu hrát strategii, kterou v předešlém kole hrál soupeř.
- Budu hrát kooperativní strategii, pokud mě soupeř zradí, budu ho trestat po dobu k kol hraním nekoop. strategie, a pak budu hrát opět kooperativní strategii (trest a pozdější odpuštění).

Soutěž v opakovaném Vězňově dilematu.

Preference v opakované hře, ohodnocovací faktor δ

Každý hráč má v každém kole t hry užitek $U_i(s^t), s^t \in S$ a definuje ohodnocovací faktor (discounting factor) $\delta_i \in \langle 0, 1)$, kterým ohodnocuje postupné výhry při hraní profilů $(s^1, s^2, ..., s^T)$, tedy $s^j \in S$:

$$U_i(s^1) + \delta_i U_i(s^2) + ... + \delta_i^{T-1} U_i(s^T) = \sum_{t=1}^{T} \delta_i^{t-1} U_i(s^t)$$

Pozor, δ_i^t je $t-$mocnina δ_i.

- Pokud je δ_i blízko 0, hráč i je velmi necitlivý k budoucnosti a naopak.
- Pokud je $\delta_i = 0$, pak hráč i neví o budoucnosti (jednotahová hra).
- Budeme nadále klást $\delta_i = \delta; \forall i \in Q$.

Preference v opakované hře

Chceme ohodnotit posloupnost zisků a vyjádřit ji jedním číslem. Pomocí něho budeme porovnávat posloupnosti zisků (preference nad posloupnostmi).

Posloupnost $W = (w^1, w^2, ..., w^T)$ má sumu $V = \sum_{t=1}^{T} \delta_i^{t-1} w^t$. Existuje nyní takové číslo c, že by byl hráč indiferentní mezi W a $(c, c, ...)$?

Suma pro $(c, c, ...)$ je $\frac{c}{1-\delta}$ (suma geometrické řady). Hráč je indiferentní mezi W a $(c, c, ...)$, pokud $c = (1-\delta)V$. Proto nazýváme $(1-\delta)V$ průměrným ziskem, tzn.

$$\pi_i = (1-\delta) \sum_{t=1}^{T} \delta_i^{t-1} w^t$$

Definice hry v rozšířené formě (opakování)

Definition

Konečná sekvenční hra s dokonalou informací je Γ^E je čtveřice (Q, H, p, U), kde:

- Q je množina hráčů,
- H je množina historií (kontextů, uzlů). H^T označuje množinu terminálních historií. $H^0 = \emptyset$ označuje počáteční uzel.
- $p(h) : H \setminus H^T \rightarrow Q$ přiřazuje každé neterminální historii h hráče, který je aktuálně na tahu.
- U je vektor užitkových funkcí $U_i(h) : H^T \rightarrow \mathbb{R}^1$ pro všechny $i \in Q$.

Pozn.: Miminum pro definici hry s dokonalou informací je (Q, H, p). Pozn.: Zavedeme: $A(h)$ je množina akcí pro každé $h \in H \setminus H^T$, tzn. množina proveditelných akcí po dosažení historie h.

Definice opakované hry

Definition

Mějme $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$. T-krát opakovaná hra Γ^T s bázovou hrou Γ a ohodnocovacím faktorem δ je hra v rozšířené formě s dokonalou informací a současnými tahy hráčů, kde:

- množina terminálních historií je množina sekvencí $(s^1, s^2, ..., s^T)$ profilů hry Γ.
- $p(h) = Q$ pro všechny historie $h = (s^1, s^2, ..., s^t)$ $(\forall t \in \{1, ..., T-1\})$.
- $A_i(h) = S_i$ pro všechny historie $h = (s^1, s^2, ..., s^t)$ $(\forall t \in \{1, ..., T-1\})$.
- každý hráč i ohodnocuje každou terminální historii $(s^1, s^2, ..., s^T)$ dle svého průměrného zisku $(1-\delta)\sum_{t=1}^{T} \delta^{t-1} U_i(s^t)$.

Poznámka k definici: nekonečně opakované hry

V případě nekonečně opakované hry Γ^∞:

- Terminální historie jsou nekonečné sekvence $(s^1, s^2, \ldots)$.
- Průměrný zisk je $(1-\delta)\sum_{t=1}^{\infty} \delta^{t-1} U_i(s^t)$.

V obou případech (konečné, nekonečné opakování) je terminální historie nazývána taky "outcome path" (výsledková posloupnost).

Path ve smyslu cesta grafem hry v rozšířené formě.

Zpátky k vězňům

	C	D
C	1,1	-1,2
D	2,-1	0,0

Hraje se T-krát:

- $T = 1$, pak je $NE = (D, D)$
- $T > 1$ a T je konečné, zpětnou indukcí získáme SPNE (Subgame Perfect NE), které je opět $NE = (D, D)$

Zpátky k vězňům

Hraje se ∞-krát:

Proposition

Je-li $\delta \geq \frac{1}{2}$, pak má ∞-krát opakované Vězňovo dilema (dle předchozího zadání) SPNE (C, C), které se hraje v každém tahu.

Proof.

Předpokládejme grim trigger strategii.
Pokud hráč i bude hrát stále C, je jeho zisk $(1-\delta)[1+\delta+\delta^2+...] = 1$. Zahraje-li D v libovolném, např. prvním, tahu, pak je jeho zisk $(1-\delta)[2+0+0+...] = 2(1-\delta)$. Pokud je $\delta = \frac{1}{2}$, pak je to srovnatelný zisk, pokud $\delta > \frac{1}{2}$, pak je horší.

$\square$

Folk theorems (General feasibility theorems, R. Myerson)

Forma "solution concept" v opakovaných hrách.

Folk theorémy jsou třída vět, které v opakovaných hrách ukazují, že kterýkoliv výsledek hry je přijatelné řešení, jestli je pro hráče alespoň tak dobrý jako jejich minimax zisk.

Definition

Mějme hru $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$. Minimax zisk hráče $i \in Q$ ve hře Γ je:

$$\underline{v_i} = \min_{s_{-i} \in S_{-i}} \left(\max_{s_i \in S_i} [U_i(s_i, s_{-i})] \right)$$

$\underline{v_i}$ je tedy nejnižší BR_i. Předpokládáme, že hráči jsou schopni si kolektivně uvědomit míru přijatelného výsledku hry (feasible outcome). Ten představuje jakousi sociální normu.

Co je Nashovo ekvilibrim v opakovaných hrách?

- Obecný význam NE známe (je to profil, kde žádný hráč změnou své strategie nezíská pro sebe lepší výsledek).
- Definice strategie v opakovaných hrách byla dána (je to chování hráče přes všechna opakování hry).
- Zisk hráče v opakované hře je suma všech dosažených zisků přes všechna opakování.
- Hráči hledají dlouhodobou strategii, která jim přinese optimální zisk – v NE.
- V rámci opakované hry, hráči chápají svůj minimax užitek a všechny lepší chápou jako feasible.

Jako vždy u NE, může existovat více NE. Dlouhodobé strategie nejsou explicitně definovány (jejich výčet), pokud by byly, pak vytvoříme superhru (supergame), která je pro nás obvyklá nekooperativní hra v normální formě.

Ukázka superhry nad Vězňovým dilematem

Uvažujeme superhru Γ nad Vězňovým dilematem, které bude T-krát/∞-krát opakováno, kde $S_i = \{trigger, C, D\}$.

Užitky jsou:

A/B	trigger	C	D
trigger	(1,1)	(1,1)	(0,0)
C	(1,1)	(1,1)	(-1,2)
D	(0,0)	(2,-1)	(0,0)

Nash folk theorem pro nekonečně opakované Vězňovo dilema

Nechť G je hra Vězňovo dilema.

- Pro libovolný ohodnocovací faktor δ, $0 < \delta < 1$, průměrný ohodnocovací zisk každého hráče i v libovolném NE nekonečně opakované hry G je přinejmenším $U_i(D, D)$.
- Nechť (x_1, x_2) je feasible zisk v G takový, že $x_i > U_i(D, D); \forall i \in Q$. Pak existuje $\delta' < 1$ takový, že pokud $\delta > \delta'$, pak nekonečně opakovaná hra G vede k průměrnému ohodnocovacímu zisku x_i.
- Pro libovolný ohodnocovací faktor δ, nekonečně opakovaná hra G vede k průměrnému ohodnocovacímu zisku $U_i(D, D)$.

Zobecněnou podobu tohoto Folk theorému získáme náhradou $U_i(D, D)$ za minimax zisk hráčů.

Náznak důkazu Nash folk theoremu v ∞-opakované hře

Důkaz je založen na použití *grim trigger* strategie.

Definition

Řekneme, že výplatní vektor v je striktně individuálně racionální, pokud platí pro všechny hráče $i \in Q : v_i > \underline{v_i}$.

Definition

Minimax strategií proti hráči i nazveme sub-profil m^i_{-i} (ve dvouhráčové hře strategii protivníka), která mu přinese $\underline{v_i}$. Je to:

$$m^i_{-i} = arg \min_{s_{-i} \in S_{-i}} \left[\max_{s_i \in S_i} U_i(s_i, s_{-i}) \right]$$

Plyne z toho, že $BR_i(m^i_{-i}) = s'_i$ taková, že $U_i(s'_i, m^i_{-i}) = \underline{v_i}$.

Jiné znění Folk theoremu

Doplňme, že $V^* \subset \mathbb{R}^N$ je množina přijatelných (stejných jako $\underline{v_i}$) a striktně individuálně racionálních zisků.

Folk theorems ukazují, že pokud jsou hráči dostatečně citliví k budoucnosti (mají dostatečně vysoké δ_i), pak jsou schopni dosáhnout libovolného dosažitelného stritně individuálně racionálního zisku formou NE.

Theorem

Pro všechny $v = (v_1, v_2, ..., v_N) \in V^*$ existuje $\underline{\delta} < 1$ takové, že pro všechny $\delta > \underline{\delta}$, existuje NE v $\Gamma^\infty(\delta)$ s vektorem zisků v.

Jiné znění Folk theoremu, důkaz

Předpokládejme nejdříve, že existuje profil $s = (s_i)_{i \in Q}$ takový, že $U_i(s) = v_i$, $(v_i)_{i \in Q} \in V^*$ (jinak bychom museli zavést smíšené strategie, jde to, ale to bychom si to značně zkomplikovali).

Zavedeme si trigger strategii:

1. Hrej od nultého kola s_i tak dlouho, dokud protihráč hraje s_{-i}. Pokud protihráč(i) vybočí, přesuň se do stavu 2.
2. Hrej už jenom m_i^j (strategie, která garantuje $\underline{v_i}$).

Pokud hráči hrají kooperativně, mají ve hře zisk v. Hráč i, který vybočí v kole t, aby získal $\overline{v_i} = BR_i(s_{-i})$, získá maximálně:

$$(1-\delta)\left[v_i + \delta v_i + \cdots + \delta^{t-1} v_i + \delta^t \overline{v_i} + \delta^{t+1} \underline{v_i} + \delta^{t+2} \underline{v_i} + \cdots\right]$$

$$(1-\delta)\left[v_i + \delta v_i + \cdots + \delta^{t-1} v_i + \delta^t \overline{v_i} + \delta^{t+1} \underline{v_i} + \delta^{t+2} \underline{v_i} + \cdots\right]$$

$$= (1-\delta)[v_i + \delta v_i + \cdots + \delta^{t-1} v_i + \delta^t \overline{v_i} + \frac{\delta^{t+1} \underline{v_i}}{1-\delta}] =$$

$$= (1-\delta) v_i + (1-\delta)[\delta v_i + \cdots + \delta^{t-1} v_i] + (1-\delta)\delta^t \overline{v_i} + \delta^{t+1} \underline{v_i}$$

Druhý člen:

$$(1-\delta)[\delta v_i + \cdots + \delta^{t-1} v_i] = v_i(1+\delta+\cdots+\delta^{t-1}-\delta-\delta^2-\cdots+\delta^t) = v_i(\delta-\delta^t)$$

První a druhý člen:

$$(1-\delta) v_i + v_i(\delta - \delta^t) = v_i - \delta v_i + \delta v_i - v_i \delta^t = v_i(1-\delta^t)$$

Celkově (zisk hráče i, který v kole t vybočil z kooperativního profilu):

$$= (1-\delta^t)v_i + \delta^t(1-\delta)\overline{v_i} + \delta^{t+1}\underline{v_i}$$

Aby navrhovaná strategie (tzn. hrát kooperativně) byla optimální, musí platit:

$$v_i \geq (1-\delta^t)v_i + \delta^t(1-\delta)\overline{v_i} + \delta^{t+1}\underline{v_i}$$

$$v_i \geq v_i - \delta^t[v_i - \overline{v_i} + \delta\overline{v_i} - \delta\underline{v_i}]$$

$$v_i \geq v_i - \delta^t[v_i + (\delta-1)\overline{v_i} - \delta\underline{v_i}]$$

To bude platit, pokud:

$$v_i + (\delta-1)\overline{v_i} + \delta\underline{v_i} \geq 0$$

Přelomová je hodnota

$$\underline{\delta} = \max_{i \in Q} \left[\frac{\overline{v_i} - v_i}{\overline{v_i} - \underline{v_i}} \right]$$

$$\underline{v_i} < v_i < \overline{v_i} \Rightarrow \overline{v_i} - v_i < \overline{v_i} - \underline{v_i} \Rightarrow \underline{\delta} < 1$$

Pro všechny $\delta \geq \underline{\delta}$ je racionální hrát zvolenou strategii a neuhýbat. Resp. při $\delta = \underline{\delta}$ je deviující hráč na stejném zisku, jako by kooperoval, při $\delta > \underline{\delta}$ má striktně horší zisk.

$\square$

Tento Folk theorem ukazuje, že pro jistou míru paměti (nebo citlivosti k budoucnosti) je protihráč racionálně nucen ke kooperativní strategii.

Příklad

	L	R
U	6,6	0,-100
D	7,1	0,-100

$NE = (D, L)$, $\underline{v_1} = 0$, $\underline{v_2} = 1$. Minimax strategie hráče 2 je hrát R.

Nash Folk Theorem připouští hraní $(U, L) \Rightarrow (6, 6)$, ovšem hráč 2 musí hrát R kdykoliv hráč 1 uhne (a to je značně bolestivé).

$$\underline{\delta_i} = \frac{\overline{v_i} - v_i}{\overline{v_i} - \underline{v_i}} = \frac{7-6}{7-0} = \frac{1}{7}$$

Další Folk Theoremy v rámci projektů.

Jak lidé hrají v opakovaných hrách

Podle laboratorních experimentů (Asu Ozdaglar):

- Lidé hrají v opakovaných vězňových dilematech kooperativněji, než by se očekávalo.
- Dokonce i u konečně opakovaných hrách.
- Lidé jsou konfrontováni s maticí zisků, ale to není jejich veškerý užitek.
- Další složky užitku pramení ze sociálních preferencí.

Typy sociálních preferencí lidí:

- Altruismus – mají užitek z toho být hodní na druhé.
- Férovost – mají užitek z toho být fair na druhé.
- Pomstichtivost – lidé rádi trestají ty, co vybočí z férovosti nebo jiných sociálních norem.

Korelované ekvilibrium

- Korelované ekvilibrium je herní koncept, který zobecňuje Nashovo ekvilibrium.
- Poprvé definováno Robertem Aumannem (Nobel. cena, 2005) v článku "Aumann, R. (1974) *Subjectivity and correlation in randomized strategies*. Journal of Mathematical Economics 1:67-96."

Korelované ekvilibrium

- Korelované ekvilibrium (CE) je založeno na myšlence, že hráč volí svou strategii na základě pozorování nějakého veřejně známého signálu (signál je common knowledge). NE předpokládá, že hráč se rozhoduje pouze dle specifikace hry.
- Signál hráči "napoví" tah. Pokud hráč uvěří, že všichni protihráči hrají dle CE a vidí, že by si pohoršil hraním jiného tahu (než je ten napovězený), pak je profil tvořený těmito nápovědami korelované ekvilibrium.
- Existuje množina takových profilů. Budeme dále předpokládat racionalitu hráčů a volbu pareto dominantního profilu.

Co je to ta nápověda?

- Nečekáme, že existuje "centrální autorita", která by hráčům vyřešila jejich problém a navrhla jim řešení.
- Spíše tím modelujeme fakt, že hráči žijí ve společném informačním kontextu (čtou stejné noviny).

Příklad

Uvažujme hru:

	a	b
A	9,9	6,10
B	10,6	0,0

MNEs:

- (A, b) zisk $(6, 10)$
- (B, a) zisk $(10, 6)$
- $((\frac{6}{7}, \frac{1}{7}), (\frac{6}{7}, \frac{1}{7}))$ zisk $(8.57, 8.57)$

Příklad

Přepodpokládejme, že někdo hodí mincí (padne H nebo T) – veřejně pozorovatelný signál. Hráči souhlasí, že budou hrát

- (A, b) v případě H
- (B, a) v případě T
- očekávaný zisk na $(8, 8)$

Co Best-response?

- Padne-li H, hráč 1 ví, že hráč 2 bude hrát b (A je tedy Best-response)
- Padne-li T, hráč 1 ví, že hráč 2 bude hrát a (B je tedy Best-response)

Očekávané zisky nyní mohou výt vyšší než v MNE (pokud např. změníme v (A, a) zisky na $(8, 8)$).

Příklad - další experiment

Signál nemusí být až tak informačně úplný. Požádáme arbitra, aby hodil kostkou (výsledek n žádný z hráčů nevidí).

- Oznámíme hráči 1, zda-li n padlo do $\{1, 2\}$ nebo do $\{3, 4, 5, 6\}$.
- Oznámíme hráči 2, zda-li n padlo do $\{1, 2, 3, 4\}$ nebo do $\{5, 6\}$.
- Hráč 1 hraje B, pokud $n \in \{1, 2\}$ nebo A při $n \in \{3, 4, 5, 6\}$.
- Hráč 2 hraje a, pokud $n \in \{1, 2, 3, 4\}$ nebo b při $n \in \{5, 6\}$.
- očekávaný zisk je $(8.333, 8.333)$.

BR_1:

- pokud $n \in \{1, 2\}$, hráč 2 bude hrát a, pak BR_1 je hrát B
- pokud $n \in \{3, 4, 5, 6\}$, hráč 2 bude hrát $(\frac{1}{2}, \frac{1}{2})$, BR_1 je hrát A

$\pi_1 = \frac{2}{6}10 + \frac{4}{6}(\frac{1}{2}9 + \frac{1}{2}6) = 8.333$

Definice korelovaného ekvilibria

Mějme hru $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$.

- CE přiřazuje každému profilu $s \in S$ pravděpodobnostní ohodnocení $\pi(s)$.
- Řešením je tedy vektor $\pi = (\pi(s))_{s \in S}$.

Definition

Mějme hru $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$ a vektor $\pi = (\pi(s))_{s \in S}$ takový, že:

1. $\pi(s) \in \langle 0, 1 \rangle \quad \forall s \in S$
2. $\sum_{s \in S} \pi(s) = 1$

Vektor π je korelovaným ekvilibriem ve hře Γ, pokud platí pro všechny hráče $i \in Q$ a každé dvě strategie $s_i, d_i \in S_i, d_i \neq s_i$:

$$\sum_{s_{-i} \in S_{-i}} \pi(s)(U_i(s_i, s_{-i}) - U_i(d_i, s_{-i})) \geq 0$$

Kompatibilita Korelovaného ekvilibria

Theorem

Je-li σ^* MNE, pak $\pi = \left(\Pi_{i=1}^{N}\sigma^*(s_i)\right)_{s\in S}$ je korelované ekvilibrium.

Důkaz jako domácí příprava (odkaz na "perchta.fit").

Výpočet Korelovaného ekvilibria

Dobré zprávy o CE:

- CE je nadmnožina MNE, tzn. všechny MNE jsou současně CE. Plyne z toho, že *CE vždy existuje* (důvod: Nash dokázal, že MNE vždy existuje).
- Ch. Papadimitriou dokázal (článek "Computing correlated equilibria in multi-player games", In: 37th ACM Symposium on Theory of computing), že výpočet CE je v polynomické časové složitostní třídě.
- M. Hrubý navrhl algoritmus efektivního výpočtu CE se zapojením paralelismu (článek "Algorithmic Approaches to Game-Theoretical Modeling and Simulation, In: AUCO Czech Ecomomic Review, on-line: auco.cuni.cz).
  - Nástroj CE-Solver, http://perchta.fit.vutbr.cz/CE-Solver/
- Výpočet CE je založen na lineárním programování.

Výpočet CE obecně (LP-úloha)

Budeme hledat Pareto efektivní profil, který má CE vlastnost. To znamená, že maximalizujeme souhrnný užitek $Z(s)$ v CE profilu Z:

$$MAX : Z = \sum_{s \in S} \pi(s) Z(s)$$

$$Z(s) = \sum_{i \in Q} w_i U_i(s)$$

Váhování $(w_i)_{i \in Q}$:

- Na to v literatuře není moc názor...
- Nejjednodušší je nastavit $w_i = 1; \forall i \in Q$.
- Váhování vybírá Pareto optimální CE mezi všemi CE.

Výpočet CE obecně (LP-úloha)

Proměnné:

$$\pi(s) \in \langle 0, 1 \rangle \,; \forall s \in S$$

Omezující podmínky (constraints) jsou (dále pak "matice G"):

$$\sum_{s \in S} \pi(s) = 1$$

Výpočet CE obecně (matice G)

Omezení dané maticí užitků (obecně):

$$G \cdot \pi^T \geq 0$$

Matice G:

- Řádky matice G jsou indexovány i, j, k, kde $i \in Q$, $j, k \in S_i$. Máme tudíž $\sum_{i=1}^{N} |S_i| \cdot (|S_i| - 1)$ řádků.
- Sloupce jsou indexovány strategickými profily, je tudíž S sloupců v matici.
- Prvek matice $G_{i,j,k}$ v profilu $s \in S$ je dán:

$$G_{i,j,k}(s) = \begin{cases} U_i(s_i, s_{-i}) - U_i(k, s_{-i}) & s_i = j \\ 0 & jinak \end{cases}$$

Výpočet CE obecně

Řešíme LP-úlohu s proměnnými $\pi(s)$ a omezujícími podmínkami $(\sum_{s \in S} \pi(s) = 1, G \cdot \pi^T \geq 0)$. Maximalizujeme Z.

Výsledkem je naplnění proměnných $\pi(s)$ pravděpodobnostmi profilů. *Důležité je si uvědomit, že tento postup výpočtu vyhodnocuje jedno Pareto efektivní CE*. Není tudíž pravda, že existuje pouze jedno CE a to je jednoznačnou odpovědí na otázku "co hráči udělají?".

Tento algoritmus nijak není omezen počtem hráčů nebo jejich strategií – pouze výslednou časovou a paměťovou složitostí.

Demo

A/B	c	d
a	10,2	8,4
b	5,4	7,3

$$S = \{(a, c), (a, d), (b, c), (b, d)\}$$

G-matice (řádky modelují možné deviace hráčů, sloupce jsou všechny profily ve hře):

$j \to k/s$	ac	ad	bc	bd
$a \to b$	5	1		
$b \to a$			-5	-1
$c \to d$	-2		1	
$d \to c$		2		-1

Vede na LP-úlohu

Maximalizace $Z = \sum_{s \in S} \pi(s) Z(s)$.

$$\begin{gathered} 5\pi(ac) + \pi(ad) \geq 0 \\ -5\pi(bc) - \pi(bd) \geq 0 \\ -2\pi(ac) + \pi(bc) \geq 0 \\ 2\pi(ad) - \pi(bd) \geq 0 \end{gathered}$$

Druhá nerovnice jasně vede k $\pi(bc) = \pi(bd) = 0$. Z toho plyne: $-2\pi(ac) \geq 0$ platí pouze pro $\pi(ac) = 0$.

Zbývá $\pi(ad) \geq 0 \wedge \sum \pi(s) = 1$, tzn. $\pi(ad) = 1$.

Princip CE obecně – analytický efekt G-matice

- Množina nerovnic $G \cdot \pi^T \geq 0$ modeluje hráčovy tendence přecházet mezi strategiemi.
- Pokud je řádek $G_{i,j,k}$ celý záporný (resp. $G_{i,j,k}(s) \leq 0; \forall s \in S$), pak strategie k striktně dominuje nad j.
- Pokud je řádek nulový, pak jsou strategie ekvivalentní. Takový řádek může být z LP-constraints vypuštěn, neboť nemá žádný efekt.
- Pokud existují kladné i záporné koeficienty $G_{i,j,k}(s)$, pak nelze nic o strategiích j, k říct.
- Pokud existují pouze kladné a nulové koeficienty, pak strategie j slabě dominuje nad k (nelze říct, že dominuje striktně).

Matice G zde vystupuje jako zajímavý analytický nástroj s velmi jednoduchou datovou strukturou.

Optimalizovaný výpočet CE ve hře (Hrubý, 2008)

Aplikovatelnost CE-Solveru:

- Redukce prostoru profilů hry na menší hru – vede na menší LP-úlohu.
- V případě zatím nevyhodnocených $U(s)$ redukujeme počet profilů, které musí mít vyhodnocený užitek.

Mějme hru $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$ a počítačový model (proceduru) $cellModel : S \rightarrow \mathbb{R}^N$ pro výpočet užitků $U(s) := cellModel(s)$.

- Předpokládejme, že výpočet $cellModel(s)$ je hlavní časovou zátěží na výpočet.
- Chceme omezit počet invokací *cellModel* na minimum, tzn. hledáme minimální $S_r \subseteq S$ potřebnou pro vyhodnocení CE.
- Sestavujeme redukovanou hru, tzn. redukovanou $S_r \subseteq S$.

Redukce G-matice v našem příkladě

	ac	ad	bc	bd
$a \longrightarrow b$	5	1		
$b \longrightarrow a$			-5	-1
$c \longrightarrow d$	-2		1	
$d \longrightarrow c$		2		-1

$$-5\pi(bc) - \pi(bd) \geq 0$$

	ab	ad
$a \longrightarrow b$	5	1
$c \longrightarrow d$	-2	
$d \longrightarrow c$		2

$\Rightarrow$

	ad
$a \longrightarrow b$	1
$d \longrightarrow c$	2

CE-Solver: Parallel algorithm solving CE

1. For all players $i \in Q$:
2. For all $j \in S_i$:
3. For all $k \in S_i, j \neq k$:
4. Compute G_{ijk} in parallel
5. If negative, disable related profiles and propagate backwards in G.

Result:

- Remaining profiles determine $S_i, \forall i \in Q$
- G is already computed
- LP-task can be generated

Data types and structures

- *valid* : $S \rightarrow Boolean$ is a boolean array displaying what profiles are valid or non-valid (zero/non-zero probability). Initially, all profiles are valid.
- *umap* : $S \rightarrow \mathbb{R}^N$ is a dynamic dictionary (items can be added and removed during the computation) assigning payoff vectors to particular profiles. Initially, *umap* is empty.
- *dominated* : $Q \times S_i \rightarrow Boolean$ is a boolean array displaying what strategies are already found to be dominated. Initially, all are set *False*.
- type *profList* is a list of tuples $(S, \mathbb{R})$.
- type $gRow = (pos : profList, neg : profList)$
- $G[i, j, k]$ is a dynamic list of *gRow* indexed by (i, j, k). G represents the G-matrix.

Gsolve – main part

```
for i in Q:
for j in S[i]:
for k in S[i]:
 if (j!=k):
  if (not(dominated[i][j]
  or dominated[i][k])):
     row = solveRow(i,j,k);
     if (row != ([],[])):
        if (row.pos != []):
          G[i,j,k] := row;
        else:
          dominated[i][k] := true
          for (s,r) in row.neg:
            disableProf(i,s);
```

Smysl CE (versus NE)

- CE na věc nahlíží jinak než (M)NE.
- NE je stavební kámen celé Teorie her (ve všech jejích podoborech).
- CE je obecnější a snadno spočitatelné ekvilibrium. Přesto není jaksi bráno v komunitě herních teoretiků vážně.
- Pochopení pozadí definice CE není snadné.

Z algoritmického simulačního hlediska:

- Pokud je více NE (MNE), pak nikdo neví, jaké hráči budou hrát. Obvykle je pravděpodobné, že to Pareto optimální.
- U CE, vyhodnocením soustavy lineárních rovnic získáme podprostor smíšených profilů, kde každé z nich je splňuje CE. Najdu-li na tomto podprostoru smíšený profil, který je pro hráče Pareto optimálním výsledkem hry, pak tento profil oznámím hráčům. Hráči pak souhlasí, že budou-li všichni tento profil hrát, nemá smysl z něj vybočit.
- CE na rozdíl od MNE dává odpovědi.

CE versus NE

Bylo řečeno, že CE je obecnější než NE. Co to znamená?

- CE je vymezeno řešením lineárních nerovnic. Ty tvoří mnohoúhelník (polytope). Všechny body mnohoúhelníku jsou CE. Můžeme pak najít optimum na mnohoúhelníku.
- Podstata MNE: smíšená strategie σ_i^* je best-response na σ_{-i}^*. MNE je diskrétní množina. Nalezení smíšených profilů tvoří algoritmickou náročnost MNE.

Theorem

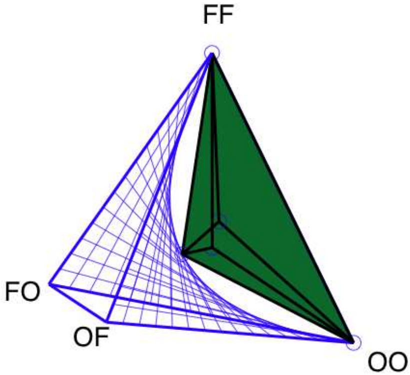
V nedegenerované dvouhráčové hře, MNE jsou vrcholy CE mnohoúhelníku.

Neznamená to ovšem, že všechny vrcholy CE mnohoúhelníku jsou MNE (minimálně však jeden).

CE versus NE

R. Nau, S.G. Canovas and P. Hansen: On the geometry of Nash equilibria and correlated equilibria, IJGT

	F	O
F	3,4	0,0
O	0,0	4,3



Výpočet herních ekvilibrií, optimalizace heuristikami

Mějme hru $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$ a počítačový model (proceduru) *cellModel* : $S \to \mathbb{R}^N$ pro výpočet užitků $U(s) := cellModel(s)$.

- $cellModel(s)$ modeluje důsledky akcí hráčů $(s_i)_{i \in Q}$.
- Nechceme hru zkoumat analyticky, ale simulačně (výpočet $U(s) = \{cellModel(s) | s \in S\}$, výpočet ekvilibria).
- Nechceme zkoumat podobu *cellModel*.
- Hledáme mechanismus (algoritmus), který při minimálním počtu kroků a s použitím minimální paměti najde ekvilibrium pouze při znalosti $\Gamma = (Q; \{S_i\}_{i \in Q}; \{U_i\}_{i \in Q})$ a dostupnosti $cellModel(s)$.
- Zajímavé zadání algoritmického problému.
- FDDS – Fast Detection of Dominant Strategies.

Příště

Máme uzavřenu část definice her, herních konceptů a ekvilibrií.

Následují aplikace a druhá strana THE:

- Mechanism design – teorie veřejné volby a aukce.
- Evoluční teorie her.
- Rozbor rozsáhlého modelu z elektroenergetiky.
- Závěrečné opakování (???).