ORF 569: Special Topics in Statistics and Operations Research - Topics in Game Theory

René Carmona

Contents

1 Game on graphs

1.1 Basic notations

An **undirected graph** G consists of a non-empty finite set V of elements called **vertices** (or **nodes**), and a finite set $E \subset V \times V$ of unordered pairs of (not necessarily distinct) vertices called **edges**. An undirected graph with a vertex set and an edge set E is written like $G = (V, E)$. For example, a graph for n agents, labeled by $1, 2, \ldots, N$, can be represented by a vertex set $V = \{1, 2, \ldots, N\}$ and an edge set E of pairs (i, j) for $i, j \in V$ such that agent i is said to be "connected" to agent j if $(i, j) \in E$. The term "set" in the edge set E emphasizes that there is at most one edge joining any two vertices in V.

We say that two vertices i and j of a graph G are **adjacent** if there is an edge joining them, namely $(i, j) \in E$, and the vertices i and j are then **incident** with such an edge. The **degree** of a vertex i of G is the number of edges incident with i.

A useful representation of a graph $G = (V, E)$ is with its **adjacency matrix** $A^{(G)} = [a_{ij}^{(G)}]_{i,j \in V}$. If a graph G has n vertices, its adjacency matrix $A^{(G)}$ is defined as an $n \times n$ matrix whose $ij-th$ entry represents the strength of connection between vertex i and vertex j. For example, we can define

$$a_{ij}^{(G)} = \begin{cases} 1 & \text{if } (i, j) \in E \\ 0 & \text{o.w.} \end{cases} \tag{1.1.1}$$

The **spectrum** of a matrix A, denoted by $\Sigma(A)$, is the collection of all eigenvalues of A. By definition of an undirected graph, $A^{(G)}$ is symmetric, so that its eigenvalues are all real numbers. Thus, the spectrum of the adjacency matrix $A^{(G)}$ associated to graph G is finite and is a subset of $\mathbb{R}$. We denote by $\lambda_{max}(A^{(G)}) = \max\{\Sigma(A^{(G)})\}$ and $\lambda_{min}(A^{(G)}) = \min\{\Sigma(A^{(G)})\}$ the maximum and minimum eigen values of $A^{(G)}$.

If we assume that $a_{ij}^{(G)} \neq 0$ if and only if $(i, j) \in E$, then a simple computation shows that

$$\left[\left(A^{(G)}\right)^2\right]_{i,j} = \sum_{k=1}^{N} a_{ik}^{(G)} a_{kj}^{(G)},$$

and thus $\left[\left(A^{(G)}\right)^2\right]_{i,j} \neq 0$ if and only if there exists a path joining vertex i to vertex j with exactly **one single hop**. Similarly $\left[\left(A^{(G)}\right)^l\right]_{i,j} \neq 0$ if and only if there exists a path joining vertex i to vertex j with exactly l **hops**.

1.2 Measure of Centrality

One fundamental problem in network analysis is to identify important nodes in complex networks. Different measures of centrality have been proposed to achieve this objective, for example, PageRank centrality measure [Page et al., 1999] that ranks the importance of web pages for online search.

- Katz centrality [Katz, 1953]:

  Definition 1. Let G be an undirected graph $G = (V, E)$ with n nodes and with an adjacency matrix $A^{(G)}$. Let a be an arbitrary positive parameter strictly smaller than $1/\lambda_{max}(A^{(G)})$. Then the **Katz centrality** of G is defined as

  $$k_a(i) = \sum_{k=1}^{\infty} \sum_{j=1}^{N} \left(a^k \left[\left(A^{(G)}\right)^k \right]_{ij} \right), \quad \text{for all } i \in V, \tag{1.2.1}$$

  where $\mathbb{1}$ represent a vector with elements all equal to 1.

  Since $0 < a < 1/\lambda_{max}(A^{(G)})$, $I - aA^{(G)}$ is then invertible. Here I stands for the identity matrix in $\mathbb{R}^{n\times n}$. Thus, we drive a matrix notation for Katz centrality as follow:

  $$k_a = \left(I - aA^{(G)}\right)^{-1} \mathbb{1} - \mathbb{1} = a\left(I - aA^{(G)}\right) A^{(G)}\mathbb{1}.$$

- Bonacich centrality and Degree centrality [Bonacich, 1987]:

  Definition 2. Let a, b be arbitrary parameters such that $b < 1/\lambda_{max}(A^{(G)})$. The **Bonacich centrality** is defined as

  $$\mathbb{b}_{a,b}(i) = \sum_{j=1}^{N} (a + b\,\mathbb{b}_{a,b}(j)) \cdot a_{ij}^{(G)}. \tag{1.2.2}$$

  In matrix notation

  $$\mathbb{b}_{a,b} = a\left(I - bA^{(G)}\right)^{-1} A^{(G)}\mathbb{1} \tag{1.2.3}$$

  The parameter b in Bonacich centrality allows us to scale up or down and change the direction (positive or negative) of the dependence of one vertex's score on the scores of other vertices.

  If $a = 1, b = 0$, we then have $\mathbb{b}_{1,0} = A^{(G)}\mathbb{1}$, so that $\mathbb{b}_{1,0}(i)$ equals to the degree of vertex i if $A^{(G)}$ takes the form (1.1.1). We call $\mathbb{b}_{1,0}$ the **Degree centrality**.

- Eigenvector Centrality [Bonacich, 1972a] [Bonacich, 1972b]

  This idea behind the Eigenvector centrality is to find a measure on V so that a node is important if it is connected to other important nodes.

  Definition 3. We define the **Eigenvector centrality** as the vector v satisfying

  $$v(i) = \frac{1}{\lambda} \sum_{j=1}^{N} a_{ij}^{(G)} v(j) \tag{1.2.4}$$

  for some real non-zero value λ and for any $i \in V$. In matrix notation:

  $$\lambda v = A^{(G)} v. \tag{1.2.5}$$

We intend to find a desirable solution v to equation (1.2.5), especially we want that v is positive (i.e. $v(i) > 0$ for any $i \in V$). If the adjacency matrix $A^{(G)}$ is positive, namely if $a_{ij}^{(G)} > 0$ for all $i, j \in V$, the Perron-Frobenius theorem tells us that the existence and uniqueness (up to a scalable multiple factor) of a positive eigenvector v and the corresponding eigenvalue is positive and largest in absolute norm, i.e. $\lambda = \max\{|\lambda(A^{(G)})| : \lambda(A^{(G)}) \text{ eigenvalues } \} = \lambda_{max}(A^{(G)}) > 0$.

1.3 Network games : static situation

Let us consider a graph $G = (V, E)$ with a vertex set $V = \{1, \ldots, N\}$ representing N different agents/players, and a edge set $E \subset V \times V$. Its adjacency matrix is denoted by $A^{(G)}$. As an example, we consider here a static network game with N players whose connections are described by graph $G = (V, E)$ and adjacent matrix $A^{(G)}$. In this static network game, **all players act simultaneously and they play only one round.**

For each player i, the set of actions/controls/strategies that she is allowed to choose is denoted by $A^{(i)}$, which is a subset of $\mathbb{R}^{k_i}$ for some $k_i > 0$. We call $A^{(i)}$ the set of admissible controls/strategies associated to player i. We denote by $A = A^{(1)} \times A^{(2)} \times \cdots \times A^{(N)}$ the set of admissible controls for all N players. For any $\boldsymbol{\alpha} \in A$, we denote by $\boldsymbol{\alpha} = (\alpha^1, \alpha^2, \ldots, \alpha^N)$ with $\alpha^i \in A^{(i)}$ and call it a **strategy profile** for all players in V.

We also associate to each player $i \in V$ a cost function

$$J^i : A \ni \boldsymbol{\alpha} \mapsto J^i(\boldsymbol{\alpha}) = J^i(\alpha^i, \alpha^{-i}) \in \mathbb{R}$$

where $\alpha^{-i} = (\alpha^1, \ldots, \alpha^{i-1}, \alpha^{i+1}, \ldots, \alpha^N)$.

Definition 4. For each player $i \in V$, the **best responses of player** i to other players using strategy $\alpha^{-i} = (\alpha^1, \ldots, \alpha^{i-1}, \alpha^{i+1}, \ldots, \alpha^N)$ is a subset in $A^{(i)}$, denoted by $br^i(\alpha^{-i})$ such that any $\alpha^{i,*} \in br^i(\alpha^{-i})$ minimizes the associate cost function $J^i(\cdot, \alpha^{-i})$ of player i, namely

$$\alpha^{i,*} \in \underset{\alpha^i \in A^{(i)}}{\arg\min}\, J^i(\alpha^i, \alpha^{-i}). \tag{1.3.1}$$

We define the notion of Nash equilibrium as follow:

Definition 5. Let $\boldsymbol{\alpha}^* \in A$ be a strategy profile for N players. The best response function br for all players in V is defined as

$$br : A \ni \boldsymbol{\alpha} = (\alpha^1, \ldots, \alpha^N) \mapsto br(\boldsymbol{\alpha}) = br^1(\alpha^{-1}) \times br^2(\alpha^{-2}) \times \cdots \times br^N(\alpha^{-N}) \subset A. \tag{1.3.2}$$

A strategy profile $\boldsymbol{\alpha}^*$ is said to be a fixed point of the map br if

$$\boldsymbol{\alpha}^* \in br(\boldsymbol{\alpha}^*). \tag{1.3.3}$$

We say that $\boldsymbol{\alpha}^*$ is a **Nash equilibrium** if it is a fixed point for the best response function br. Consequently, for each $i \in V$ we can have

$$J^i(\boldsymbol{\alpha}^*) \leq J^i(\alpha^i, \alpha^{-i,*}), \qquad \text{for all } \alpha^i \in A^{(i)}. \tag{1.3.4}$$

The collection of all Nash equilibria is denoted by a set $\mathcal{N} \subset A$:

$$\mathcal{N} = \{\boldsymbol{\alpha} \in A : \boldsymbol{\alpha} \in br(\boldsymbol{\alpha})\}. \tag{1.3.5}$$

1.4 Social cost and Price of Anarchy

Definition 6. We define the **social cost** of a strategy profile $\boldsymbol{\alpha} \in A$ the sum of cost $(J^i)_{i\in V}$ evaluated at point $\boldsymbol{\alpha}$, i.e.

$$J(\boldsymbol{\alpha}) = \sum_{i=1}^{N} J^i(\boldsymbol{\alpha}) = \sum_{i=1}^{N} J^i(\alpha^i, \alpha^{-i}). \tag{1.4.1}$$

A strategy profile α that minimizes the social cost among all admissible strategies A of N players is called the **social optimal**, and it is denote by

$$\boldsymbol{\alpha}^{*,SC} = \arg\min_{\boldsymbol{\alpha}\in A} J(\boldsymbol{\alpha}) \tag{1.4.2}$$

We will give a definition of the Price of Anarchy that measures the inefficiency of selfish behavior among players in the above static network game. In another words, the extra cost (in ratio) raised in the situation that every player optimizes her own strategy (through the corresponding cost function) without considering the effect of such strategy on her peers.

Definition 7. The **Price of Anarchy** (PoA) in a static network game described above is defined as the ratio between the worst social cost of a Nash equilibrium and the optimal social cost among all admissible strategy profiles, namely

$$PoA = \frac{\sup_{\boldsymbol{\alpha}\in\mathcal{N}} J(\boldsymbol{\alpha})}{\inf_{\boldsymbol{\alpha}\in A} J(\boldsymbol{\alpha})} = \frac{\sup\{J(\alpha^*) : \alpha^* \in \mathcal{N}\}}{J(\boldsymbol{\alpha}^{*,SC})} \tag{1.4.3}$$

A similar notion to the Price of Anarchy is the **Price of Stability** (PoS). It is defined the ratio between the best social cost of a Nash equilibrium and the optimal social cost among all admissible strategy profiles:

$$PoS = \frac{\inf_{\boldsymbol{\alpha}\in\mathcal{N}} J(\boldsymbol{\alpha})}{\inf_{\boldsymbol{\alpha}\in A} J(\alpha)} = \frac{\inf\{J(\boldsymbol{\alpha}^*) : \boldsymbol{\alpha}^* \in \mathcal{N}\}}{J(\boldsymbol{\alpha}^{*,SC})}. \tag{1.4.4}$$

For example, in most network applications, agents are interacting under the restriction of some protocol that will lead to a collective solutions of all participants, who can choose either accept it or defect from it. As a result, it will become more interesting for the protocol designer to find the best Nash equilibrium, which can be view as an optimal stable solution subject to the protocol.

2 Search for Nash equilibrium

2.1 overview

1. Fixed Point theorem

    - Kakutani (1941) theorem. Used by John Nash.

      If the best response function br is not set value function

    - Banach fixed point theorem

      If the best response function is a strict contraction, then there is a unique fixed point.

    - Brouder, Brower, Schauder fixed point theorem (used to show existence of a fixed point)

    - special structure of the Game

        - Tarski (unimodular game)
        - potential game

2. Variational Inequalities

    - Graphon Games

2.2 An example for Nash equilibrium

Assume that $A^i = \mathbb{R}$ the real line. For each player $i \in \{1, \ldots, N\}$, We consider a cost function

$$J^i(\alpha^i, \alpha^{-i}) = f_i(\alpha^i - \alpha^{i,0} + \delta \cdot \sum_{j \neq i} a_{ij}^{(G)} \alpha^j) + g_i(\alpha^{-i})$$

where $\alpha^{i,0}$ is an (autistic) action preferred by player i.

Assume that f_i is symmetric and descending on $(-\infty, 0]$, and increasing on $[0, \infty)$.

- $f_i(x) = f_i(-x)$ for any $x \in \mathbb{R}$
- $f_i(0) = \min_x f_i(x)$

For any $i \in \{1, \ldots, N\}$, if α^{-i} is given, the best response obtained for player i is

$$br^i(\alpha^{-i}) = \alpha^{i,0} - \delta \cdot \sum_{j \neq i} a_{ij}^{(G)} \cdot \alpha^j \tag{2.2.1}$$

A fixed point for the best response function will satisfy

$$\alpha^i = \alpha^{i,0} - \delta \cdot \sum_{j \neq i} a_{ij}^{(G)} \cdot \alpha^j$$

Assume that $a_{ii}^{(G)} = 0$, we can have

$$\boldsymbol{\alpha} = (\mathbf{I} + \delta A^G)^{-1} \cdot \boldsymbol{\alpha}^0$$

If the game is symmetric: $\alpha^{i,0} = \alpha^0 \in \mathbb{R}$ for all $i \in \{1, \dots, N\}$, then

$$\begin{aligned} \boldsymbol{\alpha} &= \alpha^0 \cdot (\mathbf{I} + \delta A^{(G)})^{-1} \mathbb{1} \\ &= \alpha^0 \left(\mathbf{I} - [\mathbf{I}_+ \delta A^{(G)}]^{-1} [\mathbf{I} + \delta A^{(G)}] + [\mathbf{I} + \delta A^{(G)}]^{-1} \right) \mathbb{1} \\ &= \alpha^0 (\mathbf{I} - \delta [\mathbf{I} + \delta A^{(G)}]^{-1} A^{(G)}) \mathbb{1} \\ &= \alpha^0 (\mathbb{1} - \mathbb{b}_{\delta, -\delta}) \end{aligned} \tag{2.2.2}$$

where $\mathbb{b}_{a,b}$ is the Bonacich centrality measure.

Remark 1. In this example, we can see that the Nash equilibrium only depends on the graph structure (through the Bonacich centrality measure), bu not the functions f_i nor g_i.

Remark 2. We can have a more general setting for the action space A^i, for example:

- if $A^i = [0, \infty)$, the Nash equilibrium becomes

$$\alpha^{*,i} = \max \left\{ 0, \alpha^{i,0} - \delta \sum_{j \neq i} a_{ij}^{(G)} \alpha^{*,j} \right\};$$

- if $A^i = [0, L]$ for some $L > 0$, the Nash equilibrium becomes

$$\alpha^{*,i} = \min \left\{ L, \max \left\{ 0, \alpha^{i,0} - \delta \sum_{j \neq i} a_{ij}^{(G)} \alpha^{*,j} \right\} \right\}.$$

Remark 3. In equation (2.2.1),

- if $\delta < 0$, br^i is increasing in actions of neighbors. We call the game $(A^1 \times \dots \times A^N, \{J^i\}_{i=1,\dots,N})$ is a game with **strategic complements**, as the actions chosen by among players mutually reinforce one another.
- if $\delta > 0$, br^i is decreasing in actions of neighbors. We call the game $(A^1 \times \dots \times A^N, \{J^i\}_{i=1,\dots,N})$ is a game with **strategic substitutes**, as the actions chosen among players mutually offset one another.

3 Potential Games

For the sake of simplicity, we denote by $A = A^1 \times \dots A^N$ the set of strategy profiles, and for any $i \in \{1, \dots, N\}$, the couple $(\alpha, \boldsymbol{\alpha}^{-i}) \in A$ with $\alpha \in A^i$ and $\boldsymbol{\alpha}^{-i} \in A^1 \times \dots A^{i-1} \times A^{i+1} \times \dots \times A^N$ stands for a strategy profile $(\alpha^1, \dots, \alpha^{i-1}, \alpha, \alpha^{i+1}, \dots, \alpha^N)$. Also, a function $\varphi : A \ni \boldsymbol{\alpha} \mapsto \varphi(\boldsymbol{\alpha}) \in \mathbb{R}$ evaluated at a strategy profile $\boldsymbol{\alpha} = (\alpha^i, \boldsymbol{\alpha}^{-i})$ is denoted by $\varphi(\alpha^i, \boldsymbol{\alpha}^{-i})$.

Definition 8 (Monderer and Shapley 1996)**.** A Game $(A^1 \times \dots \times A^N, \{J^i\}_{i=1\dots N})$ is said to be a **Potential Game** if there exists a map $\varphi : A \ni \boldsymbol{\alpha} \mapsto \varphi(\boldsymbol{\alpha}) \in \mathbb{R}$, called potential function, such that for every player $i \in \{1, \dots, N\}$, and for any given $\boldsymbol{\alpha}^{-i} \in A^1 \times \dots A^{i-1} \times A^{i+1} \times \dots \times A^N$, we have

$$J^i(\beta, \boldsymbol{\alpha}^{-i}) - J^i(\alpha, \boldsymbol{\alpha}^{-i}) = \varphi(\beta, \boldsymbol{\alpha}^{-i}) - \varphi(\alpha, \boldsymbol{\alpha}^{-i}), \qquad \forall\, \alpha, \beta \in A^i \tag{3.0.1}$$

In words, the change in a single player's cost function due to her own strategy deviation results in exactly the same amount of change of the potential function φ.

Assume that $(A^i)_{i=1,\dots,N}$ are intervals in $\mathbb{R}$ and each cost function J^i is continuous and differentiable. If $(A, \{J^i\}_{i=1,\dots,N})$ is a potential game, then for every $i \in \{1, \dots, N\}$ we have

$$\frac{\partial}{\partial \alpha^i} J^i(\alpha, \boldsymbol{\alpha}^{-i}) = \frac{\partial}{\partial \alpha^i} \varphi(\alpha, \boldsymbol{\alpha}^{-i}), \quad \forall\, \alpha \in A^i, \boldsymbol{\alpha}^{-i} \in A^{-i}.$$

Theorem 1. Let $\Gamma = (A^1 \times \dots \times A^N, {J^i}_{i=1,\dots N})$ be a game in which the strategy sets $(A^i)_{i=1,\dots N}$ are intervals of real numbers. Suppose the cost functions are twice continuously differentiable. Then Γ is a potential game iff

$$\frac{\partial^2 J^i}{\partial \alpha^i \partial \alpha^j} = \frac{\partial^2 J^j}{\partial \alpha^i \partial \alpha^j}, \quad \text{for every } i, j \in \{1, \dots, N\}. \tag{3.0.2}$$

Moreover, if the cost functionals satisfy (3.0.2) and $\boldsymbol{\alpha}^0$ is an arbitrary strategy profile in A, then a potential for Γ is given by

$$\varphi(\boldsymbol{\alpha}) = \sum_{i=1}^{N} \int_0^1 \frac{\partial J^i}{\partial \alpha^i}(\beta(t)) \cdot (\beta^i)'(t) dt \tag{3.0.3}$$

where $\beta : [0, 1] \to A$ is a piecewise continuously differentiable path in A that connects $\boldsymbol{\alpha}^0$ to $\boldsymbol{\alpha}$ (i.e. $\beta(0) = \boldsymbol{\alpha}^0$ and $\beta(1) = \boldsymbol{\alpha}$).

For any $i \in \{1, \dots, N\}$ and any $\boldsymbol{\alpha}^{-i} \in A^1 \times \dots A^{i-1} \times A^{i+1} \times \dots \times A^N$, if a game $(A^1 \times \dots \times A^N, \{J^i\}_{i=1,\dots,N})$ is a potential game, then the best response function br^i for player i satisfies that for every $\alpha \in A^i$ and $\alpha^{*,i} \in br^i(\boldsymbol{\alpha}^{-i})$:

$$\varphi(\alpha, \boldsymbol{\alpha}^{-i}) - \varphi(\alpha^{*,i}, \boldsymbol{\alpha}^{-i}) = J^i(\alpha, \boldsymbol{\alpha}^{-i}) - J^i(\alpha^{*,i}, \boldsymbol{\alpha}^{-i}) \geq 0. \tag{3.0.4}$$

Thus,

$$br^i(\boldsymbol{\alpha}^{-i}) = \arg\inf_{\alpha} \varphi(\alpha, \boldsymbol{\alpha}^{-i}).$$

So if $\boldsymbol{\alpha}^* \in A$ is a minimum of the mapping φ, namely for any $\boldsymbol{\alpha} \in A$,

$$\varphi(\boldsymbol{\alpha}^*) \leq \varphi(\boldsymbol{\alpha}),$$

then we must have for every $i \in \{1, \dots, N\}$ and every $\alpha \in A^i$,

$$\varphi(\alpha^{*,i}, \boldsymbol{\alpha}^{*,-i}) \leq \varphi(\alpha, \boldsymbol{\alpha}^{*,-i}),$$

so that $\alpha^{*,i} \in br^i(\boldsymbol{\alpha}^{*,-i})$ for every $i \in \{1, \dots, N\}$. Thus, $\boldsymbol{\alpha}^*$ is a Nash equilibrium. Does the inverse hold even if we don't have a unique minimum $br^i(\boldsymbol{\alpha}^{-i})$ for every $\boldsymbol{\alpha}^{-i}$?

Moreover, if the Nash equilibrium is unique, then

$$\boldsymbol{\alpha}^* \in A \text{ is a Nash equilibrium} \quad \Longleftrightarrow \quad \boldsymbol{\alpha}^* = \arg\inf_{\boldsymbol{\alpha}} \varphi(\boldsymbol{\alpha}) \tag{3.0.5}$$

where the uniqueness of minimum is deduced from the uniqueness of Nash equilibrium.

Remark 4.

- If A^i is finite for every $i \in \{1, \dots, N\}$, so as the number of admissible strategies $\boldsymbol{\alpha} \in A$. Then there is at least one Nash equilibrium when the game is potential.
- If A is compacts, and if the potential function φ is continuous, then there exists a Nash equilibrium. Moreover, if φ is strictly convex, then the Nash equilibrium is unique.

Definition 9. A sequence $(\boldsymbol{\alpha}_n)_{n \geq 0}$ of strategy profiles is called **path** if for every $n \geq 0$, $\boldsymbol{\alpha}_{n+1}$ is obtained from $\boldsymbol{\alpha}_n$ by allowing one player, say $i_n \in \{1, \dots, N\}$ to change her strategy.

For example, if $\alpha_n^{i_n}$ changes into $\beta \in A^{i_n}$, then the strategy profile $\boldsymbol{\alpha}_n$ changes to a new strategy profile $\boldsymbol{\alpha}_{n+1}$:

$$\boldsymbol{\alpha}_n = (\alpha_n^1, \dots, \alpha_n^N) \longrightarrow \boldsymbol{\alpha}_{n+1} := (\alpha_n^1, \dots, \alpha_n^{i_n - 1}, \beta, \alpha_n^{i_n+1}, \dots, \alpha_n^N).$$

A path is called an **improvement path** if

$$J^{i_n}(\boldsymbol{\alpha}_{n+1}) < J^{i_n}(\boldsymbol{\alpha}_n).$$

Proposition 1. The end-point of any finite improvement path is a Nash equilibrium.

Proof. By contradiction. □

Proposition 2. The potential function is determined up to a constant.

If φ_1 and φ_2 are potential functions for the same potential game $(A^1 \times \ldots \times A^N, \{J^i\}_{i=1,\dots,N})$, then there exists a constant $c \in \mathbb{R}$ such that

$$\varphi_1 = \varphi_2 + c.$$

Proof. Pick an arbitrary $\boldsymbol{\alpha}^0 = (\alpha^{0,1}, \dots, \alpha^{0,N}) \in A$, then define a function $H : A \ni \boldsymbol{\alpha} = (\alpha^1, \dots, \alpha^N) \mapsto H(\boldsymbol{\alpha}) \in \mathbb{R}$ by

$$H(\boldsymbol{\alpha}) = \sum_{i=1}^{N} J^i(\boldsymbol{\alpha}_i) - J^i(\boldsymbol{\alpha}_{i+1}) \tag{3.0.6}$$

where $\boldsymbol{\alpha}_1 = \boldsymbol{\alpha}$, and $\boldsymbol{\alpha}_{i+1} = (\alpha^{0,1}, \dots, \alpha^{0,i}, \alpha^{i+1}, \dots, \alpha^N)$ for $i = 1, \dots N-1$, and $\boldsymbol{\alpha}_{N+1} = \boldsymbol{\alpha}^0$. If φ_1 and φ_2 are potential functions, we can have

$$H(\boldsymbol{\alpha}) = \varphi_1(\boldsymbol{\alpha}) - \varphi_1(\boldsymbol{\alpha}^0), \quad \text{and} \quad H(\boldsymbol{\alpha}) = \varphi_2(\boldsymbol{\alpha}) - \varphi_2(\boldsymbol{\alpha}^0).$$

Let $c = \varphi_1(\boldsymbol{\alpha}^0) - \varphi_2(\boldsymbol{\alpha}^0)$, the conclusion then holds. □

3.1 Example: Cournot Competition

Let us consider N firms enumerated by numbers $\{1, \dots, N\}$. They all produce a same kind of product to sell on the market. The quantity produced by firm i is denoted by a positive real number $q^i \in [0, \infty)$. The productions profile of these firms is then denoted by $\mathbf{q} = (q^1, \dots, q^N) \in [0, \infty)^N$. The cost for producing q units of products (q can be real numbers) is identical among firms and is depicted by a function $c : [0, \infty) \ni q \mapsto c(q) \in \mathbb{R}$. We define the total production Q of these N firms as

$$Q = \sum_{i=1}^{N} q^i.$$

The price p of the product on the market depends on the amount of total production through a function $f : [0, \infty) \to \mathbb{R}$, namely $p(\mathbf{q}) = f(Q)$. We choose

$$f(Q) = a - bQ,$$

for some constants $a, b > 0$. For each firm $i \in \{1, \dots, N\}$, we associate her with a cost function J^i representing the negative of net revenue generated from her production and her selling activities. More precisely, the cost function $J^i : [0, \infty)^N \to \mathbb{R}$ is given by

$$J^i(\mathbf{q}) = -\left[p(\mathbf{q}) \cdot q^i - c(q^i)\right] = c(q^i) - aq^i + bq^i \sum_{j=1}^{N} q^j.$$

Each firm wants to minimize her associate cost function J^i among the production profiles $\mathbf{q} \in [0, \infty)^N$.

To show that the game between N firms under the Cournot competition framework is a potential game, we compute the partial derivative $\partial^2 J^i / \partial q^i \partial q^j$ for every $i, j \in \{1, \dots, N\}$. It can be shown that

$$\frac{\partial^2 J^i(\mathbf{q})}{\partial q^i \partial q^j} = b = \frac{\partial^2 J^j(\mathbf{q})}{\partial q^i \partial q^j}, \qquad \forall\, i, j \in \{1, \dots N\}, \ \forall\, \mathbf{q} \in [0, \infty)^N.$$

Thus, the game derive from Cournot competition is a potential game.

Moreover, one way to construct a potential function is by considering the mapping

$$\varphi(\mathbf{q}) = \sum_{i=1}^{N} \int_0^1 \frac{\partial^2 J^i}{\partial q^i}(\beta^i(t), \boldsymbol{\beta}^{-i}(t)) \cdot (\beta^i(t))' dt$$

where $\boldsymbol{\beta} : [0,1] \ni t \mapsto \boldsymbol{\beta}(t) = t\mathbf{q} \in [0,\infty)^N$.

A straight computation gives

$$(\beta(t)^i)' = q^i, \qquad \frac{J^i(\boldsymbol{\beta}(t))}{\partial q^i} = \frac{\partial c(tq^i)}{\partial q^i} - a + 2btq^i + bt\sum_{j\neq i}^{N} q^j$$

so that

$$\varphi(\mathbf{q}) = \sum_{i=1}^{N} \left(c(q^i) - aq^i + b \cdot (q^i)^2 + \frac{b}{2} q^i \sum_{j\neq i}^{N} q^j \right)$$

3.2 Congestion Game

There are several components in a discrete Congestion Game:

- N players enumerated by $\{1, \dots, N\}$.
- A set of resources denoted by E.
- A feasible strategy sets A^i of every player $i \in \{1, \dots, N\}$ such that $A^i \subseteq 2^E$, a collection of sets of elements in E. Namely, for player i, her strategy $\alpha^i \in A^i$ is a set of resources $\alpha^i \subseteq E$. Let $A = A^1 \times \dots A^N$ denotes the collection of all feasible strategy profiles for N players.
- We associate to each resource $e \in E$ a load function $k_e : A \to \{0, \dots, N\}$ defined as $k_e(\boldsymbol{\alpha}) = |\{i : e \in \alpha^i\}|$ for every strategy profile $\boldsymbol{\alpha} \in A$.
- We associate to each resource $e \in E$ a cost function $c_e : \{0, \dots, N\} \to \mathbb{R}$.
- For each strategy profile $\boldsymbol{\alpha} \in A$, player i experiences a cost $J^i(\boldsymbol{\alpha})$ defined by

$$J^i(\boldsymbol{\alpha}) = \sum_{e\in\alpha^i} c_e(k_e(\boldsymbol{\alpha})). \tag{3.2.1}$$

We can denote a congestion game of N players by $(E, A^1 \times \dots \times A^N, \{c_e\}_{e\in E}, \{k_e\}_{e\in E}, \{J^i\}_{i=1,\dots,N})$.

Example A Network Congestion Game is define with a graph $G = (V, E)$ such that for each player $i \in \{1, \dots, N\}$, there is a pair of vertices $(S_i, T_i) \in V \times V$ such that the feasible strategy set for player i is defined by

$$A^i = \{\alpha \subset E : \alpha \text{ is a path from } S_i \text{ to } T_i\}.$$

Proposition 3. (Rosental 73) Every congestion game has a pure Nash equilibrium

Idea: consider the Rosental potential function

$$\varphi(\boldsymbol{\alpha}) = \sum_{e\in E}\sum_{k=1}^{k_e(\boldsymbol{\alpha})} c_e(k) \tag{3.2.2}$$

Proof. We consider a potential function (3.2.2). For every $\boldsymbol{\alpha} \in A$ and $i \in \{1, \dots, N\}$, consider $\tilde{\alpha} \in A^i$,

$$\begin{aligned}
\varphi(\tilde{\alpha}, \boldsymbol{\alpha}^{-i}) &= \sum_{e\in E}\sum_{k=1}^{k_e(\tilde{\alpha},\boldsymbol{\alpha}^{-i})} c_e(k) \\
&= \sum_{e\in E}\left(\sum_{k=1}^{k_e(\boldsymbol{\alpha})} c_e(k) + \mathbb{1}_{\{e\in\tilde{\alpha}\setminus\alpha^i\}} c_e(k_e(\tilde{\alpha}, \boldsymbol{\alpha}^{-i})) - \mathbb{1}_{\{e\in\alpha^i\setminus\tilde{\alpha}^i\}} c_e(k_e(\boldsymbol{\alpha}))\right) \\
&= \sum_{e\in E}\sum_{k=1}^{k_e(\boldsymbol{\alpha})} c_e(k) + \sum_{e\in\tilde{\alpha}\setminus\alpha^i} c_e(k_e(\tilde{\alpha}, \boldsymbol{\alpha}^{-i})) - \sum_{e\in\alpha^i\setminus\tilde{\alpha}^i} c_e(k_e(\boldsymbol{\alpha})) \\
&= \sum_{e\in E}\sum_{k=1}^{k_e(\boldsymbol{\alpha})} c_e(k) + \sum_{e\in\tilde{\alpha}} c_e(k_e(\tilde{\alpha}, \boldsymbol{\alpha}^{-i})) - \sum_{e\in\alpha^i} c_e(k_e(\boldsymbol{\alpha})) \\
&= \varphi(\boldsymbol{\alpha}) + J^i(\tilde{\alpha}, \boldsymbol{\alpha}^{-i}) - J^i(\boldsymbol{\alpha}) \qquad (3.2.3)
\end{aligned}$$

Hence the conclusion. ☐

3.3 Example: Braess Paradox

Consider a congestion game on a Network Graph $G = (V = \{s, t, a, b\}, E = \{e_{sa}, e_{at}, e_{sb}, e_{bt}\})$. Assume that there are $N = 100$ cars, each of them has an identical feasible strategy set $A_0 = \{\{e_{sa}, e_{at}\}, \{e_{sb}, e_{bt}\}\}$. For edges $e \in E$, we define the cost functions:

$$c_{e_{sa}}(k) = c_{e_{bt}}(k) = k, \quad c_{e_{at}}(k) = c_{e_{sb}}(k) = c, \qquad \forall\, k \in \{0, \dots, 100\}$$

In words, the cost incurred for routing on edges $\{e_{sa}, e_{bt}\}$ is proportionate to the number of cars on this edge, whilst the cost incurred for routing on edges $\{e_{sb}, e_{at}\}$ is constant. For every player $i \in \{1, \dots, N\}$, her cost function is given by

$$J^i(\boldsymbol{\alpha}) = \sum_{e\in\alpha^i} c_e(k_e(\boldsymbol{\alpha})) = \begin{cases} c_{e_{sa}}(k_{e_{sa}}(\boldsymbol{\alpha})) + c_{e_{at}}(k_{e_{at}}(\boldsymbol{\alpha})) & if\; \alpha^i = \{e_{sa}, e_{at}\} \\ c_{e_{sb}}(k_{e_{sb}}(\boldsymbol{\alpha})) + c_{e_{bt}}(k_{e_{bt}}(\boldsymbol{\alpha})) & if\; \alpha^i = \{e_{sb}, e_{bt}\} \end{cases} \tag{3.3.1}$$

Thus,

$$\begin{aligned}
J^i(\boldsymbol{\alpha}) &= c + \text{ the number of cars going through the same route as car } i \\
&= c + \sum_{j=1}^{N} \mathbb{1}_{\alpha^j=\alpha^i} \qquad (3.3.2)
\end{aligned}$$

Lemma 1. $\boldsymbol{\alpha}^* \in A$ is a Nash equilibrium if and only if 50% of cars routes on path $\{e_{sa}, e_{at}\}$ and the other 50% on path $\{e_{sb}, e_{bt}\}$. More precisely

$$\boldsymbol{\alpha}^* \in A \text{ is a Nash equilibrium} \quad iff \quad \left|\{i : \alpha^{*,i} = \{e_{sa}, e_{at}\}\}\right| = \frac{N}{2}$$

Proof. Let $\boldsymbol{\alpha}^* \in A$ such that $|\{i : \alpha^{*,i} = \{e_{sa}, e_{at}\}\}| = \frac{N}{2}$. Then for any car $i \in \{1, \ldots, N\}$, assume w.l.o.g. that it routes on path $\alpha^{*,i} = \{e_{sa}, e_{at}\}$. The cost it that experiences is

$$J^i(\boldsymbol{\alpha}^*) = c + \sum_{j=1}^{N} \mathbb{1}_{\alpha^{*,j} = \alpha^{*,i}} = c + \frac{N}{2}.$$

If it changes to path $\tilde{\alpha} = \{s_{sb}, e_{bt}\}$, then the new cost associate to it becomes

$$J^i(\tilde{\alpha}, \boldsymbol{\alpha}^{*,-i}) = c + \left(\sum_{j \neq i} \mathbb{1}_{\alpha^{*,j} = \tilde{\alpha}} + 1 \right) = c + \frac{N}{2} + 1.$$

This show that $\boldsymbol{\alpha}^*$ is a Nash equilibrium.

Inversely, for any Nash equilibrium $\boldsymbol{\alpha}^* \in A$, we must have for every $i \in \{1, \ldots, N\}$ and every $\tilde{\alpha} \in A_0$,

$$J^i(\boldsymbol{\alpha}^*) \leq J^i(\tilde{\alpha}, \boldsymbol{\alpha}^{*,-i}),$$

namely

$$\sum_{j=1}^{N} \mathbb{1}_{\alpha^{*,j} = \alpha^{*,i}} \leq \sum_{j=1}^{N} \mathbb{1}_{\alpha^{*,j} = \tilde{\alpha}}$$

Since there are only two strategies in A_0, we denote by $\tilde{\alpha}_-$ the alternative strategy in the set A_0 compared to $\tilde{\alpha}$. We must have for every $i \in \{1, \ldots, N\}$

$$2 \sum_{j=1}^{N} \mathbb{1}_{\alpha^{*,j} = \alpha^{*,i}} \leq \sum_{j=1}^{N} \mathbb{1}_{\alpha^{*,j} = \tilde{\alpha}} + \sum_{j=1}^{N} \mathbb{1}_{\alpha^{*,j} = \tilde{\alpha}_-} = N. \tag{3.3.3}$$

Thus, there must exists another car $k \in \{1, \ldots, N\}$ such that $\alpha^{*,k} \neq \alpha^{*,i}$, so that we also have

$$\sum_{j=1}^{N} \mathbb{1}_{\alpha^{*,j} \neq \alpha^{*,i}} = \sum_{j=1}^{N} \mathbb{1}_{\alpha^{*,j} = \alpha^{*,k}} \leq \frac{N}{2}.$$

On the other hand

$$\sum_{j=1}^{N} \mathbb{1}_{\alpha^{*,j} = \alpha^{*,i}} + \sum_{j=1}^{N} \mathbb{1}_{\alpha^{*,j} \neq \alpha^{*,i}} = N$$

so that the inequality (3.3.3) holds for an arbitrary $i \in \{1, \ldots, N\}$, namely

$$\left|\{j : \alpha^{*,j} = \alpha^{*,i}\}\right| = \frac{N}{2}.$$

Again, because there is only two strategy in A_0, we have

$$\left|\{i : \alpha^{*,i} = \{e_{sa}, e_{at}\}\}\right| = \frac{N}{2}.$$

□

Now we add a new route (edge) between vertex a and b, denoted by e_{ab}, and we assign not cost for cars routing on this new route, i.e. $c_{e_{ab}}(k) = 0$ for every $k \in \mathbb{N}$. We only allows car to route from a to b, so the new feasible strategy set, still denoted by A_0, becomes

$$A_0 = \{\{e_{sa}, e_{at}\}, \{e_{sb}, e_{bt}\}, \{e_{sa}, e_{ab}, e_{bt}\}\}.$$

We assume an additional condition on the constant cost c on edge e_{at} and e_{sb}:

$$c > N = 100 \tag{3.3.4}$$

Lemma 2. $\alpha^* \in A$ is a Nash equilibrium if and only if all cars route on $\{e_{sa}, e_{ab}, e_{bt}\}$.

3.4 Strategic Game

Definition 10. A game $(A^1 \times \ldots A^N, \{J^i\}_{i=1,\ldots,N})$ is said to be a **strategic game** if it belongs to one of the two following situations:

- it is a coordination game :

$$J^i(\boldsymbol{\alpha}) = J^j(\boldsymbol{\alpha}), \qquad \forall i, j \in \{1, \ldots, N\}, \forall \boldsymbol{\alpha} \in A. \tag{3.4.1}$$

- it is a dummy game:

$$J^i(\boldsymbol{\alpha}) = J^i(\tilde{\alpha}, \boldsymbol{\alpha}^{-i}), \qquad \forall i \in \{1, \ldots, N\}, \forall \boldsymbol{\alpha} \in A, \text{ and } \forall \tilde{\alpha} \in A^i \tag{3.4.2}$$

Proposition 4. A game $(A^1 \times \ldots \times A^N, \{J^i\}_{i=1,\ldots,N})$ is a potential game if and only if there exists functions $\{\varphi_i^c\}_{i=1,\ldots,N}$ and $\{\varphi_i^d\}_{i=1,\ldots,N}$ such that for every $i \in \{1, \ldots, N\}$:

$$J^i = \varphi_i^c + \varphi_i^d$$

and

- the game $(A^1 \times \ldots \times A^N, \{\varphi_i^c\}_{i=1,\ldots,N})$ is a coordination game;
- the game $(A^1 \times \ldots \times A^N, \{\varphi_i^d\}_{i=1,\ldots,N})$ is a dummy game.

Proof. ($\Rightarrow$): We can chose a potential function $\varphi = \varphi_i^c$ as one of the cost functions in the coordination game.
($\Leftarrow$): Let φ be a potential function, and we define $\forall i \in \{1, \ldots, N\}$, $\forall \boldsymbol{\alpha} \in A$ the functions

$$\begin{cases} \varphi_i^c(\boldsymbol{\alpha}) = \varphi(\boldsymbol{\alpha}) \\ \varphi_i^d(\boldsymbol{\alpha}) = J^i(\boldsymbol{\alpha}) - \varphi(\boldsymbol{\alpha}). \end{cases}$$

Then it is easy to check that $(A, \{\varphi_i^c\}_{i=1,\ldots,N})$ and $(A, \{\varphi_i^d\}_{i=1,\ldots,N})$ are a coordination game and a dummy game respectively. $\square$

Definition 11. Two games $(A^1 \times \ldots \times A^N, \{J_1^i\}_{i=1,\ldots,N})$ and $(\tilde{A}^1 \times \ldots \times \tilde{A}^N, \{\tilde{J}_2^i\}_{i=1,\ldots,N})$ are said to be isomorphic if there exists N bijection mappings $\{\phi_i\}_{i=1\ldots,N}$:

$$\phi_i : A^i \ni \alpha^i \mapsto \phi_i(\alpha^i) \in \tilde{A}^i$$

satisfying that for every $i \in \{1, \ldots, N\}$ and for every $\boldsymbol{\alpha} = (\alpha^1, \ldots, \alpha^n) \in A$,

$$J^i(\alpha^1, \ldots, \alpha^N) = \tilde{J}^i(\phi_1(\alpha^i), \ldots, \phi_N(\alpha^N)).$$

Proposition 5.

1. Every coordination game is isomorphic to a congestion game.

2. Every dummy game is isomorphic to a congestion game.

3. Every potential game is isomorphic to a congestion game.

Proof.

1. Let $(A^1 \times \ldots \times J^N, \{J^i\}_{i=1\ldots,N})$ be a coordination game. We can denote the common cost function for players $i \in \{1, \ldots, N\}$ by

$$J(\boldsymbol{\alpha}) = J^i(\boldsymbol{\alpha}).$$

   Now we construct a congestion game $(E, \tilde{A}^1 \times \ldots \tilde{A}^N, \{c_e\}_{e\in E}, \{k_e\}_{e\in E}, \{\tilde{J}^i\}_{i=1,\ldots,N})$ by the following steps:

   - each strategy profile $\boldsymbol{\alpha} \in A$ is associated to a different resource $e(\boldsymbol{\alpha})$. The collection of resources is denoted by $E = \{e(\boldsymbol{\alpha}) : \boldsymbol{\alpha} \in A\}$. In another words, the resource set E can be indexed by the set of strategy profiles;
   - for each resource $e = e(\boldsymbol{\alpha}) \in E$, we define its cost function $c_e : \{0, \ldots, N\} \to \mathbb{R}$ by

$$c_e(k) = c_{e(\boldsymbol{\alpha})}(k) = \mathbb{1}_{\{k=N\}} \cdot J(\boldsymbol{\alpha}), \qquad \forall\, k \in \{0, \ldots, N\};$$

   - for each player i, her feasible strategy set $\tilde{A}^i$ is defined by

$$\tilde{A}^i = \{\phi_i(\alpha^i) : \alpha^i \in A^i\}$$

     where the mapping $\phi_i : A^i \to 2^E$ takes the form

$$\phi_i(\alpha^i) = \bigcup_{\boldsymbol{\alpha}^{-i}\in A^{-i}} \left\{e(\alpha^i, \boldsymbol{\alpha}^{-i})\right\}, \qquad \forall\, \alpha^i \in A^i.$$

     The collection of strategy profiles for N players is denoted by $\tilde{A} = \tilde{A}^1 \times \ldots \tilde{A}^N$.

- for each player i, her cost function $\tilde{J}^i : \tilde{A} \to \mathbb{R}$ is defined by

$$\tilde{J}^i(\tilde{\boldsymbol{\alpha}}) = \sum_{e \in \tilde{\alpha}^i} c_e(k_e(\tilde{\boldsymbol{\alpha}})),$$

  where $k_e(\tilde{\boldsymbol{\alpha}}) = |\{i : e \in \tilde{\alpha}^i\}| \in \{0, \ldots, N\}$ is the load function associated to resource $e \in E$ evaluated at strategy profile $\tilde{\boldsymbol{\alpha}} \in \tilde{A}$.

We can see that the mapping $\{\phi_i\}_{i=1,\ldots,N}$ are bijective from A^i to $\tilde{A}^i$ for every player i. Moreover, for a given $i \in \{1, \ldots, N\}$, for every resource $e = e(\boldsymbol{\beta}) \in E$ with some $\boldsymbol{\beta} \in A$, and for every strategy $\tilde{\alpha}^i = \phi_i(\alpha^i)) \in \tilde{A}^i$ with some $\alpha^i \in A^i$, we have

$$e(\boldsymbol{\beta}) = e \in \tilde{\alpha}^i \iff \beta^i = \alpha^i.$$

Thus, if a resource $e = e(\boldsymbol{\beta}) \in E$ is fully loaded with a strategy profiles $\tilde{\boldsymbol{\alpha}} = (\phi_1(\alpha^1), \ldots, \phi_N(\alpha^N) \in \tilde{A}$ with some $\boldsymbol{\alpha} = (\alpha^1, \ldots, \alpha^N) \in A$, we must have

$$\begin{aligned} k_e(\tilde{\boldsymbol{\alpha}}) = N &\iff e = e(\boldsymbol{\beta}) \in \tilde{\alpha}^i, \qquad \forall\, i = 1, \ldots, N \\ &\iff \beta^i = \alpha^i, \qquad \forall\, i = 1, \ldots, N \\ &\iff \boldsymbol{\beta} = \boldsymbol{\alpha} \end{aligned}$$

Hence, we can have for every $\boldsymbol{\alpha} \in A$, the cost function for player i evaluated at the strategy profile $\tilde{\boldsymbol{\alpha}} = (\phi_1(\alpha^1), \ldots, \phi_N(\alpha^N)) \in \tilde{A}$ satisfies

$$\tilde{J}^i(\tilde{\boldsymbol{\alpha}}) = \sum_{e \in \tilde{\alpha}^i} c_e(k_e(\tilde{\boldsymbol{\alpha}})) = \sum_{e(\boldsymbol{\beta}) \in \phi_i(\alpha^i)} c_e(k_e(\tilde{\boldsymbol{\alpha}})) = \sum_{\{e(\boldsymbol{\beta}) \in E : \beta^i = \alpha^i\}} \mathbb{1}_{k_{e(\boldsymbol{\beta})}(\tilde{\boldsymbol{\alpha}}) = N} \cdot J(\boldsymbol{\beta})$$

so that

$$\tilde{J}^i(\phi_1(\alpha^1), \ldots, \phi_N(\alpha^N)) = \sum_{\{e(\boldsymbol{\beta}) \in E : \beta^i = \alpha^i\}} \mathbb{1}_{\boldsymbol{\beta} = \boldsymbol{\alpha}} \cdot J(\boldsymbol{\beta}) = J(\boldsymbol{\alpha}) = J^i(\boldsymbol{\alpha}).$$

This shows that the congestion game constructed above is isomorphic to the coordination game $(A^1 \times \ldots \times A^N, \{J^i\}_{i=1,\ldots,N})$.

2. (todo)

3. The result follows directly by applying Proposition 4 and the previous two cases.

□

3.5 Example: Network Design Game

This example is a special case of of Network Congestion Game (Shapley).

Let us consider a graph (or network) $G = (V, E)$. There are N players enumerated by $\{1, \ldots, N\}$. We view each edge $e \in E$ as a resource, and we define a mapping $c : E \ni e \mapsto c(e) \in \mathbb{R}$ representing the cost for using edge $e \in E$.

For each player $i \in \{1, \ldots, N\}$, we associate it with two vertices $(s_i, t_i) \in V \times V$ such that the set of feasible strategy for player i are collection of paths on graph G from vertex s_i to t_i:

$$A^i = \{ \text{ path from } s_i \text{ to } t_i \text{ on graph } G\}.$$

(when we talk about paths, we should clarify if we are on a directed graph or we are talking about simple paths on undirected graph)

For every $e \in E$, we define a load function $k_e : A \ni \boldsymbol{\alpha} \mapsto k_e(\boldsymbol{\alpha}) \in \{0, \ldots, N\}$ such that

$$k_e(\boldsymbol{\alpha}) = \left\{i : e \in \alpha^i\right\}$$

For every player $i \in \{1, \ldots, N\}$, we associate her a cost function $J^i : A \ni \boldsymbol{\alpha} \mapsto J^i(\boldsymbol{\alpha}) \in \mathbb{R}$ given by

$$J^i(\boldsymbol{\alpha}) = \sum_{e \in \alpha^i} \frac{c(e)}{k_e(\boldsymbol{\alpha})}.$$

The network with these cost functions $\{J^i\}_{i=1\ldots,N}$ is called a **cost sharing network**.

For each edge $e \in E$, if we associate it with a cost function $c_e : \{0, \ldots, N\} \mapsto \mathbb{R}$ defined as

$$c_e(k) = \frac{c(e)}{k}, \qquad \forall k = 1, \ldots, N, \qquad \text{and } c_e(0) = 0,$$

then we can see that the game is indeed a congestion game. We call it a **network congestion game**.

We define the social cost for N players employing a strategy profile $\boldsymbol{\alpha} \in A$ as the sum of individual costs of each player incurred by taking collectively the strategy profile $\boldsymbol{\alpha}$, namely

$$J(\boldsymbol{\alpha}) = \sum_{i=1}^{N} J^i(\boldsymbol{\alpha}).$$

Consider a set value function $e : A \ni \boldsymbol{\alpha} \to e(\boldsymbol{\alpha}) \in E$ such that

$$e(\boldsymbol{\alpha}) = \bigcup_{i=1,\ldots,N} \alpha^i.$$

Thus, we can have

$$J(\boldsymbol{\alpha}) = \sum_{i=1}^{N} J^i(\boldsymbol{\alpha}) = \sum_{i=1} \sum_{e \in \alpha^i} \frac{c(e)}{k_e(\boldsymbol{\alpha})} = \sum_{e \in e(\boldsymbol{\alpha})} k_e(\boldsymbol{\alpha}) \frac{c(e)}{k_e(\boldsymbol{\alpha})} = \sum_{e \in e(\boldsymbol{\alpha})} c(e)$$

Proposition 6. The game $(A^1 \times \ldots \times A^N, \{J^i\}_{i=1,\ldots,N})$ defined above is a potential game.

Proof. consider a potential function

$$\varphi(\boldsymbol{\alpha}) = \sum_{e \in E} \sum_{k=1}^{k_e(\boldsymbol{\alpha})} \frac{c(e)}{k} = \sum_{e \in E} c(e) \cdot h(k_e(\boldsymbol{\alpha})) \tag{3.5.1}$$

where

$$k_e(\boldsymbol{\alpha}) = \sum_{i=1}^{k_e(\boldsymbol{\alpha})} \frac{1}{i}.$$

□

4 Graphon games

4.1 Erdos-Renyi Graph

We present 2 types of simple random graph:

- Given two integers N and M, consider a set of graphs

$$\Omega = \{\text{set of graphs with } N \text{ vertices } \{1, \dots, N\} \text{ and } M \text{ edges}\}.$$

  such that each of these graphs is sampled with the same probability equals to

$$p = \frac{1}{|\Omega|} = \binom{\binom{N}{2}}{M}^{-1}$$

  This kind of random graph is denoted by $G(N, M)$.

- (introduced by Gillbert firstly) Consider a graph with N vertices. An edge is in the graph with probability $p \in [0, 1]$. This kind of random graph is denoted by $G(N, p)$. Its set of edges is denoted by $E(G(N, p))$.

  Thus, the number of edges is random in $G(N, p)$, with expectation

$$\mathbb{E}[E(G(N, p))] = p \begin{pmatrix} N \\ 2 \end{pmatrix}.$$

  A sample of random graph $G(N, p)$ has M edges with probability

$$\mathbb{P}(|E(G(N, p)))| = M) = p^M (1 - p)^{N(N-1)/2 - M}.$$

4.2 Stochastic Block model

Definition 12. (adapted from Eabbe, ref) For positive integers N, n, a probability vector p of dimension n, and a symmetric matrix W of dimension $n \times n$ with entries in $[0, 1]$, **the model** $SBM(N, p, W)$ defines an $N-$vertex random graph with vertices split in n communities, where each vertex is assigned a community label in $\{1, \dots, n\}$ independently under the **community prior** p, and pair of vertices with labels k and l connect independently with probability $W_{k,l}$.

Definition 13. Let N be a positive integer (the number of vertices), n be a positive integer (the number of communities), $p = (p_1, \dots, p_n)$ be a probability vector on $[n] := \{1, \dots, n\}$ (the prior on the n communities, $\sum p_k = 1$.) and W be a $n \times n$ symmetric matrix with entries in $[0, 1]$ (the connectivity probabilities). The pair (X, G) is drawn under $SBM(N, p, W)$ if X is a random variable in $[n]^N$ with i.i.d. components $(X_i)_{i \in \{1,\dots,N\}}$ distributed under p, and

$G = (\{1, \ldots, N\}, E(G))$ is an $N-$vertex simple graph where vertices i and j are connected with probability W_{X_i,X_j}, independently of other pairs of vertices. Namely

$$\mathbb{P}(X_i = k) = p_k, \qquad \forall i \in \{1, \ldots, N\} \text{ and } k \in [n],$$

and

$$\mathbb{P}(e_{ij} \in E(G)|X) = W_{X_i,X_j}, \qquad \forall i, j \in \{1, \ldots, N\}.$$

The community sets $\mathcal{C} = \{\mathcal{C}_1, \ldots, \mathcal{C}_n\}$ is defined by $\mathcal{C}_k = \mathcal{C}_k(X) := \{i \in \{1, \ldots, N\} : X_i = k\}$, $k \in [n]$.

We define a vector $y \in \{0,1\}^{\binom{N}{2}}$ such that $y_{ij} = \mathbb{1}_{e_{ij} \in E(G)}$ for any $i, j \in \{1, \ldots, N\}$. Thus, the distribution of (X, G) is defined as follow: for each $x \in [n]^N$ and $y \in \{0,1\}^{\binom{N}{2}}$:

$$\mathbb{P}(X = x) = \prod_{i=1}^{N} p_{x_i} = \prod_{k=1}^{n} p_k^{|\mathcal{C}_k(x)|} \tag{4.2.1}$$

$$\begin{aligned} \mathbb{P}(E(G) = y|X = x) &= \prod_{1 \le i \le j \le N} W_{x_i,x_j}^{y_{ij}} (1 - W_{x_i,x_j})^{1-y_{ij}} && (4.2.2) \\ &= \prod_{1 \le i \le j \le N} W_{k,l}^{N_{k,l}(x,y)} (1 - W_{k,l})^{N_{k,l}^c(x,y)} && (4.2.3) \end{aligned}$$

where

$$\begin{aligned} N_{k,l}(x,y) &= \sum_{i<j, x_i=k, x_j=l} 1_{\{y_{ij}=1\}} && (4.2.4) \\ N_{k,l}^c(x,y) &= \sum_{i<j, x_i=k, x_j=l} 1_{\{y_{ij}=0\}} = |C_k(x)| \cdot |C_l(x)| - N_{k,l}(x,y), \quad k \neq l && (4.2.5) \\ N_{k,k}^c(x,y) &= \sum_{i<j, x_i=k, x_j=l} 1_{\{y_{ij}=0\}} = \binom{|C_k(x)|}{2} - N_{k,k}(x,y) && (4.2.6) \end{aligned}$$

Remark 5.

- if $W_{k,l} = q$ and $N = n$, then we recover the Erdos-Reyi graph $G(N, q)$
- for some $w, q \in [0,1]$ such that

$$W_{k,l} = \begin{cases} w & k = l \\ q & k \neq l \end{cases}$$

  Then we have **planted partition model**. Moreover, if $w > q$, we say that the graph has a assortative structure, and if $w < q$, it has a **disassortative** structure.

4.3 Graphon

We begin with some notations. Let $L^2([0,1])$ be the space of square integrable real-value functions defined on $[0,1]$ and $L^2([0,1],\mathbb{R}^k)$ the space of square integrable vector valued functions defined on $[0,1]$. The norm $\|\cdot\|_{L^2}$ are defined as usual, and the norm for operator $A : L^2 \to L^2$ is defined as

$$\|A\| = \sup_{f:\|f\|_{L^2}\leq 1} \|Af\|_{L^2}$$

For any linear operator A, we denoted by $\lambda_{max}(A)$ and $\rho(A)$ the largest eigenvalue and the spectral radius of A. The symbol $\mathbb{1}_N$ denotes the ector of all ones elements in $\mathbb{R}^N$ and $1_{[0,1]}(\cdot)$ is the function constantly equaled to one on $[0,1]$. $\mathbb{I}$ is the identify operator and I the identity matrix.

Definition 14. A mapping $W : [0,1]\times[0,1] \to [0,1]$ is a **kernel** if it is measurable (w.r.t. to the Lebesgue measure on $[0,1]$) and symmetric, namely

$$W(x,y) = W(y,x).$$

Definition 15. (Graphon games, Francesca Paris et Asuman Ozdaglar) A **graphon** is a bounded symmetric measurable function $W : [0,1]\times[0,1] \to [0,1]$ encoding interactions among an infinite (continuum) number of players. By abuse of terminology, the function W is also called the kernel of graphon W. The collection of graphons is denoted by $\mathcal{W}$.

Remark 6. In the context of graphon, the continuous variable $x \in [0,1]$ generalizes the notion of player, and $W(x,y)$ represents the strength of the interaction between player x and player y.

Definition 16. For a given graphon $W \in \mathcal{W}$, we define the associate graphon operator $A^W : L^2 \to L^2$ by

$$(A^W f)(x) = \int_{[0,1]} W(x,y)f(y)dy$$

We assume that the graphon kernel is bounded:

$$\int W(x,y)dy \leq 1.$$

Proposition 7. The graphon operator A^w is a Hilbert-Schmidt operator. In other words, A^w is a bounded operator on the Hilbert space $L^2([0,1])$ with finite Hilbert-Schmidt norm

$$\|A^W\|_{HS}^2 = \sum_{s_i\in\sigma(A^w)} s_i^2 < \infty.$$

where $\sigma(A^w)$ is the spectrum of the operator A^w.

Remark 7. Recalled that a Hilbert space is a vector space equipped with a complete inner product. A Banach space is a complete normed vector space. We also recall some notions and properties related to Hilbert-Schmidt operator below (ref. functional analysis, Walter Rudin):

- The vector space $H = L^2([0,1], \mathbb{R}^k)$ equipped with the inner product $\langle f, g\rangle = \int_{[0,1]} f(x)^\top g(x) dx$ is a Hilbert space.

- A linear operator $A : H_1 \to H_2$ from a Banach space H_1 to another Banach space H_2 is said to be **bounded** if there exists an $M > 0$ such that for all $f \in H_1$,

$$\|Af\|_{H_2} \leq M\|f\|_{H_1}.$$

  The collection of all bounded linear operators of H_1 into H_2 is denoted by $\mathcal{B}(H_1, H_2)$. For simplicity $\mathbb{B}(H, H)$ will be denoted by $\mathcal{B}(H)$.

- Since every Banach space is a norm space, so that a linear operator $A : H_1 \to H_2$ is bounded if and only if it is **continuous**. Here, the continuity of operator A is defined with respect to the topological vector spaces (H_1, τ_1) and (H_2, τ_2) in which the topologies τ_1 and τ_2 are induced by their **natural metric** $d_1(f, g) = \|f - g\|_{H_1}$ and $d_2(f', g') = \|f' - g'\|_{H_2}$ respectively.

  As a complement, since H_1 and H_2 are **norm spaces** equipped with nature topologies τ_1 and τ_2, the above definition for bounded linear operator is equivalent to the following definition (c.f. Rudin 1.31) of bounded linear operator between topological vector spaces (H_1, τ_1) and (H_2, τ_2):

  A linear operator $A : (H_1, \tau_1) \to (H_2, \tau_2)$ is said to be bounded if it maps every τ_1-bounded set in H_1 into a τ_2-bounded set in H_2.

  (a subset B of a topological vector space (Ω, τ) is said to be $\tau-$bounded if to every neighborhood V of 0 in Ω corresponds a number $s > 0$ such that $B \subset tV$ for every $t > s$.)

- (cf Rudin 4.1) Suppose H_1 and H_2 are norm spaces. Associate to each bounded linear operator $A \in \mathcal{B}(H_1, H_2)$ the number

$$\|A\| = \sup\{\|Af\| : f \in H_1, \|f\| \leq 1\}.$$

  This definition of $\|A\|$ makes $\mathcal{B}(H_1, H_2)$ into a norm space. If H_2 is a Banach space, so is $\mathcal{B}(H_1, H_2)$.

- (cf Rudin 4.16) Suppose H_1 and H_2 are Banach spaces and U is the open unit ball in H_1, i.e. $U = \{x : \|x\|_{H_1} < 1\}$. A linear map $A : H_1 \to H_2$ is said to be **compact** if the closure of $A(U)$ is compact in H_2.

- Suppose H is a Banach space, the $\mathcal{B}(H)$ is a Banach space, and moreover, it is an algebra: if $S, T \in \mathcal{B}(H)$, one defines $ST \in \mathcal{B}(H)$ by

$$(ST)(f) = S(T(f)), \quad f \in H.$$

An operator $T \in \mathcal{B}(H)$ is said to be **invertible** if there exists $S \in \mathcal{B}(H)$ such that

$$ST = I = TS.$$

In this case, we write $S = T^{-1}$. By the opening mapping theorem (cf Rudin 2.11), this happens if and only if $ker(T) = \{f \in H : Tf = 0\} = \{0\}$ and $rg(T) = \{g \in H : \exists f \in H, \quad Tf = g\} = H$.

The **spectrum** $\sigma(A)$ of an operator $A \in \mathcal{B}H$ is the set of all scalars λ such that $A - \lambda I$ is not invertible. Thus, $\lambda \in \sigma(T)$ if and only if at least one of the following two statements is true:

1. The range of $T - \lambda I$ is not all of H (not subjective).
2. $T - \lambda I$ is not one-to-one (not injective).

If $T - \lambda I$ is not one-to-one, then λ is said to be an **eigenvalue** of A; the corresponding eigenspace is $ker(T - \lambda I)$.

- Hilbert-Schmidt operator is a compact operator.

4.4 relationship between graph and graphon

Definition 17. We say that a function $W : [0,1]^2 \to [0,1]$ is a step function if there is a partition $Q = \{Q_1, \ldots, Q_n\}$ of $[0,1]$ into measurable sets such that W is constant on every product set $Q_k \times Q_l$. The sets Q_k are the steps of W. We define a **step function graphon** $W^[n] \in \tilde{\mathcal{W}}$ corresponding to a $n \times n$ matrix $P^{[n]}$ by setting

$$W^{[n]}(x,y) = P^{[n]}_{k,l}, \qquad \forall (x,y) \in Q_k \times Q_l. \tag{4.4.1}$$

Note that $W^{[n]}$ depends on the ordering of the nodes in $P^{[n]}$. Apart from that, the relation between graphs and step functions graphons is bijective.

Definition 18. (Sampling procedure) We pick uniformly N points $\{u_i\}_{i=1,\ldots,N}$ from $[0,1]$ and define the weight adjacency matrix $P^{[N]}_w$ as follows

$$[P^{[N]}_w]_{ij} = \mathbb{1}_{i \neq j} \cdot W(u_i, u_j). \tag{4.4.2}$$

Additionally, starting from $P^{[N]}_w$, we define the $0-1$ adjacency matrix $P^{[N]}_s$ as the adjacency matrix corresponding to a graph with N nodes obtained by randomly connecting nodes i to j with Bernoulli probability $[P^{[N]}_w]_{ij}$.

So, according to the sampling procedure, for any $p \in [0,1]$, the constant graphon $W(x,y) = p$ coincides with the Erdos-Renyi random graph model with edge probability p.

Moreover, if we consider a prior probability vector $p \in [0,1]^n$ with $\sum_k p_k = 1$, any step function graphon $W^{[n]}$ gives raise to a stochastic block model $SBM(N, p, W^{[n]})$. A sample (X, G) with $x_i \in Q$ and $G = ([N], E(G))$ can be characterized by the weight adjacency matrix $P^{[N]}_w$ computed from $W^{[n]}$.

4.5 cut norms

Definition 19. The cut norm of a graphon $W \in \tilde{\mathcal{W}}$ is denoted by $\|W\|_\square$ and is defined as follows:

$$\|W\|_\square = \sup_{D_1, D_2 \in [0,1]} \left| \int_{D_1 \times D_2} W(x,y) dx dy \right| \tag{4.5.1}$$

where D_1, D_2 are measurable subsets of $[0,1]$. The cut metric between two graphons $V, W \in \tilde{\mathcal{W}}$ is given by

$$d_\square(W, V) = \inf_{\pi \in \prod_{[0,1]}} \|W^\pi - V\|_\square,$$

where $W^\pi(x,y) = W(\pi(x), \pi(y))$ and $\prod_{[0,1]}$ is the class of measure preserving permutations $\pi : [0,1] \to [0,1]$.

Remark 8. By Monotonicity argument, we can have

$$\|W\|_\square = \sup_{0 \le f \le 1, 0 \le g \le 1} \left| \int \int f(x) g(y) w(x,y) dx dy \right|.$$

Because of the permutation function ϕ, the mapping $d_\square$ is not a well defined metric in $\tilde{\mathcal{W}}$. We define the space $\mathcal{W} = \tilde{\mathcal{W}} / \sim$ where the equivalent relation $\sim$ on $\tilde{W}$ is defined by

$$V \sim W \text{ if there exists } \pi \in \prod{}_{[0,1]} \text{ such that } V = W^\pi.$$

Namely, we identify graphons up to measure preserving permutations. THus, the space $(\mathcal{W}, d_\square)$ is a well defined metriczable vector space.

Proposition 8. The space $(\mathcal{W}, d_\square)$ is complete.

Lemma 3. ([Janson, 2010] Theorem E.6) Given a graphon $W \in \mathcal{W}$ and the associating graphon operator A^W, if $|W| = \int W(x,y) dy \le 1$, then we have for any $p, q \in [1, \infty]$ and $q' = (1 - 1/q)^{-1}$,

$$\|W\|_\square \le \|A^W\|_{L^p, L^q} \le \sqrt{2} (4 \|W\|_\square)^{(1-\frac{1}{p}) \wedge \frac{1}{q}}$$

where

$$\|A^W\|_{L^p, L^q} = \sup_{\|f\|_{L^p} \le 1} \|A^W f\|_{L^q} = \sup_{\|f\|_{L^p} \le 1} \sup_{\|g\|_{L^{q'}} \le 1} \left| \int A^W f(x)\, g(x) dx \right|.$$

A special case is when $p = q = 2$, then we have

$$\|W\|_\square \le \|A^W\|_{L^2, L^2} \le \sqrt{8 \|W\|_\square}.$$

The proof of Lemma 3 relies on the Riesz-Thorin interpolation theorem and the inequality

$$\|W\|_\square \le \|A^W\|_{L^\infty, L^1} \le 4 \|W\|_\square.$$

We first recall the Riesz-Thorin interpolation theorem:

Theorem 2. (Riesz-Thorin interpolation theorem)
Suppose $(p_0, q_0), (p_1, q_1) \in [1, \infty] \times [1, \infty]$ and for some $r \in (0, 1)$ such that

$$\left(\frac{1}{p}, \frac{1}{q}\right) = r\left(\frac{1}{p_0}, \frac{1}{q_0}\right) + (1-r)\left(\frac{1}{p_1}, \frac{1}{q_1}\right).$$

Consider a linear function T that maps $L^{p_0}((X, \mu), \mathbb{C}) + L^{p_1}((X, \mu), \mathbb{C})$ into $L^{q_0}((Y, \nu), \mathbb{C}) + L^{q_1}((Y, \nu), \mathbb{C})$, where (X, μ) and (Y, ν) are two measurable spaces. If $q_0 = q_1 = \infty$, we further suppose that the measure ν is semifinite. If we assume that there exists two constants $M_0, M_1 > 0$ such that

$$\|Tf|_{L^{q_0}} \leq M_0 \|f\|_{L^{p_0}}, \qquad \forall\, f \in L^{p_0}((X, \mu), \mathbb{C})$$

and

$$\|Tf|_{L^{q_1}} \leq M_1 \|f\|_{L^{p_1}}, \qquad \forall\, f \in L^{p_1}((X, \mu), \mathbb{C}),$$

then T is bounded on L^p and furthermore,

$$\|Tf\|_{L^q} \leq M_0^{1-r} M_1^r \|f\|_{L^p}$$

for any complex-value function $f \in L^p(X, \mu, \mathbb{C})$.

Lemma 4.

$$\|W\|_\square \leq \|A^W\|_{L^\infty, L^1} \leq 4\|W\|_\square.$$

Proof.

$$\begin{aligned}
\|A^W\|_{L^\infty, L^1} &= \sup_{\|f\|_{L^\infty} \leq 1} \sup_{\|g\|_{L^\infty} \leq 1} \left| \int A^W f(x) g(x) dx \right| \\
&= \sup_{-1 \leq f \leq 1} \sup_{-1 \leq g \leq 1} \left| \int \int W(x, y) f(y) g(x) dx dy \right| \qquad (4.5.2)
\end{aligned}$$

Take $f = \mathbb{1}_{D_1}$ and $g = \mathbb{1}_{D_2}$ for $D_1, D_2 \subset [0, 1]$, then we have

$$\|A^W\|_{L^\infty, L^1} \geq \|W\|_\square.$$

We can introduce two functions f_1, f_2 valued in $[0, 1]$ such that $f = f_1 - f_2$, as well as g_1, g_2 such that $g = g_1 - g_2$. Thus,

$$\|A^W\|_{L^\infty, L^1} = \sup_{0 \leq f_1, f_2 \leq 1} \sup_{0 \leq g_1, g_2 \leq 1} \left| \int \int W(x, y) [f_1(y) - f_2(y)] \cdot [g_1(x) - g_2(x)] dx dy \right| \quad (4.5.3)$$

Since

$$\begin{aligned}
& \left| \int \int W(x, y) [f_1(y) - f_2(y)] \cdot [g_1(x) - g_2(x)] dx dy \right| \\
= & \left| \langle A^W (f_1 - f_2), (g_1 - g_2) \rangle \right| \\
= & \left| \langle A^W f_1, g_1 \rangle + \langle A^W f_2, g_2 \rangle - \langle A^W f_2, g_1 \rangle - \langle A^W f_1, g_2 \rangle \right| \\
\leq & \left| \langle A^W f_1, g_1 \rangle \right| + \left| \langle A^W f_2, g_2 \rangle \right| + \left| \langle A^W f_2, g_1 \rangle \right| + \left| \langle A^W f_1, g_2 \rangle \right|
\end{aligned}$$

We have

$$\|A^W\|_{L^\infty,L^1} \leq \sup_{0\leq f_1,f_2\leq 1} \sup_{0\leq g_1,g_2\leq 1} |\langle A^W f_1, g_1\rangle|+|\langle A^W f_2, g_2\rangle|+|\langle A^W f_2, g_1\rangle|+|\langle A^W f_1, g_2\rangle| = 4\|W\|_\square.$$

□

Proof. (for Lemma 3) We show in the first place that $\|W\|_\square \leq \|A^W\|_{L^p,L^q}$. Since $L^\infty([0,1],dx) \subset L^p([0,1],dx)$ and $L^q([0,1],dx) \subset L^1([0,1],dx)$, we have

$$\begin{aligned}
\|A^W\|_{L^p,L^q} &= \sup_{\|f\|_{L^p}\leq 1} \|A^W f\|_{L^q} \\
&\geq \sup_{\|f\|_{L^\infty}\leq 1} \|A^W f\|_{L^q} \\
&\geq \sup_{\|f\|_{L^\infty}\leq 1} \|A^W f\|_{L^1} \\
&= \|A^W\|_{L^\infty,L^1} \\
&\geq \|W\|_\square
\end{aligned}$$

For the upper bound, we let $r = (1-\frac{1}{p}) \wedge \frac{1}{q}$, then $1-r = \frac{1}{p} \vee (1-\frac{1}{q})$. We define $p_1 = (1-r)p$ and $q_1 = (1-(1-r)^{-1}(1-1/q))^{-1}$. then $\overline{p}, \overline{q} \in [1,\infty]$. Thus, with $p_0 = \infty$ and $q_0 = 1$, we have

$$\left(\frac{1}{p}, \frac{1}{q}\right) = r\left(\frac{1}{p_0}, \frac{1}{q_0}\right) + (1-r)\left(\frac{1}{p_1}, \frac{1}{q_1}\right).$$

Then by applying Riesz-Thorin interpolation theorem, we obtain

$$\|A^W\|_{L^p,L^q} \leq \|A^W\|_{L^p(\mathbb{C}),L^q(\mathbb{C})} \leq \left(\|A^W\|_{L^\infty(\mathbb{C}),L^1(\mathbb{C})}\right)^r \cdot \left(\|A^W\|_{L^{p_1}(\mathbb{C}),L^{q_1}(\mathbb{C})}\right)^{1-r}.$$

where $L^\infty(\mathbb{C}), L^1(\mathbb{C}), L^{p_1}(\mathbb{C}), L^{q_1}(\mathbb{C})$ are complex-value function spaces extended from the corresponding function spaces.

We can easily see that

$$\begin{aligned}
\|A^W\|_{L^\infty(\mathbb{C}),L^1(\mathbb{C})} &= \sup_{\substack{f:[0,1]\to\mathbb{C}\\ \|f\|_{L^\infty}\leq 1}} \sup_{\substack{g:[0,1]\to\mathbb{C}\\ \|g\|_{L^\infty}\leq 1}} \left|\int\int W(x,y)f(y)g(x)dydx\right| \\
&\leq 2\|A^W\|_{L^\infty,L^1}
\end{aligned}$$

Indeed, this inequality can be improved to a constant $\sqrt{2}$ (see [Krivine, 1979])

$$\|A^W\|_{L^\infty(\mathbb{C}),L^1(\mathbb{C})} \leq \sqrt{2}\|A^W\|_{L^\infty,L^1}$$

For any $p_1, q_1 \in [1,\infty]$,

$$\begin{aligned}
\|A^W\|_{L^{p_1}(\mathbb{C}),L^{q_1}(\mathbb{C})} &= \sup_{\|f\|_{L^{p_1}(\mathbb{C})}\leq 1} \|A^W f\|_{L^{q_1}(\mathbb{C})} \\
&\leq \sup_{\|f\|_{L^1(\mathbb{C})}\leq 1} \|A^W f\|_{L^{q_1}(\mathbb{C})} \\
&\leq \sup_{\|f\|_{L^1(\mathbb{C})}\leq 1} \|A^W f\|_{L^\infty(\mathbb{C})} \\
&= \|A^W\|_{L^1(\mathbb{C}),L^\infty(\mathbb{C})}
\end{aligned}$$

where the first inequality is justified by $L^1([0,1],dx,\mathbb{C}) \subset L^{p_1}([0,1],dx,\mathbb{C})$ and the second inequality comes from $\|f\|_{L^{q_1}(\mathbb{C})} \leq \|f\|_{L^\infty(\mathbb{C})}$. Moreover,

$$\begin{aligned}
\|A^W\|_{L^1(\mathbb{C}),L^\infty(\mathbb{C})} &= \sup_{f:[0,1]\to\mathbb{C},\, \|f\|_{L^1}\leq 1} \|A^W f\|_{L^\infty(\mathbb{C})} \\
&= \sup_{\substack{f:[0,1]\to\mathbb{C} \\ \|f\|_{L^1}\leq 1}} \sup_{\substack{g:[0,1]\to\mathbb{C} \\ \|g\|_{L^1}\leq 1}} \left| \int\int W(x,y)f(y)g(x)dydx \right| \\
&\leq \sup_{f:[0,1]\to\mathbb{C},\, \|f\|_{L^1}\leq 1} \left| \int \sup_{g:[0,1]\to\mathbb{C},\, \|g\|_{L^1}\leq 1} \left| \int W(x,y)g(x)dx \right| \cdot f(y)dy \right| \\
&\leq \sup_{f:[0,1]\to\mathbb{C},\, \|f\|_{L^1}\leq 1} \|W\|_{L^\infty(\mathbb{C})} \cdot \int |f(y)|dy \\
&= \|W\|_{L^\infty(\mathbb{C})} \\
&= ess \sup_{x\in[0,1]} \left| \int W(x,y)dy \right| \\
&\leq 1
\end{aligned}$$

where the second inequality comes from Fubini's theorem and the last inequality comes from the assumption on graphon kernel $|W| \leq 1$.

Hence, with the fact that $\sqrt{2}^r \leq \sqrt{2}$ and $\|A^W\|_{L^\infty,L^1} \leq 4\|W\|_\square$, we conclude that

$$\|A^W\|_{L^p,L^q} \leq \sqrt{2}\,(4\|W\|_\square)^{\min\{1-\frac{1}{p},\frac{1}{q}\}}.$$

$\square$

4.6 Centrality measure

An extension of Bonacich measure. Far a graphon $W \in \mathcal{W}$ associated with an operator A^W and a value $\alpha \in \mathbb{R}$. Assume that $\alpha \in [0, \frac{1}{\rho(A^W)})$, and we define a mapping b_α by

$$\begin{array}{rccl} b_\alpha: & [0,1] & \longrightarrow & \mathbb{R} \\ & x & \mapsto & b_\alpha(x) := \left([\mathbb{I} - \alpha A^W]^{-1} 1_{[0,1]} \right)(x) \end{array} \tag{4.6.1}$$

The function b_α is called Bonacich centrality function.

4.7 some words on the mean field interaction

Graphon games are with an infinity of players. We can look at N player games and consider the case when $N \to \infty$. We assume that players are symmetric. Let us denote by A the set of admissible controls for a single player where A is some subset in $\mathbb{R}$. For any strategy profile $\boldsymbol{\alpha} \in (A)^N$, recall that the cost function associated to player i is defined by

$$J^i(\boldsymbol{\alpha}) = J^i(\alpha^i, (\alpha^1, \ldots, \alpha^{i-1}, \alpha^{i+1}, \ldots, \alpha^N)) = J^i(\alpha, \boldsymbol{\alpha}^{-i}) \tag{4.7.1}$$

An important special case is when J^i is symmetric with respect to all the α^j, $j \neq i$. Indeed, in this case, we can rewrite the cost function for player i into a new cost function $\tilde{J}^i$ which involves a mean filed interaction term. More precisely, we have

$$J^i(\alpha, \boldsymbol{\alpha}^{-i}) = \tilde{J}^i\left(\alpha^i, \frac{1}{N-1}\sum_{j\neq i}\delta_{\alpha^j}\right) = \tilde{J}^i\left(\alpha^i, \mu^{\boldsymbol{\alpha}^{-i}}\right) \tag{4.7.2}$$

where $\mu^{\boldsymbol{\alpha}^{-i}} := \frac{1}{N-1}\sum_{j\neq i}\delta_{\alpha^j} \in \mathcal{P}_2(A)$ is a squared integrable measure on A and $\tilde{J}^i$: $A \times \mathcal{P}_2(A) \ni (\alpha, \mu) \mapsto J^i(\alpha, \mu) \in \mathbb{R}$ represents the new cost function associated to player i.

Example: "Where do we put the towel on the beach?"

There are N players, A is a compact set of $\mathbb{R}$ and for player i, her control $\alpha^i \in A$ represents her position on the beach. Let $d : A \times A \to \mathbb{R}$ be a distance function. For player i, her cost function J^i takes the form

$$J^i(\boldsymbol{\alpha}) = a \cdot d(\alpha^i, \alpha_0) - b \cdot \frac{1}{N-1}\sum_{j\neq i} d(\alpha^i, \alpha^j) = a \cdot d(\alpha^i, \alpha_0) - b \cdot \int_A d(\alpha^i, \alpha)\mu^{\boldsymbol{\alpha}^{-i}}(d\alpha)$$

where $a, b > 0$ and $\alpha_0 \in A$ is the preferred position for all players.

The objective is to solve for ONE typical player when she plays against the distribution of all other players.

The best response function is defined as

$$\widehat{\alpha} : \mathcal{P}_2(A) \ni \mu \mapsto \arg\inf_{\alpha\in A} J(\alpha, \mu) \subseteq A$$

and a Nash equilibrium is defined as a distribution $\widehat{\mu} \in \mathcal{P}_2(A)$ that satisfies the fixed point argument:

$$\mathrm{supp}(\widehat{\mu}) \subseteq \widehat{\alpha}(\widehat{\mu}),$$

where $\mathrm{supp}(\mu)$ stands for the support of a distribution $\mu \in \mathcal{P}_2(A)$.

4.8 Graphon game

In finite player games, players can be enumerated by a finite set of numbers, say $\{1, \dots, N\}$ for example, and, a strategy profile can be by a vector $\boldsymbol{\alpha} = (\alpha_1, \dots, \alpha_N) \in A^1 \times \dots A^N$ with $A^i \subseteq \mathbb{R}^k$. Now, when dealing with Graphon games, we consider a continuum of players such that a single player is represented by a real number $x \in [0, 1]$, whilst a strategy profile turns out to be a squared integrable mapping $\boldsymbol{\alpha} : [0, 1] \ni x \mapsto \boldsymbol{\alpha}(x) \in R^k$. The set of feasible strategies for player $x \in [0, 1]$, denote by $A(x)$ is a subset of $\mathbb{R}^k$, and the collection of all feasible strategy sets among players is denote by $A = \{A(x)\}_{x\in[0,1]}$. We can also define the set of admissible strategy profiles as

$$\mathcal{A} = \left\{\boldsymbol{\alpha} \in L^2([0, 1], \mathbb{R}^k) : \quad \forall x \in [0, 1],\, \boldsymbol{\alpha}(x) \in A(x)\right\}.$$

In finite Network game, player i feels the interaction from player j through the Adjacency matrix $A^{(G)}$, and the local aggregated perceived by player i can be modeled through a term

$$z^i(\boldsymbol{\alpha}) = \frac{1}{N-1}\sum_{j\neq i} A^{(G)}_{ij}\alpha^j.$$

Similarly, in Graphon games, we define the local aggregate of interaction perceived by player $x \in [0,1]$ with respect to a strategy profile $\boldsymbol{\alpha} = (\boldsymbol{\alpha}_1, \ldots, \boldsymbol{\alpha}_k)^\top \in L^2([0,1], \mathbb{R}^k)$ by

$$z(x|\boldsymbol{\alpha}) = \int_0^1 W(x,y)\boldsymbol{\alpha}(y)dy \in \mathbb{R}^k$$

where $z(x|\boldsymbol{\alpha})_j = \int_{[0,1]} W(x,y)\boldsymbol{\alpha}_j(y)dy = (A^W\boldsymbol{\alpha}_j)(x)$ for every $j = 1, \ldots, k$ and every $x \in [0,1]$.

We define a new graphon operator $\mathbb{A}^W : L^2([0,1], \mathbb{R}^k) \to L^2([0,1], \mathbb{R}^k)$ such that for any strategy profile $\boldsymbol{\alpha} \in L^2([0,1], \mathbb{R}^k)$ and for any $x \in [0,1]$, we have

$$\left(\mathbb{A}^W\boldsymbol{\alpha}\right)(x) := z(x|\boldsymbol{\alpha}) = \begin{bmatrix} A^W\boldsymbol{\alpha}_1(x) \\ \vdots \\ A^W\boldsymbol{\alpha}_k(x) \end{bmatrix}.$$

For player $x \in [0,1]$, her cost function is defined by

$$J^x : A(x) \times \mathcal{A} \ni (\alpha', \boldsymbol{\alpha}) \mapsto J(\alpha', z(x|\boldsymbol{\alpha}))$$

where $J : \mathbb{R}^k \times \mathbb{R}^k \ni (\beta, z) \mapsto J(\beta, z) \in \mathbb{R}$ is a predefine cost function for the graphon game.

Definition 20. A graphon game $\mathcal{G}(A, J, W)$ is defined in terms of a continuum set of players indexed by $[0,1]$, a graphon $W \in \mathcal{W}$, a cost functions $J(\cdot,\cdot)$, and a collection of feasible strategy sets $A = \{A(x)\}_{x\in[0,1]}$.

The best response function for player $x \in [0,1]$ in a graphon game is define as a set value function

$$\underline{\alpha}^x_{BR} : \mathcal{A} \ni \boldsymbol{\alpha} \mapsto \arg\inf_{\beta\in A(x)} J(\beta, z(x|\boldsymbol{\alpha}))$$

An admissible strategy $\boldsymbol{\alpha}^* \in \mathcal{A}$ is a Nash equilibrium if it satisfies a fixed point condition:

$$\boldsymbol{\alpha}^*(x) \in \underline{\alpha}^x_{BR}(\boldsymbol{\alpha}^*) \quad \textbf{for all } x \in [0,1],$$

in another words, for every $x \in [0,1]$,

$$J(\boldsymbol{\alpha}(x), z(x|\boldsymbol{\alpha}^*)) \leq J(\alpha', z(x|\boldsymbol{\alpha}^*)), \qquad \alpha' \in A(x).$$

Remark 9. If we assume that for any $z \in \mathbb{R}^k$, there exists a unique minimum for the mapping $\beta \mapsto J(\beta, z)$, then we can see that the best response function for graphon game $\mathcal{G}(A, J, W)$ as a function from $\mathcal{A}$ to $\mathcal{A}$ such that

$$\underline{\alpha}_{BR} : \mathcal{A} \ni \boldsymbol{\alpha} \mapsto \left(x \mapsto \arg\inf_{\beta\in A(x)} J\left(\beta, z(x|\boldsymbol{\alpha})\right)\right).$$

Since $\boldsymbol{\alpha} \in \mathcal{A} \subseteq L^2([0,1], \mathbb{R}^k)$, we need to show that $\underline{\alpha}_{BR}$ does not depend on the representation of $\boldsymbol{\alpha}$ and it is also squared integrable on $[0,1]$. This first point can be shown from the fact that $\mathbb{A}^W \boldsymbol{\alpha}$ is independent of the choice of representative $\boldsymbol{\alpha}$ in $L^2([0,1], \mathbb{R}^d)$. The second point can be justified assuming the feasible strategy set $A(x)$ is a subset of a compact set A_{comp} of $\mathbb{R}^k$.

Example: Min-Max graphon: let $W(x,y) = x \wedge y(1 - x \vee y)$ for any $(x,y) \in [0,1] \times [0,1]$.

Remark 10. If $f \in ker(A^W - \lambda I)$, then $(f, 0, 0, \ldots, 0)$ is an eigenvector of $\mathbb{A}^W$ corresponded to the eigenvalue λ.

4.8.1 existence result

Assumption 1.

1. The function $J : \mathbb{R}^k \times \mathbb{R}^k \ni (\alpha, z) \to J(\alpha, z) \in \mathbb{R}$ is continuously differentiable and strongly convex in α with a uniform constant $l_c > 0$ for every $z \in \mathbb{R}^k$. Namely, for every $\alpha', \alpha \in \mathbb{R}^k$ and $z \in \mathbb{R}^k$, we have for every $b \in [0,1]$
$$J(b\alpha + (1-b)\alpha', z) \leq bJ(\alpha, z) + (1-b)J(\alpha', z) - \frac{l_c}{2} b(1-b)\|\alpha' - \alpha\|^2.$$
2. $\nabla_\alpha J(\alpha, z)$ is Lipschitz in z with a uniform constant l_J for every $\alpha \in \mathbb{R}^k$. Namely, for every $z', z \in \mathbb{R}^d$ and for every $\alpha \in \mathbb{R}^k$, we have
$$\|\nabla_\alpha J(\alpha, z') - \nabla_\alpha J(\alpha, z)\| \leq l_J \|z' - z\|.$$
3. For every $x \in [0,1]$, the set of feasible strategy $A(x)$ is closed and convex.

Definition 21. We define an operator $\mathbb{B} : L^2([0,1], \mathbb{R}^k) \ni \tilde{z} \mapsto \mathbb{B}\tilde{z} \in L^2([0,1], \mathbb{R}^k)$ such that for every $x \in [0,1]$
$$\mathbb{B}\tilde{z}(x) = \underset{\alpha \in A(x)}{\arg\inf}\, J(\alpha, \tilde{z}(x)).$$

We can see that

$$\boldsymbol{\alpha} \text{ is a Nash Equilibrium} \iff \boldsymbol{\alpha} \text{ is a fixed point of } \mathbb{B}\mathbb{A}^W.$$

Remark 11. $\mathbb{B}$ is not a linear operator.

Lemma 5. Under assumption 1, we can show that $\mathbb{B}$ is Lipschitz. More precisely, for every $f, g \in L^2([0,1], \mathbb{R}^k)$,
$$\|\mathbb{B}f - \mathbb{B}g\|_{L^2} \leq \frac{l_J}{l_c} \|f - g\|_{L^2}.$$

Proof. Under the strong convexity assumption, we know that for every $f \in L^2([0,1], \mathbb{R}^k)$ and every $x \in [0,1]$, the set $\mathbb{B}f(x)$ that minimizes $J(\cdot, f(x))$ is a singleton. By the abuse of terminology, we still denote it by $\mathbb{B}f(x)$.

By variation inequality, the strong convexity of J in α implies that for every $f, g \in L^2([0,1], \mathbb{R}^k)$ and $x \in [0,1]$,

$$\begin{aligned} \langle \nabla_\alpha J(\mathbb{B}f(x), f(x)),\, \mathbb{B}g(x) - \mathbb{B}f(x) \rangle &\geq 0 \\ \langle \nabla_\alpha J(\mathbb{B}g(x), g(x)),\, \mathbb{B}f(x) - \mathbb{B}g(x) \rangle &\geq 0. \end{aligned}$$

We add up these two inequalities and obtain that

$$\begin{aligned} & \langle \nabla_\alpha J(\mathbb{B}f(x), f(x)) - \nabla_\alpha J(\mathbb{B}f(x), g(x)),\, \mathbb{B}g(x) - \mathbb{B}f(x) \rangle \\ \geq\ & \langle \nabla_\alpha J(\mathbb{B}g(x), g(x)) - \nabla_\alpha J(\mathbb{B}f(x), g(x)),\, \mathbb{B}g(x) - \mathbb{B}f(x) \rangle \end{aligned}$$

The strong convexity of J in α also implies the strong monotonicity of $\nabla_\alpha J$ in α, namely for any $\alpha, \alpha', z \in \mathbb{R}^k$,

$$\langle \nabla_\alpha J(\alpha, z) - \nabla_\alpha J(\alpha', z),\, \alpha - \alpha' \rangle \geq l_c \|\alpha - \alpha'\|^2$$

Thus, we have

$$\begin{aligned} & \|\nabla_\alpha J(\mathbb{B}f(x), f(x)) - \nabla_\alpha J(\mathbb{B}f(x), g(x))\| \cdot \|\mathbb{B}g(x) - \mathbb{B}f(x)\| \\ \geq\ & \langle \nabla_\alpha J(\mathbb{B}f(x), f(x)) - \nabla_\alpha J(\mathbb{B}f(x), g(x)),\, \mathbb{B}g(x) - \mathbb{B}f(x) \rangle \\ \geq\ & \langle \nabla_\alpha J(\mathbb{B}g(x), g(x)) - \nabla_\alpha J(\mathbb{B}f(x), g(x)),\, \mathbb{B}g(x) - \mathbb{B}f(x) \rangle \\ \geq\ & l_c \|\mathbb{B}g(x) - \mathbb{B}f(x)\|^2 \end{aligned}$$

where the first inequality is justified by the Cauchy-Schwartz inequality.

Since $\nabla_\alpha J$ is Lipschitz in z, so we can have

$$\|\mathbb{B}g(x) - \mathbb{B}f(x)\| \leq \frac{l_J}{l_c} \|f(x) - g(x)\|$$

Hence,

$$\|\mathbb{B}g - \mathbb{B}f\|_{L^2}^2 = \sum_{i=1}^{k} \int_{[0,1]} \left([\mathbb{B}f(x)]_i - [\mathbb{B}g(x)]_i\right)^2 dx \leq \left(\frac{l_J}{l_c}\right)^2 \|f - g\|_{L^2}^2$$

□

The following assumption guarantees that for every admissible strategy profile $\boldsymbol{\alpha} \in \mathcal{A}$, the best response function evaluated at $\boldsymbol{\alpha} \in \mathcal{A}$, $\underline{\alpha}_{BR}(\boldsymbol{\alpha})$, is squared integrable thus belongs to $\mathcal{A}$.

Assumption 2. There exists a $M > 0$ such that for all $x \in [0,1]$, $\sup_{\alpha \in A(x)} \|\alpha\| \leq M$.

In order to proof the existence of a Nash equlibrium, we would like to use the Schauder's fixed point theorem: let $F : B \mapsto B$ and L be a closed and convex subset of B, if $F(L) \subset L$, F continuous, and $F(L)$ belongs to a compact subset of B. then there exists a fixed point $b \in L$ such that $b = F(b)$.

Lemma 6. Suppose that the graphon game $\mathcal{G}(A, J, W)$ satisfies assumptions 1 and 2, then it admits at least one Nash equilibrium.

Proof. Let us define the set $L_A = \{f \in L^2([0,1], \mathbb{R}^k), \|f\| \leq M\}$.

- We know that the operator $\mathbb{B}\mathbb{A}^W$ maps $L^2([0,1], \mathbb{R}^k)$ to $L^2([0,1], \mathbb{R}^k)$. And since $\mathbb{B} : L^2([0,1], \mathbb{R}^k) \to L^2([0,1], \mathbb{R}^k)$ and $\mathbb{A}^W : L^2([0,1], \mathbb{R}^k) \to L^2([0,1], \mathbb{R}^k)$ are continuous operators, so as $\mathbb{B}\mathbb{A}^W$.
- L_A is non-empty, convex, and compact (bounded and closed) set of $L^2([0,1], \mathbb{R}^k)$.
- $\mathbb{A}^W$ is a compact operator. Thus $\mathbb{A}^W L_A$ is pre-compact in $L^2([0,1], \mathbb{R}^k)$. Thus, $\mathbb{B}\overline{\mathbb{A}^W L_A}$ is compact. So $\mathbb{B}\mathbb{A}^W L_A$ belongs to a compact subset of $L^2([0,1], \mathbb{R}^d)$.
- Since for every $f \in L^2([0,1], \mathbb{R}^k)$ and every $x \in [0,1]$, $\mathbb{B}f(x) = \arg\inf_{\alpha \in A(x)} J(\alpha, f(x)) \in A(x)$, so that $\mathbb{B}f \in L_A$ under assumption 2. Thus $\mathbb{B}\mathbb{A}^W L_A \subseteq L_A$.

Therefore, we can apply Schauder's fixed point theorem on L_A with respect to the function $\mathbb{B}\mathbb{A}^W$. □

4.8.2 uniqueness result

We can relax assumption 2 by

Assumption 3. There exists $\widehat{z} \in \mathbb{R}^k$ and $M > 0$ such that for every $x \in [0,1]$ and for every $\widehat{\alpha} \in \arg\min_{\alpha' \in A(x)} J(\alpha', \widehat{z})$,

$$\|\widehat{\alpha}\| \leq M.$$

Lemma 7. Under assumption 1 and 3, we have for every $f \in L^2([0,1], \mathbb{R}^k)$, $\mathbb{B}f \in L^2([0,1], \mathbb{R}^k)$.

Proof. It is easy to see that for a constant function $g(x) = \widehat{z}$ for every $x \in [0,1]$,

$$\|\mathbb{B}g\|_{L^2}^2 = \int_{[0,1]} \|\arg\min_{\alpha' \in A(x)} J(\alpha', \widehat{z})\|^2 dx \leq M^2$$

Thus, for every $f \in L^2([0,1], \mathbb{R}^k)$, we have

$$\|\mathbb{B}f\|_{L^2} \leq \|\mathbb{B}f - \mathbb{B}g\|_{L^2} + \|\mathbb{B}g\|_{L^2} \leq \frac{l_J}{l_c}\|f - g\|_{L^2} + M < \infty.$$

□

Assumption 4. Suppose that

$$\frac{l_J}{l_c}\lambda_{max}(W) \leq 1,$$

where λ_{max} is the largest eigenvalue of W.

Lemma 8. (Uniqueness) Suppose that the graphon game $\mathcal{G}(A, J, W)$ satisfies Assumption 1, 3, and 4, then it admits a unique Nash equilibrium.

Proof. We can show that the game operator $\mathbb{BA}^W$ is a contraction, then the result follows from Banach fixed point theorem. $\square$

4.9 Example: LQ graphon game

Consider a graphon game $\mathcal{G}(A, J, W)$ such that $A(x) = [0, \infty)$ for every $x \in [0,1]$ and the cost function

$$J(\alpha, z) = \frac{1}{2}\alpha^2 - \alpha(az + b)$$

with $a \in \mathbb{R}$ and $b > 0$.

We can easily check that the function J satisfies assumption 1 with a constant $l_c = 1$ and $l_J = a$. We can also see that assumption 3 is satisfied with $\widehat{z} = 0$ and $M = b$. Moreover, because

$$\frac{\partial J(\alpha, z)}{\partial \alpha} = \alpha - (az + b).$$

So that Assumption 4 holds if

$$|a|\lambda_{\max}(W) \leq 1.$$

By Lemma 8, the graphon game admits a unique Nash equilibrium.

The best response for player $x \in [0,1]$ is given by

$$\underline{\alpha}_{BR}(\boldsymbol{\alpha})(x) = \max\{0, az(x|\boldsymbol{\alpha}) + b\}$$

where $z(x|\boldsymbol{\alpha}) = \int_{[0,1]} \boldsymbol{\alpha}(y)W(x,y)dy$.

We distinguish two cases

- If $a > 0$, the game is a game of strategic complements. Indeed, for every $x \in [0,1]$, $a > 0, b > 0, z(x|\boldsymbol{\alpha}) \geq 0$ so that $\partial_\alpha J(0, z(x|\boldsymbol{\alpha})) < 0$. Hence

$$J(\alpha, z(x|\boldsymbol{\alpha})) = J(0, z(x|\boldsymbol{\alpha})) + \partial_\alpha J(0, z(x|\boldsymbol{\alpha}))\alpha + o(\alpha) \leq J(0, z(x|\boldsymbol{\alpha})).$$

  As a consequence $\underline{\alpha}_{BR}(\boldsymbol{\alpha}) > 0$. The Nash equilibrium $\boldsymbol{\alpha}^* \in L^2([0,1], \mathbb{R})$ is given by

$$\boldsymbol{\alpha}^* = b(I + |a|A^w)1_{[0,1]}.$$

  which is completely determined by the information of the underlying graph.

- if $a < 0$, the game is a game of strategic substitutes. We may still have

$$\boldsymbol{\alpha}^* = b(I + |a|A^w)1_{[0,1]},$$

  but this is not always true. Explanation?

Remark 12. The measure used in graphon operator A^W is the Lebesgue measure, which has no atoms. It ensures that the weight put on a singleton $\{x\}$ is zero, i.e. $A^W\delta_x = 0$. We will relax this condition later in the course.

5 Game with Incomplete Information

Incomplete information is different from *Imperfect information.* A Perfect information game assumes that each player knows everything perfectly (what they are doing, the cost function, etc.). In an *Imperfect information* setting, they know everything but not perfectly, for instance, there is noise in what a player sees about the other players or about the cost function. However, in the *Incomplete information* setting, a single player does not know what others are doing. Here, we look at problems with incomplete information, the examples that we are going to investigate are games in auction.

5.1 Bayesian Games

We distinguish between information that is known *a priori* and information that is known *a posteriori.* These are two elements which could be stochastic.

Definition 22. A N player Bayesian game, denoted by $\mathcal{G}(\Omega, \Theta, A, J)$, consists of the following elements:

- Let Ω be the set of the states of the world.
- Each player $i \in \{1, \dots, N\}$ has a type $\theta_i \in \Theta_i$, where the set Θ_i is the feasible set of types corresponded to player i. We denote by $\Theta = \Theta_1 \times \cdots \times \Theta_N$.
- There exists a prior probability on $\Omega \times \Theta$, denoted by $\rho \in \mathcal{P}(\Omega \times \Theta)$.
- For each player $i \in \{1, \dots, N\}$, let A_i be the set of her feasible actions. Let $A = A_1 \times \cdots \times A_N$ denote the set of admissible strategy profiles.
- For each player $i \in \{1, \dots, N\}$, we associate her with a cost function $J^i : \Omega \times \Theta \times A \to \mathbb{R}$. The collection of all cost functions $\{J^i\}_{i=1,\dots,N}$ is denoted by J.

Remark 13.

- The state space Ω will not be presented in Auctions.
- Notice that this probability is easy to define when there is a finite number of players but when the number of players is infinite, the construction is more involved. We will need to define $\sigma-$algebra on the state space Ω and the set of types Θ.

Definition 23. For every player $i \in \{1, \dots, N\}$, a pure strategy for player i, denoted by $\underline{\alpha}_i$, is a function from Θ_i to A_i. The set of pure strategy for player i is denoted by $\mathcal{A}_i$. A pure strategy profile, denoted by $\underline{\boldsymbol{\alpha}}$, is the collection of N pure strategies among players, namely

$$\underline{\boldsymbol{\alpha}} = (\underline{\alpha}_1, \dots, \underline{\alpha}_N) \in \mathcal{A}_1 \times \dots \times \mathcal{A}_N.$$

The set of pure strategy profiles is denote by $\mathcal{A} = \mathcal{A}_1 \times \dots \times \mathcal{A}_N$.

The Bayesian cost associated to player i is defined as a real valued function on $\mathcal{A}$, denoted by $J^i : \mathcal{A} \to \mathbb{R}$, such that

$$\widehat{J}^i(\underline{\boldsymbol{\alpha}}) = \int_\Omega \int_\Theta J^i(w, \boldsymbol{\theta}, \underline{\boldsymbol{\alpha}}(\theta))\rho(d\omega, d\theta)$$

where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$ and $\underline{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = (\underline{\alpha}_1(\theta_1), \ldots, \underline{\alpha}_N(\theta_N))$.

Remark 14. In the above definition of Bayesian cost associated to player i, we see that her action depends only on her own type θ_i. This is called distributed actions.

Besides, in some cases, the Bayesian cost for player i is defined with respect to the conditional distribution

$$\rho(dw, d\theta^{-i}|\theta_i).$$

The conditional cost is integrated over $\Omega \times \Theta^{-i}$ and denoted by $J^i(\boldsymbol{\alpha}|\theta^i)$.

Definition 24. In the pure strategy setting, a pure strategy profile $\underline{\boldsymbol{\alpha}}^* \in \mathcal{A}$ is said to be a Bayesian Nash equilibrium (BNE for short) of the Bayesian game $\mathcal{G}(\Omega, \Theta, A, J)$ if for every $i \in \{1, \ldots, N\}$, for every $\underline{\alpha}_i \in \mathcal{A}_i$, we have

$$\widehat{J}^i(\boldsymbol{\alpha}^*) \geq \widehat{J}^i(\underline{\alpha}_i, \underline{\boldsymbol{\alpha}}^*_{-i})$$

where $(\underline{\alpha}_i, \underline{\boldsymbol{\alpha}}^*_{-i}) = (\underline{\alpha}^*_1, \ldots, \underline{\alpha}^*_{i-1}, \underline{\alpha}_i, \underline{\alpha}^*_{i+1}, \ldots, \underline{\alpha}^*_N)$.

5.2 Auctions

Let us now turn to auctions. We consider N players (bidders) who want to buy some object. We consider here a one round sealed-bid auction in which players propose their bids individually and simultaneously (sealed-bids). Depending on all these bids, only one player, called her winner, will get the object and she will have to pay a certain amount of money in return for the object.

We model this one round sealed-bid auction with a N player Bayesian game $\mathcal{G}(\Theta, A, J, \rho)$. More precisely,

- there is no state space;
- the type of player i is her valuation of the object, denoted by v_i. We assume that there is a maximum value $\bar{v}$ so that $\theta_i \in [0, \bar{v}]$ for each i;
- the action of player i is her bid for the object , denoted by α_i. The set of feasible actions for player i is all positive bid, i.e. $A_i = [0, \infty)$. We assume that all bids are i.i.d. with distribution ρ on $[0, \infty)$.

Remark 15. The common distribution ρ on $\mathbb{R}$ reflects an assumption that all players are symmetric in the auction.

The next question to define the auction mechanism. We need to precise the way how the winner is chosen as well as the amount of money she is required to pay for. There are many types of auctions.

The most two popular types are the English auction and the Dutch auction. The English auction is an multi-round ascending auction. It starts from a very low price proposed by the auction organizer. At each round, buyers accept the proposed price remain in the auction and go into the next round. The auction organize increase the proposed price progressively and the last buyer left in the auction wins. The Dutch auction is a multi-rounds descending auction. It starts from a very high price and the proposed price keeps on descending. The first buyer who accept the price wins the auction.

Here, we only focus on a one-round sealed-bid auction. There are two cases, called the first price auction and the second price auction. In both cases, the player whose bid is the highest among all players wins the auction. The winner needs to pay either the highest bid among all players including herself (first price auction), or to pay the highest bid among all bids proposed players except herself (second price auction) in return for the object. The second price auction is called the Vickrey auction.

If ties occur, i.e., if there are multiple players propose simultaneously the highest bid, the winner is chosen among them uniformly at random. And in this case, the winner of the first price auction or the second price auction has to pay the highest bid.

In the first price auction, the cost function for player $i \in \{1, \dots, N\}$ is given by the function $J^i : \Theta \times A \to \mathbb{R}$ such that

$$J^i(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \begin{cases} 0 & if\, \alpha_i < \max_{j \neq i} \alpha_j \\ \theta_i - \alpha_i & if\, \alpha_i > \max_{j \neq i} \alpha_j \\ & \\ (\theta_i - \alpha_i) \cdot Bern\left(\dfrac{1}{|\{j|\alpha_j = a_i\}|}\right) & if\, \alpha_i = \max_{j \neq i} \alpha_j \end{cases}$$

where $Bern(p)$ is a Bernoulli random variable in $\{0, 1\}$ with $\mathbb{P}(Bern(p) = 1) = p$.

In the second price auction, the cost function for player $i \in \{1, \dots, N\}$ turns out to be

$$J^i(\boldsymbol{\theta}, \boldsymbol{\alpha}) = \begin{cases} 0 & if\, \alpha_i < \max_{j \neq i} \alpha_j \\ \theta_i - \max_{j \neq i} \alpha_j & if\, \alpha_i > \max_{j \neq i} \alpha_j \\ & \\ (\theta_i - \max\limits_{j \neq i} \alpha_j) \cdot Bern\left(\dfrac{1}{\{j|\alpha_j = a_i\}|}\right) & if\, \alpha_i = \max_{j \neq i} \alpha_j \end{cases}$$

Players intend to maximize their individual cost independently.

Definition 25. For a single player $i \in \{1, \dots, N\}$, a pure strategy $\underline{\alpha}_i \in \mathcal{A}_i$ is said to be a dominant strategy if for every $\underline{\alpha}_j \in \mathcal{A}_j$ with $j \neq i$, and for every $\underline{\tilde{\alpha}}_i \in \mathcal{A}_i$, we have

$$\widehat{J}^i(\underline{\alpha}_i, \underline{\boldsymbol{\alpha}}_{-i}) \geq \widehat{J}^i(\underline{\tilde{\alpha}}_i, \underline{\boldsymbol{\alpha}}_{-i}).$$

In words, no matter what other players are doing, the Bayesian cost incurred by taking the pure strategy $\underline{\alpha}_i$ for player i is no worse than the Bayesian cost when she picks another pure strategy $\tilde{\underline{\alpha}}_i$.

Assumption 5. We assume that all players know their own evaluation of the object, i.e. player i knows her own type $\theta_i \in [0, \bar{v}]$.

Remark 16. Under assumption 5, the prior distribution for all players for their types becomes $\prod_{i=1}^{N} \delta_{\theta_i}$. And the Bayesian cost associated to player i takes the form

$$\widehat{J}^i(\underline{\boldsymbol{\alpha}}) = J^i(\theta, \underline{\boldsymbol{\alpha}}(\theta)), \forall\, \underline{\alpha} \in \mathcal{A}$$

where $\theta = (\theta_1, \ldots, \theta_N)$ is prescribed.

Theorem 3. Under Assumption 5, in a second price auction $\mathcal{G}(\Theta, A, J, \rho)$, for any player $i \in \{1, \ldots, N\}$, the pure strategy $\underline{\alpha}_i^* \in \mathcal{A}_i$ satisfies that

$$\underline{\alpha}_i^*(\theta_i) = \theta_i, \qquad \forall \theta_i \in \Theta_i$$

is a dominant strategy for player i.

In words, if player i wants to minimize her own cost without knowing how other players behave in the auction, she should play truthfully according to her own evaluation of the object

Corollary 1. The pure strategy profile $\underline{\boldsymbol{\alpha}}^* = (\underline{\alpha}_1^*, \ldots, \underline{\alpha}_N^*)$ satisfying

$$\underline{\alpha}_i^*(\theta_i) = \theta_i, \qquad \forall \theta_i \in \Theta_i$$

for every $i \in \{1, \ldots, N\}$ is a Bayesian Nash equilibrium.

Proof. (Proof of Theorem 3) We idea of the proof relies on the fact that in second price auction, the pure strategy of player i only affects whether or not she wins the auction, but the cost associated to player i does not depend on her own strategy.

Suppose that all other players except player i chose pure strategies $\widehat{\underline{\alpha}}_j \in \mathcal{A}_j$ for $j \neq i$. Under Assumption 5, all players know their own type. Let us denote the highest bid for all other players by

$$\widehat{\beta} := \max_{j \neq i} \widehat{\underline{\alpha}}_j(\theta_j).$$

If player i chooses strategy $\underline{\alpha}_i \in \mathcal{A}_i$, the Bayesian cost associated to player i becomes

$$\widehat{J}^i(\underline{\alpha}_i, \widehat{\underline{\boldsymbol{\alpha}}}_{-i}) = \begin{cases} 0 & if\, \underline{\alpha}_i(\theta_i) < \widehat{\beta} \\ \theta_i - \widehat{\beta} & if\, \underline{\alpha}_i(\theta_i) > \widehat{\beta} \\ \\ (\theta_i - \widehat{\beta}) \cdot Bern\left(\dfrac{1}{|\{j \neq i : \widehat{\underline{\alpha}}_j(\theta_j) = \widehat{\beta}\}| + 1}\right) & if\, \underline{\alpha}_i(\theta_i) = \widehat{\beta} \end{cases}$$

We suppose that player i choose an alternative strategy $\widehat{\underline{\alpha}}_i \in \mathcal{A}_i$ than $\underline{\alpha}_i^*$. Thus, her bid, denoted by $\beta_i = \widehat{\underline{\alpha}}_i(\theta_i) \in [0, \infty)$, is different from $\theta_i = \underline{\alpha}_i^*(\theta_i)$. There are two situations:

- When $\beta_i < \theta_i$.
  - if $\widehat{\beta} > \theta_i > \beta_i$ or $\widehat{\beta} < \beta_i < \theta_i$, player i has the same Bayesian cost for both pure strategy $\underline{\widehat{\alpha}}_i$ and $\underline{\alpha}_i^*$;
  - if $\beta_i < \widehat{\beta} < \theta_i$, then the Bayesian cost for employing strategy $\underline{\alpha}_i^*$ is $\theta_i - \widehat{\beta}$, whilst the Bayesian cost incurred with the pure strategy $\underline{\widehat{\alpha}}_i$ is 0. Thus

$$\widehat{J}^i(\underline{\alpha}_i^*, \underline{\widehat{\boldsymbol{\alpha}}}_{-i}) = \theta_i - \widehat{\beta} > 0 = \widehat{J}^i(\underline{\widehat{\alpha}}_i, \underline{\widehat{\boldsymbol{\alpha}}}_{-i});$$

  - if $\widehat{\beta} = \theta_i > \beta_i$, then

$$\widehat{J}^i(\underline{\alpha}_i^*, \underline{\widehat{\boldsymbol{\alpha}}}_{-i}) = (\theta_i - \widehat{\beta}) \cdot Bern\left(\frac{1}{|\{j \neq i : \underline{\widehat{\alpha}}_j(\theta_j) = \widehat{\beta}\}|}\right) = 0 = \widehat{J}^i(\underline{\widehat{\alpha}}_i, \underline{\widehat{\boldsymbol{\alpha}}}_{-i}).$$

  - if $\widehat{\beta} = \beta_i < \theta_i$, then

$$\widehat{J}^i(\underline{\alpha}_i^*, \underline{\widehat{\boldsymbol{\alpha}}}_{-i}) = \theta_i - \widehat{\beta} \geq (\theta_i - \widehat{\beta}) \cdot Bern\left(\frac{1}{|\{j \neq i : \underline{\widehat{\alpha}}_j(\theta_j) = \widehat{\beta}\}|}\right) = \widehat{J}^i(\underline{\widehat{\alpha}}_i, \underline{\widehat{\boldsymbol{\alpha}}}_{-i}).$$

    Because the Bernoulli random variable only takes value in $\{0, 1\}$.

- When $\beta_i > \theta_i$. Idem.

Therefore, we conclude that for any pure strategies $\underline{\widehat{\alpha}}_j \in \mathcal{A}_j$ for $j = 1, \ldots, N$, we have

$$\widehat{J}^i(\underline{\alpha}_i^*, \underline{\widehat{\boldsymbol{\alpha}}}_{-i}) \geq \widehat{J}^i(\underline{\widehat{\alpha}}_i, \underline{\widehat{\boldsymbol{\alpha}}}_{-i}).$$

□

5.3 symmetric auction

Definition 26. An auction is said to be symmetric, denote by $\mathcal{G}(\Theta_0, A_0, J, \rho)$, if

- all players share a common type space $\Theta_0 = [0, \overline{v}]$, and the set of all type profile is still denote by $\Theta = (\Theta_0)^N$;
- the type of an individual player is independently distributed according to a common prior distribution $\rho \in \mathcal{P}([0, \overline{v}])$, and the prior distribution for all N players is the product distribution, denoted by $\rho^N := \rho \times \ldots \times \rho$;
- all players share a common feasible action space $A_0 = [0, \infty)$, and the set of all admissible strategy profile is still denote by $A := (A_0)^N$;
- all players follow the same common pure strategy, namely for any admissible strategy profile $\underline{\boldsymbol{\alpha}} = (\underline{\alpha}_1, \ldots, \underline{\alpha}_N) \in \mathcal{A}$, there exists an pure strategy $\underline{\alpha} : \Theta_0 \to A_0$ such that $\underline{\alpha}_i = \underline{\alpha}$ for all $i \in \{1, \ldots, N\}$. The set of all admissible common pure strategies is denote by $\mathcal{A}_0$.

- players can have different cost function $J^i : \Theta \times A \ni (\boldsymbol{\theta}, \boldsymbol{\alpha}) \mapsto J^i(\boldsymbol{\theta}, \boldsymbol{\alpha}) \in \mathbb{R}$.

A pure strategy for player i, say $\underline{\alpha}_i \in \mathcal{A}_0$, is continuous and increasing if $\underline{\alpha}_i$ is continuous in $(0, \bar{v})$, and left and right continuous on $\overline{v}$ and 0, i.e.$\lim_{\theta \to 0} \underline{\alpha}(\theta) = \underline{\alpha}(0)$ and $\lim_{\theta \to \bar{v}} \underline{\alpha}(\theta) = \underline{\alpha}(\bar{v})$, and it is also increasing on $[0, \bar{v}]$.

Assumption 6. We assume that the support of $\rho \in \mathcal{P}([0, \overline{v}])$ is $[0, \overline{v}]$, namely there is no open interval $(a, b) \subseteq [0, \overline{v}]$ such that $\rho((a, b)) = 0$.

The prior distribution of N players is denoted by $\rho^N(d\boldsymbol{\theta}) = \rho(d\theta_1) \times \ldots \times \rho(d\theta_N)$ where $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$.

In the symmetric auction, recalled that the Bayesian cost function associated to player i evaluated at a strategy profile $\underline{\boldsymbol{\alpha}} \in \mathcal{A}$ can be written as

$$\begin{aligned} \widehat{J}^i(\underline{\boldsymbol{\alpha}}) &= \int_\Theta J^i(\boldsymbol{\theta}, \underline{\boldsymbol{\alpha}}(\boldsymbol{\theta}))\rho^N(d\boldsymbol{\theta}) \\ &= \int_{[0,\bar{v}]} \left(\int_{\Theta_{-i}} J^i((\theta_i, \boldsymbol{\theta}_{-i}),\ (\underline{\alpha}(\theta_1), \ldots, \underline{\alpha}(\theta_N))) \prod_{j \neq i} \rho(d\theta_j) \right) \rho(d\theta_i) \end{aligned} \tag{5.3.1}$$

where $\underline{\alpha}$ is a the common pure strategy among players, and the second equality is justified by the fact that prior distribution ρ^N for all players is a production distribution.

Definition 27. The conditional Bayesian cost for player i with type θ_i and pure strategy $\underline{\beta} \in \mathcal{A}_0$ is defined as a function from $\mathcal{A}$ to $\mathbb{R}$ such that

$$\widehat{J}^i(\underline{\beta}, \underline{\boldsymbol{\alpha}}_{-i}|\theta_i) = \int_{\Theta_{-i}} J^i((\underline{\beta}(\theta_i), \boldsymbol{\theta}_{-i}),\ (\underline{\beta}(\theta_i), \underline{\boldsymbol{\alpha}}_{-i}(\boldsymbol{\theta}_{-i})))\rho^N(d\boldsymbol{\theta}_{-i}|\theta_i)$$

where $\rho^N(d\boldsymbol{\theta}_{-i}|\theta_i) = \prod_{j \neq i} \rho(d\theta_j)$.

Definition 28. We say that a pure strategy profile $\underline{\boldsymbol{\alpha}} \in \mathcal{A}$ with a common pure strategy $\underline{\alpha} : \Theta_0 \to A_0$ is a Bayesian Nash equilibrium for a symmetric auction if and only if for every $i \in \{1, \ldots, N\}$ and for every $\theta_i \in \Theta_0$, we have

$$\underline{\alpha}(\theta_i) \in \arg\max_{\beta \in A_0} \widehat{J}^i(\beta, \underline{\boldsymbol{\alpha}}_{-i}(\cdot)|\theta_i) \tag{5.3.2}$$

where β is a constant pure strategy equaled to $\beta \in A_0$ for all $\theta_i \in \Theta_0$, and $(\beta, \underline{\boldsymbol{\alpha}}_{-i}(\boldsymbol{\theta}_{-i})) = (\underline{\alpha}(\theta_1), \ldots, \underline{\alpha}(\theta_{i-1}), \beta, \underline{\alpha}(\theta_{i+1}), \ldots, \underline{\alpha}(\theta_N))$ for any $\boldsymbol{\theta}_{-i} \in (\Theta_0)^{N-1}$.

5.3.1 second price auction

Assume that there is no tie between players, and the cost function for player i is defined as

$$J^i(\boldsymbol{\theta}, \boldsymbol{\alpha}) = (\theta_i - \max_{j \neq i} \alpha_j)\mathbb{1}_{\alpha_i > \max_{j \neq i} \alpha_j}.$$

Proposition 9. In a symmetric auction, under Assumption 6, the admissible pure strategy profile $\underline{\boldsymbol{\alpha}}^* \in \mathcal{A}$ defined by

$$\underline{\alpha}_i^*(\theta) = \underline{\alpha}^*(\theta) = \theta, \quad \forall \theta \in [0, \infty),\, i \in \{1, \ldots, N\}$$

is the unique symmetric continuous increasing Bayesian Nash equilibrium, where $\underline{\alpha}^* : [0, \bar{v}] \ni \theta \mapsto \underline{\alpha}^*(\theta) = \theta \in [0, \infty)$ is the common pure strategy mapping for all players.

Proof. Let $\widehat{\beta} : [0, \bar{v}] \to [0, \infty)$ be the common pure strategy mapping related to a symmetric continuous increasing Bayesian Nash equilibrium. We will show that $\widehat{\beta} = \underline{\alpha}^*$.

According to the second price auction payoff, the conditional Bayesian cost for player i with type θ_i takes the form

$$\widehat{J}^i(\widehat{\beta}(\cdot), \widehat{\boldsymbol{\beta}}_{-1}(\cdot))|\theta_i) = \int_{\Theta_{-i}} \left(\theta_i - \max_{j \neq i} \widehat{\beta}(\theta_j)\right) \mathbb{1}_{\{\widehat{\beta}(\theta_i) > \max_{j \neq i} \widehat{\beta}(\theta_j)\}} \rho^N(d\boldsymbol{\theta}_{-i}|\theta_i)$$

and

$$\widehat{J}^i(\underline{\alpha}^*(\cdot), \widehat{\boldsymbol{\beta}}_{-1}(\cdot))|\theta_i) = \int_{\Theta_{-i}} \left(\theta_i - \max_{j \neq i} \widehat{\beta}(\theta_j)\right) \mathbb{1}_{\{\theta_i > \max_{j \neq i} \widehat{\beta}(\theta_j)\}} \rho^N(d\boldsymbol{\theta}_{-i}|\theta_i)$$

Since $\widehat{\beta}$ is the common pure strategy for a Bayesian Nash equilibrium, by definition, we must have

$$\widehat{J}^i(\widehat{\beta}(\theta_i), \widehat{\boldsymbol{\beta}}_{-i}(\cdot)|\theta_i) \geq \widehat{J}^i(\underline{\alpha}^*(\theta_i), \widehat{\boldsymbol{\beta}}_{-i}(\cdot)|\theta_i)$$

where $(\beta, \widehat{\boldsymbol{\beta}}_{-i}(\boldsymbol{\theta}_{-i})) = (\widehat{\beta}(\theta_1), \ldots, \widehat{\beta}(\theta_{i-1}), \beta, \widehat{\beta}(\theta_{i+1}), \ldots, \widehat{\beta}(\theta_N))$ for any $\boldsymbol{\theta}_{-i} \in (\Theta_0)^{N-1}$ and $\beta \in A_0$.

Step 1: we show that for every $\theta \in [0, \overline{v}]$, $\widehat{\beta}(\theta) \leq \theta$. Otherwise, there must exist a $\theta \in [0, \overline{v}]$ such that $\widehat{\beta}(\theta) > \theta$. If $\widehat{\beta}(0) > 0$, then by the left continuity at point 0, there must exist a $\theta > 0$ such that $\widehat{\beta}(\theta) > \theta$. Similarly, if $\widehat{\beta}(\bar{v}) > \bar{v}$, then there exists an $\theta < \bar{v}$ such that $\widehat{\beta}(\theta) > \theta$. Thus, we can always pick an $\theta \in (0, \overline{v}]$ satisfying $\widehat{\beta}(\theta) > \theta$.

For any fixed player $i \in \{1, \ldots, N\}$, suppose that her type equals to $\theta_i = \theta$.

Since $\widehat{\beta}(\theta) > \theta$, we know that for any $\boldsymbol{\theta}_{-i} \in (\Theta_0)^{N-1}$ such that $\max_{j \neq i} \widehat{\beta}(\theta_j) > \widehat{\beta}(\theta)$ or $\max_{j \neq i} \widehat{\beta}(\theta_j) < \theta$, the conditional Bayesian costs for player i with pure strategy $\widehat{\beta}$ and $\underline{\alpha}^*$ have the same value. Otherwise if $\boldsymbol{\theta}_{-i} \in (\Theta_0)^{N-1}$ such that

$$\theta < \max_{j \neq i} \widehat{\beta}(\theta_j) < \widehat{\beta}(\theta),$$

then

$$\theta - \max_{j \neq i} \widehat{\beta}(\theta_j) < 0 = \left(\theta - \max_{j \neq i} \widehat{\beta}(\theta_j)\right) \mathbb{1}_{\{\theta > \max_{j \neq i} \widehat{\beta}(\theta_j)\}}$$

Hence, if the set $\{\boldsymbol{\theta}_{-i} | \theta < \max_{j \neq i} \widehat{\beta}(\theta_j) < \widehat{\beta}(\theta)\}$ has a positive measure, then

$$\widehat{J}^i(\widehat{\beta}(\theta), \widehat{\boldsymbol{\beta}}_{-i}(\cdot)|\theta) < \widehat{J}^i(\underline{\alpha}^*(\theta), \widehat{\boldsymbol{\beta}}_{-i}(\cdot)|\theta),$$

which contradicts to the fact that $\widehat{\beta}$ is a Bayesian Nash equilibrium.

Now, we construct explicitly a subset $B_i \subseteq \{\boldsymbol{\theta}_{-i} | \theta < \max_{j\neq i} \widehat{\beta}(\theta_j) < \widehat{\beta}(\theta)\}$ with a positive measure if such a θ satisfying $\widehat{\beta}(\theta) > \theta$ exists.

By continuity of $\widehat{\beta}$ on $(0, \bar{v})$, there exists a $t \in (0, \theta)$ such that $\widehat{\beta}(t) > \theta$. For any fixed $i \in \{1, \dots, N\}$, we assume that the type of player i is $\theta_i = \theta$. under assumption 6, we know that

$$\mathbb{P}(\{\theta_j \in (t, \theta) : j \neq i\}) = \prod_{j\neq i} \rho((t, \theta)) > 0.$$

Let us denote by

$$B_i = \{\boldsymbol{\theta}_{-i} \in (\Theta_0)^{N-1} | \, \theta_j \in (t, \theta), \, \forall \, j \neq i\}.$$

For any $\boldsymbol{\theta}_{-i} \in B_i$, since $\widehat{\beta}$ is increasing, we have $\widehat{\beta}(\theta) > \widehat{\beta}(\theta_j)$ for all $j \neq i$. (no tie bids between players). This implies that

$$\widehat{\beta}(\theta) > \max_{j\neq i} \widehat{\beta}(\theta_j).$$

Moreover, since $\theta_j > t$, we have $\widehat{\beta}(\theta_j) > \widehat{\beta}(t) > \theta$. Thus

$$\max_{j\neq i} \widehat{\beta}(\theta_j) > \theta.$$

By consequence, the set B_i with positive measure is a subset of $\{\boldsymbol{\theta}_{-i} | \theta < \max_{j\neq i} \widehat{\beta}(\theta_j) < \widehat{\beta}(\theta)\}$. Together with the previous discussion, we show that for all $\theta \in [0, \overline{v}]$, we have $\widehat{\beta}(\theta) \leq \theta$.

Step 2: we show that for every $\theta \in [0, \overline{v}]$, $\widehat{\beta}(\theta) \geq 0$. Otherwise, there exists a $\theta \in (0, \bar{v})$ such that $\widehat{\beta}(\theta) < \theta$. Then, there exists $t \in (\theta, \bar{v})$ such that $\widehat{\beta}(t) < \theta$.

Fixed $i \in \{1, \dots, N\}$ and consider that player i has a type $\theta_i = \theta$.

Now consider the subset $B_i = \{\boldsymbol{\theta}_{-i} \in (\Theta_0)^{N-1} | \theta_j \in (\theta, t), \, j \neq i\}$. Then $\mathbb{P}(B_i) > 0$. By monotonicity argument, we can have that for all $j \neq i$:

$$\widehat{\beta}(\theta) \leq \widehat{\beta}(\theta_j),$$

thus

$$\widehat{\beta}(\theta) < \max_{j\neq i} \widehat{\beta}(\theta_j).$$

Meanwhile, for any $\boldsymbol{\theta}_{-i} \in B_i$, we know that $\theta_j < t$ for all $j \neq i$. Then, we have

$$\widehat{\beta}(\theta_j) \leq \widehat{\beta}(t) < \theta,$$

which implies

$$\max_{j\neq i} \widehat{\beta}(\theta_j) < \theta.$$

Thus, with truthful bidding (using strategy $\underline{\alpha}^*$), the player i will certainly win the auction with positive profit $\theta - \max_{j\neq i} \widehat{\beta}(\theta_j) > 0$. This contradicts to $\widehat{\beta}$ is a Bayesian Nash equilibrium. $\square$

5.3.2 First price symmetric auction

Let us denote by $F : [0, \overline{v}] \to [0, 1]$ the c.d.f. of the distribution $\rho \in \mathcal{P}([0, \overline{v}])$. Since players are symmetric, we assume that player 1 wins the auction. If all admissible pure strategy profile are symmetric, continuous and increasing, then for any $\underline{\alpha} \in \mathcal{A}_0$ and $\boldsymbol{\theta} \in (\Theta_0)^N$ such that $\underline{\boldsymbol{\alpha}}(\boldsymbol{\theta}) = (\underline{\alpha}(\theta_1), \ldots, \underline{\alpha}(\theta_N)) \in A$, we have

$$\underline{\alpha}(\theta_1) > \max_{j \neq i} \underline{\alpha}(\theta_j) \Longleftrightarrow \theta_1 > \max_{j \neq 1} \theta_j.$$

Let denote the highest type of all other players ($j \neq 1$) by

$$Z = \max_{j \neq 1} \theta_j.$$

The c.d.f. of the random variable Z is denote by $G : \Theta_0 \to [0, 1]$. We know that

$$G(z) = \mathbb{P}(Z \leq z) = \mathbb{P}(\theta_j \leq z,\, j \neq 1) = F(z)^{N-1}.$$

We define the conditional expectation of Z on the event $\{Z < \theta\}$ for some $\theta \in [0, \overline{v}]$ by

$$\mathbb{E}[Z|Z < \theta] = \frac{1}{G(\theta)} \int_0^\theta zG(dz).$$

Definition 29. The payment from player i to the auctioneer when the type profile is $\boldsymbol{\theta} \in \Theta$ is denoted by $p_i(\boldsymbol{\theta})$.

Suppose that the player i know her own type θ_i and she believes that the other players types is drawn independently from a distribution ρ, then the expected payment for player i is denote by $\mathbb{E}[p_i(\boldsymbol{\theta})|\theta_i]$.

In the second price auction, the expected payment to the auctioneer by the winner at Bayesian equilibrium equals to the probability that she wins the auction times the conditional expectation of second highest bid. If the auction is symmetric and with increasing and continuous admissible pure strategies, and let us say player 1 wins the auction, then

$$\begin{aligned}
\mathbb{E}[p_1(\boldsymbol{\theta})|\theta_1] &= \mathbb{P}\left(\underline{\alpha}^*(\theta_1) > \max_{j \neq 1} \underline{\alpha}^*(\theta_j)\right) \cdot \mathbb{E}\left[\max_{j \neq 1} \underline{\alpha}^*(\theta_j) | \max_{j \neq 1} \underline{\alpha}^*(\theta_j) < \underline{\alpha}^*(\theta_1)\right] \\
&= \mathbb{P}(\theta_1 > Z)\mathbb{E}[Z|Z < \theta_1] \\
&= G(\theta_1)\mathbb{E}[Z|Z < \theta_1]
\end{aligned}$$

where the second equality is justified by the truthful bidding in second price auction.

Now let us look at the first price auction. The cost function for player i takes the form

$$J^i(\boldsymbol{\theta}, \boldsymbol{\alpha}) = (\theta_i - \alpha_i)\mathbb{1}_{\alpha_i \geq \max_{j \neq i} \alpha_j}.$$

for all $\boldsymbol{\theta} \in (\Theta_0)^N$ and $\boldsymbol{\alpha} \in A$.

Proposition 10. If $\mathcal{G}(\Theta_0, A_0, J, \rho)$ is a symmetric first price auction, and if any admissible pure strategy $\underline{\alpha} \in \mathcal{A}_0$ satisfies $\underline{\alpha}(0) = 0$, then

$$\underline{\alpha}^*(\theta) = \mathbb{E}[Z|Z < \theta]$$

is the unique symmetric continuous increasing Bayesian Nash equilibrium.

Proof. Assume that player 1 plays with a bid $b \in A_0$. If she wins the auction, then

$$b > \max_{j \neq 1} \underline{\alpha}^*(\theta_j) = \underline{\alpha}^*(\max_{j \neq 1} \theta_j)$$

where the second equality comes from the monotonicity of pure strategy $\underline{\alpha}^*$.

Since an increasing function is invertible, let us denote by

$$\theta^* := \inf_{\theta \in [0, \overline{v}]} (\underline{\alpha}^*)^{-1}(b)$$

Then, we can have

$$\max_{j \neq 1} \theta_j < \theta^*.$$

more precision should be made for the definition of θ^* when the pure strategy is not strictly increasing

Then for any type $\theta_1 \in \Theta_0$,

$$\begin{aligned} \mathbb{P}(\text{player 1 wins}|\theta_1) &= \mathbb{E}\left[\mathbb{1}_{\{\underline{\alpha}^*(\theta_1) > \max_{j \neq 1} \underline{\alpha}^*(\theta_j))\}}\right] \\ &= \mathbb{P}(\theta_1 \geq \theta^*) \qquad (5.3.3) \\ &= G(\theta^*) \qquad (5.3.4) \end{aligned}$$

Since $\underline{\alpha}^*(\theta_1)$ is the best response for player 1, so that with this bid, player 1 should maximizes her expected winning. For any given θ_1, and any other alternative bid b, we have

$$\mathbb{E}[her\ winning|\theta_1] = \widehat{J}^1(\underline{\boldsymbol{\alpha}}^*|\theta_1) = (\theta_1 - b) \cdot \mathbb{P}(\text{ player 1 wins}|\theta_1) = (\theta_1 - b)G((\underline{\alpha}^*)^{-1}(b)).$$

By taking the derivative w.r.t. b, the derivative equals to 0 when $\theta^* = \theta_1$ (FOC), namely

$$-G(\theta_1) + (\theta_1 - b)\frac{G'(\theta_1)}{(\underline{\alpha}^*)'(\theta_1)} = 0.$$

where $G'(z) = \frac{dG(z)}{dz}$ and $(\underline{\alpha}^*)'(\theta) = \frac{d\underline{\alpha}^*(\theta)}{d\theta}$.

Then for every $\theta_1 \in \Theta_0$, we can have

$$\begin{aligned} 0 &= -\left[G(\theta_1) \cdot (\underline{\alpha}^*)'(\theta_1) + G(\theta_1)' \cdot (\alpha^*)'(\theta_1)\right] + \theta_1 G'(\theta_1) \\ &= -\left(G\underline{\alpha}^*\right)'(\theta_1) + \theta_1 \cdot G'(\theta_1) \end{aligned}$$

If $\underline{\alpha}^*(0) = 0$, then we can have for every $\theta \in \Theta_0$,

$$-G(\theta)\underline{\alpha}^*(\theta) + \int_0^\theta zG(z)dz = 0.$$

In conclusion, the pure strategy corresponded to a Bayesian Nash equilibrium takes the form

$$\underline{\alpha}^*(\theta) = \frac{1}{G(\theta)} \int_0^{\theta} zG(z)dz = \mathbb{E}[Z|Z < \theta].$$

$\square$

Remark 17. If player i wins by bidding $\underline{\alpha}^*(\theta_1)$, then the expected payment from player 1 to the auctioneer becomes

$$\mathbb{E}[p_1(\boldsymbol{\theta})|\theta_1] = \underline{\alpha}^*(\theta_1)G(\theta_1)$$

where the first term corresponded to the amount paid to the auctioneer, and the second term corresponded to the probability that player 1 wins the auction.

Theorem 4 (Revenue Equivalence). For any sealed-bid auction where the object goes to the highest bidder. If the values are i.i.d., and if the players are risk neutral (each maximizes her own cost function). any symmetric strategy which is increasing gives the same expected payment to the auctioneer.

Remark 18. For the theorem of Revenue Equivalence, we need to assume that $0-$valued players pay nothing, i.e. for any pure strategy $\underline{\alpha} : [0, \overline{v}] \to A_0$, we have $\underline{\alpha}(0) = 0$.

6 Model for Games with a continuum of players

Definition 30.

- Player: $(I, \mathcal{J}, \nu)$ is a probability space where I represents the set of players, $\mathcal{J}$ is a σ-algebra of coalitions, and ν is a probability distribution that models the weights put on coalitions;
- E is a Banach space (for example $\mathcal{C}_0([0,1])$, $\mathbb{R}^k$, etc);
- there exists a function $A : I \to 2^E$ such that for $\nu-$almost every $i \in I$, $A(i)$ is the set of feasible actions for player i.
- $(\Omega, \mathcal{F}, \mathbb{P})$ is a probability space where Ω is the set of state of nature, $\mathcal{F}$ represents the events, and $\mathbb{P}$ is a probability distribution.

We define the set of admissible controls by

$$L_A = \{\alpha \in L^1((I, \mathcal{J}, \nu), E) |\, \alpha(i) \in A(i) \quad \nu - a.e.\, i \in I\}.$$

We define the "Bochner integral" on the Banach space E of an admissible control $\underline{\alpha} \in L_A$ by

$$\mathbb{E}[\alpha] = \int_\Omega \alpha(i) d\nu(i) \in E.$$

Definition 31. We define a preference function $\Pi : I \times \Omega \times L_A \ni (i, \omega, \alpha) \mapsto \Pi(i, \omega, \alpha) \in 2^E$ such that

$$\Pi(i, \omega, \alpha) \subset A(i), \quad \text{for } \nu - a.e.\, i \in I \text{ and for } \mathbb{P} - a.e.\, \omega \in \Omega.$$

The set $\Pi(i, \omega, \alpha)$ represents the set of all actions that are feasible for player i and preferred to the actions $\alpha(i) \in A(i)$, given that the state of the world is $\omega \in \Omega$ and the other players $j \neq i$ who play with actions $(\alpha(j))_{j \neq i}$.

We assume that players observe the state of nature $\omega \in \Omega$, and they decide to take actions according to a decision rule:

$$\underline{\alpha} : \Omega \ni \omega \mapsto \underline{\alpha}(\omega) \in L_A.$$

Definition 32. A decision rule $\underline{\alpha}^*$ is a (random) Cournot-Nash equilibrium (CNE for short) if for $\mathbb{P}$-a.e. $\omega \in \Omega$ and $\nu-$a.e. $i \in I$

$$\Pi(i, \omega, \underline{\alpha}^*(\omega)) = \emptyset.$$

Example: If we have associate to player i a cost function $J^i : \Omega \times A(i) \times L_A(\omega, \alpha_i, \alpha) \mapsto J^i(\omega, \alpha_i, \alpha) \to \mathbb{R}$. Then $\underline{\alpha}^*$ is a CNE if for $\mathbb{P}$-a.e. $\omega \in \Omega$ and $\nu-$a.e. $i \in I$, we have for all $\alpha_i \in A(i)$,

$$J^i(\omega, \underline{\alpha}^*(\omega)(i), \underline{\alpha}^*(\omega)) \leq J^i(\omega, \alpha_i, \underline{\alpha}^*(\omega)).$$

And in this case, for any decision rule $\underline{\alpha} : \Omega \to L_A$, we have for ν-a.e. $i \in I$ and $\mathbb{P}$-a.e. $\omega \in \Omega$,

$$\Pi(i, \omega, \underline{\alpha}) = \left\{ \alpha_i \in A(i) | \, J^i(\omega, \alpha_i, \underline{\alpha}) < J^i(\omega, \underline{\alpha}(\omega)(i), \underline{\alpha}(\omega)) \right\}.$$

Remark 19. For a continuous measure ν, if $\forall i \in I$, $\nu(\{i\}) = 0$, then I cannot be countable.

Definition of an atom: a set A is said to an atom if $A \in \mathcal{J}$, $\nu(A) > 0$, and for any $B \subseteq A$ and $B \in \mathcal{J}$, we have

$$\nu(B) = 0 \quad or \quad \nu(B) = \nu(A).$$

Remark 20. Case of finite games: $I = \{1, \dots, N\}$. We have $\nu(A) = \frac{1}{N}|A|$ for any $A \subset I$, and

$$\int_I f(x)dx = \frac{1}{N} \sum_{i=1}^{N} f(i).$$

Remark 21. For more references, see Aumann and Mas Colell. And also G. Carmona for non-atomic games and atomic games.

6.1 Coordination and Acquisition: Beauty contest

Example: “Beauty contest” (Morris and Shin)

Background description:
we assume that there is a continuum of players. The demands are

- uncertainty (size of customer base);
- supplier price;
- average price of the other supplier.

The improvement of decision making relies on the survey that collections the market information (acquisition part). The information consists of different signals, and can be exogenous or indigenous.

The Game is a one-shot game, in which players take actions simultaneously.

A player is represented by a real number $l \in [0, 1]$. Player l chooses n sources of information acquisitions $z_l \in \mathbb{R}^n_+$. For each source $i \in \{1, \dots, n\}$, z_{li} represents the attention paid by player i to the source i.

After the choice z_l, player i observes signals $x_l = (x_{l1}, \dots, x_{ln}) \in \mathbb{R}^n$ about an unobservable quantity θ. The attention z_{li} is the precision of signal x_{li}.

Player l then takes action α_l.

The cost function to player $l \in [0, 1]$ is defined by

$$u_l = C(z_l) + (1 - \gamma)(\alpha_l - \theta)^2_\gamma (\alpha_l - \overline{\alpha})^2$$

where $C : \mathbb{R}_+ \to \mathbb{R}$ is differentiable and increasing, $\gamma \in [-1, 1]$ and $\overline{\alpha}$ represents the average action, namely

$$\overline{\alpha} = \int_0^1 \alpha_l dl \in \mathbb{R}.$$

Remark 22.

- The term $(\alpha_l - \overline{\alpha})^2$ represents the coordination part between players. If $\gamma > 0$, each player wants to choose a similar action as her peers . Otherwise, if $\gamma < 0$, each player intends to behave differently than others.
- In order to make the average action $\overline{\alpha}$ well defined, we need to assume that the function $[0, 1] \ni l \mapsto \alpha_l$ is measurable and integrable.

6.1.1 signals

Assume that player don't know the value of θ (i.e. player starts with no knowledge of θ). The signal observed by player l to the source i is modeled by

$$x_{li} = \theta + \eta_i + \epsilon_{li}$$

where $\eta_i \sim \mathcal{N}(0, \kappa_i^2)$ and $\epsilon_{li} \sim \mathcal{N}(0, \frac{\xi_i^2}{z_{li}})$, and k_i, ξ_i are constant.

The term η_i stands for the noise displayed by the source i when player l approaches to it. The term ϵ_{li} represents the noise brought by the action α_l.

We also notice that the variance of ϵ_{li}, say $\frac{\xi_i^2}{z_{li}}$, is inversely proportional to the precision z_{li}. In another words, the more attention player l gives to signal i, the less the noise incurred by the action of player l.

Assumption 7. Suppose that η_i and ϵ_{li} for all $i \in \{1 \ldots, N\}$ and $l \in [0, 1]$ are all independent.

6.1.2 correlation

Let us look at the covariance of observations of the same source from different players. Suppose $l \neq l'$, then

$$\begin{aligned} cov(x_{li}, x_{l'i}) &= \mathbb{E}\left[(x_{li} - \mathbb{E}[x_{li}])(x_{l'i} - \mathbb{E}[x_{l'i}])\right] \\ &= \mathbb{E}[(\eta_i + \epsilon_{li})(\eta_i + \epsilon_{l'i})] \\ &= \kappa_i^2 \\ &= \rho_{ll'i}\sigma_{li}\sigma_{l'i} \end{aligned} \tag{6.1.1}$$

where

$$\sigma_{li}^2 = \kappa_i^2 + \frac{\xi_i^2}{z_{li}}, \qquad \sigma_{l'i}^2 = \kappa_i^2 + \frac{\xi_i^2}{z_{l'i}}, \qquad \text{and} \qquad \rho_{ll'i} = \frac{k_i^2}{\sigma_{li}\sigma_{l'i}}.$$

Remark 23.

- If a player pays more attention to information, then the precisions of the corresponding observed signals increase (z_{li} increases), which will also lead to an increase of the correlation $\rho_{ll'i}$ between this player and the others.

- When ξ_0 or $\zeta_{li} \to \infty$, then $\rho_{ll'i} \to 1$. In this case, we say that the signals are public, i.e. the observation x_{li} does not depends on player l.

- When $\kappa_i = 0$, then $\rho_{lli} = 0$, we say in this situation that the signals are private.

6.1.3 look for symmetric equilibrium

Definition 33. For player $l \in [0,1]$, a strategy is a couple (z_l, a_l) where $z_l \in \mathbb{R}_+^n$ and $a_l : \mathbb{R}^n \to \mathbb{R}$. A strategy (z_l, a_l) is of feedback form if the action taken by player l satisfies

$$\alpha_l = a_l(x_l).$$

where x_l is the observation of the precision z_l.

Definition 34. The game is said to be symmetric if all player $l \in [0,1]$ use a common strategy (z, a).

For any given $l \in [0,1]$, suppose that player l choses a strategy (z_l, a_l) and for all other players $l' \neq l$, they follow a strategy $(z_{l'}, a_{l'}) = (z, a)$, then the cost function for player l with strategy (z_l, a_l) is given by

$$J^l((z_l, a_l), (z_{-l}, a_{-l})) = \mathbb{E}[u_l] = C(z_l) + (1-\gamma)\mathbb{E}[(\alpha_l - \theta)^2] + \gamma\mathbb{E}[(\alpha_l - \overline{\alpha})^2]$$

where

$$\overline{\alpha} = \int_0^1 a_{l'}(x_{l'})dl'.$$

Remark 24. The definition of $\overline{\alpha}$ seems to say that it should be a random variable because x_l are random variables for all $l \in [0,1]$.

Remark 25. Think about the case in N players game, with observations $x_1, \dots, x_N$ and $l \in \{1, \dots, N\}$. We can still assume that $x_{li} = \theta + \eta_i + \epsilon_{li}$. Then

$$\overline{\alpha}^N = \frac{1}{N}\sum_{l=1}^{N} a_l(x_l).$$

Suppose that $(z_l, a_l) = (z, a)$ for all $l \in \{1, \dots, N\}$, then by law of large numbers, when $N \to \infty$,

$$\overline{\alpha}^N \longrightarrow \mathbb{E}[a(x)] \in \mathbb{R}$$

where $x = (x_1, \dots, x_n)$ with $x_i \sim \mathcal{N}(0, \kappa_i^2 + \frac{\xi_i^2}{z})$.

However, we expect that $\overline{\alpha}$ to be random variable instead of a real number. Here, we need to clarify the Law of large number for continuum of random variables.

For each $l \in [0,1]$, Let $X_l = a(x_l)$ be a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$ valued in $\mathbb{R}$. Then, for every $\omega \in \Omega$,

$$\overline{\alpha} = \int_0^1 X_l(\omega) dl.$$

Thus, assume that we can apply the Fubini's theorem on $\Omega \times [0,1] \ni (\omega, l) \mapsto X_l(\omega) = a(x_l(\omega)) \in \mathbb{R}$, then

$$\begin{aligned}
\mathbb{E}[\overline{\alpha}] &= \int_\Omega \int_0^1 X_l(\omega) dl d\mathbb{P}(\omega) \\
&= \int_0^1 \int_\Omega X_l(\omega) dP(\omega) dl \\
&= \int_0^1 \mathbb{E}[X_l] dl \\
&= \mathbb{E}[X_l]
\end{aligned}$$

This means that the Law of large number for continuum random variables seems to be equivalent to the Fubini's theorem.

However, the function $[0,1] \times \Omega \ni (l, \omega) \mapsto X_l(\omega) = a(x_l(\omega)) \in \mathbb{R}$ is not necessarily jointly measurable on the product space $[0,1] \times \Omega$ with respect to the product $\sigma-$algebra $\mathcal{J} \times \mathcal{F}$.

Remark 26.

- I want all the random variable X_l for $l \in [0,1]$ to be independent.
- I want X_l to have the same distribution μ.

Let us denote $\Omega = (\mathbb{R}^n)^I$ with $I = [0,1]$, the set of all function from I to $\mathbb{R}^n$. Consider a probability measure μ on $(\mathbb{R}^n, \mathcal{B}(\mathbb{R}^n))$ where $\mathcal{B}(\mathbb{R}^n)$ is the Borel $\sigma-$algebra on $\mathbb{R}^n$. Consider p real numbers $J_p = (l_1, \dots, l_p) \in I^p$, and a projection function $\pi_{J_p} : \Omega \ni x = (x_l)_{l \in I} \to (x_{l_1}, \dots, x_{l_p}) = x_{J_p} \in (\mathbb{R}^n)^p$, Let us denote by μ_{J_p} the distribution of $\pi_{J_p}(x)$ on $(\mathbb{R}^n)^p$. If $(x_j)_{j \in J_p}$ are independent following the distribution μ, then $\mu_{J_p} = \mu \otimes \dots \otimes \mu$.

For any finite subsets J_{p_1}, J_{p_1} of I such that $J_{p_1} \subset J_{p_2}$, the i.i.d assumption guarantees the consistency condition, namely

$$\mu_{J_{p_1}} = \mu_{J_{p_2}} \circ \pi^{-1}_{J_{p_2} \to J_{p_1}}$$

where $\pi_{J_{p_2} \to J_{p_1}}$ is the projection mapping from J_{p_2} to J_{p_1}.

The Kolmogorov extension theorem tells us that there must exists a unique probability measure on Ω, denote by $\mathbb{P}_\mu$ such that

$$\mathbb{P}_\mu(x_{J_p} \in A) = \mu_{J_p}(A), \qquad \forall\, A \in (\mathcal{B}(\mathbb{R}^n))^p.$$

Hence, the probability space $(\Omega, (\mathcal{B}(\mathbb{R}^n))^I, \mathbb{P}_\mu)$ is well defined. However, we do not have measurability for the function

$$[0,1] \times \Omega \ni (l, \omega) \mapsto X_l(\omega) = \omega(l) \in \mathbb{R}^n.$$

This is usually called the White Noise.

6.1.4 Linear strategy in "Beauty contest"

Assume that for all $l \in [0,1]$, the strategy $a_l : \mathbb{R}^n \to \mathbb{R}$ is linear in x and coefficient sum up to 1, i.e. $a_l(x) = \sum_{i=1}^n a_{li} x_i$ with $\sum_{i=1}^n a_{li} = 1$. Then

$$\overline{\alpha} = \int_0^1 a_{l'}(x_{l'})dl' = \int_0^1 \left(\sum_{i=1}^n a_{l'i} x_{l'i} \right) dl'. \tag{6.1.2}$$

We look for symmetric equilibrium, name for strategy $a : \mathbb{R}^n \ni x \mapsto \sum_{i=1}^n a_i x_{li} \in \mathbb{R}$. Since $x_{li} = \theta + \eta_i + \epsilon_{li}$, we deduce that

$$\begin{aligned}
\overline{\alpha} &= \int_0^1 \left(\sum_{i=1}^n a_i(\theta + \eta_i + \epsilon_{l'i}) \right) dl' \\
&= \theta + \sum_{i=1}^n a_i \eta_i + \int_0^1 \left(\sum_{i=1}^n a_i \epsilon_{l'i} \right) dl' \\
&= \theta + \sum_{i=1}^n a_i \eta_i
\end{aligned}$$

where the last equality is justified by the fact that

$$\int_0^1 \epsilon_{l'i} dl' = 0.$$

Hence, for a fixed $l \in [0,1]$, player l with strategy (z_l, a_l), and for all $l' \neq l$ such that $(z_{l'}, a_{l'}) = (z, a)$, the cost function of player l becomes

$$\begin{aligned}
J^l((z_l, a_l), (z_{-l}, a_{-l})) &= C(z_l) + (1-\gamma)\mathbb{E}[(a_l(x_l) - \theta)^2] + \gamma \mathbb{E}[(\alpha_l(x_l) - \overline{\alpha})^2] \\
&= C(z_l) + \sum_{i=1}^n a_{li}^2 \left[(1-\gamma)\kappa_i^2 + \frac{\xi_i^2}{z_{li}} \right] + \gamma \sum_{i=1}^n (a_{li} - a_i)^2 \kappa_i^2
\end{aligned}$$

7 2019 03 26 – 28

7.1 Kolmogorov existence theorem

Definition 35. Let E be a Polish space (homeomorphic to a separable complete metric space.) Let $\mathcal{E} = \mathcal{B}(E)$ be a Borel $\sigma-$field of E. (Think of E as $\mathbb{R}$, $\mathbb{R}^d$, separable Hilbert space L^p with $1 \leq p < \infty$ etc.)

Consider a arbitrary index set I. For each finite tuple $(i_1, \ldots, i_k)$ of I, we associate it with a probability measure $\mu_{(i_1,\ldots,i_k)}$ on $(\underbrace{E \times \ldots \times E}_{\text{k times}}, \mathcal{E}_{i_1} \otimes \ldots \otimes \mathcal{E}_{i_k})$ where $\mathcal{E}_{i_j}$ is a $\sigma-$field on E for $j = 1, \ldots, k$. Then the family of probability measures $\left\{\mu_{(i_1,\ldots,i_k)}\right\}$ indexed by the finite tuples of I is said to be consistent if

1. for each $\{i_1, \ldots, i_k\} \subseteq I$ and each permutation π of $\{i_1, \ldots, i_k\}$, we have

$$\mu_{(i_1,\ldots,i_k)} = \mu_{(\pi(i_1),\ldots,\pi(i_k))} \circ \varphi_\pi^{-1}$$

   where $\varphi_\pi(x_{\pi(i_1)}, \ldots, x_{\pi(i_k)}) = (x_{i_1}, \ldots, x_{i_k})$.

2. for each $\{i_1, \ldots, i_k, i_{k+1}\} \subset I$, for any $A \in \underbrace{\mathcal{E} \otimes \ldots \otimes \mathcal{E}}_{\text{k times}}$,

$$\mu_{(i_1,\ldots,i_k)}(A) = \mu_{(i_1,\ldots,i_k,i_{k+1})}(A \times E).$$

Theorem 5 (Kolmogorov existence theorem). Consider the space $(\Omega, \mathcal{F})$, where $\Omega = E^I$ is the product of copies of E or set of functions from I into E, and $\mathcal{F} = \mathcal{E}^I$ with $\mathcal{E}^I = \bigotimes_{i \in I} \mathcal{E}_i$ is the product $\sigma-$field (generated by the cylinders with finite base). If the family of probability measures $\underline{\mu} = \{\mu_{(i_1,\ldots i_k)}\}$ indexed by finite tuples of I is consistent, then there exists a probability measure $\mathbb{P}_{\underline{\mu}}$ on $(\Omega, \mathcal{F})$ such that for every $i \in I$ and every $\omega \in \Omega$, if we set

$$X^i(\omega) = \omega(i)$$

as a random variable valued in $(E, \mathcal{E}_i)$, then for every $\{i_1, \ldots, i_k\} \subseteq I$ we have

$$\mathcal{L}_{\mathbb{P}_{\underline{\mu}}}(X^{i_1}, \ldots, X^{i_k}) = \mu_{(i_1,\ldots,i_k)}.$$

where $\mathcal{L}_{\mathbb{P}_{\underline{\mu}}}$ is the marginal distribution of $P_{\underline{\mu}}$ on $(\underbrace{E \times \ldots, \times E}_{\text{k times}}, \mathcal{E}_{i_1} \otimes \ldots \otimes \mathcal{E}_{i_k})$.

Definition 36. The random variable X^i for every $i \in I$, viewed as a function from Ω to E, is called the $i-$coordinate projection function.

Remark 27. Given a set of *consistent* finite dimensional probability distributions over an arbitrary index set I, ONE CAN ALWAYS construct a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ and a process $(X^i)_{i \in I}$ such that the finite dimensional marginals of $\mathbb{P}$ are the probabilities we started from.

Examples:

- Construction of the Brownian Motion.
- Construction of the Poisson process.
- White noise: $(X^i)_{i\in I}$ i.i.d with distribution $\mu \in \mathcal{P}(E)$.

7.2 structure of product sigma field

We want discuss a bit more about the structure of the product σ-field $\mathcal{E}^I = \bigotimes_{i\in I} \mathcal{E}_i$ with $\mathcal{E}_i = \mathcal{E}$ for all $i \in I$.

Informally speaking, a sigma field generated by the cylinders with finite base has the following property: let $\Omega = E^I$ be a set of functions from I to E, then for any finite subset $\{i_1, \dots, i_k\}$ of I and $A_{i_1}, \dots, A_{i_k} \in \mathcal{E}$, we have

$$\left\{\omega \in \Omega : \omega(i_j) \in A_{i_j},\, j = 1, \dots, k\right\} \in \mathcal{E}^I.$$

We start by clarifying some notations.

Definition 37. For an arbitrary space Ω, an arbitrary index set A, and a measurable space $(E, \mathcal{E})$, let us denote by $\{X_\alpha\}_{\alpha\in A}$ the set of functions $X_\alpha : \Omega \to E$. We denote by $\mathcal{F} = \sigma\{X_\alpha : \alpha \in A\}$ the smallest sigma field on Ω for which X_α is $(\mathcal{F}, \mathcal{E})$ measurable for all $\alpha \in A$. In another words, for every $\alpha \in A$ and every $B \in \mathcal{E}$,

$$(X_\alpha)^{-1}(B) \in \mathcal{F}.$$

Definition 38. Consider an arbitrary index set I and a family of measurable spaces $(E_i, \mathcal{E}_i)$. Let us consider here a special case when $E_i = E$ and $\mathcal{E}_i = \mathcal{E}$ for a given measurable space $(E, \mathcal{E})$.

The *production space* is denoted by $\prod_{i\in I} E_i$, and it can be represented as a collection of all functions from I into E, denoted by E^I.

Consider a set of functions $(X^i)_{i\in I}$ where $X^i : E^I \to E$ is defined by

$$X^i(\omega) = \omega(i), \qquad \forall\, \omega \in E^I.$$

For every $i \in I$, the function is called *$i-$coordinate projection function.*

A one-dimensional cylinder set is the set of the form $\prod_{i\in I} B_i$ where there exists only one $i_0 \in I$ and one $B \in \mathcal{E}$ such that $B_{i_0} = B$ and $B_i = E$ for all $i \neq i_0$.

We define the *product $\sigma-$field*, denoted by $\bigotimes_{i\in I} \mathcal{E}_i$ or $\mathcal{E}^I$, as the one generated by all the one-dimensional cylinder sets.

Lemma 9. The product $\sigma-$field $\bigotimes_{i\in I} \mathcal{E}_i$ is the minimal $\sigma-$field on E^I that makes coordinate projection function X_i measurable for all $i \in I$. Namely

$$\mathcal{E}^I = \bigotimes_{i\in I} \mathcal{E} = \sigma\{X^i : i \in I\}.$$

Lemma 10. Let $(\Omega, \mathcal{F})$ be a measurable space and I be an arbitrary index set. Let $(E, \mathcal{E})$ be a measurable space. Let $(Y^i)_{i\in I}$ be a family of functions from Ω into E. We define $\xi : \Omega \ni \omega \mapsto \xi(\omega) \in E^I$ such that

$$\xi(\omega)(i) = Y^i(\omega).$$

Then ξ is $(\mathcal{F}, \mathcal{E}^I)$-measurable if and only if Y^i is $(\mathcal{F}, \mathcal{E})-$measurable.

Theorem 6. Let Ω be an arbitrary set and let I be an arbitrary index set. Let $(E, \mathcal{E})$ be a measurable space. Let $(Y^i)_{i\in I}$ be a family of functions from Ω into E. Then we have the following two properties:

1. For every $A \in \sigma\{Y^i : i \in I\}$ and every $\omega \in A$, if there is another $\omega' \in \Omega$ satisfying

$$Y^i(\omega) = Y^i(\omega'), \qquad \forall\, i \in I,$$

then $\omega' \in A$.

2. For any $A \in \sigma\{Y^i : i \in I\}$, there exists a subset $J \subset I$ such that J is at most countable and

$$A \in \sigma\{Y^j : j \in J\}.$$

Proof. Let $\mathcal{F} = \sigma\{Y^i, i \in I\}$ be the smallest $\sigma-$field on Ω that makes all Y^i measurable. We define $\xi : \Omega \ni \omega \mapsto \xi(\omega) \in E^I$ such that

$$\xi(\omega)(i) = Y^i(\omega).$$

Then ξ is $(\mathcal{F}, \mathcal{E}^I)$ measurable. So $\mathcal{F}$ contains $\{\xi^{-1}(A) : A \in \mathcal{E}^I\}$ and in fact they are equal.

1. For every $A \in \mathcal{F}$ and every $\omega \in A$, if $\omega' \in \Omega$ such that

$$Y^i(\omega) = Y^i(\omega'), \qquad \forall i \in I.$$

Then $\xi(\omega) = \xi(\omega')$, so that $\omega' \in A$.

2. the claim means that

$$\mathcal{F} = \sigma\{Y^i : i \in I\} = \bigcup_{J\subseteq I, J \text{ countable}} \sigma\{Y^j : j \in J\} =: \mathcal{G}.$$

It is obvious that $\mathcal{G} \subseteq \mathcal{F}$. To have the other inclusion, we will show that $\mathcal{G}$ is indeed a $\sigma-$field. The only non-trivial thing is to show the countable union property. Let $(A_n)_{n\in\mathbb{N}}$ be a countable sequence of elements in $\mathcal{G}$, then there exists a sequence of countable subsets of I, denoted by $(J_n)_{n\in\mathbb{N}}$, such that for every $n \in \mathbb{N}$ $A_n \in \sigma\{Y^j : j \in J_n\}$. Let $J = \cup_n J_n$, then J is also a countable subset of I. Moreover, for all $n \in \mathbb{N}$,

$$A_n \in \sigma\{Y^j, j \in J_n\} \subseteq \sigma\{Y^j : j \in J\}.$$

Since $\sigma\{Y^j : j \in J\}$ is a $\sigma-$field, we deduce that

$$\bigcup_n A_n \in \sigma\{Y^j : j \in J\} \subseteq \mathcal{G}.$$

□

7.3 Doob lemma

Let $(I, \mathcal{I}, \lambda)$ is a probability space. Let E be a polish space and $\mu \in \mathcal{P}(E)$ be a probability measure on a Borel $\sigma-$field of E.

Kolmogorov says that (even if I is only a set) for the space $(\Omega, \mathcal{F}) = (E^I, \bigotimes_{i\in I} \mathcal{B}(E))$, there exists a probability measure $\mathcal{P}_\mu$ on $(\Omega, \mathcal{F})$ such that if we define X^i for $i \in I$ as coordinate projection, namely

$$\begin{array}{rccc} X^i: & \Omega & \longrightarrow & E \\ & \omega & \mapsto & \omega(i) \end{array}$$

then X^i is $(\mathcal{F}, \mathcal{B}(E))$-measurable, the push forward distribution of X^i is

$$\mathcal{L}(X^i) = \mathbb{P}_\mu \circ (X^i)^{-1} = \mu,$$

and $(X^i)_{i\in I}$ are independent (i.e. any finite subset $X^{i_1}, \dots, X^{i_k}$ are independent).

We denote by $\mathbb{P} = \mathbb{P}_\mu$ in the following.

Now I want to think of I as part of $(I, \mathcal{I}, \lambda)$. The following lemma, due to Doob (1953 or 1963), indicates the inadequacy of the product $\sigma-$filed $\mathcal{F}$. We recalled that $\mathcal{F}$ is generated by the cylinders.

Lemma 11. Assume that $I = [0,1]$ and $(E, \mathcal{E}) = (\mathbb{R}, \mathcal{B}(\mathbb{R}))$. Consider a function $h : [0,1] \to E$ and the set

$$\mathfrak{M}_h = \{\omega \in \Omega : X^i(\omega) = h(i) \text{ except for at most countably many } i \in I\}.$$

Then $\mathfrak{M}_h$ has outer measure 1.

Proof. Let $A \in \mathcal{F} = \bigotimes_{[0,1]} \mathcal{B}(E)$. Then A is determined by a countable subsets $I_A \subseteq I$ in the sense that if $\omega \in \Omega$ and $\omega' \in \Omega$ satisfying $\omega(i) = \omega'(i)$ for all $i \in I_A$, then $\omega \in A$ if and only if $\omega' \in A$.

Now suppose that $\mathfrak{M}_h \subset A$, and pick any $\omega \in \Omega$, we construct $\omega' \in \Omega$ such that

$$\omega'(i) = \begin{cases} \omega(i) & i \in I_A \\ h(i) & i \notin I_A \end{cases}$$

Since $\omega'(i) = h(i)$ except for countably many $i \in I$, then we have $\omega' \in \mathfrak{M}_h$. Since $\mathfrak{M}_h \subset A$, so that $\omega' \in A$. Moreover, we know that $\omega'(i) = \omega(i)$ for all $i \in I_A$, then $\omega \in A$ from the property of the $\sigma-$field $\mathcal{F}$. This means that $A = \Omega$. Hence, the outer measure of $\mathfrak{M}_h$ takes the value

$$\mathbb{P}^*(\mathfrak{M}_h) = \inf_{\mathfrak{M}_h \subseteq A} \mathbb{P}(A) = \mathbb{P}(\Omega) = 1.$$

□

Remark 28.

- Space of continuous functions $\mathcal{C}^0 = \mathcal{C}(I; E)$ is not in $\mathcal{F}$.

- A set $A \subseteq \Omega$ cannot belongs to $\mathcal{F}$ unless there exists a countable set $I_A \subset I$ such that for any $\omega, \omega' \in \Omega$ satisfying $\omega(i) = \omega'(i)$ for all $i \in I_A$ implies the property that $\omega \in A$ if and only if $\omega' \in A$.

- set of sample paths of Poisson process are not in $\mathcal{F}$.

***Remark* 29.** Assume that $I = [0, 1]$ and $\mathcal{I} = \mathcal{B}(I)$ and λ is the Lebesgue measure on $[0, 1]$.

- Let $h : [0, 1] \to \mathbb{R}$ be non-Lebesgue measurable (no measurable function equals to h $\lambda-$almost everywhere), then all the elements of $\mathfrak{M}_h$ are non-Lebesgue measurable.

  If $\mathcal{N}$ represents the set of non-Lebesgue measurable functions, then
$$\mathbb{P}^*(\mathcal{N}) = 1$$
  so that
$$\mathbb{P}_*(\mathcal{N}^c) = 0$$
  where $\mathcal{N}^c$ is the set of Lebesgue measurable functions.

- Let $h : [0, 1] \to \mathbb{R}$ be a Lebesgue measurable function, then all elements of $\mathfrak{M}_h$ are Lebesgue measurable, so that
$$\mathbb{P}^*(\mathcal{N}^c) = 1.$$

Thus, $\mathcal{N}^c$ is not $\mathbb{P}-$measurable.

7.4 Games with a Continuum of Players[cf. Carmona section 3.7]

The rationale for the mean field game models studied in this book is based on the limit as $N \to \infty$ of N-player games with mean field interactions. One of the justifications given in Chapter 1 for the formulation of the mean field game paradigm is that the influence on the game of each individual player vanishes in this limit. Mathematical physicists and economists have been using game models in which the impact of each single player is *insignificant*. They do just that by considering games for which the players are labelled by elements i of an uncountable set I, accounting for a continuum of agents. This set I is equipped with a σ-field $\mathcal{I}$ and a probability measure λ which is assumed to be continuous (i.e. non-atomic). In this way, if $i \in I$ represents a player, the fact that $\lambda(\{i\}) = 0$ accounts for the insignificance of the players in the model. This section is thus intended to be a quick introduction to the framework of games with a continuum of players.

The classical Glivenko-Cantelli form of the Law of Large Numbers (LLN) states that if F denotes the cumulative distribution function of a probability measure on $\mathbb{R}$, if $(X^n)_{n\geq 1}$ is an infinite sequence of independent identically distributed random variables on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$ with common distribution μ, and if we use the notation:

$$F_\omega(x) = \limsup_{N\to\infty} \frac{1}{N} \#\big\{n \in \{1, \cdots, N\} : X^n(\omega) \leq x\big\}, \quad x \in \mathbb{R},\ \omega \in \Omega, \tag{7.4.1}$$

for the proportion of $X^n(\omega)$'s not greater than x, then this lim sup is in fact a limit for all $x \in \mathbb{R}$ and $\mathbb{P}$-almost all $\omega \in \Omega$, and $\mathbb{P}[\{\omega \in \Omega : F_\omega(\cdot) = F\}] = 1$.

Switching gears momentarily, recall that, over fifty years ago, economists suggested that the appropriate model for perfectly competitive markets is a model with a continuum of traders represented as elements of a measurable space. In such a set-up, the insignificance of individual traders is captured by the idea of a set with zero measure, and summation or aggregation is generalized by the notion of integral. In games with a continuum of players, the latter are labelled by the elements $i \in I$ of an arbitrary set I (often assumed to be uncountable, and most often chosen to be the unit interval $[0,1]$) equipped with a σ-field $\mathcal{I}$ and a probability measure λ. In this set-up, if the state of each player $i \in I$ is given by a random variable X^i on a probability space $(\Omega, \mathcal{F}, \mathbb{P})$, in analogy with the countable case leading to formula (7.4.1), the quantity:

$$F_\omega(x) = \lambda\big(\{i \in I : X^i(\omega) \le x\}\big) \tag{7.4.2}$$

appears as a natural generalization of the *proportion of* $(X^i(\omega))_{i \in I}$*'s not greater than* x, in other words of the cumulative distribution function of the empirical distribution, and if the $(X^i)_{i\in I}$'s were to be independent with the same distribution, we could have a reasonable generalization of the Law of Large Numbers to this setting. However, as we explain in the next subsection, measurability issues get in the way and such a generalization is, when it does exist, far from trivial.

7.5 The Exact Law of Large Numbers

If E is a Polish space, for each probability measure $\mu \in \mathcal{P}(E)$, Kolmogorov's theorem can be used to construct on the product space $\Omega = E^I$ equipped with the σ-field $\mathcal{F}$ obtained as the product of copies of the Borel σ-field of E, the product probability measure $\mathbb{P}$ for which the coordinate projections $(X^i : \Omega \ni \omega \mapsto X^i(\omega) = \omega(i) \in E)_{i \in I}$ become independent and identically distributed random variables with common distribution μ on the probability space $(\Omega, \mathcal{F}, \mathbb{P})$. It is also well known that the sample paths $I \ni i \mapsto X^i(\omega) \in E$ are pretty rough functions since they are (for $\mathbb{P}$-almost $\omega \in \Omega$) nowhere continuous and not even measurable.

Hence, this construction of a continuum of independent identically distributed random variables leads to irregular structures lacking measurability properties. The following definition offers an alternative which keeps most of what is needed from the independence.

Definition 39. If E is a Polish space, the E-valued random variables $(X^i)_{i\in I}$ are said to be essentially pairwise independent if, for λ-almost every $i \in I$, the random variable X^i is independent of X^j for λ-almost every $j \in I$. Accordingly, if the real valued random variables $(X^i)_{i\in I}$ are square integrable, we say that the family $(X^i)_{i\in I}$ is essentially pairwise uncorrelated if, for λ-almost every $i \in I$, the correlation coefficient of X^i with X^j is 0 for λ-almost every $j \in I$.

One may wonder if essentially pairwise independent families $(X^i)_{i\in I}$ can be constructed on probability spaces $(\Omega, \mathcal{F}, \mathbb{P})$ so that the process $\mathbf{X} : I \times \Omega \ni (i, \omega) \mapsto X^i(\omega)$ satisfies relevant measurability properties. To do so, we shall construct such processes on extensions of the product space $(I \times \Omega, \mathcal{I} \otimes \mathcal{F}, \lambda \otimes \mathbb{P})$, which are called Fubini's extensions. sub]Fubini extensionsub]extension!Fubini

Definition 40. If $\mathcal{I} \boxtimes \mathcal{F}$ is a σ-field containing $\mathcal{I} \otimes \mathcal{F}$ and $\lambda \boxtimes \mathbb{P}$ is a probability measure on $(I \times \Omega, \mathcal{I} \boxtimes \mathcal{F})$, then $(I \times \Omega, \mathcal{I} \boxtimes \mathcal{F}, \lambda \boxtimes \mathbb{P})$ is said to be a Fubini extension of $(I \times \Omega, \mathcal{I} \otimes \mathcal{F}, \lambda \otimes \mathbb{P})$ if, for every measurable and $\lambda \boxtimes \mathbb{P}$-integrable $X : I \times \Omega \ni (i, x) \mapsto X^i(\omega) \in \mathbb{R}$, we have:

1. for λ-a.e. $i \in I$, $\Omega \ni \omega \mapsto X^i(\omega)$ is a $\mathbb{P}$-integrable random variable, and for $\mathbb{P}$-a.e. $\omega \in \Omega$, $I \ni i \mapsto X^i(\omega)$ is measurable and λ-integrable;
2. $I \ni i \mapsto \int_\Omega X^i(\omega) d\mathbb{P}(\omega)$ is measurable and λ-integrable, and $\Omega \ni \omega \mapsto \int_I X^i(\omega) d\lambda(i)$ is a $\mathbb{P}$-integrable random variable, and:

$$\begin{aligned} \int_I \left(\int_\Omega X^i(\omega) d\mathbb{P}(\omega) \right) d\lambda(i) &= \int_\Omega \left(\int_I X^i(\omega) d\lambda(i) \right) d\mathbb{P}(\omega) \\ &= \int_{I \times \Omega} X^i(\omega) d\big(\lambda \boxtimes \mathbb{P}\big)(i, \omega). \end{aligned} \tag{7.5.1}$$

In the sequel, we shall use the standard symbol $\mathbb{E}$ for denoting the expectation under the sole probability $\mathbb{P}$.

Measurable essentially pairwise independent processes $\mathbf{X}$ are first constructed in such a way that, for each $i \in I$, the law of X^i is the uniform distribution on the unit interval $[0, 1]$. Then, using the tools we develop in Chapter **??**, see for example Lemma **??**, we easily construct measurable essentially pairwise independent Euclidean-valued processes with any given prescribed marginals. So the actual problem is to construct rich product probability spaces in the sense of the following definition.

Definition 41. A Fubini extension $(I \times \Omega, \mathcal{I} \boxtimes \mathcal{F}, \lambda \boxtimes \mathbb{P})$ is said to be rich if there exists a real valued $\mathcal{I} \boxtimes \mathcal{F}$-measurable essentially pairwise independent process $\mathbf{X}$ such that the law of X^i is the uniform distribution on $[0, 1]$ for every $i \in I$.

We refer to the Notes & Complements at the end of the chapter for references to papers giving the construction of essentially pairwise independent measurable processes on Fubini extensions.

The following gives a simple property of rich Fubini extensions.

Lemma 12. If the Fubini extension $(I \times \Omega, \mathcal{I} \boxtimes \mathcal{F}, \lambda \boxtimes \mathbb{P})$ is rich, then λ is necessarily atomless.

Proof. We shall argue by contradiction. If $A \in \mathcal{I}$, with $\lambda(A) > 0$, is an atom of $(I, \mathcal{I}, \lambda)$, then, for $\mathbb{P}$-a.e. $\omega \in \Omega$, the function $I \ni i \mapsto X^i(\omega)$ is λ-a.e. constant on A. So for $\mathbb{P}$-a.e. $\omega \in \Omega$ and λ-a.e. $i \in A$,

$$X^i(\omega) = \int_A X^j(\omega) \frac{d\lambda(j)}{\lambda(A)},$$

and using the Fubini property (7.5.1), we deduce that for λ-a.e. $i \in A$, the random variable $\Omega \ni \omega \mapsto X^i(\omega)$ is $\mathbb{P}$-a.e. equal to the random variable $\theta : \Omega \ni \omega \mapsto \int_A X^j(\omega) d\lambda(j)/\lambda(A)$. Also, for any event $B \in \mathcal{F}$,

$$\begin{aligned}\mathbb{P}[\theta \in B] &= \lambda \boxtimes \mathbb{P}\big[(i,\omega) \in I \times \Omega : X^i(\omega) \in B\big] \\ &= \int_I \mathbb{P}[X^i \in B] d\lambda(i) = \mathrm{Leb}_1(B),\end{aligned}$$

proving that θ, as a random variable on $(\Omega, \mathcal{F}, \mathbb{P})$, has the uniform distribution. In particular, $\mathbb{E}[\theta^2] = 1/3$.

On the other hand, we know that, for almost every $i \in I$, the function $I \times \Omega \ni (j,\omega) \mapsto X^i(\omega)X^j(\omega)$ is $\mathcal{I} \boxtimes \mathcal{F}$-measurable. Also, by the Fubini property, the function $I \ni j \mapsto \mathbb{E}[X^i X^j]$ is integrable with respect to λ and

$$\int_I \mathbb{E}[X^i X^j] d\lambda(j) = \mathbb{E}[X^i \theta]. \tag{7.5.2}$$

Now, we observe that the function $I \times \Omega \ni (i,\omega) \mapsto X^i(\omega)\theta(\omega)$ is also $\mathcal{I} \boxtimes \mathcal{F}$-measurable. Hence, $I \ni i \mapsto \mathbb{E}[X^i\theta]$ is integrable with respect to λ and

$$\int_I \mathbb{E}[X^i \theta] d\lambda(i) = \mathbb{E}[\theta^2] = \frac{1}{3}.$$

The contradiction comes from the fact that, for almost every $i \in I$, X^i is independent to X^j for almost every $j \in I$. In other words, the left-hand side in (7.5.2) is equal to:

$$\int_I \mathbb{E}[X^i X^j] d\lambda(j) = \frac{1}{\lambda A} \int_A \mathbb{E}[X^i]\mathbb{E}[X^j] d\lambda(j) = \frac{1}{4},$$

which gives the desired contradiction.

Using Lemma 5.29 from Chapter 5 when E is a Euclidean space and an extension of it when E is a more general Polish space, see for instance the Notes & Complements at the end of Chapter 5, we get the following result which we already announced.

We recall the following lemma

Lemma 13 (see lemma 5.29). Let E be a Polish space, $\exists \varphi : [0,1] \times \mathcal{P}(E) \to E$ measurable such that $\forall \nu \in \mathcal{P}(E)$

$$Leb \circ \psi(\cdot, \nu)^{-1} = \nu.$$

Proposition 11. If the Fubini extension $(I \times \Omega, \mathcal{I} \boxtimes \mathcal{F}, \lambda \boxtimes \mathbb{P})$ is rich, if E is a Polish space, and if $\mu : I \mapsto \mathcal{P}(E)$ is $\mathcal{I}$-measurable, then there exists a $\mathcal{I} \boxtimes \mathcal{F}$-measurable E-valued essentially pairwise independent process $\mathbf{Y} : I \times \Omega \mapsto E$ such that for λ-a.e. $i \in I$, $\mathbb{P} \circ Y_i^{-1} = \mu_i$.

Proof. Define $Y(i,\omega) = \psi(X^i(\omega), \mu_i)$ where $X = (X^i)_{i \in I}$ is an essential white noise whit distribution uniform $[0,1]$. For $i \in I$ fixed, the mapping $\omega \mapsto X^i(\omega)$ has a uniform distribution on $[0,1]$, i.e. $\mathcal{L}(X^i) = Leb([0,1])$. And $\psi(\cdot, \mu_i)$ takes the uniform distribution on $[0,1]$ into μ_i.

Since for $\lambda-$a.e. $i \in I$, X^i is independent of X^j for $\lambda-$a.e. $j \in I$, so that for $\lambda-$a.e. $i \in I$, $\psi(X^i, \mu_i)$ is independent of $\psi(X^j, \mu_j)$ for $\lambda-$a.e. $j \in I$.

Done. $\square$

An exact law of large numbers can be proven on Fubini's extensions. In a weak form, this law can be given in the following wasy.

Theorem 7. Let $\mathbf{X} = (X^i)_{i \in I}$ be a measurable square integrable process on a Fubini extension $(I \times \Omega, \mathcal{I} \boxtimes \mathcal{F}, \lambda \boxtimes \mathbb{P})$. The following are equivalent:

(i) The random variable $(X^i)_{i \in I}$ are essentially pairwise uncorrelated;

(ii) For every $A \in \mathcal{I}$ with $\lambda(A) > 0$, one has for $\mathbb{P}$-almost surely in $\omega \in \Omega$:

$$\int_A X^i(\omega) d\lambda(i) = \int_A \mathbb{E}[X^i] d\lambda(i).$$

Proof. First Step: We first check that if $\mathbf{Y} = (Y^i)_{i \in I}$ and $\mathbf{Z} = (Z^i)_{i \in I}$ are measurable and square integrable processes on the Fubini extension $(I \times \Omega, \mathcal{I} \boxtimes \mathcal{F}, \lambda \boxtimes \mathbb{P})$, and if we set $\tilde{X}^{i,j}(\omega) = Y^i(\omega) Z^j(\omega)$ for $i, j \in I$ and $\omega \in \Omega$, then $\Omega \ni \omega \mapsto \tilde{X}^{i,j}$ is $\mathbb{P}$-integrable for λ-a.e. $i \in I$ and $j \in I$. Now, proceeding as in the proof of Lemma 12 and using the Fubini property of the space, we easily check that, for λ-a.e. $i \in I$, the function $I \ni j \mapsto \mathbb{E}[\tilde{X}^{i,j}]$ is λ-integrable, that the function $I \ni i \mapsto \int_I \mathbb{E}[\tilde{X}^{i,j}] d\lambda(j) = \mathbb{E}[Y^i \int_I Z^j d\lambda(j)]$ is λ-integrable, that the function $\Omega \ni \omega \mapsto (\int_I Y^i(\omega) d\lambda(i))(\int_I Z^j(\omega) d\lambda(j))$ is $\mathbb{P}$-integrable and that:

$$\mathbb{E}\Big[\Big(\int_I Y^i(\omega) d\lambda(i)\Big)\Big(\int_I Z^j(\omega) d\lambda(j)\Big)\Big] = \int_I \Big(\int_I \mathbb{E}[\tilde{X}^{i,j}] d\lambda(i)\Big) d\lambda(j). \tag{7.5.3}$$

Second Step: Let $A, B \in \mathcal{I}$, and let us define the processes $\mathbf{Y} = (Y^i)_{i \in I}$ and $\mathbf{Z} = (Z^i)_{i \in I}$ by $(Y^i = \mathbf{1}_A(i)(X^i - \mathbb{E}[X^i]))_{i \in I}$ and $(Z^i = \mathbf{1}_B(i)(X^i - \mathbb{E}[X^i]))_{i \in I}$ respectively. Applying (7.5.3) from the first step we get:

$$\begin{aligned} &\int_A \int_B \mathbb{E}\Big[\Big(X^i - \mathbb{E}[X^i]\Big)\Big(X^j - \mathbb{E}[X^j]\Big)\Big] d\lambda(i) d\lambda(j) \\ &\qquad = \mathbb{E}\Big[\int_A \Big(X^i - \mathbb{E}[X^i]\Big) d\lambda(i) \int_B \Big(X^j - \mathbb{E}[X^j]\Big) d\lambda(j)\Big], \end{aligned} \tag{7.5.4}$$

and the implication *(i)* $\Rightarrow$ *(ii)* follows by taking $B = A$. On the other hand, if we assume that *(ii)* holds, equation (7.5.4) implies that:

$$\int_A \int_B \mathbb{E}\Big[\Big(X^i - \mathbb{E}[X^i]\Big)\Big(X^j - \mathbb{E}[X^j]\Big)\Big] d\lambda(i) d\lambda(j) = 0$$

for all $A, B \in \mathcal{I}$. The set $A \in \mathcal{I}$ being arbitrary, we conclude that:

$$\int_B \mathbb{E}\Big[\Big(X^i - \mathbb{E}[X^i]\Big)\Big(X^j - \mathbb{E}[X^j]\Big)\Big] d\lambda(j) = 0,$$

for λ-a.e. $i \in I$. So for λ-a.e. $i \in I$, $B \in \mathcal{I}$ being arbitrary, we conclude that:

$$\mathbb{E}\Big[\Big(X^i - \mathbb{E}[X^i]\Big)\Big(X^j - \mathbb{E}[X^j]\Big)\Big] = 0$$

for λ-a.e. $j \in I$ which completes the proof. $\square$

Theorem 7 provides a form of the weak law of large numbers for essentially pairwise uncorrelated uncountable families of random variables. Here is a stronger form for essentially pairwise independent families of random variables.

Theorem 8. Let E be a Polish space and $\mathbf{X} = (X^i)_{i\in I}$ be a measurable E-valued process on a Fubini extension $(I \times \Omega, \mathcal{I} \boxtimes \mathcal{F}, \lambda \boxtimes \mathbb{P})$ such that the random variables $(X^i)_{i\in I}$ are essentially pairwise independent. Then, for $\mathbb{P}$ almost every $\omega \in \Omega$ and for any B in the Borel σ-field $\mathcal{B}(E)$,

$$\lambda\big[\{i \in I : X^i(\omega) \in B\}\big] = \int_I \mathbb{P}\big[X^i \in B\big] d\lambda(i).$$

Of course, we may choose E as a Euclidean space, in which case we get a strong form of the exact law of large numbers for essentially pairwise independent families of random variables with values in $\mathbb{R}^d$, for some $d \geq 1$. By choosing E as a functional space, the same holds true for a continuum of essentially pairwise independent random processes.

Finally, we can also derive conditional versions of these exact laws. We do not give the details here because we want to keep the presentation to a rather non-technical level since our motivation is merely to connect our approach to mean field games to the existing literature on games with a continuum of players. The interested reader is referred to the Notes & Complements at the end of the chapter for references.

7.6 Question

Question: Can the product probability space $(\Omega, \mathcal{F}, \mathbb{P}_\mu)$ be a component of a Fubini extension?

Proposition 12. There is no atomless probability space $(I, \mathcal{I}, \lambda)$ satisfying

- $(I \times \Omega, \mathcal{I} \boxtimes \mathcal{F}, \lambda \boxtimes \mathbb{P})$ is a Fubini extension with marginals $(I, \mathcal{I}, \lambda)$ and $(\Omega, \mathcal{F}, \mathbb{P})$.
- $(X^i)_{i\in I}$ is $\mathcal{I} \boxtimes \mathcal{F}$ measurable.

Proof. Suppose such a Fubini extension exists. Take $E = \mathbb{R}$. Let B be an interval in E and $\mu(B) < 1$. Let $c \in B$ and we define two functions $h : I \ni i \mapsto h(i) = c \in E$ and $g = \mathbb{1}_B \circ X : I \times \Omega \to E$, namely for every $(i, \omega) \in I \times \Omega$,

$$g(i,\omega) = \begin{cases} 1 & if \quad X^i(\omega) \in B \\ 0 & o.w. \end{cases}$$

Then, g is bounded and $\mathcal{I} \boxtimes \mathcal{F}$ measurable, $(g(i,\cdot))_{i\in I}$ are independent.

By the exact law of large number,

$$\lambda \circ g_\omega^{-1} = (\lambda \boxtimes \mathbb{P}) \circ g^{-1}$$

for $\mathbb{P}$-almost every $\omega \in \Omega$, where $g_\omega = g(\cdot, \omega)$ for $\omega \in \Omega$.

Applied thus to the whole space $E = \mathbb{R}$, we get

$$\lambda(g_\omega^{-1}(E)) = (\lambda \boxtimes \mathbb{P})(g^{-1}(E))$$

or equivalently

$$\lambda(\{i \in I : g(i,\omega) \in B\}) = \int \lambda(di)\mathbb{P}(d\omega)g(i,\omega) = \int \lambda(di)\mu(B) = \mu(B) < 1.$$

So for $\mathbb{P}$-a.e. $\omega \in \Omega$, $X^{\cdot}(\omega) = \omega(\cdot) \notin \mathfrak{M}_h$. This is because if $\omega \in \Omega$ such that $X^i(\omega) = h(i) = c$, then $X^i(\omega) \in B$. If that were to happen except for countably many i, then $\lambda-$measure of those i's would be 1 since λ is atomless.

Hence $\mathbb{P}(\mathfrak{M}_h) = 0$. Contradiction. $\square$

8 2019 04 02 Multi-stage games

Definition 42. A multi-stage game consists of the following components:

- A sequence $(0, \dots, K)$ represents the stage of the game;
- there are N players who play repeatedly and chose their actions simultaneously at each stage of the game;
- Each player observes the actions of all N players in all previous stages;
- For each stage $k \in \{0, \dots, K\}$, the action profile of players is given by

  $$\underline{\alpha}_k = (\alpha_k^1, \dots, \alpha_k^N)$$

  where α_k^i is the action chosen by player i at stage k;
- The available information prior to stage k, called *history*, is defined as

  $$h^k = (\underline{\alpha}_0, \dots, \underline{\alpha}_{k-1});$$

- For every player i at stage k, the feasible set of actions is denoted by $A^i(h^k)$, which depends on the information of history prior to stage k. For example, the resources that have been used in the past can affect the current actions of players. So, at stage k, the action chosen by player i can be viewed as a function of history h^k;
- We would like to include *"no action"* as part of $A^i(h^k)$ for every $i = 1, \dots, N$ and for every $k = 0, \dots, K$. When player i choses "no action", she is inactive at this stage.

Here are some examples for multi-stage games:

1. $K = 0$, Cournot competition.
2. $K = 1$, Stackelberg game. For example $N = 2$, one player is called the leader, and the other called the follower. Then
   - at stage $k = 0$, leader chooses an action $\alpha_0 \in A^{leader}(\emptyset)$. The follower is inactive, namely $A^{follower}(\emptyset) = \{\text{``}no\,action\text{''}\}$.
   - at stage $k = 1$, leader is inactive $A^{leader}(h^1) = \{\text{``}no\,action\text{''}\}$ and the follower chooses an action $\alpha^{follower} \in A^{follower}(h^1)$ where $h^1 = ((\alpha_0, \text{``}no\,action\text{''}))$.
3. $K = 2$, Entry-deterrence game.
   - at stage $k = 0$, incumbent raises money α_0, entrant "no action";
   - at stage $k = 1$, incumbent "no action", entrant choses between "enter" or "not enter";

- at stage $k = 2$, Cournot competition.

Definition 43. A pure strategy for player $i \in \{1, \dots, N\}$ at stage $k \in \{0, \dots, K\}$ is denoted by $\widehat{\alpha}^i_k$.
A pure strategy profile, denoted by $\widehat{\alpha}$, is the collection of pure strategies of all players at all stages except the last one $k = K$, namely

$$\widehat{\alpha} = (\widehat{\alpha}^i_k)_{i=1,\dots,N,\, k=0,\dots,K-1}.$$

It is said to be an admissible pure strategy profile if $\widehat{\alpha}^i_k \in A^i(h^k)$ for all $k = 0, \dots, K-1$ and $i = 1, \dots, N$.

Remark 30. A stage $k = K$, we collect everything and decide the reward or cost for players.

Definition 44. The set of histories up to stage $k \in \{0, \dots, K\}$ is denoted by $H(k)$. So $H(K)$ represents the set of all histories. For each player $i \in \{1, \dots, N\}$, we associate her with a cost function $J^i : H(K) \to \mathbb{R}$.

Very often, the cost function J^i is addictive, i.e.

$$J^i(h^K) = \sum_{k=0}^{K-1} c_k f^i(\alpha^i_k)$$

The coefficients $(c_k)_{k=0,\dots,K-1}$ can be discounted factors, for example, $c_k = c_K \delta^k$ for some $c_K \in \mathbb{R}$ and $\delta \in [0, 1]$.
When $K = \infty$, we need to make sure $\delta < 1$.

Definition 45. A pure strategy profile $\widehat{\alpha}$ is said to be a pure strategy Nash equilibrium if for every $i \in \{1, \dots, N\}$, for every admissible sequence of actions $\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_{k-1})$,

$$J^i(\underline{\widehat{\alpha}}_0, \dots, \underline{\widehat{\alpha}}_{K-1}) \le J^i((\alpha_0, \underline{\widehat{\alpha}}^{-i}_0), \dots, (\alpha_{K-1}, \underline{\widehat{\alpha}}^{-i}_{K-1}))$$

where $(\alpha_k, \underline{\widehat{\alpha}}^{-i}_k) = (\widehat{\alpha}^1_k, \dots, \widehat{\alpha}^{i-1}_k, \alpha_k, \widehat{\alpha}^{i+1}_k, \dots, \widehat{\alpha}^N_k)$ is the action profile at stage k for $k \in \{0, \dots, K-1\}$.

Remark 31. Given a strategy profile $\widehat{\alpha}$ and a player i, a sequence of actions $\boldsymbol{\alpha} = (\alpha_0, \dots, \alpha_{K-1})$ is said to be admissible if for every $k = 1, \dots, K-1$,

$$\alpha_k \in A^i(h^k_{\boldsymbol{\alpha}})$$

where $h^0_{\boldsymbol{\alpha}} = \emptyset$ and $h^k_{\boldsymbol{\alpha}} = ((\alpha_0, \underline{\widehat{\alpha}}^{-i}_0), \dots, (\alpha_{k-1}, \underline{\widehat{\alpha}}^{-i}_{k-1}))$

Definition 46. For every stage $k = 0, \dots, K$, let $G(h^k)$ be a new game depending on the history $h^k = (\underline{\alpha}_0, \dots, \underline{\alpha}_{k-1})$.

- The new strategy profile for N players in game $G(h^k)$ is denoted by $\alpha_{h^k} = (\alpha^i_l)_{i=1,\dots,N,\, l=k,k+1,\dots,K-1}$ (similarly, $\widehat{\alpha}_{h^k}$ for pure strategy profile);

- The new history is denoted by $(h^k, \underline{\alpha}_k, \dots, \underline{\alpha}_{K-1})$;
- The new cost function is a mapping $(\underline{\alpha}_k, \dots, \underline{\alpha}_{K-1}) \mapsto J^i(h^k, \underline{\alpha}_k, \dots, \underline{\alpha}_{K-1})$

A strategy profile $\widehat{\alpha}$ is said to be a *sub-game perfect Nash equilibrium* if for every $k \in \{0, \dots, K\}$, for every $h^k \in H(k)$, the strategy profile $(\widehat{\alpha}^i_l)_{i=1,\dots,N,\, l=k,\dots,K-1}$ is a Nash equilibrium for the sub game $G(h^k)$.

In words, at every stage k, whatever the history is, if players play according to the strategy profile $(\widehat{\alpha}^i_l)_{i=1,\dots,N,\, l=k,\dots,K-1}$ afterwards, then they are in a Nash equilibrium associated to the new game.

Remark 32. We need to clarify the notion of closed loop strategy and open loop strategy. Let α^i_k be the action taken by player i at stage k. For the sake of simplicity, we assume that $A^i(h^k) = \mathbb{R}$ for every $h^k \in H(k)$ for every $k = 0, \dots, K$.

- α^i_k is said to be an open loop strategy if there exists a function $\varphi^i : \{0, \dots, K\} \to \mathbb{R}$ such that $\alpha^i_k = \varphi^i(k)$ for every $k = 0, \dots K$. The function φ^i is decided before the game.
- α^i_k is said to be a closed loop strategy if there exists a function $\varphi^i_k : H(k) \to \mathbb{R}$ such that $\alpha^i_k = \varphi^i_k(h^k)$ for some $h^k \in H(k)$.

It will be rare for open loop strategy profile to have a sub-game perfect Nash equilibrium.

8.1 Prisoner's dilemma

Assume that there are only two players $N = 2$, and the feasible set of actions are $A_1 = A_2 = \{C, D\}$ for every stage k of the game. Here C stands for cooperation and D stands for defection.

Let $T < R < P < S$ where $T, R, P, S \in \mathbb{R}$ and T stands for *temptation*, R stands for *negative rewards* or *cost* , P stands for *punishment*, and S stands for *sucker*. Assume that the cost functions are defined as follow (see also the table 1):

$$J^1(\alpha_1, \alpha_2) = \mathbb{1}_{\alpha_1=C}(R\mathbb{1}_{\alpha_2=C} + T\mathbb{1}_{\alpha_2=D}) + \mathbb{1}_{\alpha_1=D}(S\mathbb{1}_{\alpha_2=C} + P\mathbb{1}_{\alpha_2=D})$$

and

$$J^2(\alpha_1, \alpha_2) = \mathbb{1}_{\alpha_2=C}(R\mathbb{1}_{\alpha_1=C} + T\mathbb{1}_{\alpha_1=D}) + \mathbb{1}_{\alpha_2=D}(S\mathbb{1}_{\alpha_1=C} + P\mathbb{1}_{\alpha_1=D}).$$

The objective of players are to minimize their own cost function. We observe that the choice of relationships $T < R$ and $P < S$ imply that the defection is a preferable strategy for both player 1 and 2.

It can be verified that the strategy (D, D) is a Nash equilibrium.

Lemma 14. The strategy profile $((D, D), \dots, (D, D))$ is a sub-game perfect Nash equilibrium.

player 2\palyer 1	C	D
C	(R, R)	(T, S)
D	(S, T)	(P, P)

Table 1: the cost function table. The tuple $(J^1(\cdot), J^2(\cdot))$ represents the costs of player 1 and player 2 in different situation. The column names represent the strategies chosen by player 1 and the row names are for player 2.

Proof. Step 1: If I am looking for a sub-game perfect Nash equilibrium, the last choice should be (D, D) whatever the history is.

Step 2: One step perturbation principle (for sub-game perfect Nash equilibrium) for the induction part of the proof. □

***Remark* 33.** The result may be different if the number of stages equals to infinity, i.e. $K = \infty$.

References

[Bonacich, 1972a] Bonacich, P. (1972a). Factoring and weighting approaches to status scores and clique identification. *Journal of mathematical sociology*, 2(1):113–120.

[Bonacich, 1972b] Bonacich, P. (1972b). Technique for analyzing overlapping memberships. *Sociological methodology*, 4:176–185.

[Bonacich, 1987] Bonacich, P. (1987). Power and centrality: A family of measures. *American journal of sociology*, 92(5):1170–1182.

[Janson, 2010] Janson, S. (2010). Graphons, cut norm and distance, couplings and rearrangements. Technical report, Department of Mathematics, Uppsala University.

[Katz, 1953] Katz, L. (1953). A new status index derived from sociometric analysis. *Psychometrika*, 18(1):39–43.

[Krivine, 1979] Krivine, J.-L. (1979). Constantes de grothendieck et fonctions de type positif sur les spheres. *Advances in Mathematics*, 31(1):16–30.

[Page et al., 1999] Page, L., Brin, S., Motwani, R., and Winograd, T. (1999). The pagerank citation ranking: Bringing order to the web. Technical report, Stanford InfoLab.