Sampling techniques for the approximation of solutions for TU games

Alejandro Saavedra-Nieves

(based on joint works with E. Algaba, B. Casas-Méndez, M. G. Fiestras-Janeiro,

I. García-Jurado and P. Saavedra-Nieves)

USC
UNIVERSIDADE
DE SANTIAGO
DE COMPOSTELA

Saint-Étienne, 15th April 2022

TU-games

Game theory: mathematical theory of interactive decision problems

A TU-game is a pair (N, v):

- N is the set of players and,
- $v : 2^N \longrightarrow \mathbb{R}$ is the characteristic function with $v(\emptyset) = 0$.

A main goal: definition and analysis of rules to allocate $v(N)$

- Point-valued solutions: the Shapley value, the Banzhaf-value, the Owen value, the Banzhaf-Owen value...
- Set-valued solutions: the imputation set, the core,...

Their exact computation is a difficult task for large sets of players!

The Galician milk conflict

- After the suppression of the European milk quotas in March 2015...

Year	Jan.	Feb.	Mar.	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
2013	32.65	32.66	32.76	32.84	33.06	33.03	34.44	34.86	35.59	38.57	38.93	39.17
2014	39.24	38.90	38.65	36.05	35.61	35.37	33.72	33.73	33.64	32.24	32.10	31.95
2015	30.52	30.60	30.30	28.80	28.40	27.90	27.60	27.70	28.30	28.70	28.70	28.80

Table: Averaged prices of the milk in Galicia (in euros per 100 litres) in the period 2013-2015.

How to increase the price of milk?

A low production scenario

- Reducing the milk production in Galicia?
- If we know a maximum production per municipality, how this reduction affect each of the 190 involved?
- We build a new system of quotas for councils.

*Data source: Consellería de Medio Rural da Xunta de Galicia.

The Spanish hotel industry

The hotel industry in Spain*

- The hotel industry generated 12% of GDP.
- 83.7 million annual travellers.
- Spain was one the World's favourite destinations in 2019.

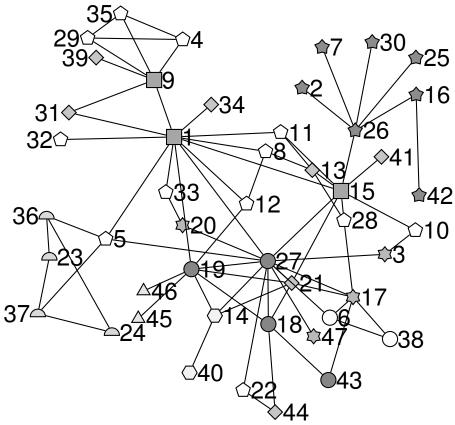
The crisis of the hotel industry

- After corona crisis, reservations collapse.
- Join resources minimizes the impact of crisis on the economy.
- The UN World Tourism Organization seeks a more efficient tourism.

*Data source: United Nations World Tourism Organization (UNWTO, `https://www.unwto.org/`).

Saavedra-Nieves, A., Fiestras-Janeiro, M. G. (2022). Analysis of the impact of DMUs on the overall efficiency in the event of a merger. Expert Systems With Applications, 195, 116571.

The Zerkani network



- Ranking the members of the Zerkani network, responsible for the attacks in Paris (2015) and Brussels (2016).
- It contains 47 members.
- Hamers et al. (2019) use the Shapley value.

Hamers, H., Husslage, B., Lindelauf, R. (2019). Analysing ISIS Zerkani Network using the Shapley Value. Handbook of the Shapley Value, pp. 463-481.

TU-games

Game theory: mathematical theory of interactive decision problems

A TU-game is a pair (N, v):

- N is the set of players and,
- $v : 2^N \longrightarrow \mathbb{R}$ is the characteristic function with $v(\emptyset) = 0$.

A main goal: definition and analysis of rules to allocate $v(N)$

- Point-valued solutions: the Shapley value, the Banzhaf-value, the Owen value, the Banzhaf-Owen value...
- Set-valued solutions: the imputation set, the core,...

Sampling techniques are considered as a solution!

Table of contents

The Shapley value

The Shapley value (Shapley, 1953) for (N, v), for every $i \in N$, is

$$Sh_i(N, v) = \frac{1}{|\Pi(N)|} \sum_{\sigma \in \Pi(N)} (v(P_i^\sigma \cup \{i\}) - v(P_i^\sigma)),$$

being $\Pi(N)$ the set of all orders of N and $P_i^\sigma = \{j \in N : \sigma(j) < \sigma(i)\}$ for any $\sigma \in \Pi(N)$.

Shapley, L. S. (1953). A value for n-person games. In: Kuhn, H. W. and Tucker, A. W. (Eds.), Contributions to the Theory of Games II, Princeton University Press, Princeton, NJ, 307-317.

A procedure based on simple random sampling with replacement

1. We generate a sample with replacement $\mathcal{S} = \{\sigma_1, \ldots, \sigma_\ell\}$ of ℓ elements in $\Pi(N)$.
2. For each $\sigma \in \mathcal{S}$, $x(\sigma)_i = v(P_i^\sigma \cup \{i\}) - v(P_i^\sigma)$, for all $i \in N$.
3. The estimation of Sh_i is

$$\hat{Sh}_i = \frac{1}{\ell} \sum_{\sigma \in \mathcal{S}} x(\sigma)_i, \text{ for all } i \in N.$$

Unbiased and consistent estimator

Castro, J., Gómez, D., and Tejada, J. (2009). Polynomial calculation of the Shapley value based on sampling. Computers & Operations Research, 36(5), 1726-1730.

Analysis of the error

A confidence interval for the Shapley value

$$\mathbb{P}(|\hat{Sh}_i - Sh_i| \leq \varepsilon) \geq 1 - \alpha, \text{ with } \varepsilon > 0 \text{ and } \alpha \in (0, 1].$$

What is the required sample size?

Using Chebyshev's inequality

$$\ell \geq \frac{\theta^2}{\alpha\varepsilon^2},$$

being θ^2 the variance of $x(\sigma)_i$.

Using Hoeffding's inequality

$$\ell \geq \frac{ln(2/\alpha) w_i^2}{2\varepsilon^2},$$

being w_i the range of $x(\sigma)_i$.

Thus, by Popoviciu's inequality,

$$\ell \geq \min\left\{\frac{ln(2/\alpha)}{2\varepsilon^2}, \frac{1}{4\alpha\varepsilon^2}\right\} w_i^2.$$

- For $\alpha \leq 0.23$, we use Hoeffding's inequality.

Maleki, S. (2015). Addressing the computational issues of the Shapley value with applications in the smart grid. PhD Thesis, Southampton University, Southampton.

Table of contents

The Banzhaf value

The Banzhaf value (Banzhaf, 1964) for (N, v), for every $i \in N$, is

$$Bz_i(N, v) = \frac{1}{2^{n-1}} \sum_{S \subseteq N \setminus \{i\}} (v(S \cup \{i\}) - v(S)).$$

Banzhaf, J. F. (1964). Weighted voting doesn't work: A mathematical analysis. Rutgers L. Rev., 19, 317.

A procedure based on simple random sampling with replacement

Bachrach, Y., Markakis, E., Resnick, E., Procaccia, A. D., Rosenschein, J. S., Saberi, A. (2010). Approximating power indices: theoretical and empirical analysis. Autonomous Agents and Multi-Agent Systems, 20(2), 105-122.

1. Fixed $i \in N$, we obtain a sample $\mathcal{S} = \{S_1, \ldots, S_\ell\}$ of S coalitions in $N \setminus \{i\}$.
2. For each $S \in \mathcal{S}$, we do $x(S)_i = v(S \cup \{i\}) - v(S)$.
3. The estimation of Bz_i, for all $i \in N$, is

$$\overline{Bz}_i = \frac{1}{\ell} \sum_{S \in \mathcal{S}} x(S)_i.$$

Unbiased and consistent estimator

Analysis of the error

A confidence interval for the Banzhaf value

$$\mathbb{P}(|\overline{Bz}_i - Bz_i| \leq \varepsilon) \geq 1-\alpha, \text{ with } \varepsilon > 0 \text{ and } \alpha \in (0,1].$$

What is the required sample size?

Using Chebyshev's inequality

$$\ell \geq \frac{\theta^2}{\alpha\varepsilon^2},$$

being θ^2 the variance of $x(\sigma)_i$.

Using Hoeffding's inequality

$$\ell \geq \frac{ln(2/\alpha) w_i^2}{2\varepsilon^2},$$

being w_i the range of $x(S)_i$.

Thus, by Popoviciu's inequality,

$$\ell \geq \min\left\{\frac{ln(2/\alpha)}{2\varepsilon^2}, \frac{1}{4\alpha\varepsilon^2}\right\} w_i^2.$$

- For $\alpha \leq 0.23$, we use Hoeffding's inequality.

Maleki, S. (2015). Addressing the computational issues of the Shapley value with applications in the smart grid. PhD Thesis, Southampton University, Southampton.

An application: the Galician milk conflict

- After the European milk quotas in March 2015...

Year	Jan.	Feb.	Mar.	April	May	June	July	Aug.	Sept.	Oct.	Nov.	Dec.
2013	32.65	32.66	32.76	32.84	33.06	33.03	34.44	34.86	35.59	38.57	38.93	39.17
2014	39.24	38.90	38.65	36.05	35.61	35.37	33.72	33.73	33.64	32.24	32.10	31.95
2015	30.52	30.60	30.30	28.80	28.40	27.90	27.60	27.70	28.30	28.70	28.70	28.80

Table: Averaged prices of the milk in Galicia (in euros per 100 litres) in the period 2013-2015.

How to increase the price of milk?

A low production scenario

- Reducing the milk production in Galicia?
- If we know a maximum production per municipality, how this reduction affect each of the 190 involved?
- We build a new system of quotas for councils.

*Data source: Consellería de Medio Rural da Xunta de Galicia.

A low milk production scenario

A new bankruptcy problem

- Set of agents: the 190 most representative councils.
- Estate: the tons of milk in 2014-2015 for Galicia reduces by $\rho\%$, $\rho \in (0, 100]$
- Claims: the capabilities of milk production (individual production of the councils, March 2015).

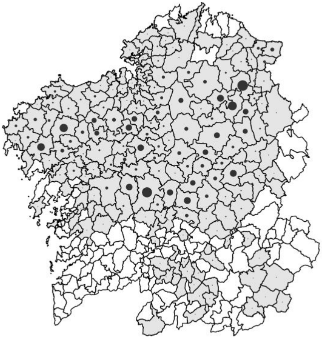
*Data source: Consellería de Medio Rural da Xunta de Galicia.

Computing the random arrival rule is a difficult task!

190! permutations to be evaluated

Saavedra-Nieves, A., Saavedra-Nieves, P. (2020). On systems of quotas from bankruptcy perspective: the sampling estimation of the random arrival rule. European Journal of Operational Research, 285(2), 655-669.

A low milk production scenario: a case study



What about variability?

- We approximate the *RA*-rule for the most representative councils in Galicia.
- We obtain 100 estimations and we do some basic statistics.
- Differences in the results seem bearable.

Top 10 of councils with the largest milk production for Galicia

Council	A Pastoriza	Lalín	Castro de Rei	Santa Comba	Mazaricos	Chantada	Cospeito	Sarria	Silleda	Arzúa
Maximum	52499.22	49633.35	43188.07	40002.56	36723.09	32731.26	32668.26	32068.42	31319.75	30539.21
Average	52472.49	49593.97	43161.96	39973.33	36696.91	32708.51	32645.75	32050.23	31300.44	30519.81
Minimum	52425.39	49561.72	43135.33	39942.24	36668.44	32681.45	32626.93	32030.53	31281.85	30502.95

Table: Summary of 100 estimations of the milk quotas for the councils with $\rho = 40\%$.

Table of contents

The Owen value

Let (N, v) be a TU-game and $P = \{P_1, \ldots, P_m\}$ is a partition of N.

The Owen value (Owen, 1977) for (N, v, P), for every $i \in N$, is

$$O_i(N, v, P) = \frac{1}{|\Pi_P(N)|} \sum_{\sigma \in \Pi_P(N)} (v(P_i^\sigma \cup \{i\}) - v(P_i^\sigma)),$$

being $\Pi_P(N)$ the set of all compatible orders of N with P and
$P_i^\sigma = \{j \in N : \sigma(j) < \sigma(i)\}$ for any $\sigma \in \Pi_P(N)$.

- $\sigma \in \Pi_P(N)$ *compatible with* P:

$$\forall i, j \in T \in P,\ \forall k \in N,\ \sigma(i) < \sigma(k) < \sigma(j) \Rightarrow k \in T.$$

- It is a variation of the Shapley value (Shapley, 1953).

Shapley, L. S. (1953). A value for n-person games. In: Kuhn, H. W. and Tucker, A. W. (Eds.), Contributions to the Theory of Games II, Princeton University Press, Princeton, NJ, 307-317.

Owen, G. (1977). Values of games with a priori unions. In R. Henn and O. Moeschlin (Eds.), Mathematical economics and game theory (pp. 76–88). Springer.

Estimating the Owen value

Saavedra-Nieves, A., García-Jurado, I., Fiestras-Janeiro, M. G. (2018). Estimation of the Owen value based on sampling. In The Mathematics of the Uncertain (pp. 347-356). Springer, Cham.

A procedure based on simple random sampling with replacement

1. We generate a sample with replacement $\mathcal{S} = \{\sigma_1, \ldots, \sigma_\ell\}$ of ℓ elements in $\Pi_P(N)$.
2. For each $\sigma \in \mathcal{S}$, $x(\sigma)_i = v(P_i^\sigma \cup \{i\}) - v(P_i^\sigma)$, for all $i \in N$.
3. The estimation of O_i is

$$\hat{O}_i = \frac{1}{\ell} \sum_{\sigma \in \mathcal{S}} x(\sigma)_i, \text{ for all } i \in N.$$

Unbiased and consistent estimator

A confidence interval for the Owen value

$$\mathbb{P}(|\hat{O}_i - O_i| \leq \varepsilon) \geq 1 - \alpha, \text{ with } \varepsilon > 0 \text{ and } \alpha \in (0, 1].$$

Popoviciu's inequality:

$$\ell \geq \min\left\{\frac{ln(2/\alpha)}{2\varepsilon^2}, \frac{1}{4\alpha\varepsilon^2}\right\} w_i^2.$$

Table of contents

The Banzhaf-Owen value

The Banzhaf-Owen value (Owen, 1982) for (N, v, P), for every $i \in N$, is

$$BzO_i(N,v,P) = \sum_{R \subseteq P \setminus P_{(i)}} \frac{1}{2^{m-1}} \sum_{S \subseteq P_{(i)} \setminus \{i\}} \frac{1}{2^{p_i-1}} \left(v(\underset{P_l \in R}{\cup} P_l \cup S \cup \{i\}) - v(\underset{P_l \in R}{\cup} P_l \cup S) \right),$$

where $i \in P_{(i)} \in P$ and with $p_i = |P_{(i)}|$.

- For $i \in N$, $T \subseteq N \setminus \{i\}$ is *compatible with* P for i:

$$T = \underset{P_l \in R}{\cup} P_l \cup S \text{ for } R \subseteq P \setminus P_{(i)} \text{ and } S \subseteq P_{(i)} \setminus \{i\}.$$

- It is a variation of the Banzhaf value (Banzhaf, 1964).

Banzhaf, J. F. (1964). Weighted voting doesn't work: A mathematical analysis. Rutgers L. Rev., 19, 317.

Owen, G. (1982). Modification of the Banzhaf-Coleman index for games with a priori unions. In M. J. Holler (Ed.), Power, voting, and voting power (pp. 232–238). Physica-Verlag HD.

Estimating the Banzhaf-Owen value

Saavedra-Nieves, A., Fiestras-Janeiro, M. G. (2021). Sampling methods to estimate the Banzhaf–Owen value. Annals of Operations Research, 301(1), 199-223.

Let (N, v, P) be a game with $P = \{P_1, \dots, P_m\}$.

- We take $N^* = \{k : P_k \in P \setminus P_{(i)}\} \cup \{j : j \in P_{(i)}\}$.
- Each $T \subseteq N^*$ is given by $T = \bigcup_{P_l \in R} P_l \cup S$ with $R \subseteq P \setminus P_{(i)}$ and $S \subseteq P_{(i)}$.

A procedure based on simple random sampling without replacement

1. We generate a sample without replacement $\mathcal{T} = \{T_1, \dots, T_\ell\}$ of ℓ coalitions in $N^* \setminus \{i\}$.
2. For each $T_j \in \mathcal{T}$,

$$x(R_j, S_j)_i = v(\bigcup_{P_l \in R_j} P_l \cup S_j \cup \{i\}) - v(\bigcup_{P_l \in R_j} P_l \cup S_j),$$

   being $R_j \subseteq P \setminus P_{(i)}$ and $S_j \subseteq P_{(i)} \setminus \{i\}$ such that $T_j = \{k : P_k \in R_j\} \cup S_j$.

3. The estimation of BzO_i is $\overline{BzO}_i = \frac{1}{\ell} \sum_{j=1}^{\ell} x(R_j, S_j)_i$, for all $i \in N$.

Unbiased and consistent estimator

Analysis of the error

A confidence interval for the Banzhaf-Owen value

$$\mathbb{P}\big(|\overline{BzO}_i - BzO_i| \leq \varepsilon\big) \geq 1-\alpha, \text{ with } \varepsilon > 0 \text{ and } \alpha \in (0,1].$$

What is the required sample size?

- No replacement implies the dependence of the sampled units.

Using Chebyshev's inequality

$$\ell \geq \frac{\theta^2 2^{m-1} 2^{p_i-1}}{\alpha\varepsilon^2 2^{m-1} 2^{p_i-1} + \theta^2},$$

being θ^2 the variance of $x(R,S)_i$.

Using Serfling's inequality

$$\ell \geq \min\left\{ \frac{ln(2/\alpha) w_i^2 (2^{m-1} 2^{p_i-1} + 1)}{ln(2/\alpha) w_i^2 + 2\varepsilon^2 2^{m-1} 2^{p_i-1}}, 2^{m-1} 2^{p_i-1} \right\},$$

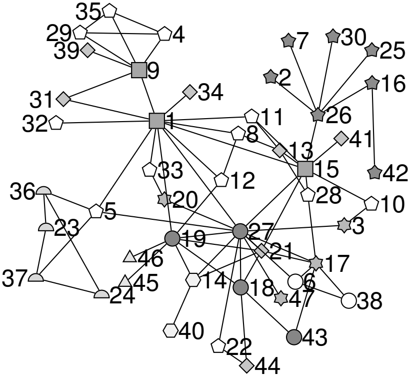
being w_i the range of $x(R,S)_i$.

Thus, by Popoviciu's inequality,

$$\ell \geq \min\left\{ \frac{w_i^2 2^{m-1} 2^{p_i-1}}{4\alpha\varepsilon^2 2^{m-1} 2^{p_i-1} + w_i^2}, \frac{ln(2/\alpha) w_i^2 (2^{m-1} 2^{p_i-1} + 1)}{ln(2/\alpha) w_i^2 + 2\varepsilon^2 2^{m-1} 2^{p_i-1}} \right\}.$$

An application: the analysis of covert networks

The Zerkani network



- We rank the members of the Zerkani network, responsible for the attacks in Paris (2015) and Brussels (2016).
- It contains 47 members.
- We consider TU-games with a priori unions.

E. Algaba, A. Prieto, A. Saavedra-Nieves, H. Hamers (2022). Analyzing the Zerkani network with the Owen value. Collective Decisions - Interdisciplinary Perspectives for the 21st Century. Studies in Choice and Welfare. *To appear*.

E. Algaba, A. Prieto, A. Saavedra-Nieves (2022). Ranking the Zerkani network by sampling methods based on the Banzhaf value. Submitted to Applied Mathematics & Computation.

An application: the analysis of covert networks

The effectiveness of a coalition

$$f(S,\mathcal{I},\mathcal{R}) = \begin{cases} \left(\sum_{i\in S} w_i\right) \cdot \max_{ij\in E_S} k_{ij}, & \text{if } |S| > 1,\\ w_S, & \text{if } |S| = 1. \end{cases}$$

- w_i, individual weights for all $i \in N$.
- k_{ij}, weight of the link ij, with $i, j \in N$.

Σ_S denotes the set of connected subcoalitions of S in the graph.

The weighted connectivity TU-game and the additive weighted connectivity game

$$v^{\text{wconn}}(S) = \begin{cases} f(S,\mathcal{I},\mathcal{R}), & \text{if } S \text{ connected},\\ \max_{T\in\sum_S} v^{\text{wconn}}(T), & \text{if } S \text{ disconnected}. \end{cases}$$

$$v^{\text{awconn}}(S) = \begin{cases} f(S,\mathcal{I},\mathcal{R}), & \text{if } S \text{ connected},\\ \sum_{T\in\sum_S} v^{\text{awconn}}(T), & \text{if } S \text{ disconnected}. \end{cases}$$

Husslage, B., Borm, P., Burg, T., Hamers, H., and Lindelauf, R. (2015). Ranking terrorists in networks: A sensitivity analysis of al qaeda's 9/11 attack. Social Networks, 42, 1–7.

Lindelauf, R., Hamers, H. J., and Husslage, B. (2013). Cooperative game theoretic centrality analysis of terrorist networks: The cases of jemaah islamiyah and al qaeda. European Journal of Operational Research, 229(1), 230–238

An application: the analysis of covert networks

Rankings of terrorists in Zerkani network

Ranking R_{wconn}				Ranking R_{awconn}			
Terrorist	$\overline{Sh}$	Terrorist	$\overline{O}$	Terrorist	$\overline{Sh}$	Terrorist	$\overline{O}$
Ab. Abaaoud	17.108	Khalid Zerkani	39.242	Mohamed Belkaid	13.987	Mohamed Belkaid	28.460
Khalid Zerkani	15.026	Ab. Abaaoud	36.129	Khalid Zerkani	12.332	Khalid Zerkani	27.677
Salah Abdeslam	14.741	Mohamed Belkaid	29.236	Ab. Abaaoud	11.850	Mohamed Bakkali	27.168
Mohamed Belkaid	14.249	Mohamed Bakkali	27.845	Salah Abdeslam	11.453	Ab. Abaaoud	26.157
Najim Laachraoui	7.918	Salah Abdeslam	27.042	Fabien Clain	8.295	Salah Abdeslam	22.439
Mohamed Bakkali	7.356	Fabien Clain	13.661	Mohamed Bakkali	7.625	Fabien Clain	16.404
Fabien Clain	5.884	Reda Kriket	11.642	Najim Laachraoui	7.549	Reda Kriket	11.395
Reda Kriket	3.696	Ahmed Dahmani	10.754	Reda Kriket	4.923	Ahmed Dahmani	6.142
Ahmed Dahmani	3.369	Khaled Ledjeradi	4.702	Mohamed Abrini	2.996	Miloud F.	6.093
Mohamed Abrini	2.917	Miloud F.	4.209	Miloud F.	2.827	Khaled Ledjeradi	5.271
Terrorist	$\overline{Bz}$	Terrorist	$\overline{BzO}$	Terrorist	$\overline{Bz}$	Terrorist	$\overline{BzO}$
Ab. Abaaoud	38.372	Khalid Zerkani	39.328	Mohamed Belkaid	31.333	Mohamed Belkaid	32.274
Salah Abdeslam	34.993	Ab. Abaaoud	35.702	Salah Abdeslam	26.114	Khalid Zerkani	27.854
Khalid Zerkani	33.992	Salah Abdeslam	33.639	Khalid Zerkani	25.752	Salah Abdeslam	27.010
Mohamed Belkaid	33.144	Mohamed Belkaid	33.400	Ab. Abaaoud	24.206	Ab. Abaaoud	25.426
Najim Laachraoui	18.827	Mohamed Bakkali	22.473	Mohamed Bakkali	18.360	Mohamed Bakkali	22.181
Mohamed Bakkali	18.367	Fabien Clain	12.298	Najim Laachraoui	17.381	Fabien Clain	15.892
Fabien Clain	11.903	Ahmed Dahmani	9.900	Fabien Clain	16.538	Reda Kriket	10.830
Reda Kriket	8.316	Reda Kriket	9.166	Reda Kriket	10.620	Miloud F.	5.916
Ahmed Dahmani	8.111	Najim Laachraoui	4.740	Mohamed Abrini	6.242	Ahmed Dahmani	5.712
Mohamed Abrini	6.924	Mohamed Abrini	4.621	Miloud F.	5.833	Khaled Ledjeradi	5.379

- Rankings based on the estimations of the Shapley value, the Banzhaf value, the Owen value, and the Banzhaf-Owen value.

Table of contents

The Spanish hotel industry

The hotel industry in Spain*

- The hotel industry generated 12% of GDP.
- 83.7 million annual travellers.
- Spain was one the World's favourite destinations in 2019.

The crisis of the hotel industry

- After corona crisis, reservations collapse.
- Join resources minimizes the impact of crisis on the economy.
- The UN World Tourism Organization seeks a more efficient tourism.

*Data source: United Nations World Tourism Organization (UNWTO, `https://www.unwto.org/`).

Saavedra-Nieves, A., Fiestras-Janeiro, M. G. (2022). Analysis of the impact of DMUs on the overall efficiency in the event of a merger. Expert Systems With Applications, 195, 116571.

Assessing the efficiency of DMUs

Data environment analysis (DEA)

- The efficiency of a set of Decision Making Units (DMUs) is evaluated.

$$
\begin{array}{lll}
\textbf{MP} & \max & \eta_{i_0} \\
& \text{subject to} & \\
& & -\sum_{i\in N} y_{ri}\lambda_i + y_{ri_0}\eta_{i_0} \le 0,\ r = 1,\dots,s \\
& & \sum_{i\in N} x_{ji}\lambda_i \le x_{ji_0},\ j = 1,\dots,m \\
& & \lambda_i \ge 0,\ \forall i \in \{1,\dots,n\} \\
& & \eta_{i_0} \in \mathbb{R}.
\end{array}
$$

Charnes, A., Cooper, W. W., and Rhodes, E. (1978). Measuring the efficiency of decision making units. European Journal of Operational Research, 2(6), 429–444.

Let $N = \{1,\dots,n\}$ be a system of DMUs. Each of them is characterized by:

- m inputs. x_{ki} is the amount of input k, with $k = 1,\dots,m$, of DMU i, for every $i \in N$.
- s outputs. y_{ki} is the amount of output k, with $k = 1,\dots,s$, produced by DMU i, for every $i \in N$.

This problem is known as a multi-agent DEA problem and denoted by $(N; X; Y)$.

A new approach of cooperation

- As a novelty, the overall efficiency of a merger of DMUs is influenced by the organization of the remainder of DMUs and their mergers.
- That is, the existence of coalition structure describing the affinities of $N \setminus S$ influences into the relative efficiency of $[i_S]$.

Take $P \in \Pi(N)$ a coalition structure for the players in N.

- We define the artificial DMUs $[i_S]$, for every $S \in P$.
- The new set of DMUs is given by $N^P = \{[i_S] : \ S \in P\}$.
- $x_{[i_S]}$ and $y_{[i_S]}$ are the inputs and outputs resulting from their aggregation.

Thus, this *DEA problem* will be denoted by $(N^P; X^P; Y^P)$.

Kritikos, M. N. (2017). A full ranking methodology in data envelopment analysis based on a set of dummy decision making units. Expert Systems with Applications, 77, 211–225.

On the cooperation of DMUs under externalities

A multi-agent DEA problem under externalities is denoted by $(N; X; Y)$.

DEA sum games

Thrall, R. M., and Lucas, W. F. (1963). N-person games in partition function form. Naval Research Logistics Quarterly, 10(1), 281–298.

A *DEA partition function form game* $(N; X; Y; \boldsymbol{e})$ (or simply, $\boldsymbol{e}$) is defined as follows:

$$\boldsymbol{e}(S; P) = \begin{cases} \frac{1}{\eta^*_{[i_S],P}}, & \text{if } \emptyset \neq S \subseteq N,\ P \in \Pi(N \setminus S), \\ 0, & \text{otherwise,} \end{cases}$$

being $\eta^*_{[i_S],P}$ the optimal value of Problem MP to $(N^{P\cup\lceil S\rceil}, X^{P\cup\lceil S\rceil}, Y^{P\cup\lceil S\rceil})$, with $P\cup\lceil S\rceil \in \Pi(N)$ having S as a block[1], for DMU $i_0 = [i_S]$.

[1] $\lceil S\rceil$ means that S is an element of the partition.

Ranking DMUs under externalities

- Task: ranking DMUs in multi-agent DEA problems with externalities.
- DMUs' efficiency is used as criterion.
- To this aim, values or solutions for TU games are used.

The Shapley value of (N, v)

$$Sh_i(N,v) = \sum_{T \subseteq N \setminus \{i\}} \frac{|T|!\,(|N|-|T|-1)!}{|N|!}(v(T \cup \{i\}) - v(T)), \text{ for every } i \in N.$$

Solutions for a game with externalities e

- de Clippel and Serrano (2008): $Sh(N, e_{min})$.
- McQuillin (2009): $Sh(N, e_{max})$.
- Albizuri et al. (2005): $Sh(N, \bar{e})$.
- Hu and Yang (2010): $Sh(N, \bar{\bar{e}})$.

Ranking DMUs under externalities

Let e be a partition function form game. For every $S \subseteq N$,

$$e_{\max}(S) = \max_{P \in \Pi(N \setminus S)} e(S; P) \text{ and } e_{\min}(S) = \min_{P \in \Pi(N \setminus S)} e(S; P).$$

The TU game $(N, \bar{e})$ of Albizuri

$$\bar{e}(S) = \frac{1}{|\Pi(N \setminus S)|} \sum_{Q \in \Pi(N \setminus S)} e(S; Q), \text{ for each } S \subseteq N.$$

The TU game $(N, \bar{\bar{e}})$ of Hu & Yang

$$\bar{\bar{e}}(S) = \frac{1}{|\Pi(N)|} \sum_{P \in \Pi(N)} e(S; P_{-S}), \text{ for each } S \subseteq N.$$

- Albizuri, M. J., Arin, J., and Rubio, J. (2005). An axiom system for a value for games in partition function form. International Game Theory Review, 7(01), 63–72.
- Hu, C.-C., and Yang, Y.-Y. (2010). An axiomatic characterization of a value for games in partition function form. SERIEs, 1(4), 475–487.

Estimating the TU game of Hu and Yang

- The obtaining of the characteristic function of the TU game of Hu and Yang complicates for a “large” amount of players.

A procedure based on simple random sampling with replacement

Take $S \subseteq N$.

1. We generate a sample with replacement $\mathcal{S}_{\mathcal{P}} = \{P_1, \ldots, P_\ell\}$ of ℓ partitions of N.
2. For each $P \in \mathcal{S}_{\mathcal{P}}$, we obtain
$$e(S; P_{-S}),$$
being P_{-S} the partition induced by P for $N \setminus S$.
3. The estimation of $\overline{\overline{e}}(S)$ is
$$\widehat{\widehat{e_S}} = \frac{1}{\ell} \sum_{P \in \mathcal{S}_{\mathcal{P}}} e(S; P_{-S}).$$

Estimating the TU game of Hu and Yang

- The obtaining of the characteristic function of the TU game of Hu and Yang complicates for a "large" amount of players.

A procedure based on simple random sampling with replacement

Take $S \subseteq N$.

1. We generate a sample with replacement $\mathcal{S}_{\mathcal{P}} = \{P_1, \ldots, P_\ell\}$ of ℓ partitions of N.
2. For each $P \in \mathcal{S}_{\mathcal{P}}$, we obtain
$$e(S; P_{-S}),$$
being P_{-S} the partition induced by P for $N \setminus S$.
3. The estimation of $\overline{\overline{e}}(S)$ is
$$\widehat{\widehat{e_S}} = \frac{1}{\ell} \sum_{P \in \mathcal{S}_{\mathcal{P}}} e(S; P_{-S}).$$

Unbiased and consistent estimator

Estimating the TU game of Albizuri

- The task of obtaining the TU game of Albizuri et al. is also complicated.

A procedure based on simple random sampling with replacement

Take $S \subseteq N$.

1. We generate a sample with replacement $\mathcal{S}_{\mathcal{P}} = \{P_1, \ldots, P_{\ell_S}\}$ of ℓ_S partitions of $N \setminus S$.
2. For each $P \in \mathcal{S}_{\mathcal{P}}$, we obtain
$$e(S; P),$$
being P a sampled partition for $N \setminus S$.
3. The estimation of $\overline{e}(S)$ is
$$\widehat{e_S} = \frac{1}{\ell_S} \sum_{P \in \mathcal{S}_{\mathcal{P}}} e(S; P).$$

Estimating the TU game of Albizuri

- The task of obtaining the TU game of Albizuri et al. is also complicated.

A procedure based on simple random sampling with replacement

Take $S \subseteq N$.

1. We generate a sample with replacement $\mathcal{S}_{\mathcal{P}} = \{P_1, \ldots, P_{\ell_S}\}$ of ℓ_S partitions of $N \setminus S$.
2. For each $P \in \mathcal{S}_{\mathcal{P}}$, we obtain
$$e(S; P),$$
being P a sampled partition for $N \setminus S$.
3. The estimation of $\overline{e}(S)$ is
$$\widehat{e_S} = \frac{1}{\ell_S} \sum_{P \in \mathcal{S}_{\mathcal{P}}} e(S; P).$$

Unbiased and consistent estimator

Table of contents

Estimating the Shapley value

- The Shapley value of the estimated TU games are natural estimators.

The case of Hu and Yang

$$\widehat{\widehat{Sh}}_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N|-|S|-1)!}{|N|!} \left(\widehat{\widehat{e}}_{S\cup\{i\}} - \widehat{\widehat{e_S}}\right), \text{ for all } i \in N.$$

The case of Albizuri et al.

$$\widehat{Sh}_i = \sum_{S \subseteq N \setminus \{i\}} \frac{|S|!\,(|N|-|S|-1)!}{|N|!} \left(\widehat{e}_{S\cup\{i\}} - \widehat{e_S}\right), \text{ for all } i \in N.$$

- Both estimators are unbiased and consistent.

But we have no bounds of the error!

The hotel industry in Spain

DMUs, the set of 17 autonomous communities and 2 autonomous cities.

Inputs, for measuring the hotel industry's potential in Spain.

- Number of hotels.
- Number of occupied bed places.
- Number of employees.

Outputs, as result of managing the existing resources.

- Number of hotel guests.
- Number of the overnight stays.
- Number of the occupied accommodations.

Data set extracted from the *Instituto Nacional de Estadística (INE)*

Task: rank the 17 regions and the 2 autonomous cities using DEA sum games.

- We measure the capability for increasing the overall efficiency in the event of a merger.

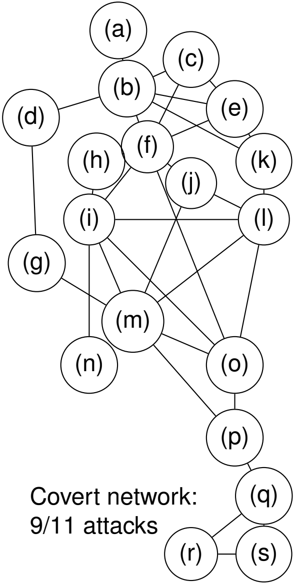
Saavedra-Nieves, A., Fiestras-Janeiro, M. G. (2022). Analysis of the impact of DMUs on the overall efficiency in the event of a merger. Expert Systems With Applications, 195, 116571.

The hotel industry in Spain

Region	Exact rankings				Estimated rankings			
	(CS)	Rank	(MQ)	Rank	(A)	Rank	(HY)	Rank
Andalucía	0.04017	14	0.05232	11	0.04237	15	0.04133	14
Aragón	0.05117	7	0.05181	15	0.05287	8	0.05202	8
Principado de Asturias	0.03417	17	0.05194	14	0.03704	17	0.03606	17
Illes Balears	0.10943	1	0.05596	1	0.09270	1	0.09902	1
Canarias	0.05665	6	0.05031	19	0.05545	6	0.05620	6
Cantabria	0.03790	16	0.05223	12	0.04122	16	0.03985	16
Castilla y León	0.04094	13	0.05131	17	0.04371	13	0.04266	13
Castilla - La Mancha	0.04737	9	0.05156	16	0.04882	9	0.04791	9
Cataluña	0.07229	4	0.05518	3	0.07094	3	0.07228	3
Comunitat Valenciana	0.07726	3	0.05389	4	0.07076	4	0.07227	4
Extremadura	0.03001	18	0.05205	13	0.03341	18	0.03225	18
Galicia	0.02533	19	0.05094	18	0.02817	19	0.02707	19
Comunidad de Madrid	0.10352	2	0.05537	2	0.08920	2	0.09413	2
Región de Murcia	0.04227	11	0.05243	9	0.04697	11	0.04587	11
Com. Foral de Navarra	0.04212	12	0.05234	10	0.04543	12	0.04426	12
País Vasco	0.05668	5	0.05262	6	0.05802	5	0.05753	5
La Rioja	0.04391	10	0.05245	8	0.04718	10	0.04607	10
Ceuta	0.05090	8	0.05263	5	0.05306	7	0.05214	7
Melilla	0.03793	15	0.05261	7	0.04271	14	0.04109	15

Table: Rankings under the approaches of (CS) and (MQ), and under the ones of (A) and of (HY) with sampling.

An application: the analysis of covert networks



Covert network:
9/11 attacks

The effectiveness of a coalition

$$f(S,\mathcal{I},\mathcal{R}) = \begin{cases} \left(\sum_{i\in S} w_i\right) \cdot \max_{ij\in E_S} k_{ij}, & \text{if } |S|>1,\\ w_S, & \text{if } |S|=1. \end{cases}$$

- w_i, individual weights for all $i \in N$.
- k_{ij}, weight of the link ij, with $i, j \in N$.

Σ_S denotes the set of connected subcoalitions of S in the graph.

A *covert nwetwork game* $v_{G,f}(S;P)$ is defined as follows:

$$v_{G,f}(S;P) = \begin{cases} 1, & \text{if } \max_{T\in\Sigma_S} f(T,\mathcal{I},\mathcal{R}) \geq \max_{T\in\Sigma_{S'}} f(T,\mathcal{I},\mathcal{R}),\\ & \forall S' \in P, \text{ with } \emptyset \neq S \subseteq N,\ P \in \Pi(N\setminus S),\\ 0, & \text{otherwise,} \end{cases}$$

We identify the most effective coalitions by using f.

We rank the members of a covert network using solutions for games with externalities.

An application: the analysis of covert networks

Rankings of hijackers involved in 9/11 attacks

	Shapley value (SH)		Gen. Eg. Shapley value (GES)		Solidarity value (S)		Banzhaf value (BZ)	
	Hijacker	Alloc.	Hijacker	Alloc.	Hijacker	Alloc.	Hijacker	Alloc.
1	Salem Alhazmi	0.12290	Salem Alhazmi	0.08735	Salem Alhazmi	0.05957	Salem Alhazmi	0.33510
2	Khalid Al-Mihdhar	0.11923	Khalid Al-Mihdhar	0.08551	Khalid Al-Mihdhar	0.05923	Khalid Al-Mihdhar	0.32247
3	Ziad Jarrah	0.11258	Ziad Jarrah	0.08226	Ziad Jarrah	0.05856	Ziad Jarrah	0.30589
4	Mohamed Atta	0.10337	Mohamed Atta	0.07762	Mohamed Atta	0.05765	Mohamed Atta	0.28439
5	Hani Hanjour	0.09800	Hani Hanjour	0.07499	Hani Hanjour	0.05714	Hani Hanjour	0.26511
6	Ahmed Al-Haznawi	0.07050	Ahmed Al-Haznawi	0.06141	Ahmed Al-Haznawi	0.05441	Majed Moqed	0.19389
7	Majed Moqed	0.07017	Majed Moqed	0.06129	Majed Moqed	0.05436	Ahmed Al-Haznawi	0.19310
8	Marwan Al-Shehhi	0.05154	Marwan Al-Shehhi	0.05211	Marwan Al-Shehhi	0.05252	Marwan Al-Shehhi	0.14298
9	Hamza Alghamdi	0.04792	Hamza Alghamdi	0.05030	Hamza Alghamdi	0.05216	Hamza Alghamdi	0.13330
10	Nawaf Alhazmi	0.04179	Nawaf Alhazmi	0.04723	Nawaf Alhazmi	0.05155	Nawaf Alhazmi	0.11959
11	Saeed Alghamdi	0.03880	Saeed Alghamdi	0.04586	Saeed Alghamdi	0.05125	Saeed Alghamdi	0.10737
12	Fayez Ahmed	0.02810	Fayez Ahmed	0.04053	Fayez Ahmed	0.05020	Fayez Ahmed	0.07893
13	Mohand Alshehri	0.02360	Mohand Alshehri	0.03830	Mohand Alshehri	0.04976	Mohand Alshehri	0.06639
14	Ahmed Alnami	0.02251	Ahmed Alnami	0.03780	Ahmed Alnami	0.04965	Ahmed Alnami	0.06381
15	Abdul Aziz Al-Omari	0.01839	Abdul Aziz Al-Omari	0.03572	Abdul Aziz Al-Omari	0.04924	Abdul Aziz Al-Omari	0.05420
16	Satam Suqami	0.01261	Satam Suqami	0.03287	Satam Suqami	0.04868	Satam Suqami	0.03651
17	Ahmed Alghamdi	0.00944	Ahmed Alghamdi	0.03129	Ahmed Alghamdi	0.04837	Ahmed Alghamdi	0.02716
18	Waleed Alshehri	0.00428	Waleed Alshehri	0.02878	Waleed Alshehri	0.04786	Waleed Alshehri	0.01268
19	Wail Alshehri	0.00428	Wail Alshehri	0.02878	Wail Alshehri	0.04786	Wail Alshehri	0.01268

Table: Rankings based on the estimation of the TU-games of Albizuri.

- Rankings based on the TU-games of Hu and Yang, de Clippel and Serrano, and McQuillin.

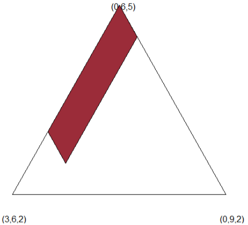
Saavedra-Nieves, A., Casas-Méndez, B. (2022). On the centrality analysis of covert networks using games with externalities. Submitted to European Journal of Operational Research.

Table of contents

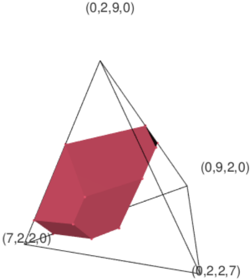
The core of a TU-game

The core

$$C(N, v) = \left\{ x \in \mathbb{R}^{|N|} \ : \ \sum_{i \in N} x_i = v(N) \text{ and } \sum_{i \in T} x_i \geq v(T), \text{ for each } T \subset N \right\}.$$



3 players



4 players

Saavedra-Nieves, A., Saavedra-Nieves, P. (2021). On core reconstruction for TU games from nonparametric set estimation techniques. *Preprint.*

The core of a TU-game

The core

$$C(N, v) = \left\{ x \in \mathbb{R}^{|N|} \ : \ \sum_{i \in N} x_i = v(N) \text{ and } \sum_{i \in T} x_i \geq v(T), \text{ for each } T \subset N \right\}.$$

Related literature

- Avis and Fukuda (1992) computes the core vertices in time $O(h^2 n\nu)$, h being the number of inequalities and ν, the number of vertices.
- The exact computation of the core vertices reaches exponential time complexity.
- Derks and Kuipers (2002) upper-bounded the number of vertices by $n!$.

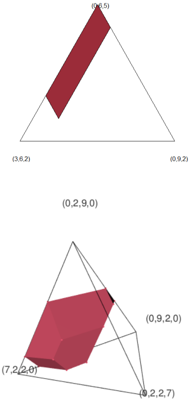
Set estimation for reconstructing the core in polynomial time, $\hat{C}(N, v)$.

Avis, D., Fukuda, K. (1992). A pivoting algorithm for convex hulls and vertex enumeration of arrangements and polyhedra. Discrete & Computational Geometry, 8(3), 295–313.

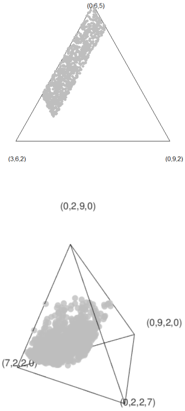
Derks, J., Kuipers, J. (2002). On the number of extreme points of the core of a transferable utility game. In Chapters in game theory (pp. 83–97). Springer.

Core reconstruction

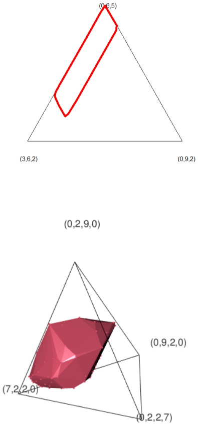
Estimation of a set (or its characteristic features such as its vertices or its volume) from a random sample of points.



Exact core



Sample of points



Core reconstruction

Core reconstruction: an algorithm

Take (N, v) a TU-game.

A procedure based on sampling for reconstructing the core $C(N, v)$

1. The set to be estimated is $C(N, v)$, the set of stable allocations for N.
2. $C(N, v)$ is considered as the support of a uniform distribution.
3. A uniform sample of size m supported on $C(N, v)$, $\mathcal{X}_m$, has to be generated.
4. Compute the convex hull of $\mathcal{X}_m$, $H(\mathcal{X}_m)$, as the reconstruction of $C(N, v)$, $\hat{C}(N, v)$.

Note: If (N, v) is convex, we will use a sample of vectors of marginal contributions (vertices) in (3).

Some comments

- This algorithm allows to approximate the core in polynomial time.
- The resulting core estimator is mathematically consistent (Dümbgen and Walther, 1996)

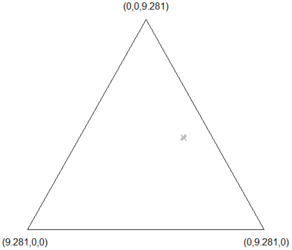
Dümbgen, L., Walther, G. (1996). Rates of convergence for random approximations of convex sets. Advances in applied probability, 28(2), 384–393.

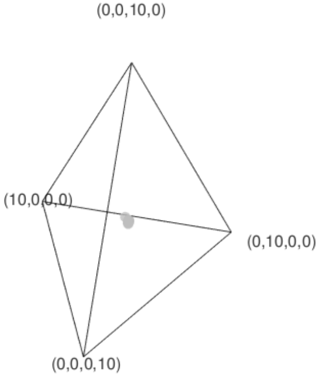
Table of contents

A geometrical application: estimating the core-center

Take (N, v) a TU-game.

- The core-center of (N, v), $cc(v)$, is the barycenter of the convex and compact polyhedron determined by $C(N, v)$.
- The natural estimator $\hat{cc}_H(v)$ of this allocation rule emerges by considering the centroid of the core reconstruction.





Some comments on reconstructing the core

- We illustrate the performance of this proposal os 3 and 4 player situations.
- However, it can be applied in a general setting.
- The main limitation is imposed by the characteristic of the used computers.
- We have been able to apply it until 16 players.
- R software has specific libraries for this purpose.

Concluding remarks

- We have reviewed different proposals based on sampling for the approximation of TU game solutions.
- These allow us to provide solutions to problems where there is no alternative in polynomial time for their computation.
- The proposed methodologies are sufficiently robust to guarantee the quality of the estimates obtained.
- The main computational difficulties are in the handling of the information when the number of players is very large.
- Most of the implemented methodologies are available in R software, although not specifically for this context.
- The characteristics of the machine to be used determine the computational speed.

Sampling techniques for the approximation of solutions for TU games

Alejandro Saavedra-Nieves

(based on joint works with E. Algaba, B. Casas-Méndez, M. G. Fiestras-Janeiro, I. García-Jurado and P. Saavedra-Nieves)

USC UNIVERSIDADE DE SANTIAGO DE COMPOSTELA

Saint-Étienne, 15th April 2022