A Computational Measure of Heterogeneity on Mathematical Skills

Eugenio M. Fedriani Rafael Moyano

Dept of Economics, Quantitative Methods and Economic History

Introduction

Education:

- Necessity of analyze and evaluate
- Many variables are involved (multivariate)
- Subdivisions of analysis units

Introduction

Education:

- Necessity of analyze and evaluate
- Many variables are involved (multivariate)
- Subdivisions of analysis units

Study of multivariate phenomena:

- Statistically
- Using topological tools (new)

Introduction

Study of multivariate phenomena:

- Economic and Business Management (Kaufmann, Gil-Aluja, 1995)
- Economic study of poverty (Martín, Mínguez, 2005)
- Input-Output Analysis (Fedriani, Tenorio, 2006)
- Economic study of tourism (Ballesteros, Hernández, Frdriani, 2008)

Introduction

Study of multivariate phenomena:

- Economic and Business Management (Kaufmann, Gil-Aluja, 1995)
- Economic study of poverty (Martín, Mínguez, 2005)
- Input-Output Analysis (Fedriani, Tenorio, 2006)
- Economic study of tourism (Ballesteros, Hernández, Frdriani, 2008)

Our proposal: APPLICATIONS ON EDUCATION

An application

Objective: To study the homogeneity in mathematical skills of the different groups of students in 3rd course from two Andalusian comprehensive secondary schools

An application

Objective: To study the homogeneity in mathematical skills of the different groups of students in 3rd course from two Andalusian comprehensive secondary schools

Data:

- CENTER 1: Groups $3A$, $3B$, $3C$, $3D$
- CENTER 2: Groups $3A'$, $3B'$, $3C'$, $3D'$, Div_1, Div_2
- Mathematical skills: diagnostic tests in Maths by the Junta de Andalucía. Academic year 2007/08 (census)

Intensity Topological Indicator

Elements:

- Universe U

Intensity Topological Indicator

Elements:

- Universe U
- Analysis Units: A finite partition $\{C_1, ..., C_n\}$ of U (i.e. $\bigcup_{i=1}^{n} C_i = U$, $C_i \cap C_j = \emptyset$, $i \neq j$)

Intensity Topological Indicator

Elements:

- Universe U
- Analysis Units: A finite partition $\{C_1, ..., C_n\}$ of U (i.e. $\bigcup_{i=1}^{n} C_i = U$, $C_i \cap C_j = \emptyset$, $i \neq j$)
- Components: Finite partitions of every analysis unit.

Intensity Topological Indicator

Elements:

- Universe U
- Analysis Units: A finite partition $\{C_1, ..., C_n\}$ of U (i.e. $\bigcup_{i=1}^{n} C_i = U$, $C_i \cap C_j = \emptyset$, $i \neq j$)
- Components: Finite partitions of every analysis unit.
- Variables: X_1, X_2,... ,X_k

Intensity Topological Indicator

Elements:

- Universe U
- Analysis Units: A finite partition $\{C_1, ..., C_n\}$ of U (i.e. $\bigcup_{i=1}^{n} C_i = U$, $C_i \cap C_j = \emptyset$, $i \neq j$)
- Components: Finite partitions of every analysis unit.
- Variables: X_1, X_2,... ,X_k
- Thresholds: For variable X_k is $\mu_k \in \mathbb{R}$.

Intensity Topological Indicator

Bipartite associated graph G_C to an analysis unit C:
Suppose that components of C are $\{C_1, ..., C_p\}$. Then:

- $V(G_C) = \{C_1, ..., C_p\} \cup \{X_1, ..., X_k\}$
- $E(G_C) = \{\{C_i, X_j\} | X_j(C_i) < \mu_i\}$

Intensity Topological Indicator

Bipartite associated graph G_C to an analysis unit C:
Suppose that components of C are $\{C_1, ..., C_p\}$. Then:

- $V(G_C) = \{C_1, ..., C_p\} \cup \{X_1, ..., X_k\}$
- $E(G_C) = \{\{C_i, X_j\} | X_j(C_i) < \mu_i\}$

Intensity Topological Indicator:

$$ITI(C) = \frac{\#(E(G_C))}{p}$$

Intensity Topological Indicator

Bipartite associated graph G_C to an analysis unit C:
Suppose that components of C are $\{C_1, ..., C_p\}$. Then:

- $V(G_C) = \{C_1, ..., C_p\} \cup \{X_1, ..., X_k\}$
- $E(G_C) = \{\{C_i, X_j\} | X_j(C_i) < \mu_i\}$

Intensity Topological Indicator:

$$ITI(C) = \frac{\#(E(G_C))}{p}$$

Property: $0 \leq ITI(C) \leq k$

Intensity Topological Indicator

Partial Intensity Topological Indicator with respect to X_j: $ITI_{X_j}(C) = \dfrac{\delta(X_j)}{p}$

Intensity Topological Indicator

Partial Intensity Topological Indicator with respect to X_j: $ITI_{X_j}(C) = \dfrac{\delta(X_j)}{p}$

Property: $0 \leq ITI_{X_j}(C) \leq 1$

Intensity Topological Indicator

Partial Intensity Topological Indicator with respect to X_j: $ITI_{X_j}(C) = \dfrac{\delta(X_j)}{p}$

Property: $0 \leq ITI_{X_j}(C) \leq 1$

Practical interpretation:

- $ITI(C) \approx k$ lower general mathematical skills
- $ITI(C) \approx 0$ higher general mathematical skills
- $ITI_{X_j}(C) \approx 1$ lower mathematical skills with respect to X_j
- $ITI_{X_j}(C) \approx 0$ higher mathematical skills with respect to X_j

Intensity Topological Indicator

Elements in our study:

- Universe $U = \{\text{CENTER 1}, \text{CENTER 2}\}$
- Analysis Units: C_1 is CENTER 1, C_2 is CENTER 2.
- Components: $C_1 = \{3A, 3B, 3C, 3D\}$, $C_2 = \{3A', 3B', 3C', 3D', Div_1, Div_2\}$
- Variables: O:*organizing, understanding, and interpreting information*, E:*mathematical expression*, P:*planning and solving problems* $(O, E, P \in \{1, ..., 6\})$
- Thresholds: Mean in U, $\mu_O = 4$, $\mu_E = 3$, $\mu_P = 4$.

Intensity Topological Indicator

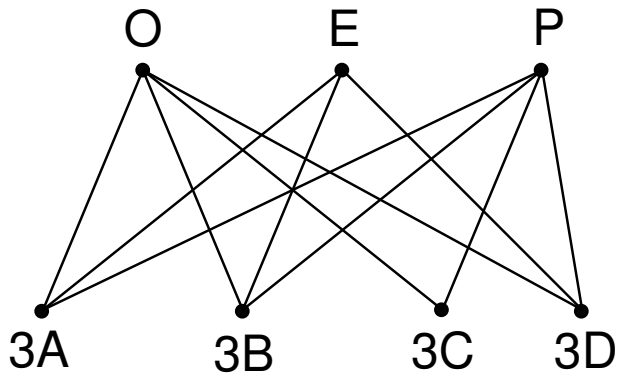
	3A	3B	3C	3D
O	5	4	5	5
E	3	3	2	3
P	4	5	4	4

	3A'	3B'	3C'	3D'	Div_1	Div_2
O	4	4	4	5	3	4
E	3	3	3	4	2	3
P	3	4	4	5	2	3

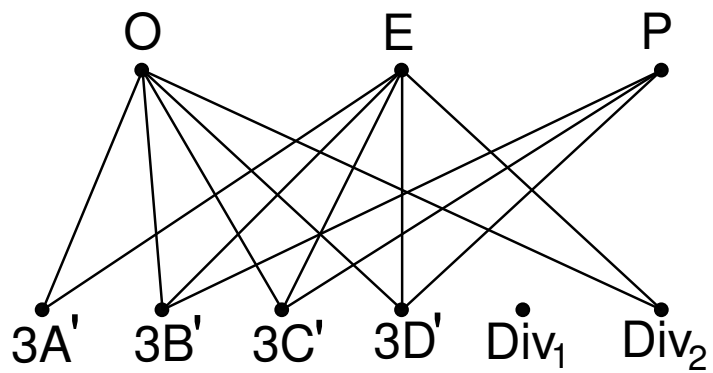
Intensity Topological Indicator

	3A	3B	3C	3D
O	5	4	5	5
E	3	3	2	3
P	4	5	4	4

	3A'	3B'	3C'	3D'	Div_1	Div_2
O	4	4	4	5	3	4
E	3	3	3	4	2	3
P	3	4	4	5	2	3



Unidad C_1



Unidad C_2

Intensity Topological Indicator

Obtained values for ITIs:

Unit	ITI_O	ITI_E	ITI_P	ITI
C_1	$\frac{4}{4} = 1{,}000$	$\frac{3}{4} = 0{,}750$	$\frac{4}{4} = 1{,}000$	$2{,}750$
C_2	$\frac{5}{6} = 0{,}835$	$\frac{5}{6} = 0{,}835$	$\frac{3}{6} = 0{,}500$	$2{,}170$

Intensity Topological Indicator

Obtained values for ITIs:

Unit	ITI_O	ITI_E	ITI_P	ITI
C_1	$\frac{4}{4} = 1{,}000$	$\frac{3}{4} = 0{,}750$	$\frac{4}{4} = 1{,}000$	$2{,}750$
C_2	$\frac{5}{6} = 0{,}835$	$\frac{5}{6} = 0{,}835$	$\frac{3}{6} = 0{,}500$	$2{,}170$

Unit	ITI_O	ITI_E	ITI_P	ITI
C_1	100 %	75 %	100 %	91.67 %
C_2	83.5 %	83.5 %	50 %	72.23 %

Intensity Topological Indicator

Obtained values for ITIs:

Unit	ITI_O	ITI_E	ITI_P	ITI
C_1	$\frac{4}{4} = 1{,}000$	$\frac{3}{4} = 0{,}750$	$\frac{4}{4} = 1{,}000$	$2{,}750$
C_2	$\frac{5}{6} = 0{,}835$	$\frac{5}{6} = 0{,}835$	$\frac{3}{6} = 0{,}500$	$2{,}170$

Unit	ITI_O	ITI_E	ITI_P	ITI
C_1	100 %	75 %	100 %	91.67 %
C_2	83.5 %	83.5 %	50 %	72.23 %

Induced order by ITIs: $C_2 \prec C_1$, $C_1 \prec_E C_2$, $C_2 \prec_{O,P} C_1$

Disparity Topological Indicator

Associated multidigraph (directed with multiple edges) $G(C)$ to an analysis unit C: Suppose that components of C are $\{C_1, ..., C_p\}$. Then:

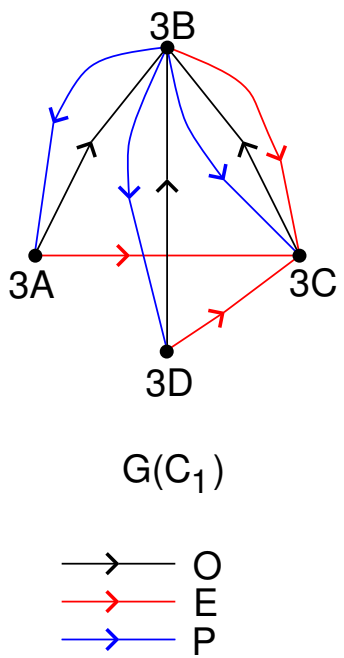
- $V(G_C) = \{C_1, ..., C_p\}$
- $E(G_C) = \{(C_i, C_j) | \exists X_k \text{such that} X_k(C_i) \geq X_k(C_j)\}$

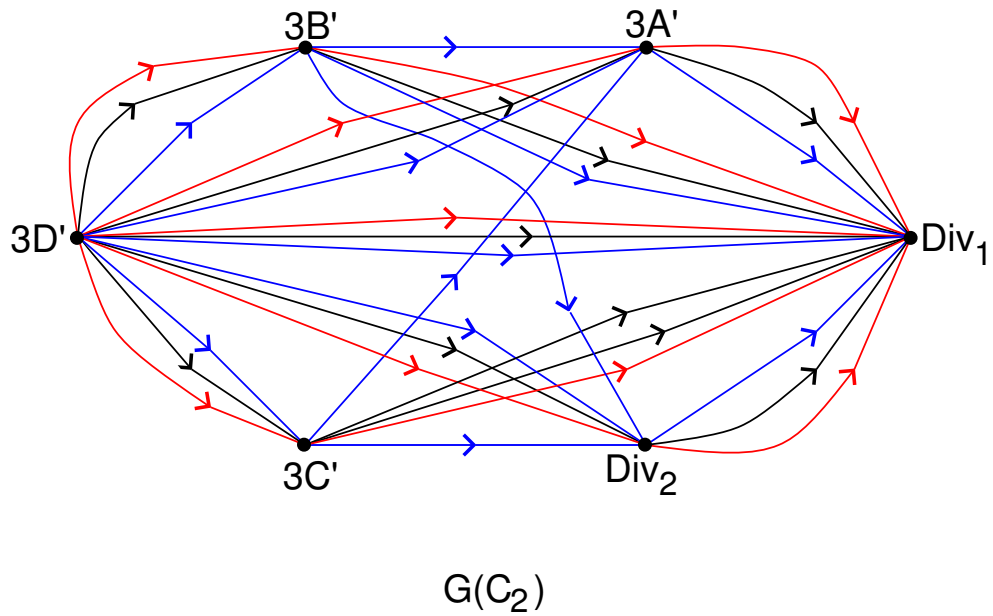
Disparity Topological Indicator

Associated multidigraph (directed with multiple edges) $G(C)$ to an analysis unit C: Suppose that components of C are $\{C_1, ..., C_p\}$. Then:

- $V(G_C) = \{C_1, ..., C_p\}$
- $E(G_C) = \{(C_i, C_j) | \exists X_k \text{such that} X_k(C_i) \geq X_k(C_j)\}$

Associated multidigraphs to C_1 and C_2:





Disparity Topological Indicator

Competence levels:

- *Adjacency matrix of* $G(C)$: $M(C) = (m_{ij})$ where m_{ij} is the number of edges from C_i to C_j

Disparity Topological Indicator

Competence levels:

- *Adjacency matrix of* $G(C)$: $M(C) = (m_{ij})$ where m_{ij} is the number of edges from C_i to C_j
- *Transitive closure of* $G(C)$: Graph whose adjacency matrix $\hat{M}(C)$ is the boolean addition of the boolean powers of $M(C)$

Disparity Topological Indicator

Competence levels:

- *Adjacency matrix of* $G(C)$: $M(C) = (m_{ij})$ where m_{ij} is the number of edges from C_i to C_j
- *Transitive closure of* $G(C)$: Graph whose adjacency matrix $\hat{M}(C)$ is the boolean addition of the boolean powers of $M(C)$
- *Strongly connected graph*: Any two distinct vertices are connected by a path.

Disparity Topological Indicator

Competence levels:

- *Adjacency matrix of* $G(C)$: $M(C) = (m_{ij})$ where m_{ij} is the number of edges from C_i to C_j
- *Transitive closure of* $G(C)$: Graph whose adjacency matrix $\hat{M}(C)$ is the boolean addition of the boolean powers of $M(C)$
- *Strongly connected graph*: Any two distinct vertices are connected by a path.
- *Maximal strongly connected subgraph*: A strongly connected subgraph not strictly contained in another strongly connected subgraph.

Disparity Topological Indicator

Some results:

- Maximal strongly connected subgraph containing vertex C_i is $S(C_i) = (\hat{\Gamma}(C_i) \cap \hat{\Gamma}^{-1}(C_i)) \cup \{C_i\}$.
- A graph G is strongly connected if and only if every entry of $\hat{M}(G)$ equals 1.

Disparity Topological Indicator

Some results:

- Maximal strongly connected subgraph containing vertex C_i is $S(C_i) = (\hat{\Gamma}(C_i) \cap \hat{\Gamma}^{-1}(C_i)) \cup \{C_i\}$.
- A graph G is strongly connected if and only if every entry of $\hat{M}(G)$ equals 1.

Notation: $\hat{\Gamma}(C_i) = \{C_j | \hat{m}_{ij} = 1\}$, $\hat{\Gamma}^{-1}(C_i) = \{C_j | \hat{m}_{ji} = 1\}$

Disparity Topological Indicator

Some results:

- Maximal strongly connected subgraph containing vertex C_i is $S(C_i) = (\hat{\Gamma}(C_i) \cap \hat{\Gamma}^{-1}(C_i)) \cup \{C_i\}$.
- A graph G is strongly connected if and only if every entry of $\hat{M}(G)$ equals 1.

Notation: $\hat{\Gamma}(C_i) = \{C_j | \hat{m}_{ij} = 1\}$, $\hat{\Gamma}^{-1}(C_i) = \{C_j | \hat{m}_{ji} = 1\}$

Practical interpretation: Components in a same maximal strongly connected subgraph are assumed to have similar mathematical skills.

Disparity Topological Indicator

The adjacency matrices associated to multidigraphs $G(C_1)$ and $G(C_2)$

$$M(C_1) = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} \quad M(C_2) = \begin{pmatrix} 0 & 0 & 0 & 0 & 3 & 0 \\ 1 & 0 & 0 & 0 & 3 & 1 \\ 1 & 0 & 0 & 0 & 3 & 1 \\ 3 & 3 & 3 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \end{pmatrix}$$

Disparity Topological Indicator

The adjacency matrices associated to multidigraphs $G(C_1)$ and $G(C_2)$

$$M(C_1) = \begin{pmatrix} 0 & 1 & 1 & 0 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 0 \\ 0 & 1 & 1 & 0 \end{pmatrix} \quad M(C_2) = \begin{pmatrix} 0 & 0 & 0 & 0 & 3 & 0 \\ 1 & 0 & 0 & 0 & 3 & 1 \\ 1 & 0 & 0 & 0 & 3 & 1 \\ 3 & 3 & 3 & 0 & 3 & 3 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 3 & 0 \end{pmatrix}$$

Matrices of the transitive closures of $G(C_1)$ and $G(C_2)$:

$$\hat{M}(C_1) = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & 1 \end{pmatrix} \quad \hat{M}(C_2) = \begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}$$

Disparity Topological Indicator

Layers: $N_k \subset G_C$ such that:

- $\{N_k\}_k$ is a partition of $V(G_C)$.
- $\{N_k\}_k$ can be totally ordered.

Disparity Topological Indicator

Layers: $N_k \subset G_C$ such that:

- $\{N_k\}_k$ is a partition of $V(G_C)$.
- $\{N_k\}_k$ can be totally ordered.

Algorithm to obtain N_k:

- G'_C is built by contracting to a point every maximal strongly connected subgraph.
- $N'_0 = \{V \in G' | \hat{\Gamma}^{-1}(V) = \emptyset\}$.
  $N'_1 = \{V \in G' - N'_0 | \hat{\Gamma}^{-1}(V) = \emptyset\}$.
  $N'_2 = \{V \in G' - (N'_0 \cup N'_1) | \hat{\Gamma}^{-1}(V) = \emptyset\}$, and so on.
- N_k is the set of vertices of G lying in a maximal strongly connected subgraph contracted to a point of N'_k in G'

Disparity Topological Indicator

Practical interpretation: Components in a same layer are assumed to have similar mathematical skills.

Disparity Topological Indicator

Practical interpretation: Components in a same layer are assumed to have similar mathematical skills.

Our practical case:

- Maximal strongly connected subgraph of $G(C_1)$: $G(C_1)$
- Maximal strongly connected subgraph of $G(C_2)$: $\{3A'\}$, $\{3B'\}$, $\{3C'\}$, $\{3D'\}$, $\{Div_1\}$ and $\{Div_2\}$
- Layers of $G(C_1)$: $N_0 = G(C_1)$
- Layers of $G(C_2)$: $N_0 = \{3D'\}$, $N_1 = \{3B', 3C'\}$, $N_2 = \{3A', Div_2\}$ and $N_3 = \{Div_1\}$

Disparity Topological Indicator

Practical interpretation: Components in a same layer are assumed to have similar mathematical skills.

Our practical case:

- Maximal strongly connected subgraph of $G(C_1)$: $G(C_1)$
- Maximal strongly connected subgraph of $G(C_2)$: $\{3A'\}$, $\{3B'\}$, $\{3C'\}$, $\{3D'\}$, $\{Div_1\}$ and $\{Div_2\}$
- Layers of $G(C_1)$: $N_0 = G(C_1)$
- Layers of $G(C_2)$: $N_0 = \{3D'\}$, $N_1 = \{3B', 3C'\}$, $N_2 = \{3A', Div_2\}$ and $N_3 = \{Div_1\}$

Induced order in C_2: $\{3D'\} \prec \{3B', 3C'\} \prec \{3A', Div_2\} \prec \{Div_1\}$

Disparity Topological Indicator

Practical conclusions: Unit C_1: All groups in 3rd course from CENTER 1 present similar mathematical competence.

Disparity Topological Indicator

Practical conclusions: Unit C_1: All groups in 3rd course from CENTER 1 present similar mathematical competence.

Practical conclusions: Unit C_2: We observe four different levels; group $3D'$ is the most advanced, Div_1 has the most difficulties and $3A'$ and Div_2 has similar skills (not expected).

Disparity Topological Indicator

Definition: $DTI(G) = \dfrac{k}{\#V(G)}$, where k is the number of layers.

Disparity Topological Indicator

Definition: $DTI(G) = \dfrac{k}{\#V(G)}$, where k is the number of layers.

Property: $DTI(G) \in (0, 1]$

Disparity Topological Indicator

Definition: $DTI(G) = \dfrac{k}{\#V(G)}$, where k is the number of layers.

Property: $DTI(G) \in (0, 1]$

Practical meaning: The more homogeneity in the analysis unit the closer to 0 the index

Disparity Topological Indicator

Definition: $DTI(G) = \dfrac{k}{\#V(G)}$, where k is the number of layers.

Property: $DTI(G) \in (0, 1]$

Practical meaning: The more homogeneity in the analysis unit the closer to 0 the index

Our case: $DTI(G(C_1)) = \dfrac{1}{4} = 0{,}25$, $DTI(G(C_2)) = \dfrac{4}{6} \approx 0{,}67$. Thus a greater homogeneity is expected in C_1

REMARK: ITIs, DTI, the associated bipartite graph and multidigraph and the algorithm to find layers has been successfully implemented in Maxima

Disparity Topological Indicator

Definition: $DTI(G) = \dfrac{k}{\#V(G)}$, where k is the number of layers.

Property: $DTI(G) \in (0, 1]$

Practical meaning: The more homogeneity in the analysis unit the closer to 0 the index

Our case: $DTI(G(C_1)) = \dfrac{1}{4} = 0{,}25$, $DTI(G(C_2)) = \dfrac{4}{6} \approx 0{,}67$. Thus a greater homogeneity is expected in C_1

REMARK: ITIs, DTI, the associated bipartite graph and multidigraph and the algorithm to find layers has been successfully implemented in Maxima