Multidimensional Techniques for Social Network Data

Giuseppe Giordano
Dept. of Economics and Statistics
University of Salerno - Italy

ggiordan@unisa.it

Cnam

March, 5th 2015

Outline

- **Setting**: Multidimensional Data Analysis (MDA) and Social Network Analysis (SNA)
- **Part 1**: to present a framework for the treatment of SNA data structures with explorative techniques of MDA
- **Methods**: Smooth factorial analysis-SFA; Factorial Analysis of Local Differences-FALD *Benali H. Escofier, B. (1990)*
- **Part 2**: To define ad hoc relational data structures highlighting the effect of external information on networks
- **Theoretical frameworks:** Homophily principle, Social Influence, ...

Mainly based on the paper by G. Giordano, M. P. Vitale (2011). On the use of external information in social network analysis. *ADVANCES IN DATA ANALYSIS AND CLASSIFICATION*, pp.95- 112, Vol. 5.

Background and Aim

Background

- **SNA** focuses on **ties** among **interacting units** (Dyad, Triad, Subgroups) to describe the pattern of the social relationships in a network

- the techniques of **MDA** consider statistical observations (at individual level) to obtain syntheses of variables and units

Aim

to present a framework for the analysis of relational data and attribute variables through **Explorative Techniques of Multidimensional Data Analysis**

Why Networks...

Network Data are present in our life beside our awareness to be part of a system

•Traditional Human interactions
(difficult to collect data)
•Telecommunication
(including online chat messenger, e-mail,…
containing rich information in form of text, photo, video, etc.)
•Online Social Network
(Facebook, Tweeter, Flicker, etc)
•Biological Networks
(Protein interactions)

Elements of SNA

Actor: social entities
(individual, corporate, or collective social units)

Triad: consists of a subset of three actors and the (possible) tie(s) among them

Dyad: consists of a pair of actors and the (possible) tie(s) between them

Subgroup: consists of a subset of actors and all ties among them

Network: is the collection of all actors on which ties are to be measured

Ties: linkages between pair of actors (friendship, business, transaction, kinship…)

Relation: is the collection of ties of a specific kind among members

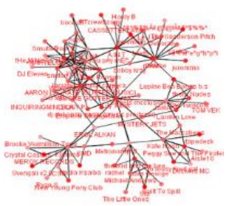
Social Network: consists of a finite set or sets of actors and the relation or relations defined on them (Wasserman, Faust, 1994)

Part 1 Multidimensional Techniques for Social Network Data

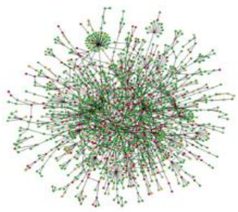
Complex data structure

large networks, thousands of vertices and lines (e.g.: genealogies, molecule, computer networks, social networks, intra/inter organizational networks...)
(de Nooy et al., 2005)

MySpace friends



Protein-Protein network



- relational data
- attribute variables

http://www.visualcomplexity.com/vc/

Complex network structures (topological and dynamic properties) and associated **contextual information** (attribute variables on nodes)

MDA could be used to deal with relational data and attribute variables arising from complex network structure…

Statistical Issues

Huge Data Sets
Sparse Data Matrix
Sampling Data: nodes, ties, snowball
Data Dependence
Missing Data

Different Frameworks

Substantive
Exploratory
Stocastic

[...] the network approach to the study of behavior has involved two commitments:
(1) it is guided by formal theory organized in mathematical terms, and
(2) it is grounded in the systematic analysis of empirical data

(L. Freeman)

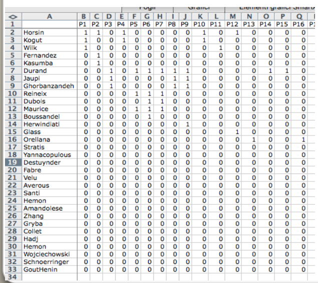
Substantive	Exploratory	Stocastic
Theories of self-interest Theory of Social capital Strength of Weak Ties Theory Transaction Cost Economics	*Network Demogrophy* Size, Density	Treat networks as realizations of random variables; Propose a model for the distribution
Contagion theories Social influence, Imitation, modeling, Learning, Mimetic behavior, Similar positions in structure and roles	Role and Positions of actors in the network, Cliques and Subgroups Cohesion and Distance	Fit the model to some observed data; Use the model to predict properties of the network
Cognitive theories Semantic Networks Knowledge Structures Homophily theories Social support theories	Visualization Topological properties Distribution of network statistics	*Single Network* ERGM – P* Latent Space Network Conditionally Uniform models
Theories of proximity Physical proximity Electronic proximity Influence of distance Influence of accessibility Theories of uncertainty	One-mode, Two-mode, Multiplex network Ego-network	*Dynamic Network* Continuous-time models Actor-oriented models Dynamic Exponential Random Graph Models Hidden Markov Models

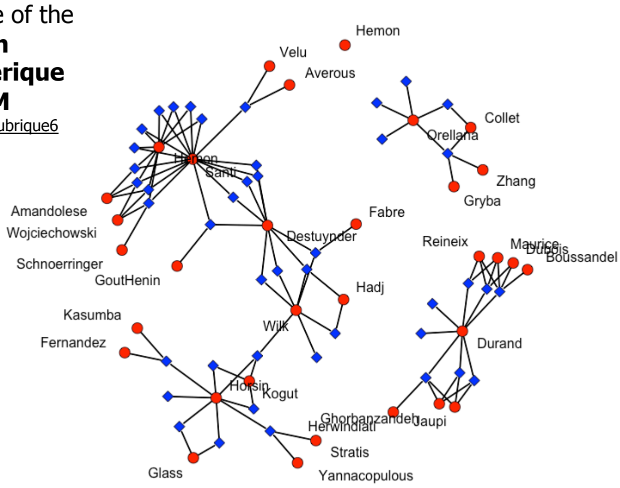
But First...

…Let me take a selfie!

Graph representation of the Affiliation matrix of Authors and Papers as appear at publications page of the **M2N - Modélisation mathématique et numérique laboratory at CNAM**
http://maths.cnam.fr/M2N/spip.php?rubrique6

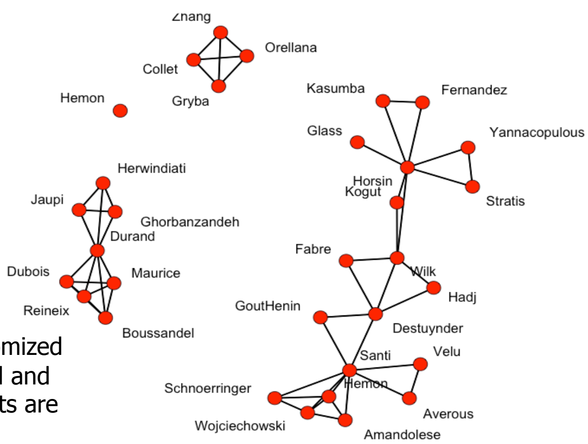
A Network @ the CNAM



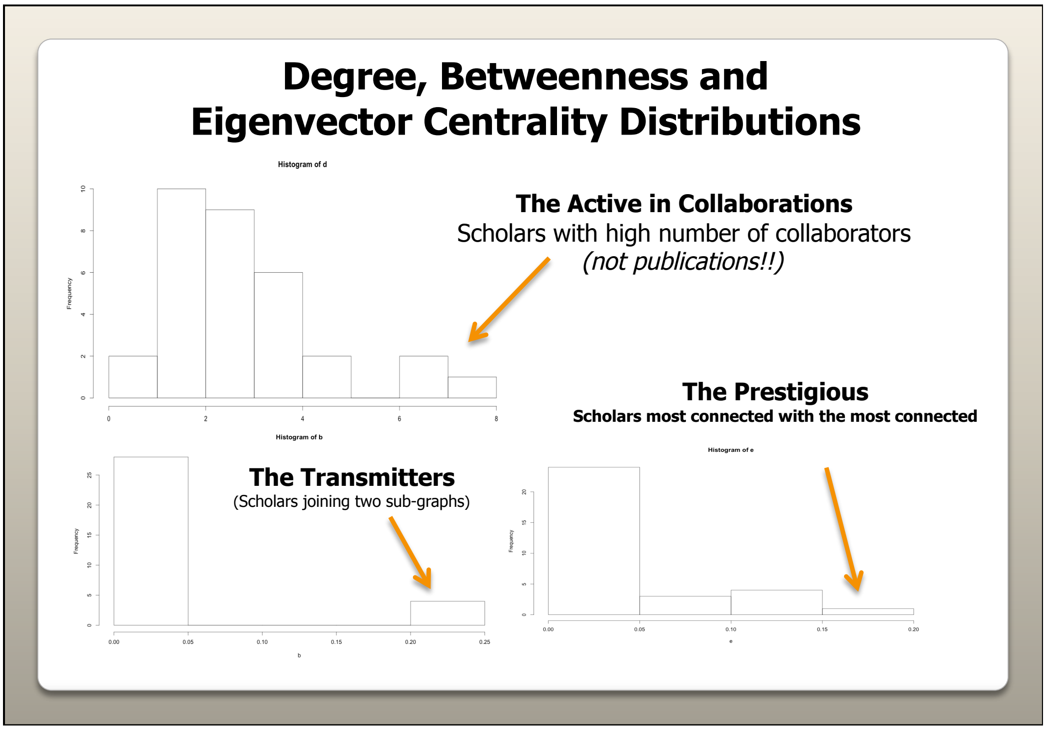


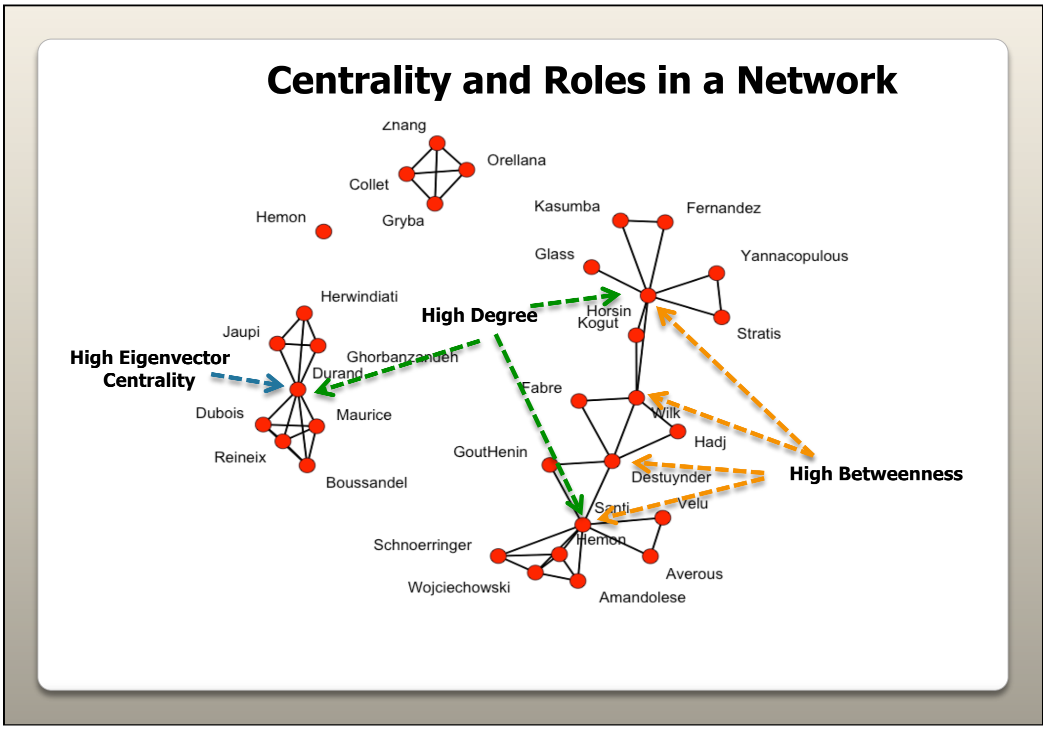
The Co-Authorship Network at M2N

G = AA'



G can be dichotomized with a threshold and diagonal elements are set to 0


Degree, Betweenness and
Eigenvector Centrality Distributions
Histogram of d
The Active in Collaborations
Scholars with high number of collaborators
(not publications!!)
Histogram of b
The Prestigious
Scholars most connected with the most connected
The Transmitters
(Scholars joining two sub-graphs)
Histogram of e


Centrality and Roles in a Network
Zhang
Orellana
Collet
Hemon
Gryba
Kasumba
Fernandez
Glass
Yannacopulous
Herwindiati
High Degree
Horsin
Kogut
Stratis
Jaupi
High Eigenvector
Centrality
Ghorbanzadeh
Durand
Fabre
Maurice
Dubois
Wilk
Hadj
Reineix
GoutHenin
Boussandel
Destuynder
High Betweenness
Santi
Velu
Hemon
Schnoerringer
Averous
Wojciechowski
Amandolese

Data Structures

Relational data (pairwise links joining two units)

=> SNA

Attribute data (qualitative or quantitative variables)

=> MDA

Two Perspectives (to put together the 2 data structures)

Multidimensional Data Analysis - MDA

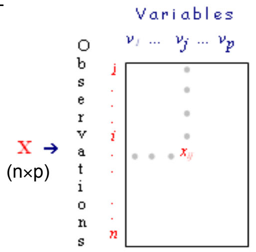
Social Network Analysis - SNA

MDA and SNA: Data structures

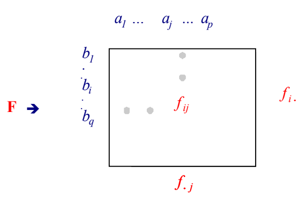
Attribute Data Matrix

Relational Data Matrix

MDA

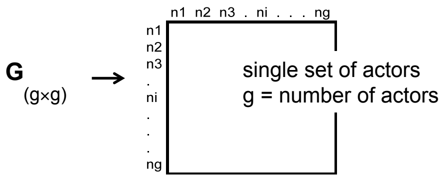


Contingency Table

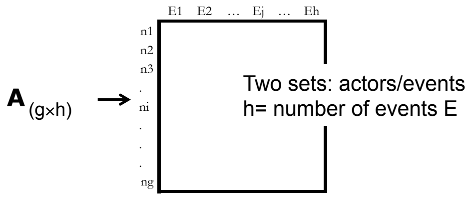


SNA

Adjacency matrix: 1-mode network



Affiliation matrix: 2-mode network



MDA for SNA

Usually different techniques of MDA have been used to visualise and explore the relationships in the net structure

such as…

- ***Multidimensional Scaling*** : representation of similarity or dissimilarity measures among the actors onto a factorial map. (Freeman, 2005)
- ***Canonical correlation***: analysis of associations among actor characteristics, (i.e. network composition) and the pattern of social relationships (i.e. network structure) (Wasserman and Faust, 1989)
- ***Correspondence Analysis***: analysis of 2-mode networks (Roberts, 2000; de Nooy, 2003 Faust, 2005)
- ***Clustering techniques for network data*** (Batagelj, Ferligoy, 1982, 2000)

A Brief Review: Contiguity analysis

How to take into account relational data in MDA?

Answer

Contiguity analysis is a generalization of linear discriminant analysis in which the partition of elements is replaced by a more general **graph structure** defined a priori on the set of observations (disjoined cliques, chain structures, undirected graph).

(Lebart 1969, 2006; Lebart *et al.*, 2000)

G (n,n) **Adjacency matrix** holding the n vertices in a contiguity graph (symmetric binary matrix);
$g_{ii'} = 1$ if i is a neighbour of i' and $g_{ii'} = 0$ otherwise

Contiguity analysis in MDA: SFA and FALD

Contiguity analysis in MDA
- Smooth factorial analysis – SFA (Benali *et al.* 1990)
- Factorial Analysis of Local Differences - FALD
(Benali, Escofier, 1990)

Smooth factorial analysis - SFA
Analysis of the general pattern in the data by removing local variations
(replacing each point with the center of gravity of its neighbors)

Factorial Analysis of Local Differences - FALD
Analysis of the local variations
(replacing each point with the differences from the barycentre of its neighbors)

SFA and FALD: matrices definition

Qualitative Variables: *Multiple Correspondence Analysis* - MCA	Quantitative Variables: *Principal Component Analysis* - PCA
$Q_{(n,k)}$ = full disjunctive matrix (0/1)	**$X_{(n,p)}$ attribute matrix information** on p characteristics of n statistical units (vertices)

$G_{(n,n)}$ contiguity matrix (network structure)
$N_{(n,n)}$ diagonal matrix [**N**= diag (**G'G**)] holding the degree of each vertices

Analysis of relational data and auxiliary information

Given the triplet **Q**, **G**, **N** for a MCA multiply the **Q** matrix by $\mathbf{N^{-1}G}$

SFA

$$\mathbf{N^{-1}GQ}$$

FALD

$$\mathbf{Q - N^{-1}GQ + \frac{q_{i.}q'_{.j}}{q_{..}}}$$

Analysis of relational data and auxiliary information

Given the triplet **X**, **G**, **N** for a PCA multiply the **X** matrix by $\mathbf{N^{-1}G}$

SFA

$$\mathbf{N^{-1}GX}$$

FALD

$$\mathbf{X - N^{-1}GX}$$

SFA and FALD in SNA

Entries in adjacency matrix can be seen as a particular case of contiguity relation among statistical units defined in **G**. It produces a fuzzy partitioning of the units.

Decomposition of the total variance/inertia into two components:

- **local variance between the adjacent units**

to discover patterns in the data → **SFA:** variability explained by the presence of a contiguity structure

- **residual variance**

analysis of cohesive sub-group variations → **FALD:** actors with a prominent role in contiguous groups

SFA and FALD for qualitative data (MCA):
Illustrative example

Data sources: students enrolled at the first academic year of the second level degree of Sociology at University of Salerno 2008/09.

Web survey: Questionnaire to collect individual characteristics and relational data about collaboration dynamics and community formation during the first level degree.

60% answers (**81** enrolled students)

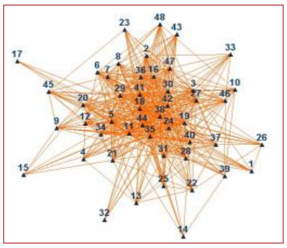
Aim: to explore the association structure between socio-demographic background, attitudinal variables, academic background, university performance (**attribute variables on actors**) and collaboration among students at university (**relational data**)

Illustrative example (continued)

Attribute variables on nodes
- Socio-demographic (*gender; residence*; …)
- Individual characteristics *(work, grade and type of secondary school,…)*
- Performance *(perceived performance of services vs/ final grade,…)*

Collaboration Network

Formal contacts
- *information exchange*
- *exams preparation*
- *classmates*



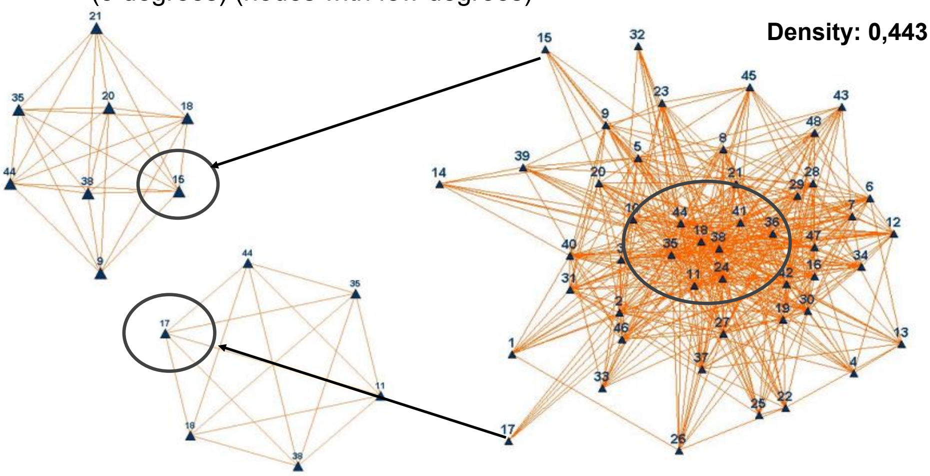
Collaboration Network: Definition

From 3 binary symmetric adjacency matrices…to one binary symmetric adjacency matrix obtained by joining the 3 above matrices

Looking at nodes position in the net structure

the neighbors of the nodes 15 (7 degrees) and 17 (5 degrees) (nodes with few degrees)

Density: 0,443



Attributes information on nodes

Gender: *Male, Female*

Work: *work, no_work during the first level degree*

Career: Regular (<=4 years); Irregular (>4 years)

Secondary School= SecondarySchool ("Lyceum");
NoSecSchool ("no lyceum")

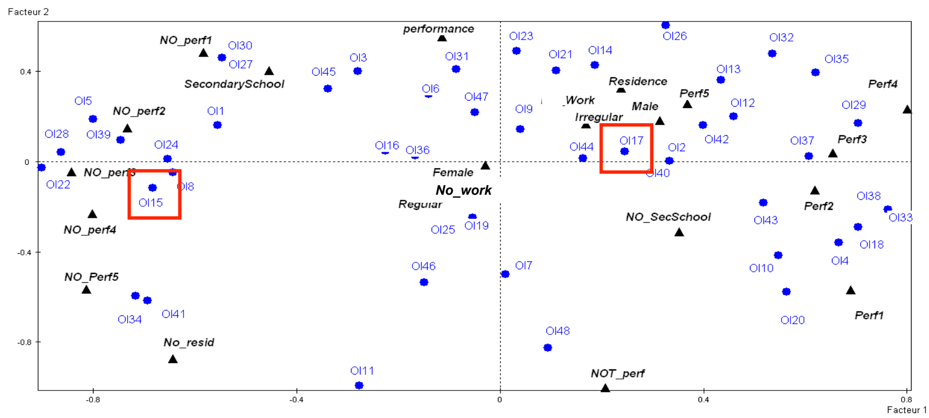
Performance = (objective indicator) final degree
(NoPerf : score <90/100; Perf : score >= 90/100)

Perceived performance= 5 subjective indicators measured on 10-scale scores (Perf1, Perf2, Perf3, Perf4 Perf5 : score >= 8;
No_perf1, No_perf2, No_perf3, No_perf4, No_perf5 : score < 8)

Part 1 Multidimensional Techniques for Social Network Data

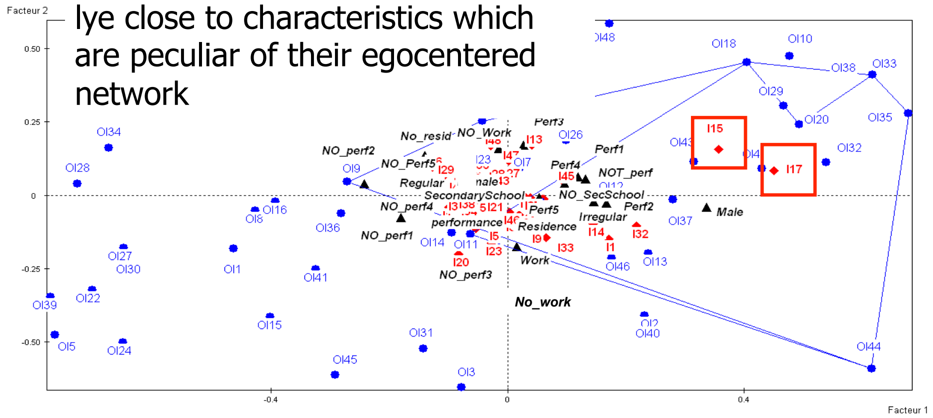
Looking at association structure on attribute variables: MCA

"Good" students against "Polemic" students



Introducing contiguity matrix G (SFA)

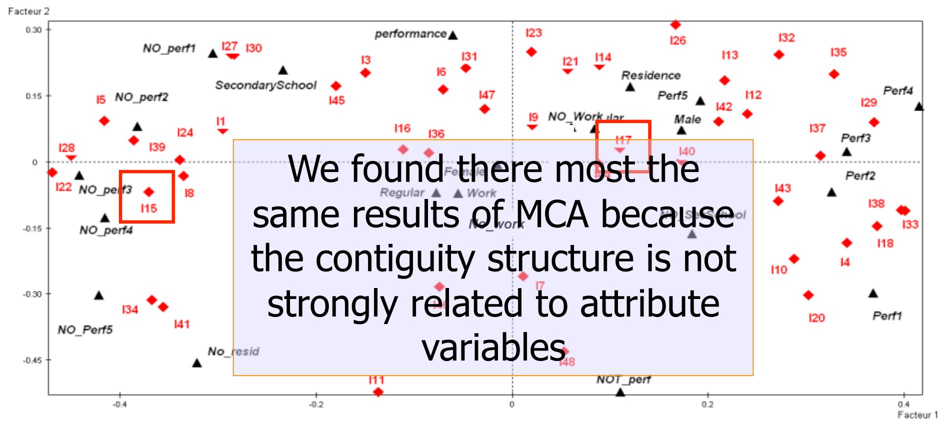
Smoothed observations are the barycenter of contiguous units and lye close to characteristics which are peculiar of their egocentered network



Part 1 Multidimensional Techniques for Social Network Data

Introducing contiguity matrix: FALD

putting emphasis on characteristics describing heterogeneity into contiguous units



REMARKS

SFA and FALD can be used to explore the relationships between the network structure and attribute variables…

The factorial maps show the actor position as a function of the attributes in X (or Q) and the contiguity structure…

Can network structure (dense vs sparse) influence the results on MDA in presence of contiguity?

Main references

Aluja Banet, T., Lebart, L. (1984) Local and Partial Principal Component Analysis and Correspondence Analysis, in: COMPSTAT Proceedings, Havranek T., Sidak Z. & Novak M. (Eds.), Phisyca-Verlag, Vienna, 113-118.

Batagelj, V., Ferligoj A., (1982) Clustering with relational constraint, Psychometrika, 47, 413-426.

Batagelj V., Ferligoj A.: Clustering relational data. Data Analysis (ed.: W. Gaul, O. Opitz, M. Schader), Springer, Berlin 2000, 3-15.

Benali H. Escofier, B. (1990) Analyse factorielle lissée et analyse factorielle des différences locales. Revue de Statistique Appliquée, 38, 55-76.

Escofier, B., H. Benali, K. Bachar (1990), How to introduce contiguity in correspondence
analysis? Application to image segmentation (French), Statistique et Analyse des Donnees, 15, 61-92.

Faust K. (2005) Using Correspondance Analysis for Joint Displays of Affiliation Networks, in: Models and Methods in Social Network Analysis, Carrington P., Scott J. & Wasserman S. (Eds.), Cambridge University Press, Cambridge, 117-147.

Freeman, L. C. (2005). Visualizing Social Networks, in: Models and Methods in Social Network Analysis,
Carrington P., Scott J. & Wasserman S. (Eds.), Cambridge University Press, Cambridge, 117-147.

Giordano, G., Vitale, M.P. (2011). On the use of external information in social network analysis. Advances in Data Analysis and Classification, 5, pp.95-112.

Lebart L. (1969) Analyse statistique de la contiguité. Publications de l'ISUP, 81–112.

Lebart L. (2006) Assessing self organizing maps via contiguity analysis, Neural Networks 19:847 – 854, Elsevier Science Ltd. Oxford, UK.

Tukey, J.W. (1977), Exploratory Data Analysis, Addison-Wesley, Reading.

Wasserman S., Faust K., (1989) Canonical Analysis of the Composition and Structure of Social Networks, Sociological Methodology: 19, pp. 1-42.

Relational data and Auxiliary information

Attribute variables on nodes (network actors):
- socio-economic and demographic characteristics of nodes
- variables under analysis (e.g. performance)

Use of attribute data in SNA
- to enrich the graph interpretation and its visualization (e.g. colouring)
- to explain the formation of ties in network modelling (e.g. p* models)

Question

How exploratory MDA techniques could take into account both relational and attribute data?

How to put together the two data structures?

SNA framework

emphasis on **relational data**

-> Attribute data are external information characterizing actors

AIM

- **descriptive** analysis - to depict properties of networks, actors and sub-sets of actors;
- **predictive** models to explain the presence or absence of ties as in ERG-Models to investigate whether actors with similar attributes are more likely to form ties.
(Robins et al., 2007)
- **classification** of actors by attribute variables with relational constraints
(Ferligoj, Kronegger, 2009)

Our Aim

To define ad hoc relational data structures highlighting the effect of external information on networks

Data Source

relational data ->

here derived from the presence of actors at events or activities (two-mode network) arranged in affiliation matrix

external information ->

represents attribute data on events and actors, arranged in data vectors

Notation and Definition

A (nxp) affiliation matrix (binary); *n* actors; *p* events

$a_{ij} = 1$ -> *i*-th actor is present at the *j*-th event

(i=1, …, n; j=1, …, p)

X (nxm) *n actors; r* nominal variables expanded into *m* dummy variables

$x_{ik} = 1$ -> *i*-th actor belongs to the *k*-th category

(i=1, …, n; k=1, …, m)

Z' (pxq) *p* events; *s* nominal variables expanded into *q* dummy variables

$z_{hj} = 1$ -> *j*-th event belong to the *h*-th category

(h=1, …, q; j=1, …,p)

… The Underlying Decomposition Model

(Takane, Shibayama, 1991)

$$\mathbf{A} \approx \mathbf{X}\mathbf{B} + \mathbf{C}'\mathbf{Z} + \mathbf{X}\Theta\mathbf{Z}$$

Relational data

Interaction Effect

Effect of event characteristics

Effect of actor characteristics

Effect of actor characteristics on relational data

$$\mathbf{A} \approx \mathbf{XB}$$

$$J(\mathbf{B}) = \underset{\mathbf{B}}{argmin} \left\| \mathbf{A} - \mathbf{XB} \right\|^2$$

$$\hat{\mathbf{B}} = (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\mathbf{A}$$

particular case of orthogonal design X :

$$\hat{\mathbf{B}} = [diag(\mathbf{X}'\mathbf{X})]^{-1}\mathbf{X}'\mathbf{A}$$

statistical interpretation as conditional frequency

$b_{k,j} \in [0,1]$

=> Affiliation effect of the *k*-th category to the *j*-th event.

Effect of events characteristics in determining the presence of actors at events

$$\mathbf{A}' \approx \mathbf{Z}'\mathbf{C}$$

$$\mathbf{J}(C) = \underset{\mathbf{C}}{argmin} \left\| \mathbf{A}' - \mathbf{Z}'\mathbf{C} \right\|^2$$

we retain the particular solution

$$\hat{\mathbf{C}} = [diag(\mathbf{ZZ}')]^{-1}\mathbf{ZA}'$$

C_{hi} affiliation effect due to the presence of actors at event given the h-th event's category

Use of the coefficients B and C in SNA

The **B** e **C** coefficients have been used to approximate the original **affiliation matrix A** and to derive the **adjacency matrices** $\mathbf{G}_x$ and $\mathbf{G}_z$ (*actors* x *actors*)

$$\hat{\mathbf{B}} \Rightarrow \hat{\mathbf{A}}_X \Rightarrow \mathbf{G}_X$$

$$\hat{\mathbf{C}} \Rightarrow \hat{\mathbf{A}}_Z \Rightarrow \mathbf{G}_Z$$

$\mathbf{G}_x$ and $\mathbf{G}_z$ are then analyzed by SNA methods specially to look for peculiar patterns of ties and homogeneous groups of actors induced by the effect of the external information.

Working Example *Collaboration network in an e-learning course*

Data structure

A affiliation matrix where the 18 actors (students) participated in 12 events (weekly practices);

X attribute matrix - 2 nominal variables: *gender* (Male, Female); *skills (technical, math, humanities*) observed on 18 students;

Z attribute matrix - 1 nominal variable: *theme of practices (reporting, logical, technical*) observed on 12 events.

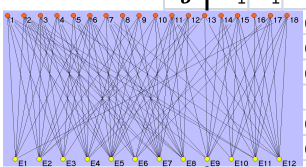
Expected results

- *Students with a given skill should attend to peculiar weekly practices*
  obs. network structure could be influenced by **actors characteristics**
- *Weekly practices requiring particular skills should attract different students*
  obs. network structure could be influenced by **events characteristics**

Data structure

Z	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
Z11	1	1	1	1	0	0	0	0	0	0	0	0
Z12	0	0	0	0	1	1	1	1	0	0	0	0
Z13	0	0	0	0	0	0	0	0	1	1	1	1

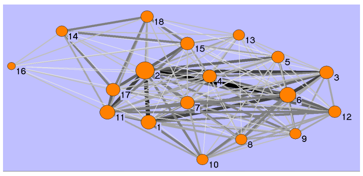
A	E1	E2	E3	E4	E5	E6	E7	E8	E9	E10	E11	E12
1	1	1	1	0	1	0	1	1	0	0	1	0
2	1	1	1	1	1	1	1	1	1	0	1	0
3	1	1	1	1	1	0	0	1	0	0	0	0
4	1	1	0	1	1	1	0	0	0	0	0	1
5	0	0	1	1	0	1	1	0	1	0	0	0
6	1	1	1	1	1	1	1	1	0	0	0	0
[illegible]	[illegible]	[illegible]	0	0	1	0	1	1	0	1	0	1
[illegible]	[illegible]	[illegible]	0	1	1	0	1	1	0	0	0	0
[illegible]	[illegible]	[illegible]	0	1	1	0	1	0	0	0	0	0
[illegible]	[illegible]	[illegible]	1	0	1	0	1	1	0	0	0	0
[illegible]	[illegible]	[illegible]	0	0	1	1	1	1	1	0	1	1
[illegible]	[illegible]	[illegible]	0	0	1	1	1	1	0	0	0	0
13	0	1	0	1	0	0	0	0	1	0	0	1
14	0	0	0	0	0	0	0	0	1	1	1	1
15	0	0	0	1	0	1	0	0	1	1	1	1
16	0	0	0	0	0	0	0	0	0	0	1	1
17	0	0	0	0	1	1	0	0	1	1	1	1
18	1	0	0	0	0	0	0	0	1	1	1	1



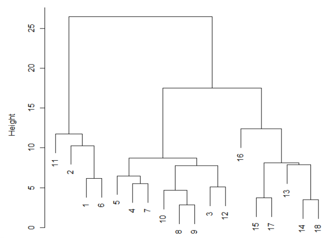
X	X11	X12	X13	X21	X22
1	1	0	0	1	0
2	1	0	0	1	0
3	1	0	0	1	0
4	1	0	0	1	0
5	1	0	0	1	0
6	1	0	0	1	0
7	0	1	0	1	0
8	0	1	0	1	0
9	0	1	0	1	0
10	0	1	0	0	1
11	0	1	0	0	1
12	0	1	0	0	1
13	0	0	1	0	1
14	0	0	1	0	1
15	0	0	1	0	1
16	0	0	1	0	1
17	0	0	1	0	1
18	0	0	1	0	1

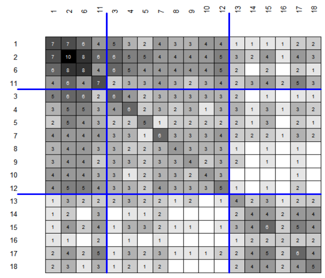
From A to G: Clustering and Blockmodeling of G

$$\mathbf{G} = \mathbf{A} \cdot \mathbf{A}'$$



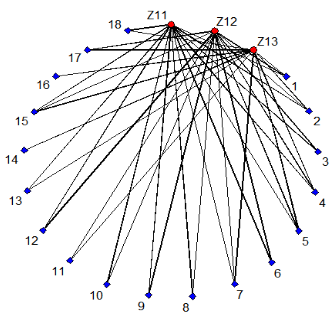
Hierarchical clustering of network positions and Generalized blockmodeling



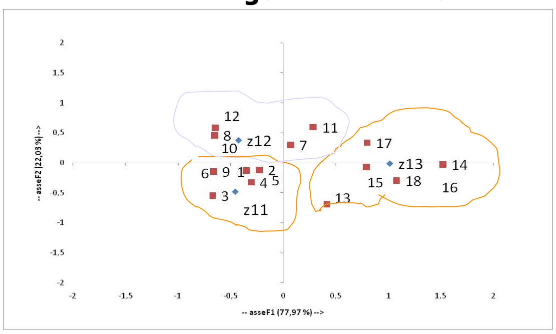


Analysis of Z and A

C (q x n)



Events' categories and actors



Results …

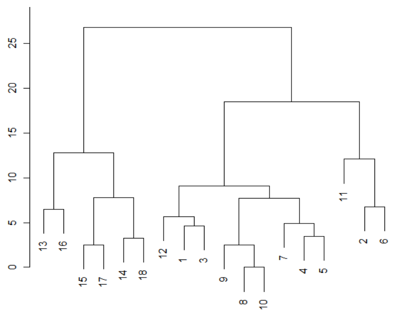
$$\hat{\mathbf{A}}_Z = \mathbf{C}'\mathbf{Z}$$

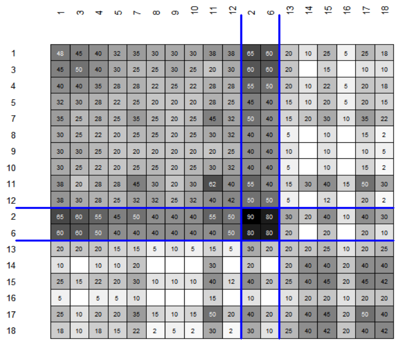
Analysis of Z and A

To derive from $\hat{\mathbf{A}}_Z$ the weighted adjacency $\mathbf{G}_Z = \hat{\mathbf{A}}_Z \cdot \hat{\mathbf{A}}'_Z$

where the generic element of $\mathbf{G_Z}$ is the weight of the tie between a pair of actors related to their presence at the same categories of events

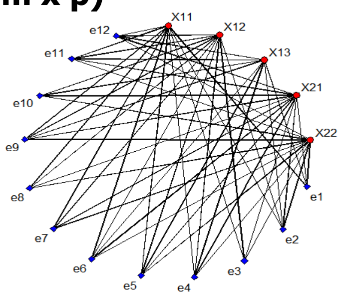
Clustering and Blockmodeling





Analysis of X and A

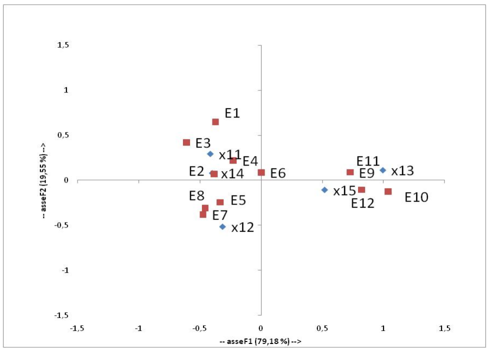
B (m x p)



r x m = maximum value of the columns marginal of **B** (all n actors are present at the j-th event)

p = maximum value of rows marginal of **B** (a category fully characterises all p events)

Events and actors' attributes



$$\hat{\mathbf{A}}_X = \mathbf{X}\hat{\mathbf{B}}$$

1 = each column maximum value
r x p = row marginal maximum theoretical value

Analysis of X and A

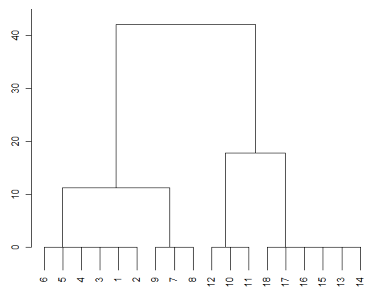
$$\mathbf{G_X} = \hat{\mathbf{A}}_{\mathbf{X}} \cdot \hat{\mathbf{A}}'_{\mathbf{X}}$$

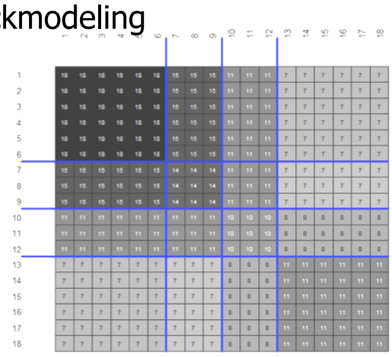
whose generic element weights the joint presence at events of actors having similar characteristics. maximum theoretical value in $\mathbf{G_X}$

$$\sum_{j=1}^{p} r^2 = p \cdot r^2$$

If a *homophily effect* is present, then $\mathbf{G_X}$ shows a weighting system coherent with the capability of all actors characteristics to jointly explain the participation to events.

Clustering and Blockmodeling





Real data set: Scientific collaboration among scientists

… the process generating ties (e.g. co-authorship) in a collaboration network is somewhat affected by attribute data on authors (*academic position, research specialty, geographical proximity* …) or on publications (*type, scientific relevance*, …).

Expected results:

- authors with different academic positions in their institution (phd. student, assistant professor, full professor) are more likely to collaborate in writing a publication than authors sharing the same position;

- authors who work in the same research specialty (e.g. statisticians) are more likely to collaborate in writing a publication than authors from different specialties.

Data source: 81 scientists involved in the Economics & Statistics field in an academic institution in Southern Italy

(MIUR database and Local Research Archive)

Relational Data

A = affiliation matrix 81 x 358
81 scientists, 358 publications and the cells report 1 if two authors co-authored a paper

Data vectors Definition (Nominal Variables)

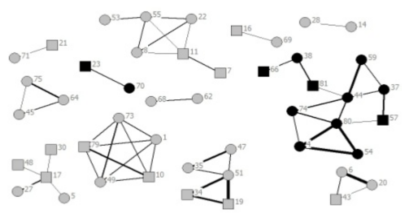
For ***authors***

X = matrix 81 x 4
4 columns dummy coding of the 2 nominal variables *academic position* and *research specialty*
(assistant professor vs full professor; statistics vs economics)

For ***publications***:

Z = matrix 4 x 358
4 rows dummy coding of the 2 nominal variables *type* and *number of author in publication*
(article – no article; single vs co-authored publication)

Graph of G



33 isolated authors dropped

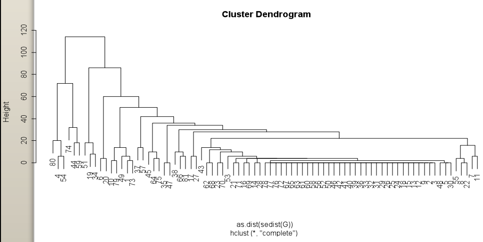
research specialty
Black= Statistics field
Grey= Economics field

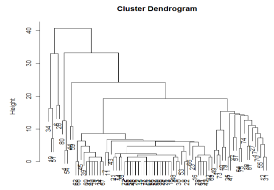
academic position
Square= Assistant professor
Circle= Full professor

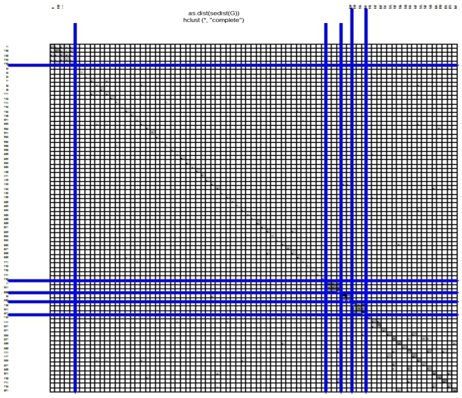
Comparison of G, G_Z and G_X

Clustering and **Blockmodeling**: to aggregate actors who present the same (or similar) pattern of ties, looking for a clustering homogeneous groups of authors

Clustering results for G

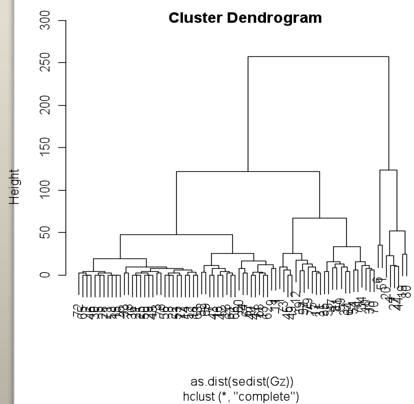


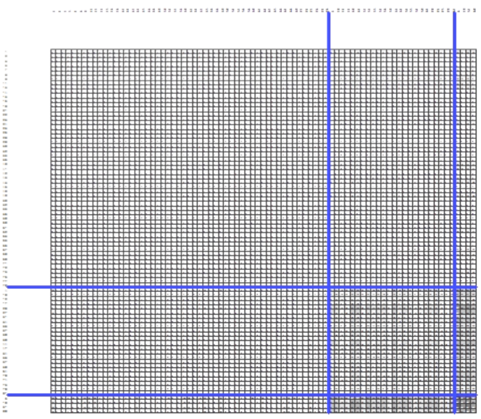




Results …

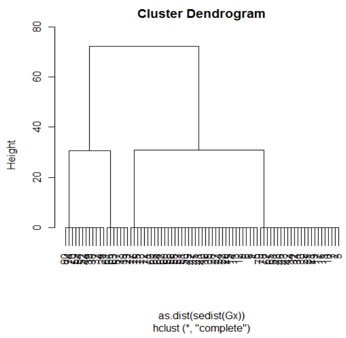
Clustering results for G_Z



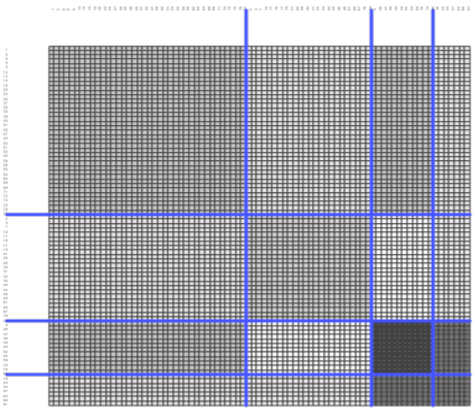


Results …

Clustering results for G_X



Results …



Comparison of partition in G, G_Z and G_X

Authors attributes

Clusters

	Total	Economists	Statisticians	Full Prof.	Assistant Prof.
G					
1	5	100,00	0,00	60,00	40,00
2	65	95,38	20,00	58,46	41,54
3	3	0,00	100,00	100,00	0,00
4	2	100,00	0,00	100,00	0,00
5	3	100,00	0,00	33,33	66,67
6	3	0,00	100,00	100,00	0,00
Gz					
1	73	78,08	21,92	58,90	41,10
2	6	50,00	50,00	83,33	16,67
3	2	100,00	0,00	100,00	0,00
Gx					
1	38	100,00	0,00	100,00	0,00
2	24	100,00	0,00	0,00	100,00
3	12	0,00	100,00	100,00	0,00
4	7	0,00	100,00	0,00	100,00

	Gz			
G	1	2	3	Total
1	5			5
2	64	1		65
3	1	2		3
4		1	1	2
5	1	1	1	3
6	2	1		3
Total	73	6	2	81

	Gx				
G	1	2	3	4	Total
1	3	2			5
2	32	20	6	7	65
3			3		3
4	2				2
5	1	2			3
6			3		3
Total	38	24	12	7	81

Preference Data Structure

(Full Profile, Metric Response)

Individual Judgements

Preference Ratings

Y_1	Y_2	...	Y_N
y_{11}	y_{12}	...	y_{1N}
y_{21}	y_{22}	...	y_{2N}
y_{31}	y_{32}	...	y_{3N}
y_{41}	y_{42}	...	y_{4N}
...	...	...	...
y_{S1}	y_{S2}	...	y_{SN}

Levels Attributes

Stimuli

x_{11}	x_{12}	x_{21}	x_{22}	x_{23}	...	x_{k1}	x_{k2}
1	0	1	0	0	...	1	0
1	0	0	1	0	...	0	1
1	0	0	0	1	...	1	0
0	1	1	0	0	...	0	1
...	...	...	...	...	...	...	...
0	1	0	0	1	...	0	1

$$\mathbf{Y} = \mathbf{XB} + err$$

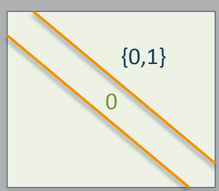
$$\mathbf{B} = (\mathbf{X'X})^{-}\mathbf{X'Y}$$

Particular case: $(\mathbf{X'X})^{-} = \mathbf{\Delta}_x^{-1} = [Diag(\mathbf{X'X})]^{-1}$

Adjacency Matrix

$a_1, a_2, \ldots\ldots\ldots, a_n$

a_1
a_2
.
.
.
a_n



The Adjacency Matrix **A** will be here considered in the case of undirected, unvalued graphs

A is Binary (presence/absence)
A is Symmetric (undirected graph)

The row marginal is the "degree" of each node

"[...] It has been found that in many networks, the distribution of actors' degrees is highly skewed, with a small number of actors having an unusually large number of ties. Simulations and analytic work have suggested that this skewness could have an impact on the way in which communities operate, including the way information travels through the network and the robustness of networks to removal of actors".

E. J. Newman, D. J. Watts, and S. H. Strogatz