Bipartite graphs and affiliation networks

Outline

Affiliation networks and bipartite graphs

Duality

What UCINET can do

Exponential random bipartite graph models

Extension to *k*-partite data structures

Civic mobilisation and the 1992 presidential impeachment in Brazil

Lattice representations for bipartite graphs

The notion of duality

duality refers to the mutually constitutive relationship between elements of two distinct sets

persons and groups in ***affiliation networks***

> Individuals come together within groups, which are collectivities based on the shared interests, personal affinities, or ascribed status of members who participate regularly in collective activities. At the same time, the particular patterning of an individual's affiliations (or the "intersection" of groups within the person) defines his points of reference and (at least partially) determines his individuality (Breiger, 1974)

cultural symbols and practice

> "neither the material world (the world of action) nor the cultural world (the world of symbols) can exist (or be coherently structured) independently. Rather each is built up through its immediate association with the other." (Mohr & Duquenne)

Exploring structural dualities

Breiger:

duality can be explored *empirically* through the structure of the relation between two sets of elements (in a similar way to network relations)

Mohr and Duquenne:

structural dualities are generalised social phenomena and can be represented formally using Galois lattices

A famous affiliation network: the Southern women

participation of 18 women in 14 social events

		1	2	3	4	5	6	7	8	9	10	11	12	13	14
		E	E	E	E	E	E	E	E	E	E	E	E	E	E
1	EVELYN	1	1	1	1	1	1	0	1	1	0	0	0	0	0
2	LAURA	1	1	1	0	1	1	1	1	0	0	0	0	0	0
3	THERESA	0	1	1	1	1	1	1	1	1	0	0	0	0	0
4	BRENDA	1	0	1	1	1	1	1	1	0	0	0	0	0	0
5	CHARLOTTE	0	0	1	1	1	0	1	0	0	0	0	0	0	0
6	FRANCES	0	0	1	0	1	1	0	1	0	0	0	0	0	0
7	ELEANOR	0	0	0	0	1	1	1	1	0	0	0	0	0	0
8	PEARL	0	0	0	0	0	1	0	1	1	0	0	0	0	0
9	RUTH	0	0	0	0	1	0	1	1	1	0	0	0	0	0
10	VERNE	0	0	0	0	0	0	1	1	1	0	0	1	0	0
11	MYRNA	0	0	0	0	0	0	0	1	1	1	0	1	0	0
12	KATHERINE	0	0	0	0	0	0	0	1	1	1	0	1	1	1
13	SYLVIA	0	0	0	0	0	0	1	1	1	1	0	1	1	1
14	NORA	0	0	0	0	0	1	1	0	1	1	1	1	1	1
15	HELEN	0	0	0	0	0	0	1	1	0	1	1	1	0	0
16	DOROTHY	0	0	0	0	0	0	0	1	1	0	0	0	0	0
17	OLIVIA	0	0	0	0	0	0	0	0	1	0	1	0	0	0
18	FLORA	0	0	0	0	0	0	0	0	1	0	1	0	0	0

UCINET options

2-mode scaling

Display of row and column entities in a common space using for example correspondence analysis (Faust, 2005)

2-mode factions

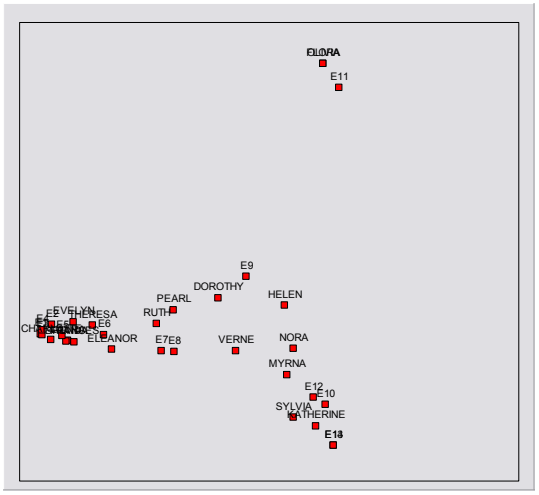
A ***pair*** of divisions of row and column categories so that concentration in diagonal blocks is maximised (uses a genetic algorithm)

Presumably, one can repeat on each faction identified

Core/periphery analysis

A pair of divisions of row and column categories so that a core-periphery structure (ie lack of ties among pairs of entities in 'peripheral' categories) is emphasised (uses a genetic algorithm)

Correspondence analysis for Southern Women



2-MODE FACTION

Blocked Adjacency Matrix

		1 2 3 4 5 6 7 8	9 0 1 2 3 4
		E E E E E E E E	E E E E E E
1	EVELYN	1 1 1 1 1 1	1	1
2	LAURA	1 1 1	1 1 1 1	
3	THERESA	1 1 1 1 1 1 1	1	
4	BRENDA	1 1 1 1 1 1 1		
5	CHARLOTTE	1 1 1	1	
6	FRANCES	1	1 1	1	
7	ELEANOR		1 1 1 1	
8	PEARL		1	1	1	1
9	RUTH		1	1 1	1	1
10	VERNE			1 1	1	1
11	MYRNA			1	1 1	1
12	KATHERINE			1	1 1	1 1 1
13	SYLVIA			1 1	1 1	1 1 1
14	NORA		1 1		1 1 1 1 1 1
15	HELEN			1 1		1 1 1
16	DOROTHY			1	1	1
17	OLIVIA				1	1
18	FLORA				1	1

Density matrix

	1	2
1	0.578	0.057
2	0.153	0.531

2-MODE CATEGORICAL CORE/PERIPHERY MODEL

		8 E	9 E	7 E	3 E	5 E	6 E	4 E	1 E	2 E	1 0 E	1 1 E	2 E	3 E	4 E
1	EVELYN	1	1		1	1	1	1	1	1					
2	LAURA	1		1	1	1	1		1	1					
3	THERESA	1	1	1	1	1	1	1		1					
4	BRENDA	1		1	1	1	1	1							
5	CHARLOTTE			1	1	1		1							
6	FRANCES	1			1	1	1								
7	ELEANOR	1		1		1	1								
12	KATHERINE	1	1								1		1	1	1
9	RUTH	1	1	1		1									
10	VERNE	1	1	1								1			
15	HELEN	1		1							1	1	1		
14	NORA		1	1			1				1	1	1	1	1
13	SYLVIA	1	1	1							1		1	1	1
8	PEARL	1	1				1								
11	MYRNA	1	1								1		1		
16	DOROTHY	1	1												
17	OLIVIA		1									1			
18	FLORA		1									1			

Density matrix

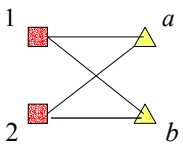
	1	2
1	0.703	0.293
2	0.500	0.093

Building exponential random bipartite graph models

1. Regard ties as variable: X_{ij} random variable signifying the relation between woman *i* and event *j*

2. Neighbourhood structure: attendances are conditionally dependent if:
   - They share a woman
   - They share an event
   - They are linked by a 4-cycle

	Events	
e.g.	*a*	*b*
woman 1	1	1
woman 2	1	0



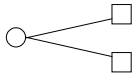
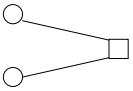
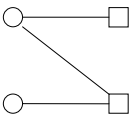
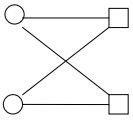
3. Distinguish women from events in homogeneity assumptions

The Southern Women: (Davis, Davis & Gardner, 1941; see also Skvoretz & Faust, 1999)

Model	*no. of parameters*	*-2LPL*	*MAR*
Bernoulli	1	327.3	.457
p_1 for two-mode relational data	31	256.2	.338
Markov (edges, 2-stars)	3	304.8	.416
Markov (edges, 2-stars, 3-stars)	5	302.6	.411
clustering (edges, 2-stars, 3-paths, $K_{2,2}$)	5	228.6	.293

The clustering model

(Key: ○ woman □ event)

neighbourhood	*parameter estimate*
	-2.65 (.60)
	1.42 (.25)
	1.06 (.17)
	-0.17 (.02)
	0.33 (.05)

Southern women: fit of clustering model

woman	*events*
1	1 1 1 1 1 1 **0** 1 **1** 0 0 0 0 0
2	1 1 1 **0** 1 1 1 1 0 0 0 0 0 0
3	**0 1** 1 1 1 1 1 1 **1** 0 0 0 0 0
4	1 **0** 1 1 1 1 1 1 0 0 0 0 0 0
5	0 0 1 **1 1 0** 1 **0** 0 0 0 0 0 0
6	0 0 **1** 0 1 1 **0 1** 0 0 0 0 0 0
7	0 0 **0** 0 1 **1** 1 1 **0** 0 0 0 0 0
8	0 0 0 0 0 **1** 0 1 1 0 0 0 0 0
9	0 0 0 0 **1** 0 **1** 1 1 0 0 0 0 0
10	0 0 0 0 0 0 **1** 1 1 0 0 **1** 0 0
11	0 0 0 0 0 0 0 1 1 **1 0 1** 0 0
12	0 0 0 0 0 0 0 1 1 1 **0** 1 1 1
13	0 0 0 0 0 0 **1** 1 1 1 **0** 1 1 1
14	0 0 0 0 0 **1 1** 0 1 1 1 1 1 1
15	0 0 0 0 0 0 1 1 **0 1 1** 1 0 0
16	0 0 0 0 0 0 0 1 1 0 0 0 0 0
17	0 0 0 0 0 0 0 **0** 1 0 **1** 0 0 0
18	0 0 0 0 0 0 0 **0** 1 0 **1** 0 0 0

bold values associated with absolute residuals ≥ .6

The convergence of the 1992 Brazilian impeachment movement (Mische)

Entities:

Groups (22 key organisations)
Projects (25 distinct projects)
Events (27 key events, divided into three main periods:
articulation, *denunciation*, *mobilisation*)

Relational data:

adoption of projects by groups
formal presence of groups at events
formal expression of projects at events

each represented by a binary array or bipartite graph

K-partite relations: the Brazilian political arena

Organisations

Youth organizations

a. National Student Union (UNE)
b. Law Students association (CAXI)
c. Catholic Youth Pastoral (PJMP)
d. Union of Socialist Youth (UJS)
e. Agronomy Students group (FEAB)
f. Black Students group (CONUN)
g. Junior Enterprises (FEJESP)
h. Business/NGOs (AIESEC)

Political parties

i. Communist Party (PCdB)
j. Workers' Party (PT)
k. Social-democratic Party (PSDB)
l. Right-wing Party (PDS)

Civic organizations

m. Bar Association (OAB)
n. Bishops' Conference (CNBB)
o. Research NGO (IBASE)
p. Unified Labor Central (CUT)

Business associations

q. Federation of Industry (FIESP)
r. "Modern" Business Association (PNBE)

Mainstream media

s. Mediate magnate (GLOBO)
t. Liberal press (VEJA.)

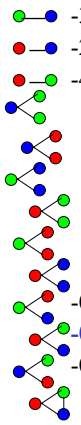
State actors

u. State Government of São Paulo (GESP)
v. Military Police (PM)

Projects

1. reforming university system
2. improving professional training
3. democratizing student movement
4. recruiting new participants
5. controlling student organizations
6. detonating revolutionary activity
7. constructing socialist society
8. empowering community organizat'n
9. defeating neo-liberal project
10. increasing social programs
11. protecting human rights
12. defending workers' rights
13. ending social discrimination
14. winning local/national elections
15. strengthening democratic instit'ns
16. expanding civic participation
17. reconstituting civic organizations
18. stabilizing economy
19. deregulating economy
20. privatizing state industries
21. expanding public ratings
22. consolidating institutional power
23. maintaining public order
24. fighting public corruption
25. defending national sovereignty
26. protecting the environment
27. conducting land reform
28. developing science and technology
29. improving education

Brazil



time 1	*time 2*	*time 3*
-3.222(.44)	-3.805(.44)	-4.678(.46)
-2.223(1.1)	-6.665(1.8)	-10.71(1.5)
-4.405(.98)	-6.333(1.5)	-9.322(1.8)
0.099(.02)	0.116(.02)	0.170(.02)
0.123(.17)	**0.734(.17)**	**1.051(.15)**
0.198(.02)	0.207(.03)	0.202(.02)
0.204(.04)	0.309(.06)	0.459(.14)
0.745(.10)	0.886(.14)	0.906(.12)
0.320(.06)	0.443(.09)	0.444(.06)
-0.177(.04)	-0.123(.05)	-0.022(.04)
-0.461(.06)	**-0.307(.06)**	**0.000(.06)**
-0.146(.07)	-0.041(.05)	-0.024(.03)
0.808(.08)	**0.472(.07)**	**0.139(.06)**

Key : organisation ● project ● event ●

What does this suggest?

Articulation phase

particularistic projects expressed in partially segmented student, labour and professional sub-contexts

Denunciation phase

Organisational diversification combined with projective suppression; the broader civic projects gained salience over more particularisitc projects in complex public settings

Mobilisation phase

Organisational convergence; particularistic projects held in suspension in order to unite diverse organisations at events

Lattice representations

Affiliation network data: membership of 5 actors in 4 groups

	group			
actor	*a*	*b*	*c*	*d*
1	1	1	1	1
2	1	1	1	0
3	0	1	1	1
4	0	1	1	0
5	0	0	0	1

Elementary comparisons

We can compare membership *profiles* across actors

e.g.	group			
	a	*b*	*c*	*d*
actor 2	1	1	1	0
actor 3	0	1	1	1

Overlaps and orderings

We can ask what groups two actors have in common:

e.g.	group			
	a	*b*	*c*	*d*
actor 2	1	1	1	0
actor 3	0	1	1	1

The ***overlap*** (or intersection) is:

0 1 1 0 *common* group memberships

For the profiles:

1 1 1 0 (actor 2)

0 1 1 0 (actor 4)

the overlap is

0 1 1 0 *same as actor 4*

Here we can say that the two profiles are ***ordered***: actor 4 profile < actor 2 profile

Representing overlaps and ordering

In general: if the overlap of profile *x* and profile *y* is equal to profile *y*, then profile *y* ≤ profile *x*

Identifying orderings and overlaps

1. Begin with the set of observed profiles
2. Compute all possible distinct overlaps between subsets of profiles
3. Add the profile that comprises all features: 1 1 1 1 … 1

The line diagram

1. Each distinct observed or overlap profile becomes a "node" in the line diagram
2. The line diagram is drawn so that profile *x* is connected to profile *y* by an ascending path of lines if and only if profile *x* < profile *y*

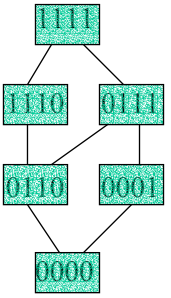
Building the line diagram

Data	**Overlaps**
groups	
abcd	
1 1111	1110 and 1111 = 1110
2 1110	1110 and 0111 = 0110
3 0111	1110 and 0110 = 0110
4 0110	1110 and 0001 = 0000
5 0001	*etc. etc.*

overlap profiles

0 0 0 0

Partial order diagram



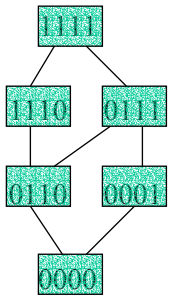
profile *x* is connected to profile *y* by an ascending path if and only if profile *x* < profile *y*

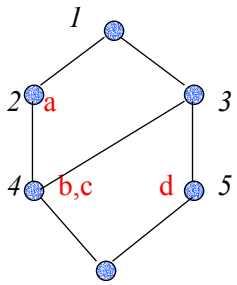
The line diagram

Data	*Partial order diagram*	*Line diagram*

abcd

1 1111
2 1110
3 0111
4 0110
5 0001



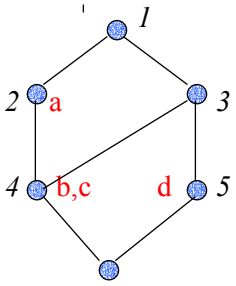


Each person is at or above the groups to which they belong
Each group is at or below the people who are members

The line diagram is in 1-1 correspondence with the data

Reading the line diagram

Line diagram



The line diagram displays:

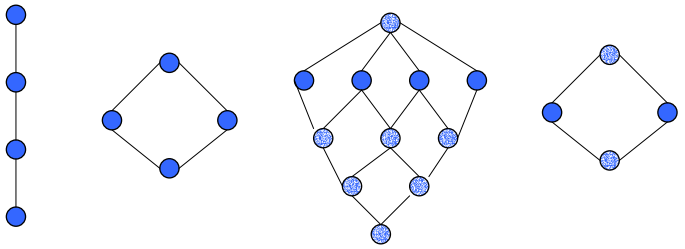
a. groups in relation to actors
b. actors in relation to groups
c. orderings/overlaps/clusters for actors
d. orderings/overlaps/cluster for groups
e. the raw data

Each actor is at or above the groups to which they belong
Each group is at or below the actors who are members

Note: the *duality* of the representation

Some simple structural forms

ordering	partial	unfolding	clustering
1111	1111	111000	1100
1110	1110	011100	1100
1100	0111	001110	0011
1000	0110	000111	0011



Note: *filled nodes indicate item locations*

The concept (Galois) lattice (Wille, 1984; Ganter & Wille, 1999)

Let

A set of n people

B set of p problems

$\mathbf{X}$ $n \times p$ data matrix, with

$X_{ij} = 1$ if person i solves problem j,

0 otherwise

For $P \subseteq A$ and $Q \subseteq B$, define

$P' = \{j \in B \mid X_{ij} = 1 \text{ for all } i \in P\}$ and $Q' = \{i \in A \mid X_{ij} = 1 \text{ for all } j \in Q\}$

The mappings $P \rightarrow P'$ and $Q \rightarrow Q'$ form a *Galois connection* between the power sets of A and B

The concept lattice and the line diagram

A *concept* is a pair (P,Q) such that $P' = Q$ and $Q' = P$

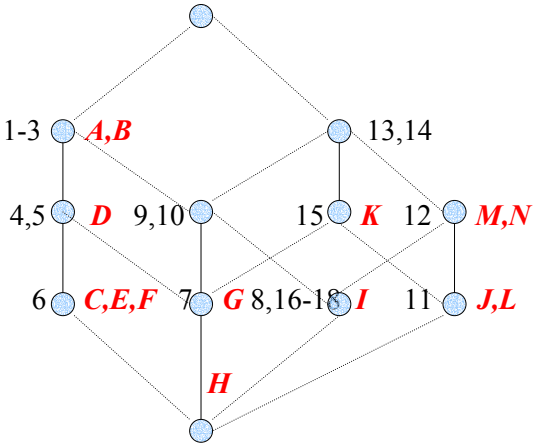
Partial ordering of concepts: $(P_1,Q_1) \leq (P_2,Q_2)$ iff $P_1 \subseteq P_2$ iff $Q_2 \subseteq Q_1$

The set of all concepts under this partial ordering form a *lattice*

The *line diagram* represents the lattice:

Each node corresponds to a concept
$(P_1,Q_1) \leq (P_2,Q_2)$ iff there is an ascending path from (P_1,Q_1) to (P_2,Q_2)
The label *i* is attached to the concept $(\{i\}'',\{i\}')$
The label *j* is attached to the concept $(\{j\}',\{j\}'')$

Southern women: rank 4 approximation



TOSCANA (Wille and others)

A serious lattice visualisation tool for bipartite graphs

Worth a look!

Finally

Many thanks

to:

Sean and Rebecca, for everything

All the people who shared their projects with us all

All of you, for great questions, much patience and excellent stamina

Stay connected!

References

Breiger, R. L. (1974) The duality of persons and groups. Social Forces,53, 181-190.

Faust, K. (2005). Using correspondence analysis for joint displays of affiliation networks. In P. Carrington, J. Scott, & S. Wasserman (Eds). Models and methods in social network analysis. Cambridge University Press.

Freeman, L., & White, D. R. (1993). Using Galois lattices to represent network data. In P. V. Marsden (Ed.), *Sociological Methodology 1993*, (pp. 127-146). Cambridge, MA: Basil Blackwell.

Mische, A., & Pattison, P. E. (2000). The plurality of civic relations: publics, projects and social settings. *Poetics*, *27*, 163-194.

Wille, R. (1984). Line diagrams of hierarchical concept systems. *International Journal of Classification*, *11*, 77-86.