Multivariate analysis and complex networks

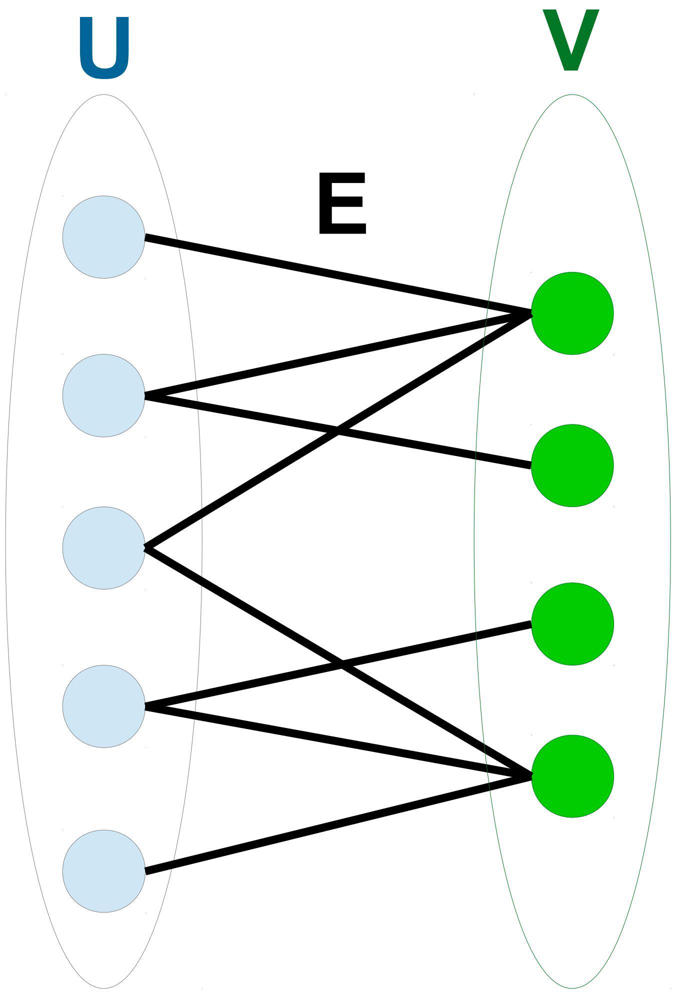
2: Bipartite Networks & Projections

Antonio Scala

CNR-ISC, the Institute for Complex Systems

INFN School of Statistics
Ischia, May 2017

Bipartite Graphs



Bipartite graph:

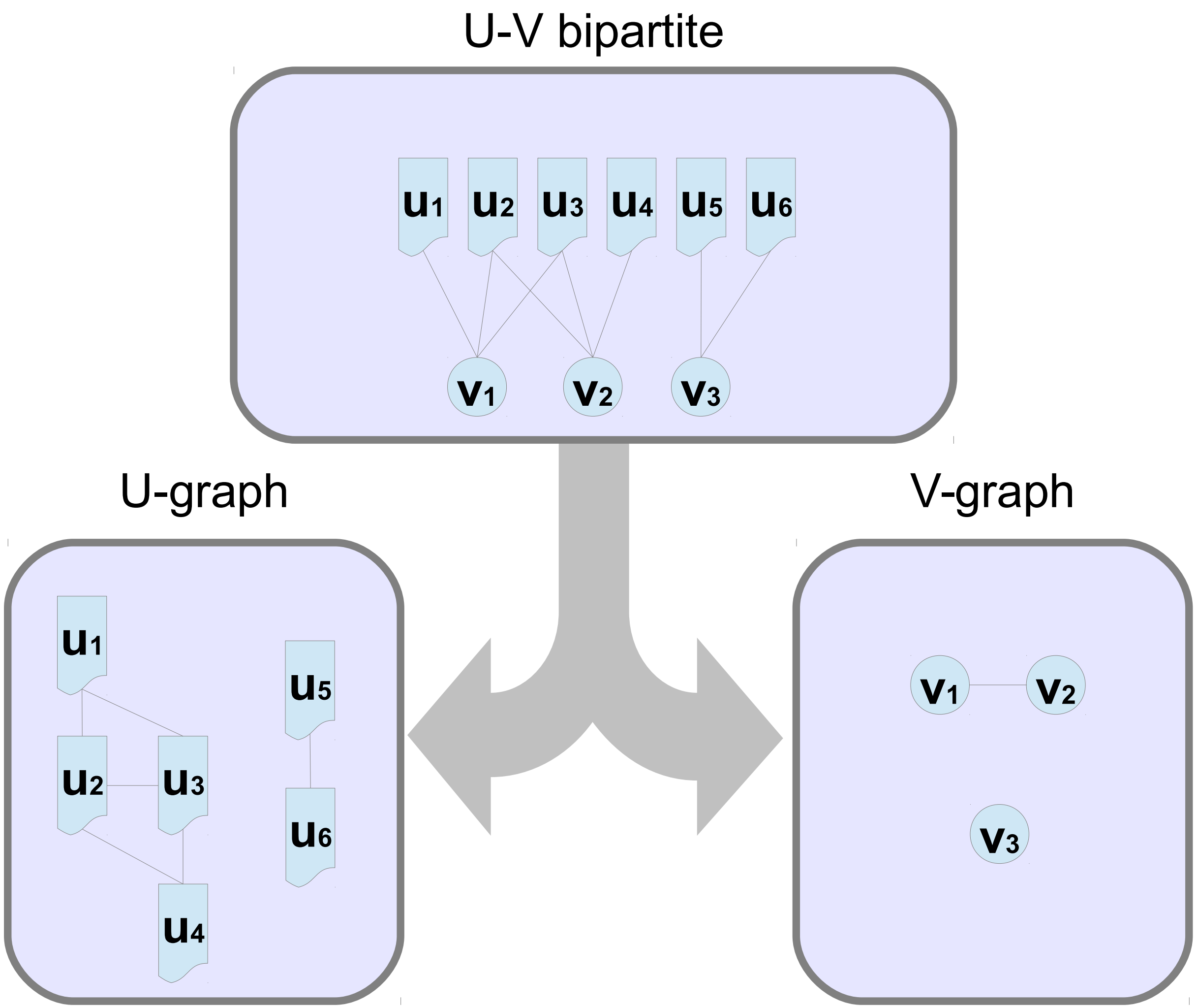
- G = (U, V, E)
- U, V nodes
- E edges among U,V

Bipartite graphs arise naturally when modelling relations between two different classes of entities

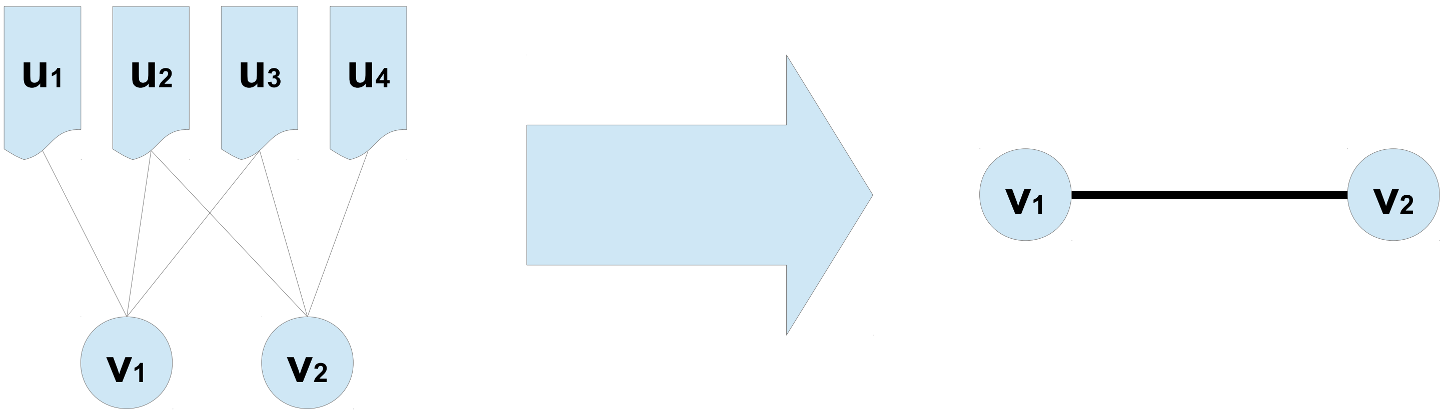
Examples

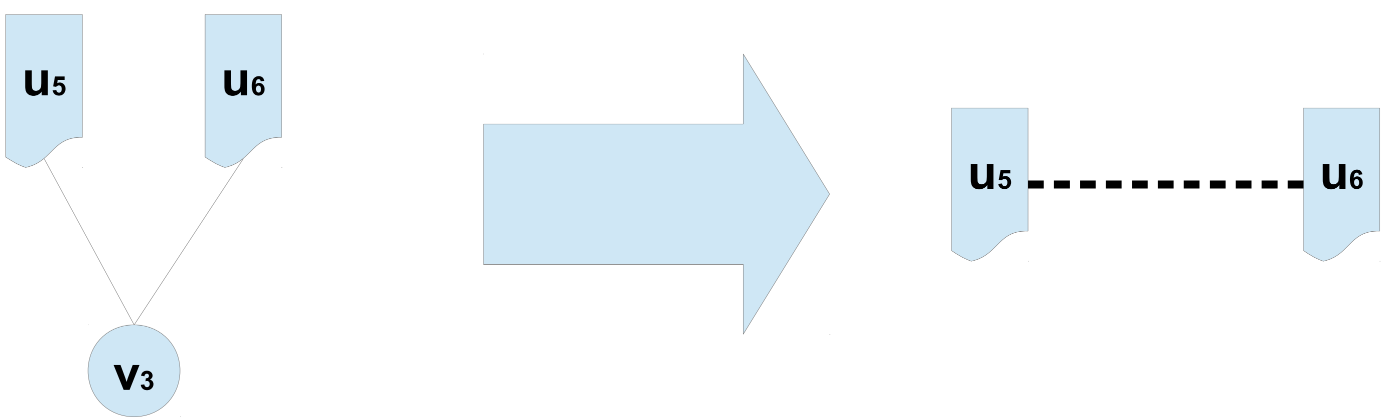
- Film – Actors
- Papers – Authors
- Nation – Products
- Text – Words
- Birds – Islands
- Plants – Pollinators
- Politician – Votes
- Users – Newsfeeds
- Patients – Sympthoms
- Proteins - xxxx

Projections



Projections



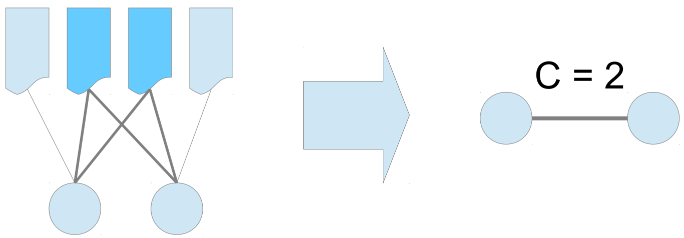


Co-occurrence matrix

- B = adjacency matrix of bipartite graph G = (U,V,E)
- B_{uv} = 1 if v has feature u , B_{uv} = 0 otherwise

The co-occurrence C_{uw} counts the number of times two features u,w occur together

$$C_{uw} = \Sigma_v B_{uv} B_{wv}$$



- $B B^T$ → weighted adjacency matrix of the projection graph on U
- $B^T B$ → weighted adjacency matrix of the projection graph on V

Null model

- C_{uw} = numbers of common neighbors of u,w
- n = maximum possible number of links
- $d_u = C_{uu}$ = degree of u
- $f_u = d_u / n$ fraction of possible links present
- If nodes were chosen at random:

$$f_{uv} = C_{uw} / n \rightarrow f_u f_v$$

$$P_{uw}(C) = \binom{C}{n} (f_u f_v)^C (1-f_u f_v)^{n-C}$$

Other Projections

- Similarity Matrix

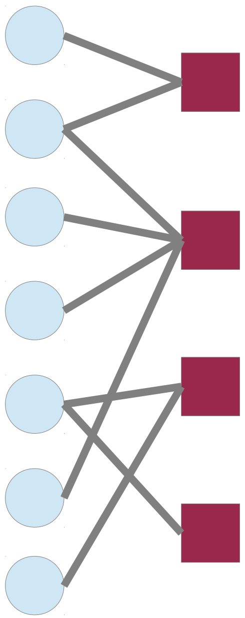
$$S_{uw} = 2\ C_{uw} / (\ C_{uu} + C_{ww}\ )$$

- Correlation matrix

$$\varphi_{uv} = (\ f_{uv} - f_u\ f_v\ ) / \sigma_u\ \sigma_v$$

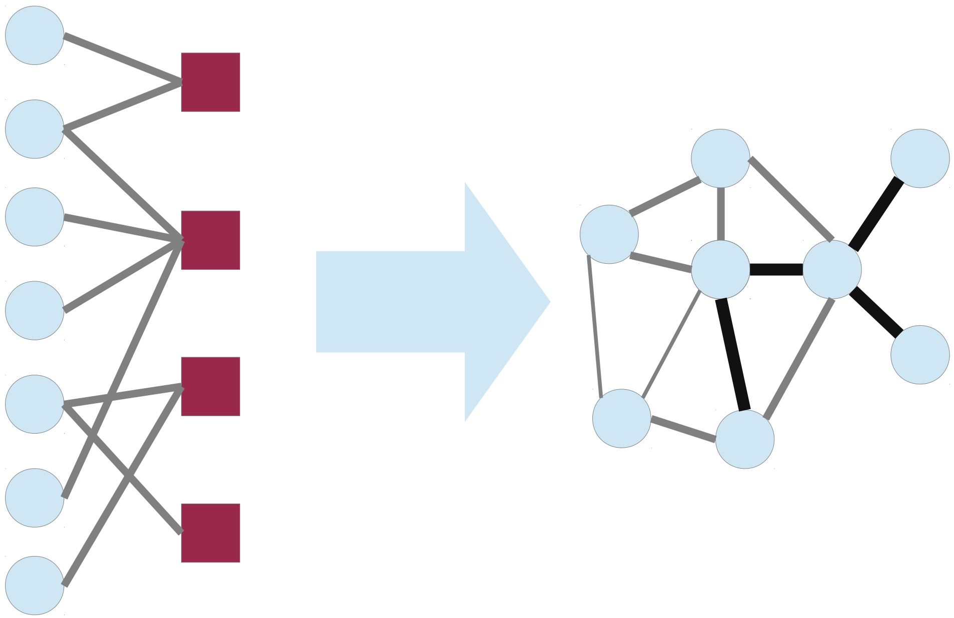
$$\sigma_u^2 = f_u\,(1-f_u\ )$$

Methodology



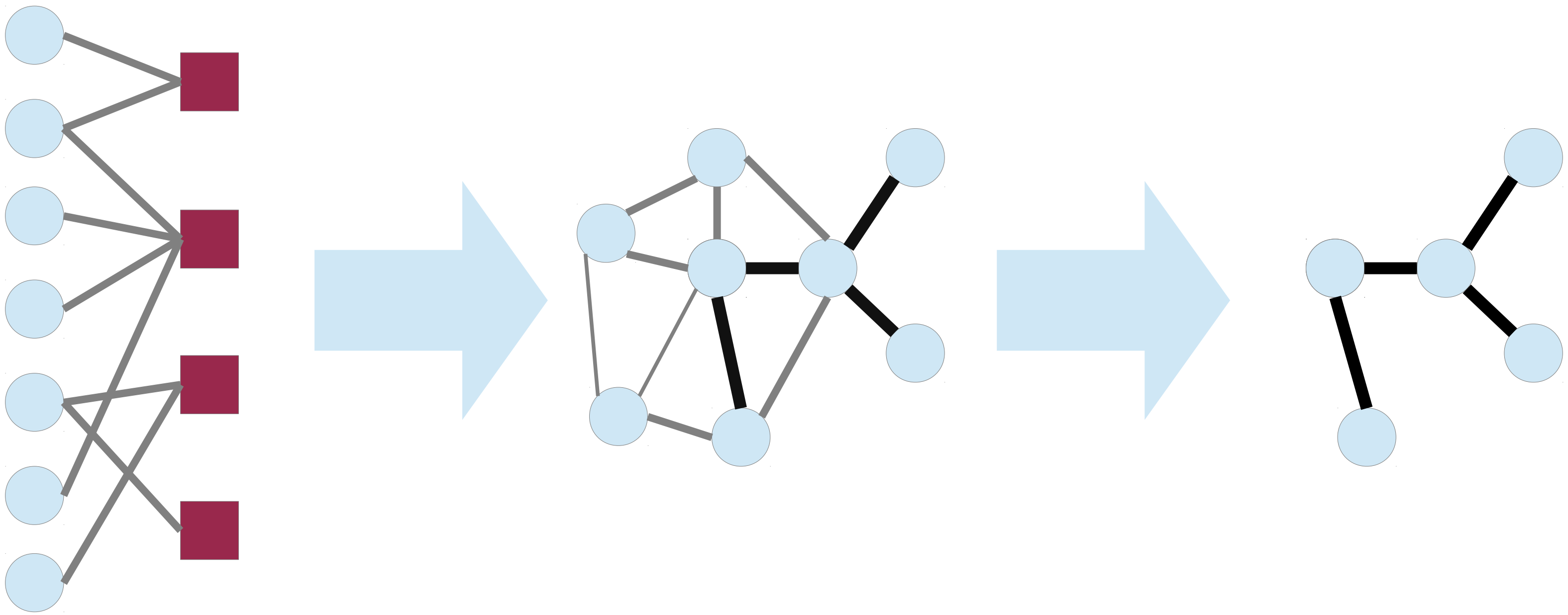
COLLECT

Methodology



PROJECT

Methodology



SELECT