

I: Introduction aux réseaux et à leur analyse

Formation Analyse des réseaux
MIREs

Outline

- 1 Premiers exemples et terminologie
- 2 Représentation des réseaux
- 3 Statistiques résumées simples
- 4 Modèles probabilistes
- 5 Covariables

Données relationnelles



Les réseaux peuvent représenter des

- relations sociales (amitié, connaissance, professionnelles),
- échanges,
- inventaires,
- ...

Les réseaux peuvent être ou non bipartites : les interactions peuvent avoir lieu exclusivement entre des nœuds appartenant à deux groupes fonctionnels différents.

Terminologie

Un réseau est constitué de :

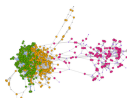
- nœuds / sommets qui représentent des individus / acteurs qui interagissent ou non,
- liens / arêtes / connexions qui représentent les interactions entre les paires de nœuds (dyade).

Un réseau peut être

- dirigé / orienté (e.g. échange),
- symétrique / non-dirigé (e.g. amitié),
- avec ou sans boucle.

Cette distinction a un sens uniquement pour les réseaux simples (pas bipartite).

Données disponibles et but



Données :

- Le réseau est fourni par :
 - une matrice d'adjacence (réseau simple) ou une matrice d'incidence (réseau bipartite),
 - une liste de dyades connectés (c'est-à-dire toutes les arêtes). Il est sous entendu que les dyades non mentionnés ne sont pas connectées...
- des covariables additionnelles sur les nœuds ou sur les dyades.

Buts :

- Révéler / décrire / modéliser la topologie du réseau.
- Découvrir des structures d'interactions particulières entre des sous-parties du réseau.
- Comprendre l'hétérogénéité du réseau.
- Pas d'inférer le réseau !

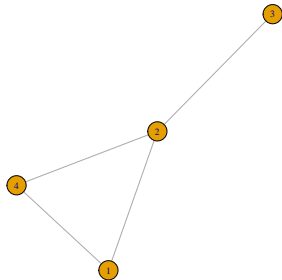
Représentation du réseau

Matrice d'adjacence :

$$X = \begin{pmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 \end{pmatrix}$$

Liste d'arêtes (si
dirigé attention à
l'ordre...)

$$\begin{pmatrix} 1 & 2 \\ 1 & 4 \\ 2 & 3 \\ 2 & 4 \end{pmatrix}$$



- n lignes et n colonnes,
- réseau non dirigé = matrice d'adjacence symétrique.

Réseau bipartite

Matrice d'incidence

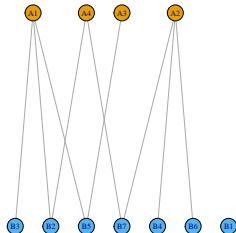
$$X = \begin{pmatrix} 0 & 1 & 1 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

Liste d'arêtes :

$$\begin{pmatrix} A1 & B2 \\ A1 & B3 \\ A1 & B5 \\ A2 & B4 \\ A2 & B6 \\ A2 & B7 \\ A3 & B5 \\ A4 & B2 \\ A4 & B7 \end{pmatrix}$$

- n lignes et m colonnes, matrice rectangulaire.
- matrice d'adjacence correspondante $(n + m) \times (n + m)$:

$$\begin{pmatrix} 0 & X \\ X^T & 0 \end{pmatrix}$$



Outline

1 Premiers exemples et terminologie

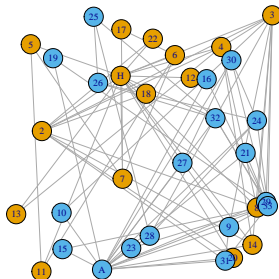
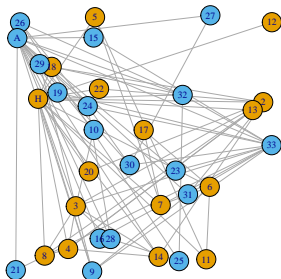
2 Représentation des réseaux

3 Statistiques résumées simples

4 Modèles probabilistes

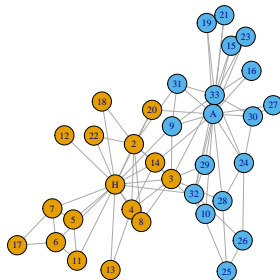
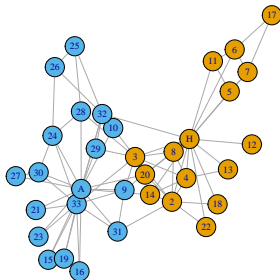
5 Covariables

Différentes représentations possibles : au hasard



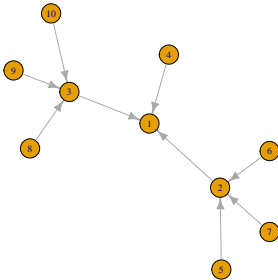
Fruchterman–Reingold

Arêtes de longueurs à peu près égales et peu de croisement. Mais représentations différentes à chaque appel.

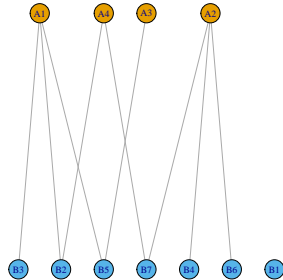


Autres représentations

Graphe dirigé



Graphe bipartite



Outline

- 1 Premiers exemples et terminologie
- 2 Représentation des réseaux
- 3 Statistiques résumées simples**
- 4 Modèles probabilistes
- 5 Covariables

Quelques statistiques résumées simples

- Degrés, sortant ou entrant si dirigés :

$$\vec{D}_i = \sum_j x_{ij}$$

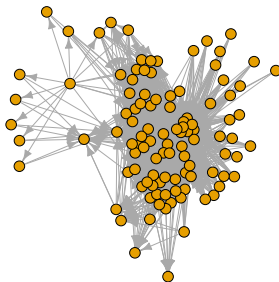
$$\overleftarrow{D}_j = \sum_i x_{ij}$$

- Nestedness : si les nœuds de plus petit degré se connectent aux nœuds de plus haut degré, [Rodríguez-Gironés & Santamaria \(2006\)](#)
- Centralité (betweenness) : pour un nœud, nombres de chemins les plus courts entre n'importe quelle paire de nœuds passant par ce nœud. [Freeman \(1979\)](#)
- Modularité : mesure pour une partition de sa tendance à favoriser l'intra-connexion par rapport à l'inter-connexion. \Rightarrow Recherche de la meilleure partition. [Clauset, Newman & Moore \(2004\)](#)

Critères à adapter aux

- réseaux dirigés,
- réseaux bipartites.

Example Chilean food web

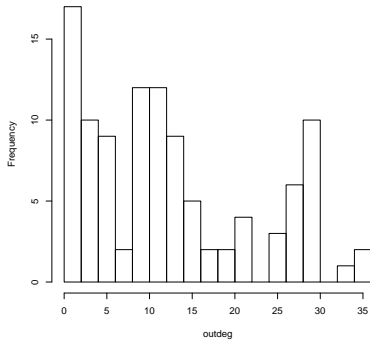


- $n = 106$ espèces / nœuds,
- densité : 12.1%.

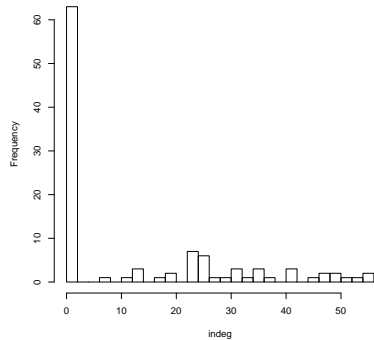
Kéfi, Miele, Wieters, Navarrete & Berlow (2016)

Distribution des degrés

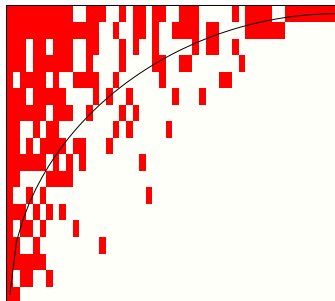
Histogram of outdeg



Histogram of indeg

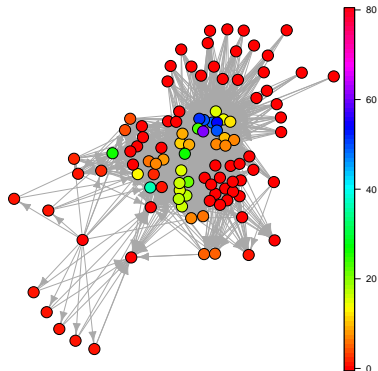


Nestedness



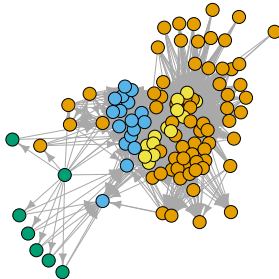
- plus généralement sur les graphes bipartites,
- ce réseau est trouvé emboîté.

Centralité (Betweenness)



Min.	1st Qu.	Median	Mean	3rd Qu.	Max.
0.000	0.000	0.000	6.604	6.929	59.570

Modularité



1	2	3	4
69	17	7	13

- modularité très faible.

Outline

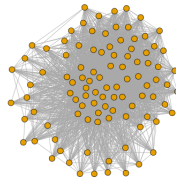
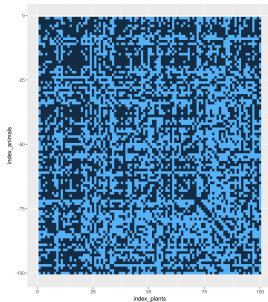
- 1 Premiers exemples et terminologie
- 2 Représentation des réseaux
- 3 Statistiques résumées simples
- 4 Modèles probabilistes**
- 5 Covariables

Modèles

- Stochastic Block Model (SBM) pour les réseaux simples / matrices d'adjacence.
- Latent Block Model (LBM) pour les réseaux bipartites / matrices d'incidence.
- Exponential Random Graph Model pour les réseaux simples.

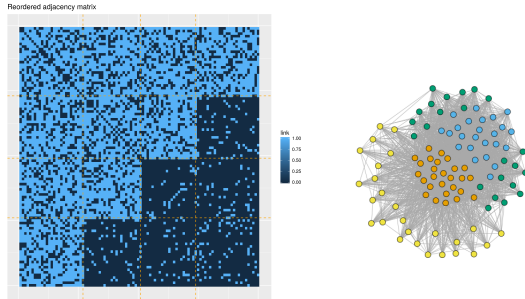
Inférence du SBM

De...



Inférence du SBM

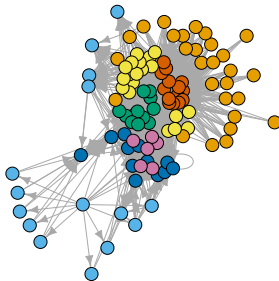
... à



Travail statistique

- Identifier les groupes
- Trouver le nombre de cluster
- Estimation des paramètres

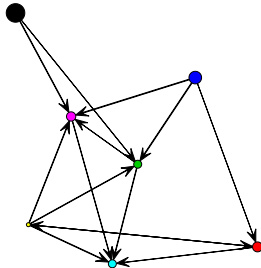
Application au réseau trophique chilien



- 7 groupes/blocs/clusters,

	1	2	3	4	5	6	7
	28	15	12	19	12	14	6

Application au réseau trophique chilien



Outline

- 1 Premiers exemples et terminologie
- 2 Représentation des réseaux
- 3 Statistiques résumées simples
- 4 Modèles probabilistes
- 5 Covariables**

Différents types de covariables

■ Nœuds

- catégorielle : ethnie, village...
- quantitative : âge, surface cultivée...

■ Dyades

- à partir des covariables sur les nœuds : différence d'âge, âge du donneur ou âge du receveur,
- distance géographique entre les exploitations.

Analyse à l'aide des covariables

Interprétation :

- Covariables au niveau des nœuds à mettre en relation avec statistiques propres au nœuds (degré, centralité ...)
- Analyser la structuration des covariables au niveau des nœuds par rapport aux groupes du SBM ou modularité...

Modélisation :

- Tester l'effet de covariables sur les nœuds ou dyades sur la présence d'une arête,
- Intégration dans un modèle pour analyser au delà de l'effet des covariables (prendre en compte un effet d'échantillonnage pour qu'il ne perturbe pas l'analyse).