Network Modeling

Viviana Amati Jürgen Lerner Mehwish Nasim Bobo Nick

Dept. Computer & Information Science
University of Konstanz

Winter 2012/2013
(version 07 February 2013)

Outline

Outline

Introduction

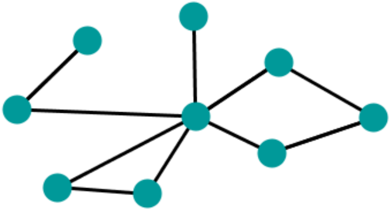
Where are we going?

The Stochastic actor-oriented model

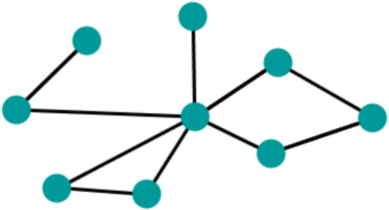
Extending the model: analyzing the co-evolution of networks and behavior

Something more on the SAOM

Where we are

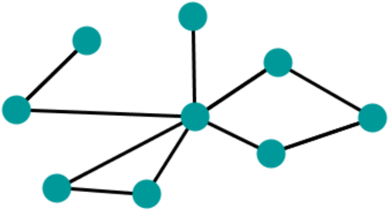


Where we are



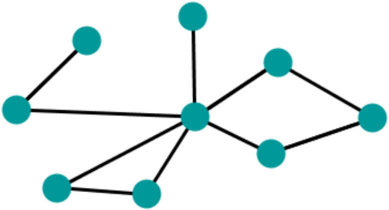
Model	Main feature	Real data
$\mathcal{G}(n, p)$	ties are independent	ties are usually dependent

Where we are



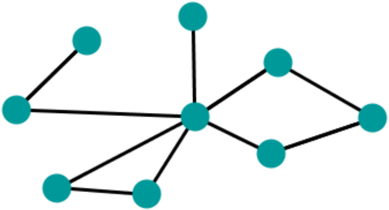
Model	Main feature	Real data
$\mathcal{G}(n, p)$	ties are independent	ties are usually dependent
Preferential attachment	based on degree distribution	there are other structural properties

Where we are



Model	Main feature	Real data
$\mathcal{G}(n, p)$	ties are independent	ties are usually dependent
Preferential attachment	based on degree distribution	there are other structural properties
ERGM	class of models	reasonable representation of the data

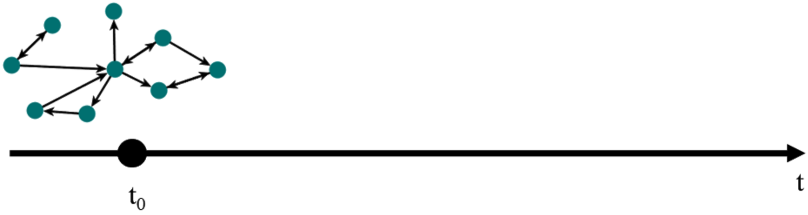
Where we are



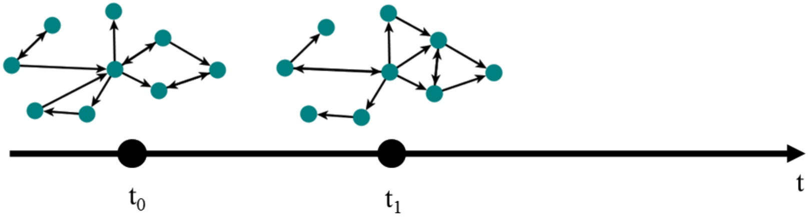
Model	Main feature	Real data
$\mathcal{G}(n, p)$	ties are independent	ties are usually dependent
Preferential attachment	based on degree distribution	there are other structural properties
ERGM	class of models	reasonable representation of the data

These are models for cross-sectional data

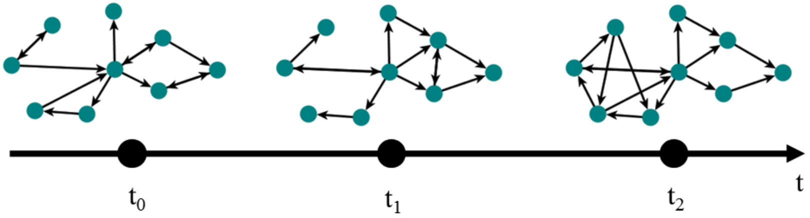
Where we are going



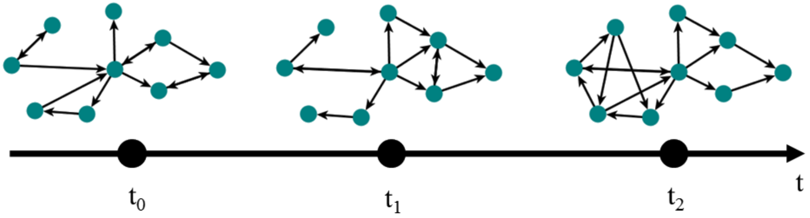
Where we are going



Where we are going

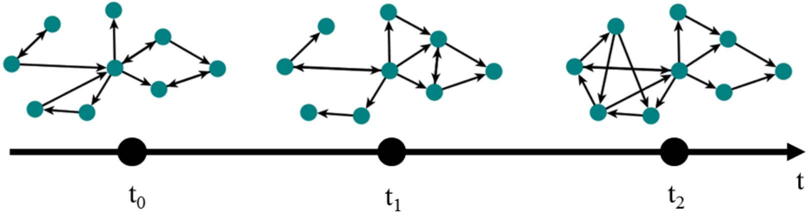


Where we are going



Network are dynamic by nature. How to model network evolution?

Where we are going



Network are dynamic by nature. How to model network evolution?

We need a model for longitudinal data

Networks are dynamic by nature: a real example

The *Teenage Friends and Lifestyle Study* analyzes smoking behavior and friendship

Data collection: (available from http://www.stats.ox.ac.uk/∼snijders/siena/)

- One school year group monitored over 3 years;
- questionnaires at approximately one year interval:
    1. Friendship relation: each pupil could name up to 12 friends
    2. Individual information and lifestyle elements: gender, age, substances use, smoking of parents and siblings etc.

Networks are dynamic by nature: a real example

The *Teenage Friends and Lifestyle Study* analyzes smoking behavior and friendship

Data collection: (available from http://www.stats.ox.ac.uk/∼snijders/siena/)

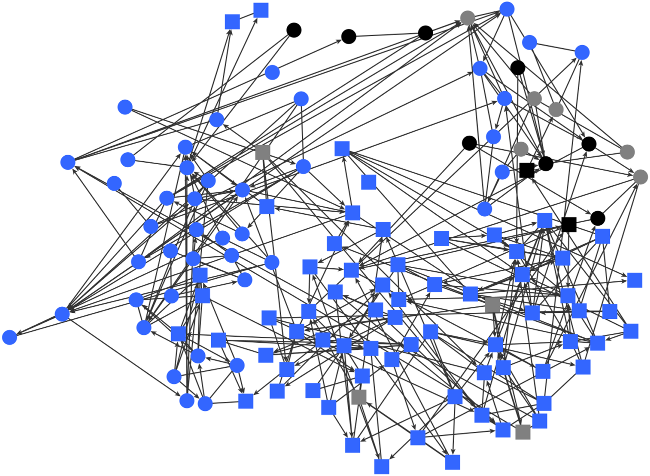
- One school year group monitored over 3 years;
- questionnaires at approximately one year interval:
    1. Friendship relation: each pupil could name up to 12 friends
    2. Individual information and lifestyle elements: gender, age, substances use, smoking of parents and siblings etc.

arrows = friendship relation

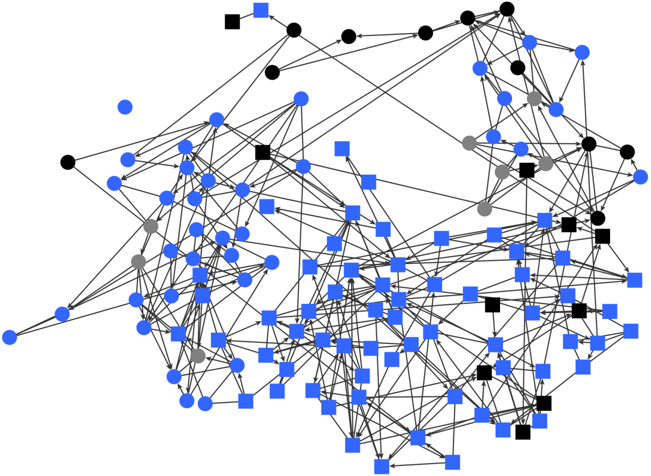
gender: circle = girl, square = boy

smoking behavior: blue = non, gray = occasional, black = regular

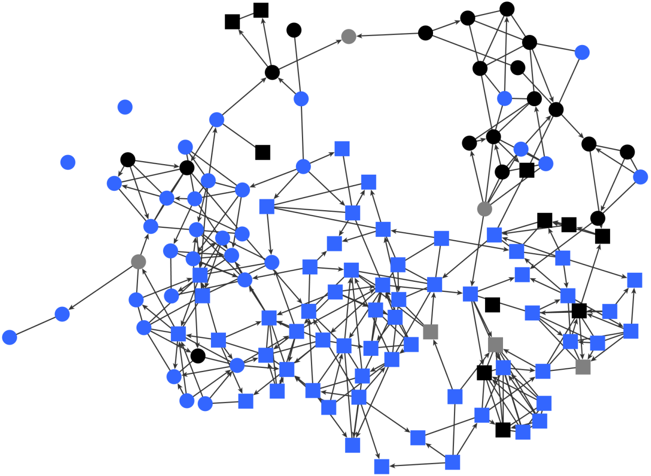
Networks are dynamic by nature: a real example



Networks are dynamic by nature: a real example



Networks are dynamic by nature: a real example



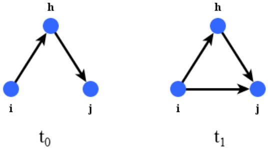
Some questions

- Is there any tendency in friendship formation towards reciprocity?

Some questions

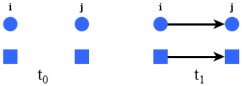
- Is there any tendency in friendship formation towards reciprocity?

- Is there any tendency in friendship formation towards transitivity?



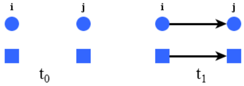
Some questions

- Is there any homophily in friendship formation with respect to gender?

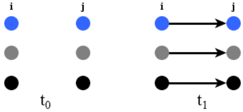


Some questions

- Is there any homophily in friendship formation with respect to gender?



- Is there any homophily in friendship formation with respect to smoking behavior?



Solution

Stochastic actor-oriented model (SAOM)

Aim

Explain network evolution as a result of

- endogenous variables: structural effects depending on the network only (e.g. reciprocity, transitivity, etc.)
- exogenous variables: actor-dependent and dyadic-dependent covariates (e.g. effect of a covariate on the existence of a tie or on homophily)

simultaneously

Background: random variable

Definition

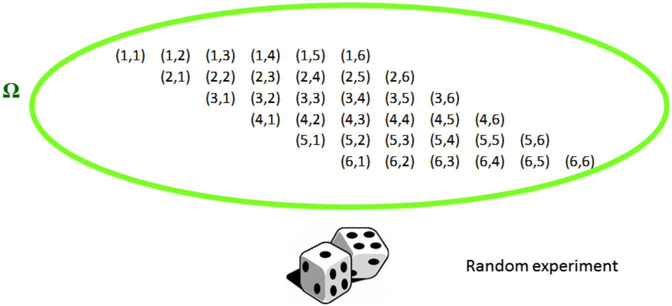
Let (Ω, P) be a probability space.
A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Background: random variable

Definition

Let (Ω, P) be a probability space.
A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Example

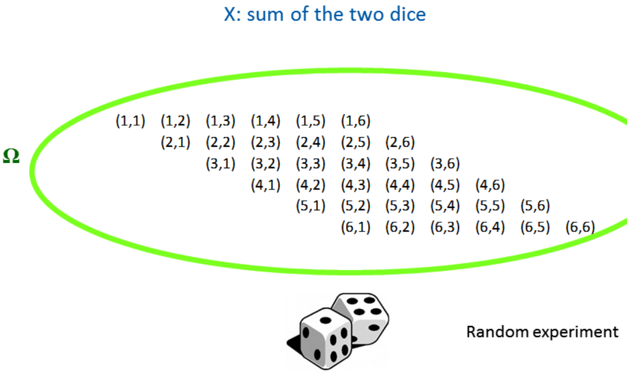


Background: random variable

Definition

Let (Ω, P) be a probability space.
A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Example

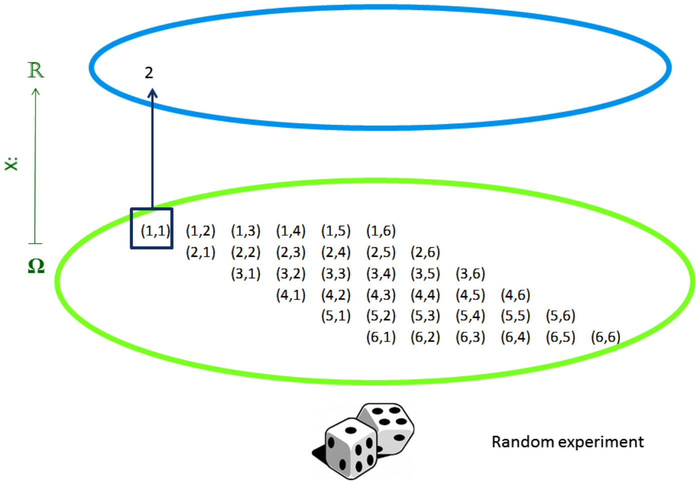


Background: random variable

Definition

Let (Ω, P) be a probability space.
A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Example

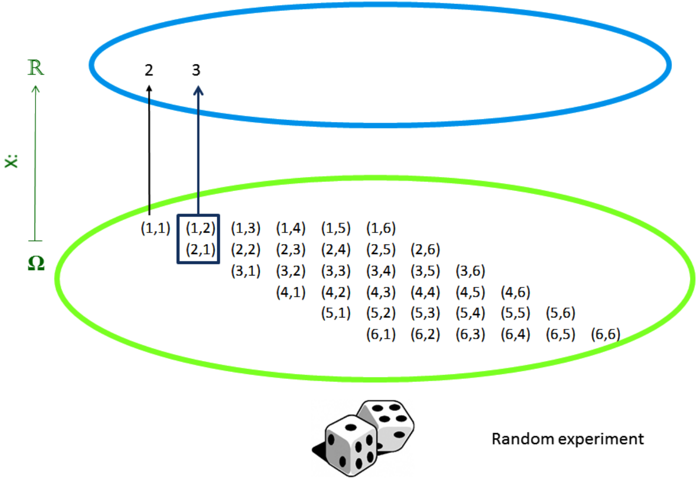


Background: random variable

Definition

Let (Ω, P) be a probability space.
A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Example

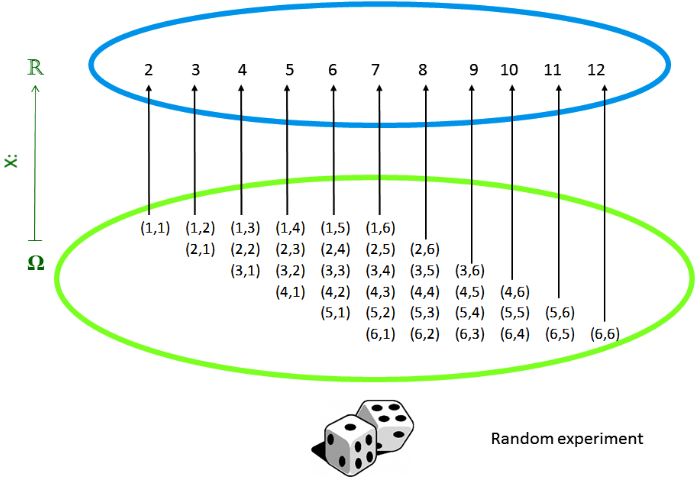


Background: random variable

Definition

Let (Ω, P) be a probability space.
A (real-valued) *random variable* (r.v.) is a function $X : \Omega \to \mathbb{R}$.

Example



Background: stochastic (or random) process

Definition

A stochastic process $\{X(t), t \in \mathcal{T}\}$ is a mapping

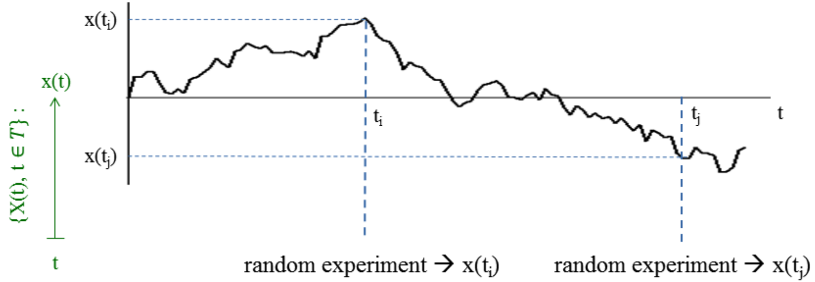
$$\forall t \in \mathcal{T} \mapsto X(t) : \Omega \to \mathbb{R}$$

Background: stochastic (or random) process

Definition

A stochastic process $\{X(t), t \in \mathcal{T}\}$ is a mapping

$$\forall t \in \mathcal{T} \mapsto X(t) : \Omega \to \mathbb{R}$$



Background: stochastic process

$\mathcal{T}$ = index set (usually interpreted as time)
$\mathcal{S}$ = state space

Different stochastic processes can be defined according to $\mathcal{S}$ and $\mathcal{T}$

$\mathcal{S}$	$\mathcal{T}$	
	Countable (discrete)	**Uncountable** (continuous)
Countable (finite)	discrete-time with finite state space	continuous-time with finite state space
Uncountable (continuous)	discrete-time with continuous state space	continuous-time with continuous state space

Background: stochastic process

Example

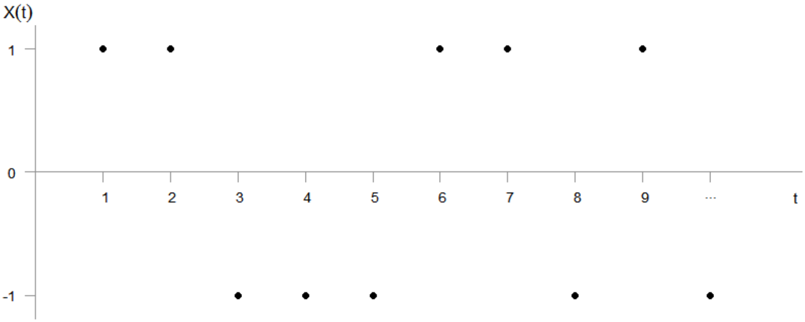
$X(t) =$ the outcome of flipping a coin

Background: stochastic process

Example

$X(t)$ = the outcome of flipping a coin

$\mathcal{S} = \{-1, 1\}$, where -1 =tail 1 =head
$\mathcal{T} = \{1, 2, \cdots\}$



$\{X(t), t \in \mathcal{T}\}$ is a discrete-time stochastic process with a finite state space

Background: stochastic process

Example

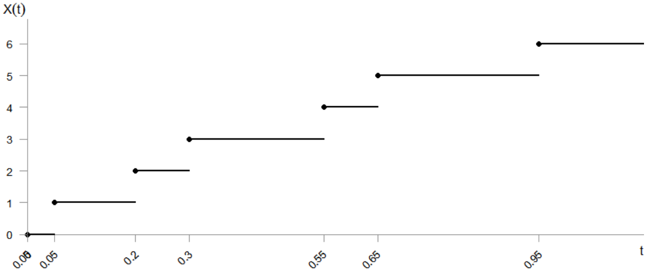
$X(t) =$ the number of telephone call at a switchboard of a company from 8 a.m. to 8 p.m.

Background: stochastic process

Example

$X(t)$ = the number of telephone call at a switchboard of a company from 8 a.m. to 8 p.m.

$\mathcal{S} = \{0, 1, 2, \cdots\}$
$\mathcal{T} = [0, 12]$



$\{X(t), t \in \mathcal{T}\}$ is a continuous-time stochastic process with a finite state space

Background: continuous-time Markov Chain

Definition

$\{X(t),\ t \in \mathcal{T}\}$ has the *Markov property* if:

$\forall\ x \in \mathcal{S}$ and $\forall\ t_i < t_j$

$$P(X(t_j) = x(t_j) \mid X(t) = x(t) \ \ \forall \ \ t \leq t_i) = P(X(t_j) = x(t_j) \mid X(t_i) = x(t_i))$$

Definition

A *continuous-time* Markov chain $\{X_t,\ t \geq 0\}$ is a stochastic process having

1. **finite state**
2. **continuous-time**
3. the **Markovian property**

Example

$X(t) = \#$ of goals that a given soccer player scores by time t (time played in official matches)

$$\{X(t),\ t \geq 0\} \text{ is a continuous-time Markov chains}$$

Why?

Background: continuous-time Markov Chain

Example

$X(t) = \#$ of goals that a given soccer player scores by time t (time played in official matches)

$$\{X(t),\ t \geq 0\} \text{ is a continuous-time Markov chains}$$

Why?

1. state space: $\mathcal{S} = \{0, 1, 2, \ldots, B\}$
   B = total number of goals scored during the career

Background: continuous-time Markov Chain

Example

$X(t) = \#$ of goals that a given soccer player scores by time t (time played in official matches)

$$\{X(t),\ t \geq 0\} \text{ is a continuous-time Markov chains}$$

Why?

1. state space: $\mathcal{S} = \{0, 1, 2, \ldots, B\}$
   B = total number of goals scored during the career
2. the time is continuous: [0,T]
   T = time of retirement

Background: continuous-time Markov Chain

Example

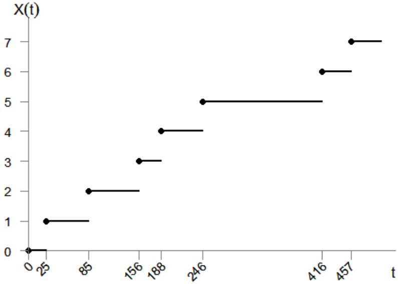
$X(t) = \#$ of goals that a given soccer player scores by time t (time played in official matches)

$$\{X(t),\ t \geq 0\} \text{ is a continuous-time Markov chains}$$

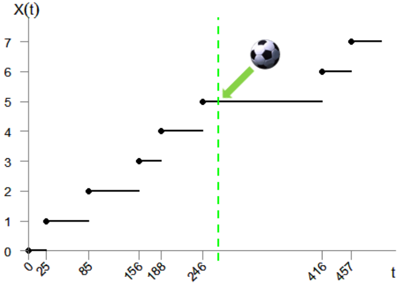
Why?

1. state space: $\mathcal{S} = \{0, 1, 2, \ldots, B\}$
   B = total number of goals scored during the career
2. the time is continuous: [0,T]
   T = time of retirement
3. the process $\{X(t),\ t \geq 0\}$ has the Markov property

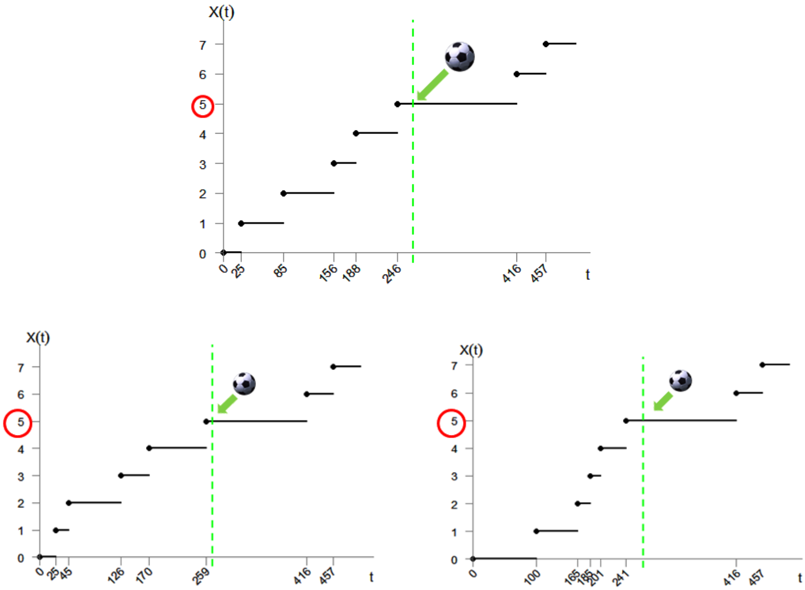
Background: Markov property



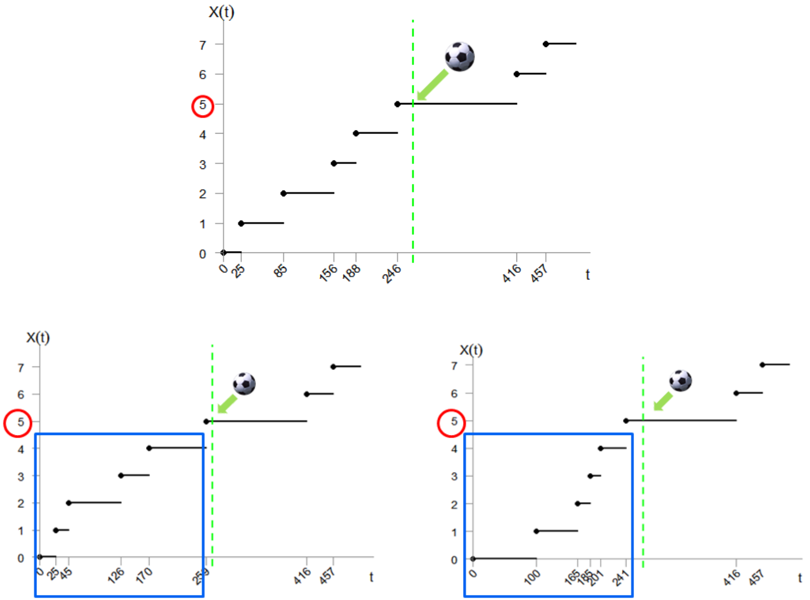
Background: Markov property



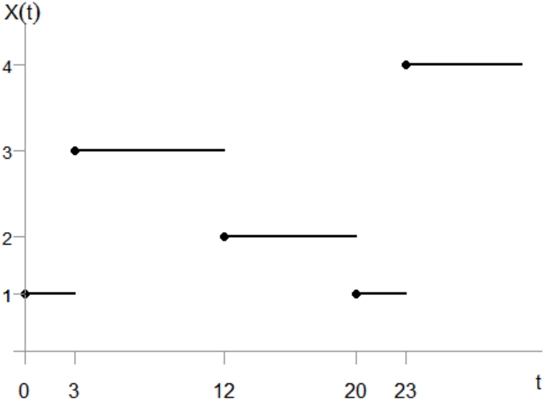
Background: Markov property



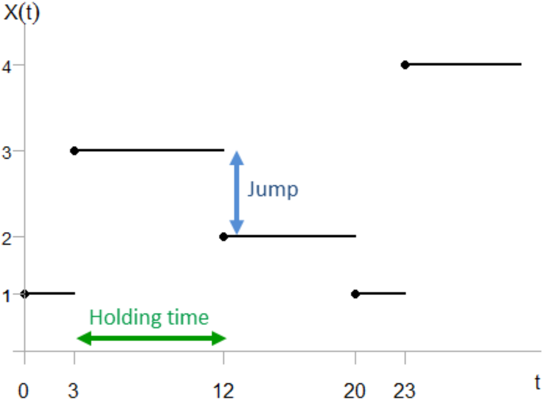
Background: Markov property



Background: describing a continuous-time Markov chain



Background: describing a continuous-time Markov chain



Background: describing a continuous-time Markov chain

Holding time

T = amount of time the chain spends in state i (Exponential r.v.)

$$f_T(t) = \lambda_i e^{-\lambda_i t}, \qquad \lambda_i > 0, \quad t > 0$$

$f_T(t) : \mathbb{R}^+ \to \mathbb{R}^+$ such that

$$P(T \leq t') = \int_0^{t'} f_T(t) dt = 1 - e^{-\lambda_i t'} \qquad \forall t \geq 0$$

Background: describing a continuous-time Markov chain

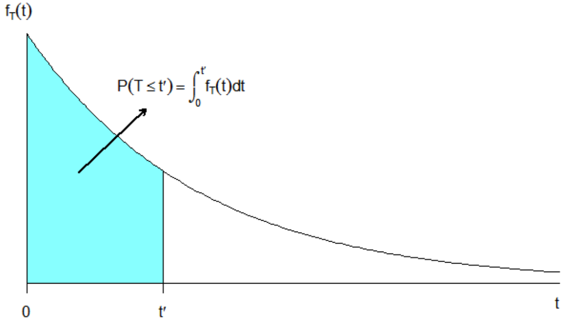
Holding time

T = amount of time the chain spends in state i (Exponential r.v.)

$$f_T(t) = \lambda_i e^{-\lambda_i t}, \qquad \lambda_i > 0, \quad t > 0$$

$f_T(t) : \mathbb{R}^+ \rightarrow \mathbb{R}^+$ such that

$$P(T \leq t') = \int_0^{t'} f_T(t)dt = 1 - e^{-\lambda_i t'} \qquad \forall t \geq 0$$



Background: describing a continuous-time Markov chain

Holding time

T = amount of time the chain spends in state i (Exponential r.v.)

$$f_T(t) = \lambda_i e^{-\lambda_i t}, \qquad \lambda_i > 0, \quad t > 0$$

λ_i is the rate parameter

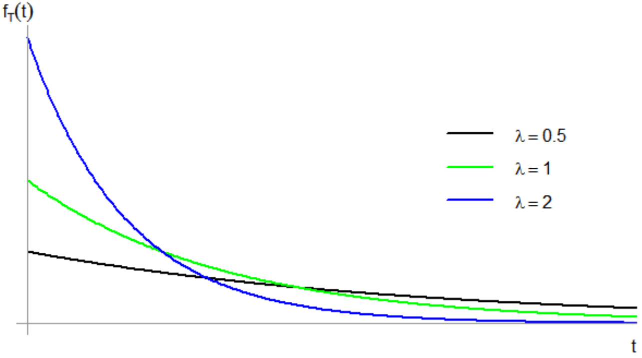
Background: describing a continuous-time Markov chain

Holding time

T = amount of time the chain spends in state *i* (Exponential r.v.)

$$f_T(t) = \lambda_i e^{-\lambda_i t}, \qquad \lambda_i > 0, \quad t > 0$$

λ_i is the rate parameter



Background: describing a continuous-time Markov chain

Holding time

The Exponential r.v. has the *memoryless property*

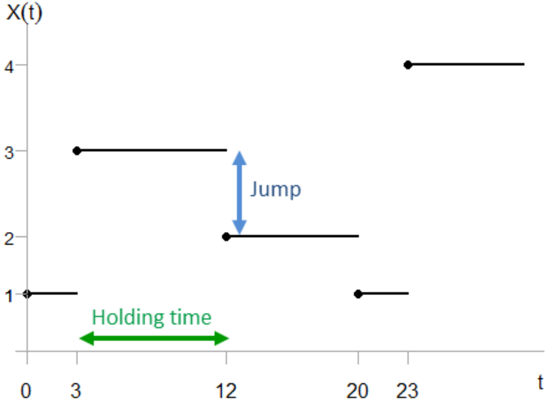
$$P(T > s + t \mid T > t) = P(T > s) \quad \forall\ s, t > 0$$

Proof.

$$P(T > s + t \mid T > t) = \frac{P(T>t+s \ \cap \ T>t)}{P(T>t)} = \frac{P(T>t+s)}{P(T>t)} =$$

$$= \frac{1-P(T\leq t+s)}{1-P(T\leq t)} = \frac{1-1+e^{-\lambda_i(t+s)}}{1-1+e^{-\lambda_i t}} = e^{-\lambda_i s} = P(T > s)$$

Background: describing a continuous-time Markov chain



Background: describing a continuous-time Markov chain

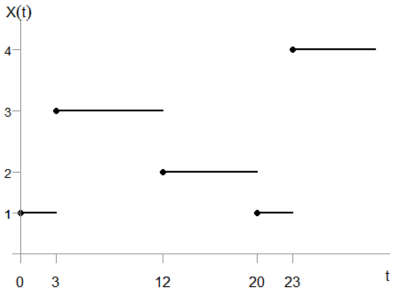
Jump chain

$P = (p_{ij} : i, j \in \mathcal{S}) = \textit{jump matrix}$

$$p_{ij} = P(X(t') = j | X(t) = i, \ \textit{the opportunity to leave } i)$$

$$p_{ij} \geq 0 \qquad \sum_{j \in \mathcal{S}} p_{ij} = 1 \qquad \forall i, j \in S$$

Background: describing a continuous-time Markov chain



Example

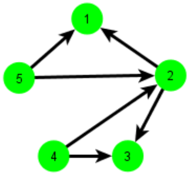
$$P = \begin{bmatrix} 0.1 & 0 & 0.6 & 0.3 \\ 0.8 & 0.1 & 0.1 & 0 \\ 0.05 & 0.5 & 0.05 & 0.4 \\ 0.6 & 0.1 & 0.15 & 0.15 \end{bmatrix}$$

Outline

Recall: adjacency matrix and directed relations

Social network: a set of actors $\mathcal{N}$ + a relation $\mathcal{R}$ existing among them

Graph $= G(\mathcal{N}, \mathcal{R})$



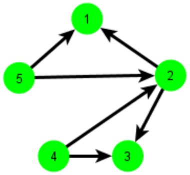
Adjacency matrix$=X$

-	0	0	0	0
1	-	1	0	0
0	0	-	0	0
0	1	1	-	0
1	1	0	0	-

Recall: adjacency matrix and directed relations

Social network: a set of actors $\mathcal{N}$ + a relation $\mathcal{R}$ existing among them

Graph $= G(\mathcal{N}, \mathcal{R})$



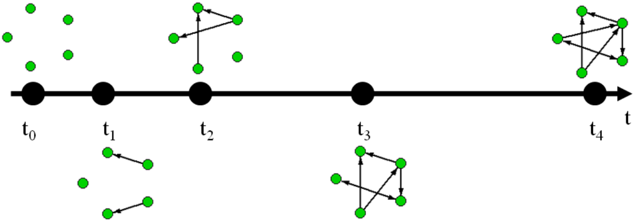
Adjacency matrix$=X$

-	0	0	0	0
1	-	1	0	0
0	0	-	0	0
0	1	1	-	0
1	1	0	0	-

Directed relation:

$$i \rightarrow j \quad \not\Rightarrow \quad j \rightarrow i$$

Data



Longitudinal (or panel) network data = M (≥ 2) repeated observations on a network

$$x(t_0),\ x(t_1),\ \ldots,\ x(t_m),\ \ldots,\ x(t_{M-1}),\ x(t_M)$$

- set of actors $\mathcal{N} = \{1, 2, \ldots, n\}$
- a non reflexive and directed relation $\mathcal{R}$
- actor covariates V (gender, age,social status, ...)

Model definition: assumptions

Network evolution is the outcome of a **Continuous-time Markov-Chain**

Model definition: assumptions

Network evolution is the outcome of a **Continuous-time Markov-Chain**

1. *Ties are state*:
   a tie is a state with a tendency to endure over time

Model definition: assumptions

Network evolution is the outcome of a **Continuous-time Markov-Chain**

1. *Ties are state*:
   a tie is a state with a tendency to endure over time

2. *Distribution of the process*:
   $\{X(t), t \in \mathcal{T}\}$ is a continuous time Markov Chain defined on:
   - the state space $\mathcal{X}$
   - the set of actors $\mathcal{N}$

Model definition: assumptions

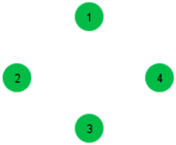
State space: $\mathcal{X}$ is the set of all possible adjacency matrices defined on $\mathcal{N}$

$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \textit{ is a countable set}$$

Model definition: assumptions

State space: $\mathcal{X}$ is the set of all possible adjacency matrices defined on $\mathcal{N}$

$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \textit{ is a countable set}$$

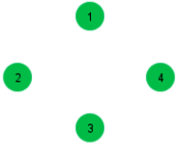


	1	2	3	4
1	-	0	0	0
2	0	-	0	0
3	0	0	-	0
4	0	0	0	-

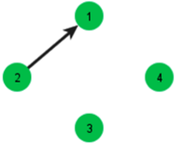
Model definition: assumptions

State space: $\mathcal{X}$ is the set of all possible adjacency matrices defined on $\mathcal{N}$

$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \textit{ is a countable set}$$



	1	2	3	4
1	-	0	0	0
2	0	-	0	0
3	0	0	-	0
4	0	0	0	-

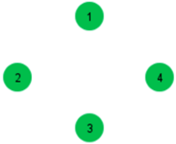


	1	2	3	4
1	-	0	0	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-

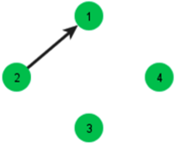
Model definition: assumptions

State space: $\mathcal{X}$ is the set of all possible adjacency matrices defined on $\mathcal{N}$

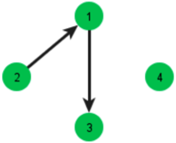
$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \textit{ is a countable set}$$



	1	2	3	4
1	-	0	0	0
2	0	-	0	0
3	0	0	-	0
4	0	0	0	-



	1	2	3	4
1	-	0	0	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-

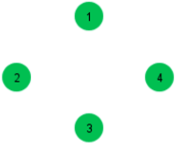


	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-

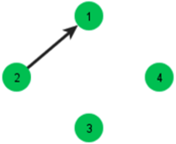
Model definition: assumptions

State space: $\mathcal{X}$ is the set of all possible adjacency matrices defined on $\mathcal{N}$

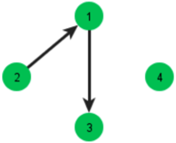
$$\mathcal{X} = 2^{n(n-1)} \Rightarrow \mathcal{X} \textit{ is a countable set}$$



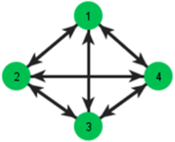
	1	2	3	4
1	-	0	0	0
2	0	-	0	0
3	0	0	-	0
4	0	0	0	-



	1	2	3	4
1	-	0	0	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-



	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	0	0	-	0
4	0	0	0	-



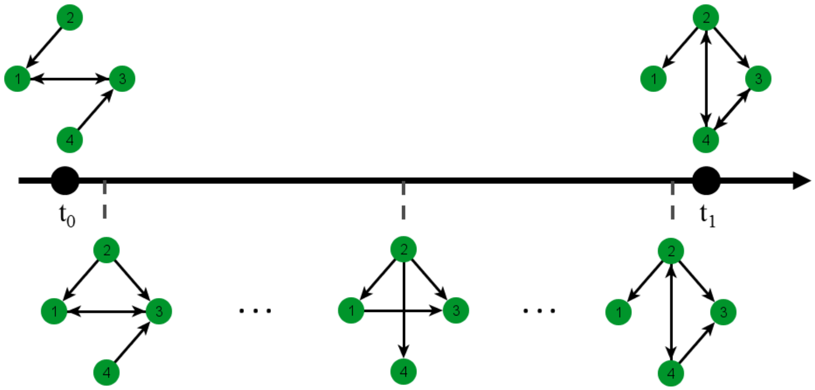
	1	2	3	4
1	-	1	1	1
2	1	-	1	1
3	1	1	-	1
4	1	1	1	-

Model definition: assumptions

Continuous-time process

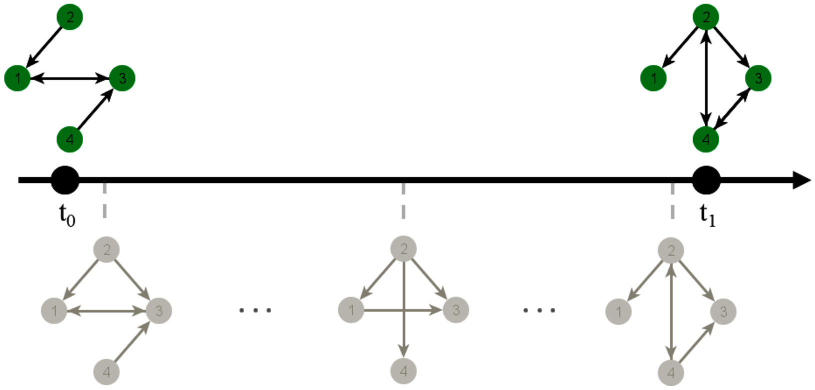
Model definition: assumptions

Continuous-time process



Model definition: assumptions

Continuous-time process



Latent process: the network evolves in continuous-time but we observed it only at discrete time points

Model definition: assumptions

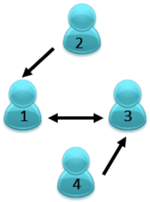
Markov property: the current state of the network determines probabilistically its further evolution

Model definition: assumptions

3. *Opportunity to change*: at any given moment t one actor has the opportunity to change

Model definition: assumptions

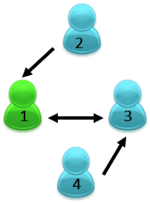
3. *Opportunity to change*: at any given moment t one actor has the opportunity to change



	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

Model definition: assumptions

3. *Opportunity to change*: at any given moment t one actor has the opportunity to change



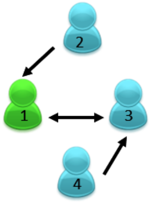
	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

Model definition: assumptions

4. *Absence of co-occurrence*: no more than one tie can change at any given moment t

Model definition: assumptions

4. *Absence of co-occurrence*: no more than one tie can change at any given moment t

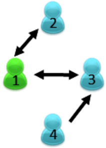


	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

x = current state of the network

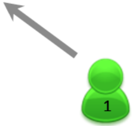
Model definition: assumptions

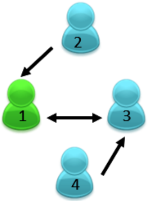
4. *Absence of co-occurrence*: no more than one tie can change at any given moment t



	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

$x(1 \rightsquigarrow 2)$



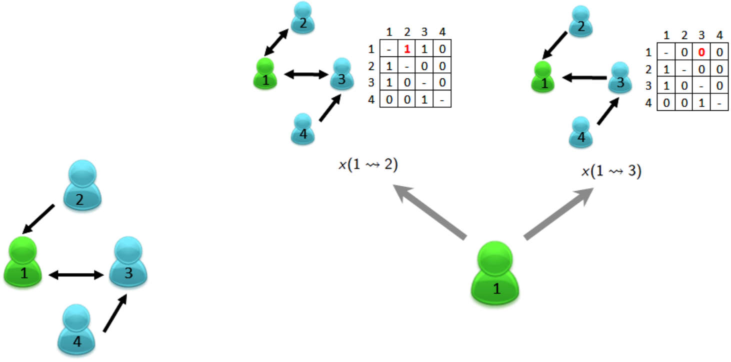


	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

x = current state
of the network

Model definition: assumptions

4. *Absence of co-occurrence*: no more than one tie can change at any given moment t



$x(1 \rightsquigarrow 2)$:

	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

$x(1 \rightsquigarrow 3)$:

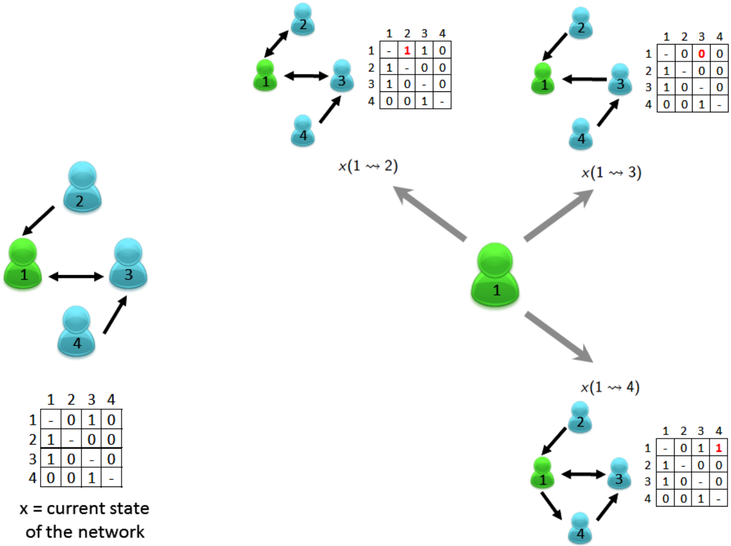
	1	2	3	4
1	-	0	0	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

x = current state of the network

Model definition: assumptions

4. *Absence of co-occurrence*: no more than one tie can change at any given moment t



Current state x:

	1	2	3	4
1	-	0	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

$x(1 \rightsquigarrow 2)$:

	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

$x(1 \rightsquigarrow 3)$:

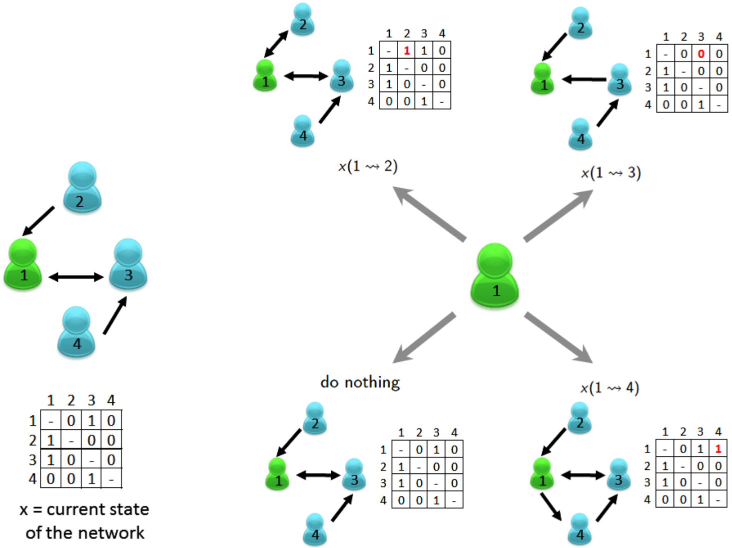
	1	2	3	4
1	-	0	0	0
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

$x(1 \rightsquigarrow 4)$:

	1	2	3	4
1	-	0	1	1
2	1	-	0	0
3	1	0	-	0
4	0	0	1	-

Model definition: assumptions

4. *Absence of co-occurrence*: no more than one tie can change at any given moment t



Model definition: assumptions

5. *Actor-oriented perspective*: actors control their outgoing ties

   - change in ties are made by the actor who sends the ties

   - decisions are made according to the position of the actor in the network, his attributes and the characteristics of the others

     Aim: maximize *a utility function*

   - actors have complete knowledge about the network and all the other actors

   - the maximization is based on immediate returns (myopic actors)

Model definition: assumptions (recap)

1. *Ties are states*
2. *The evolution process is a continuous-time Markov chain*
3. *At any given moment t one probabilistically selected actor has the opportunity to change*
4. *No more than one tie can change at any given moment t*
5. *Actor-oriented perspective*

Model definition

Consequences of the assumptions

The evolution process can be decomposed into micro-steps

Model definition

Consequences of the assumptions

The evolution process can be decomposed into micro-steps

Micro-step
- the time at which i had the opportunity to change
- the precise change i made

Model definition

Consequences of the assumptions

The evolution process can be decomposed into micro-steps

Micro-step	Continuous-time Markov chain
- the time at which *i* had the opportunity to change	- the waiting time until the next opportunity for a change made by an actor *i* (holding time)
- the precise change *i* made	- the probability of changing the link x_{ij} given that *i* is allowed to change (jump chain)

Model definition

Consequences of the assumptions

The evolution process can be decomposed into micro-steps

Micro-step	Continuous-time Markov chain
- the time at which *i* had the opportunity to change	- the waiting time until the next opportunity for a change made by an actor *i* (holding time)
- the precise change *i* made	- the probability of changing the link x_{ij} given that *i* is allowed to change (jump chain)

Distribution of the holding time: *rate function*

Transition matrix of the jump chain: *objective function*

Model definition: rate function

How fast is the opportunity for changing?

Waiting time between opportunities of change for actor $i \sim Exp(\lambda_i)$

λ_i is called the rate function

Simple specification: all actors have the same rate of change λ

$$P(i \ has \ the \ opportunity \ of \ change) = \frac{1}{n} \quad \forall i \in \mathcal{N}$$

Model definition: rate function

How fast is the opportunity for changing?

More complex specification

Actors may change their ties at different frequencies $\lambda_i(\alpha, x, v)$

Model definition: rate function

How fast is the opportunity for changing?

More complex specification

Actors may change their ties at different frequencies $\lambda_i(\alpha, x, v)$

Example

"Young girls might change their ties more frequently"

$$\lambda_i(\alpha, x, v) = \alpha_{age} * v_{age} + \alpha_{gender} * v_{gender}$$

Model definition: rate function

How fast is the opportunity for changing?

More complex specification

Actors may change their ties at different frequencies $\lambda_i(\alpha, x, v)$

Example

"Young girls might change their ties more frequently"

$$\lambda_i(\alpha, x, v) = \alpha_{age} * v_{age} + \alpha_{gender} * v_{gender}$$

It follows

$$P(i \ has \ the \ opportunity \ of \ change) = \frac{\lambda_i(\alpha, x, v)}{\sum_{j=1}^{n} \lambda_j(\alpha, x, v)}$$

Model definition: rate function

How fast is the opportunity for changing?

In the following we assume that:

- all actors have the same rate of change

  $\Longrightarrow \quad \lambda$ is constant over the actors

- the frequencies at which actors have the opportunity to make a change depends on time

  $\Longrightarrow \lambda$ is not constant over time

As a consequence, we must specify $M-1$ rate functions

$$\lambda_1, \cdots, \lambda_{M-1}$$

Model definition: objective function

Which tie is changed?

Changing a tie means turning it into its opposite:

$x_{ij} = 0$ is changed into $x_{ij} = 1$ tie creation

$x_{ij} = 1$ is changed into $x_{ij} = 0$ tie deletion

Model definition: objective function

Which tie is changed?

Changing a tie means turning it into its opposite:

$x_{ij} = 0$ is changed into $x_{ij} = 1$ tie creation

$x_{ij} = 1$ is changed into $x_{ij} = 0$ tie deletion

Given that i has the opportunity to change:

Possible choices of i	Possible reachable states
$n-1$ changes	$n-1$ networks $x(i \rightsquigarrow j)$

Model definition: objective function

Which tie is changed?

Changing a tie means turning it into its opposite:

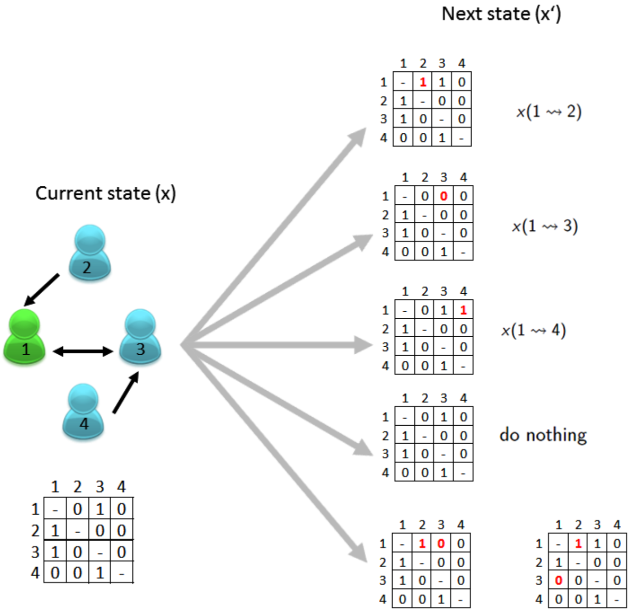
$x_{ij} = 0$ is changed into $x_{ij} = 1$ tie creation

$x_{ij} = 1$ is changed into $x_{ij} = 0$ tie deletion

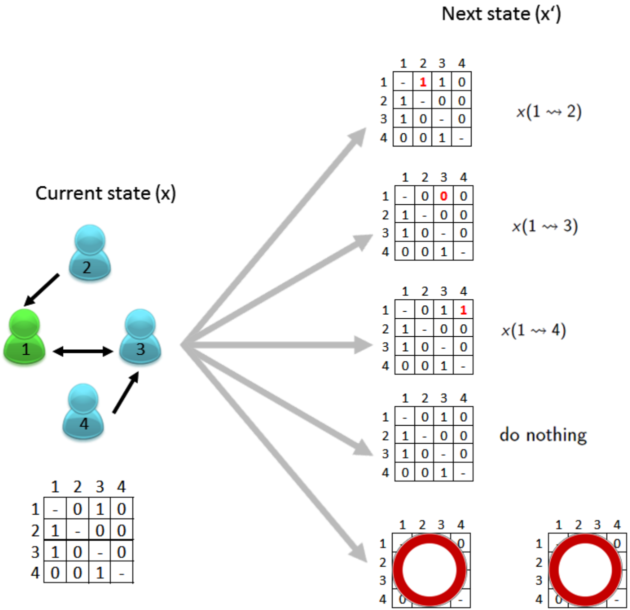
Given that i has the opportunity to change:

Possible choices of i	Possible reachable states
$n-1$ changes	$n-1$ networks $x(i \rightsquigarrow j)$
1 non-change	1 network equal to x

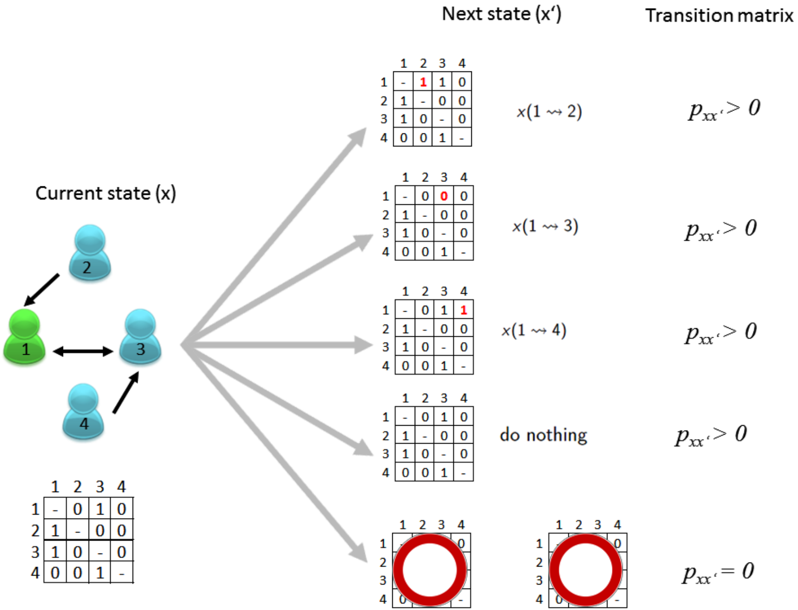
Model definition: objective function



Model definition: objective function



Model definition: objective function



Background: random utility model

Setting: decision makers who face a choice between n-alternatives

Decision rule: choose the alternative that assures the highest utility

$$U_{ij} = F_{ij} + \mathcal{E}_{ij}$$

F_{ij}: deterministic part of the utility

$\mathcal{E}_{ij}$: random term

Background: random utility model

Setting: decision makers who face a choice between n-alternatives

Decision rule: choose the alternative that assures the highest utility

$$U_{ij} = F_{ij} + \mathcal{E}_{ij}$$

F_{ij}: deterministic part of the utility

$\mathcal{E}_{ij}$: random term

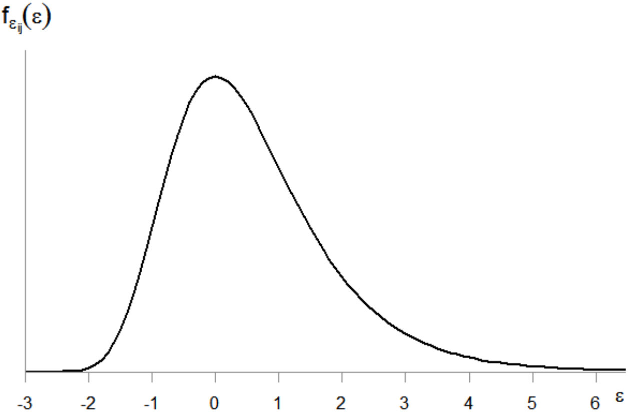
Decision probabilities: for a suitable choice of $\mathcal{E}_{ij}$

$$p_{ij} = \frac{e^{F_{ij}}}{\sum_{j=1}^{n} e^{F_{ij}}}$$

Background: random utility model

It is assumed that $\mathcal{E}_{ij}$ is Gumbel distributed

$$f_{\mathcal{E}_{ij}}(\epsilon) = e^{-\epsilon} e^{-e^{-\epsilon}} \qquad \epsilon \in \mathbb{R}$$



Model definition: objective function

Actors change their ties in order to maximize a utility function

$$u_i(\beta, x(i \rightsquigarrow j)) = f_i(\beta, x(i \rightsquigarrow j)) + \mathcal{E}_i(t, x, j)$$

- $f_i(\beta, x(i \rightsquigarrow j))$ is the *objective function*
- $\mathcal{E}_i(t, x, j))$ is a random utility term Gumbel distributed

Model definition: objective function

Actors change their ties in order to maximize a utility function

$$u_i(\beta, x(i \rightsquigarrow j)) = f_i(\beta, x(i \rightsquigarrow j)) + \mathcal{E}_i(t, x, j)$$

- $f_i(\beta, x(i \rightsquigarrow j))$ is the *objective function*
- $\mathcal{E}_i(t, x, j))$ is a random utility term Gumbel distributed

Probabilities

$$p_{ij} = \frac{exp(f_i(\beta, x(i \rightsquigarrow j))))}{\sum_{h=1}^{n} exp(f_i(\beta, x(i \rightsquigarrow h)))}$$

Probabilities interpretation:

p_{ij} is the probability that i changes the tie towards j

p_{ii} is the probability of not changing

Model definition: objective function

The objective function is defined as a linear combination

$$f_i(\beta, x(i \rightsquigarrow j)) = \sum_{k=1}^{K} \beta_k s_{ik}(x(i \rightsquigarrow j))$$

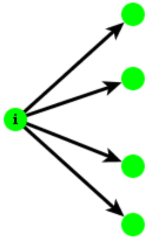
- $s_{ik}(x(i \rightsquigarrow j))$ are effects
- β_k are statistical parameters

Objective function specification

Endogenous effects = *dependent on the network structures*

- Outdegree effect

$$s_{i\_out}(x') = \sum_j x'_{ij}$$

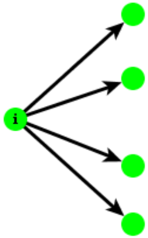


Objective function specification

Endogenous effects = *dependent on the network structures*

- Outdegree effect

$$s_{i\_out}(x') = \sum_j x'_{ij}$$



- Reciprocity effect

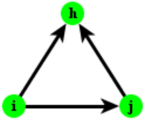
$$s_{i\_rec}(x') = \sum_j x'_{ij} x'_{ji}$$

Objective function specification

Endogenous effects = *dependent on the network structures*

- Transitive effect

$$s_{i\_trans}(x') = \sum_{j,h} x'_{ij} x'_{ih} x'_{jh}$$

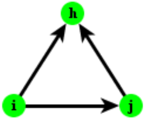


Objective function specification

Endogenous effects = *dependent on the network structures*

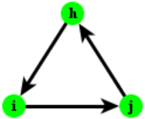
- Transitive effect

$$s_{i\_trans}(x') = \sum_{j,h} x'_{ij} x'_{ih} x'_{jh}$$



- three cycle-effect

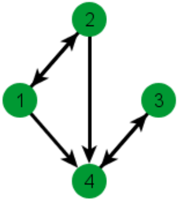
$$s_{i\_cyc}(x') = \sum_{j,h} x'_{ij} x'_{jh} x'_{hi}$$



Objective function specification

Example

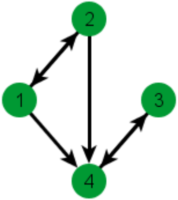
$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

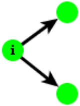
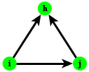


Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

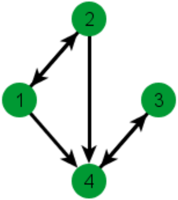


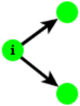
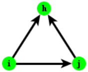
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$			
$1 \to 2$			
$1 \to 3$			
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

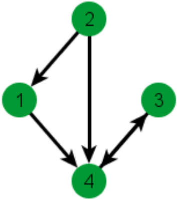


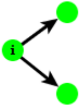
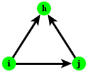
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$	2	1	1
$1 \to 2$			
$1 \to 3$			
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

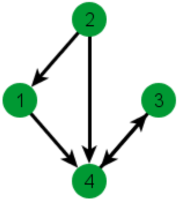


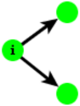
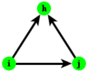
$i \rightarrow j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \rightarrow 1$	2	1	1
$1 \rightarrow 2$			
$1 \rightarrow 3$			
$1 \rightarrow 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

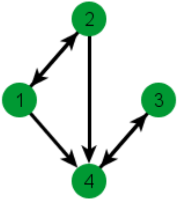


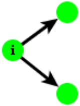
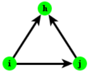
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$	2	1	1
$1 \to 2$	1	0	0
$1 \to 3$			
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

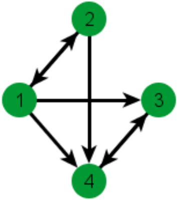


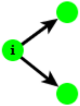
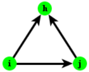
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$	2	1	1
$1 \to 2$	1	0	0
$1 \to 3$			
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

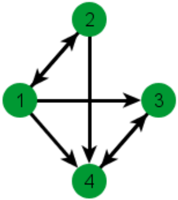


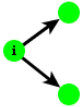
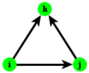
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$	2	1	1
$1 \to 2$	1	0	0
$1 \to 3$			
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

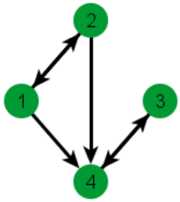


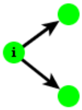
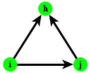
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$	2	1	1
$1 \to 2$	1	0	0
$1 \to 3$	3	1	3
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

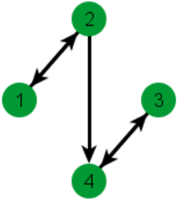


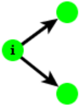
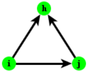
	s_{i_out}	s_{i_rec}	s_{i_trans}
$i \rightarrow j$			
$1 \rightarrow 1$	2	1	1
$1 \rightarrow 2$	1	0	0
$1 \rightarrow 3$	3	1	3
$1 \rightarrow 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

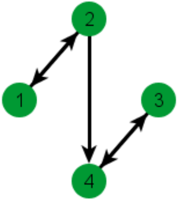


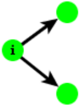
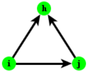
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$	2	1	1
$1 \to 2$	1	0	0
$1 \to 3$	3	1	3
$1 \to 4$			

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

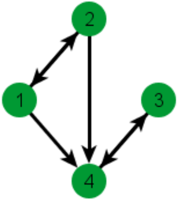


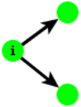
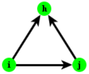
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 
$1 \to 1$	2	1	1
$1 \to 2$	1	0	0
$1 \to 3$	3	1	3
$1 \to 4$	1	1	0

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$

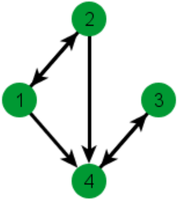


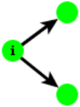
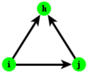
$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 	f_i
$1 \to 1$	2	1	1	-1.75
$1 \to 2$	1	0	0	-1
$1 \to 3$	3	1	3	-3.25
$1 \to 4$	1	1	0	-0.5

Objective function specification

Example

$$\beta_{out} = -1 \qquad \beta_{rec} = +0.5 \qquad \beta_{trans} = -0.25$$



$i \to j$	s_{i_out} 	s_{i_rec} 	s_{i_trans} 	f_i
$1 \to 1$	2	1	1	-1.75
$1 \to 2$	1	0	0	-1
$1 \to 3$	3	1	3	-3.25
$1 \to 4$	1	1	0	-0.5

$$p_{11} = 0.146 \qquad p_{12} = 0.310 \qquad p_{13} = 0.033 \qquad p_{14} = 0.511$$

Objective function specification

Exogenous effects = *related to actor's attributes*

Example

- Friendship among pupils:

  Smoking: non, occasional, regular

  Gender: boys, girls

- Trade/Trust (Alliances) among countries:

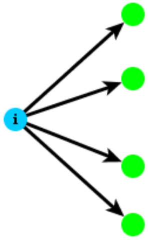
  Geographical area: Europe, Asia, North-America,...

  Worlds: first, Second, Third, Fourth

Objective function specification

- covariate-ego

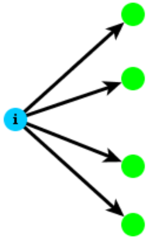
$$s_{i\_cego}(x) = \sum_j x_{ij} v_i$$



Objective function specification

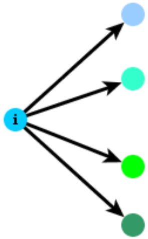
- covariate-ego

$$s_{i\_cego}(x) = \sum_j x_{ij} v_i$$



- covariate-alter

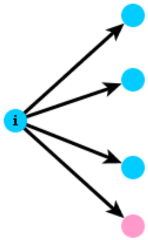
$$s_{i\_calt}(x) = \sum_j x_{ij} v_j$$



Objective function specification

- covariate-related similarity

$$s_{i\_csim}(x) = \sum_j x_{ij}\left(1 - \frac{\left|v_i - v_j\right|}{R_V}\right)$$

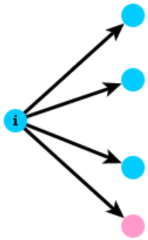


where R_V is the range of V and $\left(1 - \frac{|v_i - v_j|}{R_V}\right)$ is called *similarity score*

Objective function specification

- covariate-related similarity

$$s_{i\_csim}(x) = \sum_j x_{ij}\left(1 - \frac{\left|v_i - v_j\right|}{R_V}\right)$$



where R_V is the range of V and $\left(1 - \frac{|v_i - v_j|}{R_V}\right)$ is called *similarity score*

Remark:

when V is a binary covariate, the covariate-related similarity can be written in the following way:

$$s_{i\_csim}(x) = \sum_j x_{ij}\mathbb{I}\left\{v_i = v_j\right\}$$

Objective function specification

Which effects must be included in the objective function?

Outdegree and Reciprocity must always be included.
The choice of the other effects must be determined according to hypotheses derived from theory

Objective function specification

Which effects must be included in the objective function?

Outdegree and Reciprocity must always be included.
The choice of the other effects must be determined according to hypotheses derived from theory

Example

Friendship network

Theory		Effect
the friend of my friend is also my friend	$\Rightarrow$	transitive effect

Objective function specification

Which effects must be included in the objective function?

Outdegree and Reciprocity must always be included.
The choice of the other effects must be determined according to hypotheses derived from theory

Example

Friendship network

Theory		Effect
the friend of my friend is also my friend	$\Rightarrow$	transitive effect
girls trust girls boys trust boys	$\Rightarrow$	covariate-related similarity

Objective function specification

1. Parameter interpretation: β_k quantifies the role of $s_{ik}(x')$ in the network evolution.

    - $\beta_k = 0$: $s_{ik}(x')$ plays no role in the network dynamics
    - $\beta_k > 0$: higher probability of moving into networks where $s_{ik}(x')$ is higher
    - $\beta_k < 0$: higher probability of moving into networks where $s_{ik}(x')$ is lower

Objective function specification

1. Parameter interpretation: β_k quantifies the role of $s_{ik}(x')$ in the network evolution.

   - $\beta_k = 0$: $s_{ik}(x')$ plays no role in the network dynamics
   - $\beta_k > 0$: higher probability of moving into networks where $s_{ik}(x')$ is higher
   - $\beta_k < 0$: higher probability of moving into networks where $s_{ik}(x')$ is lower

2. The preferences driving the choice of the actors have the same intensities over time

$$\Longrightarrow \quad \beta_1, \cdots, \beta_K \text{ are constant over time}$$

Parameter interpretation

The procedures for estimating the parameters of the SAOM are implemented in a R library called ***RSiena***

(SIENA = **S**imulation **I**nvestigation for **E**mpirical **N**etwork **A**nalysis)

The R script “estimation.R” contains the R commands to implement the estimation procedure in R and the folder “tfls.zip” includes the data files.

Example data: an excerpt from the “Teenage Friends and Lifestyle Study” data set:

- Networks: relation = friendship
  actors = 129 pupils present at all three measurement points
- Covariates: gender (1 = Male, 2 = Female)
  smoking behavior (1 = no, 2= occasional, 3 = regular)

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Rate parameter: expected frequency, between two consecutive network observations, with which actors get the opportunity to change a network tie

- about 9 opportunities for change in the first period
- about 7 opportunities for change in the second period

The estimated rate parameters will be higher than the observed number of changes per actor (why?)

Parameter interpretation: a very simple model

Interpreting the objective function parameters:

The parameter β_k quantifies the role of the effect s_{ik} in the network evolution.

$\beta_k = 0$ s_{ik} plays no role in the network dynamics

$\beta_k > 0$ higher probability of moving into networks where s_{ik} is higher

$\beta_k < 0$ higher probability of moving into networks where s_{ik} is lower

Which β_k are "significantly" different from 0?

E.g. $\beta_{rec} = 0.13$ is "significantly" different from 0?

Parameter interpretation: a very simple model

Hypothesis test:

1. State the hypotheses.
   - The *null hypothesis* (H_0) states that the observed increase or decrease in the number of network configurations related to a certain effect results purely from chance.

$$H_0 : \beta_k = 0$$

Parameter interpretation: a very simple model

Hypothesis test:

1. State the hypotheses.
   - The *null hypothesis* (H_0) states that the observed increase or decrease in the number of network configurations related to a certain effect results purely from chance.

$$H_0 : \beta_k = 0$$

   - The *alternative hypothesis* (H_1) states that the observed increase or decrease in the number of network configurations related to a certain effect is influenced by some non-random cause.

$$H_1 : \beta_k \neq 0$$

Parameter interpretation: a very simple model

Hypothesis test:

2. Define a decision rule

$$\begin{cases} \left|\frac{\beta_k}{s.e.(\beta_k)}\right| \geq 2 & \textit{reject } H_0 \\ \\ \left|\frac{\beta_k}{s.e.(\beta_k)}\right| < 2 & \textit{fail to reject } H_0 \end{cases}$$

The logic behind this decision rule is based on the standard error concept.

Parameter interpretation: a very simple model

Hypothesis test:

2. Define a decision rule

$$\begin{cases} \left|\frac{\beta_k}{s.e.(\beta_k)}\right| \geq 2 & \textit{reject } H_0 \\ \left|\frac{\beta_k}{s.e.(\beta_k)}\right| < 2 & \textit{fail to reject } H_0 \end{cases}$$

The logic behind this decision rule is based on the standard error concept.

Example

Is the value $\beta_{rec} = 0.13$ far enough from 0?

If $s.e.(\beta_{rec}) = 0.9$, a more or less plausible set of values that the parameter can assume is approximately

$$[0.04, 0.22]$$

Parameter interpretation: a very simple model

Hypothesis test:

2. Define a decision rule

$$\begin{cases} \left|\frac{\beta_k}{s.e.(\beta_k)}\right| \geq 2 & \textit{reject } H_0 \\ \left|\frac{\beta_k}{s.e.(\beta_k)}\right| < 2 & \textit{fail to reject } H_0 \end{cases}$$

The logic behind this decision rule is based on the standard error concept.

Example

Is the value $\beta_{rec} = 0.13$ far enough from 0?

If $s.e.(\beta_{rec}) = 0.9$, a more or less plausible set of values that the parameter can assume is approximately

$$[0.04, 0.22]$$

$$\left|\frac{\beta_{rec}}{s.e.(\beta_{rec})}\right| = \left|\frac{0.13}{0.9}\right| = 0.14 < 2$$

β_{rec} is not significantly different from 0

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Objective function parameters:

- outdegree parameter: the observed networks have low density

Parameter interpretation: a very simple model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.5948	(0.7091)	
Rate parameter period 2	7.2115	(0.5751)	
Other parameters:			
outdegree (density)	-2.4147	(0.0387)	-62.3875
reciprocity	2.7106	(0.0811)	33.4061

Objective function parameters:

- outdegree parameter: the observed networks have low density
- reciprocity parameter: strong tendency towards reciprocated ties

Parameter interpretation: a very simple model

In more detail

$$\beta_{out} \sum_{j=1}^{n} x_{ij} + \beta_{rec} \sum_{j=1}^{n} x_{ij} x_{ji} = -2.4147 \sum_{j=1}^{n} x_{ij} + 2.7106 \sum_{j=1}^{n} x_{ij} x_{ji}$$

Adding a reciprocated tie (i.e., for which $x_{ji} = 1$) gives

$$-2.4147 + 2.7106 = 0.2959$$

while adding a non-reciprocated tie (i.e., for which $x_{ji} = 0$) gives

$$-2.4147$$

Conclusion: reciprocated ties are valued positively and non-reciprocated ties are valued negatively by actors

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- friendship relations tend to be reciprocated $\rightarrow$ reciprocity effect

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- friendship relations tend to be reciprocated $\rightarrow$ reciprocity effect

- the statement "the friend of my friend is also my friend" is almost always true

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- friendship relations tend to be reciprocated $\rightarrow$ reciprocity effect

- the statement "the friend of my friend is also my friend" is almost always true $\rightarrow$ transitive triplets effect

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- pupils prefer to establish friendship relations with others that are similar to themselves $\rightarrow$ covariate similarity

Parameter interpretation: a more complex model

Specifying the objective function

In friendship context, sociological theory suggests that:

- pupils prefer to establish friendship relations with others that are similar to themselves $\rightarrow$ covariate similarity

This effect must be controlled for the sender and receiver effects of the covariate.

- Covariate ego effect

- Covariate alter effect

Parameter interpretation: a more complex model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	10.6809	(1.0425)	
Rate parameter period 2	9.0116	(0.8386)	
Other parameters:			
outdegree (density)	-2.8597	(0.0608)	-47.0288
reciprocity	1.9855	(0.0876)	22.6765
transitive triplets	0.4480	(0.0257)	17.4558
sex alter	-0.1513	(0.0980)	-1.5445
sex ego	0.1571	(0.1072)	1.4659
sex similarity	0.9191	(0.1076)	8.5440
smoke alter	0.1055	(0.0577)	1.8272
smoke ego	0.0714	(0.0623)	1.1469
smoke similarity	0.3724	(0.1177)	3.1647

Parameter interpretation: a more complex model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	10.6809	(1.0425)	
Rate parameter period 2	9.0116	(0.8386)	
Other parameters:			
outdegree (density)	-2.8597	(0.0608)	-47.0288
reciprocity	1.9855	(0.0876)	22.6765
transitive triplets	0.4480	(0.0257)	17.4558
sex alter	-0.1513	(0.0980)	-1.5445
sex ego	0.1571	(0.1072)	1.4659
sex similarity	0.9191	(0.1076)	8.5440
smoke alter	0.1055	(0.0577)	1.8272
smoke ego	0.0714	(0.0623)	1.1469
smoke similarity	0.3724	(0.1177)	3.1647

- outdegree parameter: the observed networks have low density
- reciprocity parameter: strong tendency towards reciprocated ties
- transitivity parameter: preference for being friends with friends'friends

Parameter interpretation: a more complex model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	10.6809	(1.0425)	
Rate parameter period 2	9.0116	(0.8386)	
Other parameters:			
outdegree (density)	-2.8597	(0.0608)	-47.0288
reciprocity	1.9855	(0.0876)	22.6765
transitive triplets	0.4480	(0.0257)	17.4558
sex alter	-0.1513	(0.0980)	-1.5445
sex ego	0.1571	(0.1072)	1.4659
sex similarity	0.9191	(0.1076)	8.5440
smoke alter	0.1055	(0.0577)	1.8272
smoke ego	0.0714	(0.0623)	1.1469
smoke similarity	0.3724	(0.1177)	3.1647

- sex alter: gender does not affect actor popularity
- sex ego: gender does not affect actor activity
- sex similarity: tendency to choose friends with the same gender

Parameter interpretation: a more complex model

- Gender: coded with 1 for boys and with 2 for girls.

Parameter interpretation: a more complex model

- Gender: coded with 1 for boys and with 2 for girls.

- All actor covariates are centered: $\overline{v} = 1.434$ is the mean of the covariate

$$v_i - \overline{v} = \begin{cases} -0.434 & \textit{for boys} \\ \\ 0.566 & \textit{for girls} \end{cases}$$

Parameter interpretation: a more complex model

- Gender: coded with 1 for boys and with 2 for girls.

- All actor covariates are centered: $\overline{v} = 1.434$ is the mean of the covariate

$$v_i - \overline{v} = \begin{cases} -0.434 & \textit{for boys} \\ \\ 0.566 & \textit{for girls} \end{cases}$$

- The contribution of x_{ij} to the objective function is

$$\beta_{ego}(v_i - \overline{v}) + \beta_{alter}(v_j - \overline{v}) + \beta_{same}\left(\mathbb{I}\{v_i = v_j\} - sim_v\right) =$$

$$= 0.1571(v_i - \overline{v}) - 0.1513(v_j - \overline{v}) + 0.9191\left(\mathbb{I}\{v_i = v_j\} - 0.5048\right)$$

where sim_v is the average of the similarity score.

Parameter interpretation: a more complex model

	Male	Female
Male	0.4526	-0.618
Female	-0.309	0.4584

Table : Gender-related contributions to the objective function

Conclusions: Preference for intra-gender relationships.

Parameter interpretation: a more complex model

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	10.6809	(1.0425)	
Rate parameter period 2	9.0116	(0.8386)	
Other parameters:			
outdegree (density)	-2.8597	(0.0608)	-47.0288
reciprocity	1.9855	(0.0876)	22.6765
transitive triplets	0.4480	(0.0257)	17.4558
sex alter	-0.1513	(0.0980)	-1.5445
sex ego	0.1571	(0.1072)	1.4659
sex similarity	0.9191	(0.1076)	8.5440
smoke alter	0.1055	(0.0577)	1.8272
smoke ego	0.0714	(0.0623)	1.1469
smoke similarity	0.3724	(0.1177)	3.1647

- smoke alter: smoking behavior does not affect actor popularity
- smoke ego: smoking behavior not affect actor activity
- smoke similarity: tendency to choose friends with the same smoking behavior

Parameter interpretation: a more complex model

- Smoking behavior: coded with 1 for “no”, 2 for “occasional”, and 3 for “regular” smokers.

Parameter interpretation: a more complex model

- Smoking behavior: coded with 1 for "no", 2 for "occasional", and 3 for "regular" smokers.

- The smoking covariate is centered: $\overline{v} = 1.310$ is the mean of the covariate

$$v_i - \overline{v} = \begin{cases} -0.310 & \textit{for no smokers} \\ 0.690 & \textit{for occasional smokers} \\ 1.690 & \textit{for regular smokers} \end{cases}$$

Parameter interpretation: a more complex model

- Smoking behavior: coded with 1 for "no", 2 for "occasional", and 3 for "regular" smokers.

- The smoking covariate is centered: $\overline{v} = 1.310$ is the mean of the covariate

$$v_i - \overline{v} = \begin{cases} -0.310 & \text{for no smokers} \\ 0.690 & \text{for occasional smokers} \\ 1.690 & \text{for regular smokers} \end{cases}$$

- The contribution of x_{ij} to the objective function is

$$\beta_{ego}(v_i - \overline{v}) + \beta_{alter}(v_j - \overline{v}) + \beta_{same}\left(1 - \frac{|v_i - v_j|}{R_v} - sim_v\right) =$$

$$= 0.0714(v_i - \overline{v}) + 0.1055(v_j - \overline{v}) + 0.3724\left(1 - \frac{|v_i - v_j|}{2} - 0.7415\right)$$

Parameter interpretation: a more complex model

	no	occasional	regular
no	0.0414	-0.0734	-0.1882
occasional	-0.0393	0.2183	0.1035
regular	-0.1200	0.1376	0.3952

Table : Smoking-related contributions to the objective function

Parameter interpretation: a more complex model

	no	occasional	regular
no	0.0414	-0.0734	-0.1882
occasional	-0.0393	0.2183	0.1035
regular	-0.1200	0.1376	0.3952

Table : Smoking-related contributions to the objective function

Conclusions:

- preference for similar alters
- this tendency is strongest for high values on smoking behavior

Simulating network evolution

Reproducing a possible series of micro-steps between t_0 and t_1 according to fixed parameter value and the network $x(t_0)$.

t = time
dt = holding time between consecutive opportunities to change

---

Algorithm 1: Network evolution

Input: $x(t_0)$, λ, β, n
Output: $x^{sim}(t_1)$
$t \leftarrow 0$
$x \leftarrow x(t_0)$
while $condition = TRUE$ **do**
    $dt \sim Exp(n\lambda)$
    $i \sim Uniform(1, \ldots, n)$
    $j \sim Multinomial(p_{i1}, \ldots, p_{in})$
    if $i \neq j$ **then**
        $x \leftarrow x(i \rightsquigarrow j)$
    else
        $x \leftarrow x$
    $t \leftarrow t + dt$
$x^{sim}(t_1) \leftarrow x$
return $x^{sim}(t_1)$

Simulating network evolution

Two different stopping rules:

1. *Unconditional* simulation:

   the simulation of the network evolution carries on until a predetermined time length has elapsed (usually until $t = 1$).

Simulating network evolution

Two different stopping rules:

1. *Unconditional* simulation:

   the simulation of the network evolution carries on until a predetermined time length has elapsed (usually until $t = 1$).

2. *Conditional* simulation on the observed number of changes:

   simulation runs on until the number of different entries between $x(t_0)$ and the simulated network $x^{sim}(t_1)$ is equal to the number of entries that differ between $x(t_0)$ and $x(t_1)$

$$\sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| x_{ij}^{obs}(t_1) - x_{ij}(t_0) \right| = \sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| x_{ij}^{sim}(t_1) - x_{ij}(t_0) \right|$$

   This criterion can be generalized conditioning on any other explanatory variable.

Parameter estimation

The formulation of the SAOM depends on $M-1+K$ statistical parameters

$$\theta = (\lambda_1, \cdots, \lambda_{M-1}, \beta_1, \cdots, \beta_K)$$

Aim: estimate θ

Different estimation methods:

- the Method of Moments (MoM)
- the Maximum Likelihood Estimation (MLE)

Background: Method of Moments (MoM)

Let

- X be a r.v. with distribution $f_X(x;\theta)$

- $E_\theta[X]$ be the expected value of X

- $(x_1, \ldots, x_n)$ be n observations from the r.v. X.

Background: Method of Moments (MoM)

Let

- X be a r.v. with distribution $f_X(x;\theta)$
- $E_\theta[X]$ be the expected value of X
- $(x_1, \ldots, x_n)$ be n observations from the r.v. X.

Definition

The sample counterpart of $E_\theta[X]$ is defined as:

$$\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$$

Background: Method of Moments (MoM)

One can observe that the expected value of a certain distribution usually depends on the parameter θ

Definition

The method of moment estimator for θ is found by equating the expected value $E_\theta[X]$ to its sample counterpart μ

$$E_\theta[X] = \mu$$

and solving the resulting equation for the unknown parameter.

Background: Method of Moments (MoM)

One can observe that the expected value of a certain distribution usually depends on the parameter θ

Definition

The method of moment estimator for θ is found by equating the expected value $E_\theta[X]$ to its sample counterpart μ

$$E_\theta[X] = \mu$$

and solving the resulting equation for the unknown parameter.

In practice:

1. Compute the expected value $E_\theta[X]$
2. Compute the sample counterpart $\mu = \frac{1}{n}\sum_{i=1}^{n} x_i$
3. Solve the moment equation $E_\theta[X] = \mu$ for θ

Background: Method of Moments (MoM)

Example

10 undirected, simple, loopless graphs are generated according to $\mathcal{G}(30, p)$. The number of edges y_i in each graph is reported in the following table:

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}
y_i	37	40	35	32	39	34	25	28	41	32

Find a plausible value for p that might have generated the observed graphs.

Background: Method of Moments (MoM)

Example

10 undirected, simple, loopless graphs are generated according to $\mathcal{G}(30, p)$. The number of edges y_i in each graph is reported in the following table:

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}
y_i	37	40	35	32	39	34	25	28	41	32

Find a plausible value for p that might have generated the observed graphs.

Y = r.v. describing the number of edges

$$P(Y = y) = \binom{N}{y} p^y (1-p)^{N-y}$$

where $N = \frac{n(n-1)}{2}$

Background: Method of Moments (MoM)

Example

1. The theoretical expected value of the number of edges is:

$$
\begin{aligned}
E_\theta[Y] &= \sum_{y=0}^{N} y P(Y=y) = \sum_{y=0}^{N} y \binom{N}{y} p^y (1-p)^{N-y} \\
&= \sum_{y=1}^{N} y \binom{N}{y} p^y (1-p)^{N-y} \\
&= \sum_{y=1}^{N} y \frac{N!}{y!(N-y)!} p^y (1-p)^{N-y} \\
&= Np \sum_{y=1}^{N} \frac{(N-1)!}{(y-1)!(N-1-(y-1))!} p^{y-1} (1-p)^{N-1-(y-1)} \\
&= Np \sum_{i=0}^{N-1} \binom{N-1}{i} p^i (1-p)^{N-1-i} = Np
\end{aligned}
$$

Background: Method of Moments (MoM)

Example

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}
y_i	37	40	35	32	39	34	25	28	41	32

2. The sample counterpart is:

$$\mu = \frac{1}{10}\sum_{i=1}^{10} y_i = \frac{343}{10} = 34.3$$

3. The estimate for p is given by:

$$E_\theta[Y] = \mu$$
$$Np = \mu$$
$$\widehat{p} = \frac{\mu}{N} = \frac{34.3}{435} = 0.079$$

Background: Method of Moments (MoM)

Remark

It is easy to imagine that the estimate of the parameter can vary according to the selected sample.

Example

Other 10 generated graphs result in the following number of edges:

	g_1	g_2	g_3	g_4	g_5	g_6	g_7	g_8	g_9	g_{10}
y_i	34	28	23	31	32	36	33	41	26	39

The new estimate for p is now

$$\widehat{p} = 0.074$$

This value is close to the one obtained before, but it is not the same!
For this reason, we usually associate to an estimator its standard error.

Background: Generalizations of MoM

The principle of the MoM can be easily generalized to any function $s(X)$.

1. Expected value of $s(X)$:

$$E_\theta[s(X)] = \sum_x s(x) f_X(x; \theta)$$

2. Corresponding sample moment:

$$\gamma = \frac{1}{n} \sum_{i=1}^{n} s(x_i)$$

3. Moment equation:

$$E_\theta[s(X)] = \gamma$$

Background: Generalizations of MoM

The principle of the MoM can be easily generalized to any function $s(X)$.

1. Expected value of $s(X)$:

$$E_\theta[s(X)] = \sum_x s(x) f_X(x;\theta)$$

2. Corresponding sample moment:

$$\gamma = \frac{1}{n}\sum_{i=1}^{n} s(x_i)$$

3. Moment equation:

$$E_\theta[s(X)] = \gamma$$

The functions $s(X)$ are called *statistics*

Background: Generalizations of MoM

The MoM can be applied also in situations where $\theta = (\theta_1, \ldots, \theta_p)$.

1. Definition of p statistics $(s_1(X), \ldots, s_p(X))$
2. Definition of p moment conditions:

$$E_\theta[s_1(X)] = \gamma_1$$

$$E_\theta[s_2(X)] = \gamma_2$$

$$\ldots$$

$$E_\theta[s_p(X)] = \gamma_p$$

3. Solving the resulting equations for the unknown parameters

Estimating the parameter of the SAOM using MoM

Aim: estimate θ using the MoM

$$\theta = (\lambda_1,\ \dots,\ \lambda_{M-1},\ \beta_1,\ \dots,\ \beta_K)$$

Estimating the parameter of the SAOM using MoM

Aim: estimate θ using the MoM

$$\theta = (\lambda_1,\ \ldots,\ \lambda_{M-1},\ \beta_1,\ \ldots,\ \beta_K)$$

In practice:

1. find $M - 1 + K$ statistics
2. set the theoretical expected value of each statistic equal to its sample counterpart
3. solve the resulting system of equations with respect to θ.

For simplicity, let us assume to have observed a network at two time points t_0 and t_1 and to condition the estimation on the first observation $x(t_0)$

1. Defining the statistics

The rate parameter λ describes the frequency at which changes can potentially happen.

$$s_\lambda(X(t_1), X(t_0) | X(t_0) = x(t_0)) = \sum_{i,j=1}^{n} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

1. Defining the statistics

The rate parameter λ describes the frequency at which changes can potentially happen.

$$s_\lambda(X(t_1), X(t_0) | X(t_0) = x(t_0)) = \sum_{i,j=1}^{n} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

Reason

	$\lambda = 2$	$\lambda = 3$	$\lambda = 4$
s_λ	94	135	171

$\Rightarrow$ higher values of λ leads to higher values of s_λ

1. Defining the statistics

The parameter β_k quantifies the role played by each effect in the network evolution.

$$s_k(X(t_1)|X(t_0) = x(t_0)) = \sum_{i=1}^{n} s_{ik}(X(t_1))$$

1. Defining the statistics

The parameter β_k quantifies the role played by each effect in the network evolution.

$$s_k(X(t_1)|X(t_0) = x(t_0)) = \sum_{i=1}^{n} s_{ik}(X(t_1))$$

Example

Let us consider the outdegree:

$$s_{out}(X(t_1)|X(t_0) = x(t_0)) = \sum_{i=1}^{n} s_{i\_out}(X(t_1)) = \sum_{i=1}^{n}\sum_{j=1}^{n} x_{ij}(t_1)$$

	$\beta_{out} = -2.5$	$\beta_{out} = -2$	$\beta_{out} = -1.5$
s_{out}	195	214	234

$\Rightarrow$ higher values of β_{out} leads to higher values of s_{out}

1. Defining the statistics

Generalizing to $M-1$ periods:

- Statistics for the rate function parameters

$$s_{\lambda_1}(X(t_1), X(t_0) | X(t_0) = x(t_0)) = \sum_{i,j=1}^{n} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

...

$$s_{\lambda_{M-1}}(X(t_M), X(t_{M-1}) | X(t_{M-1}) = x(t_{M-1})) = \sum_{i,j=1}^{n} \left| X_{ij}(t_M) - X_{ij}(t_{M-1}) \right|$$

1. Defining the statistics

Generalizing to $M-1$ periods:

- Statistics for the rate function parameters

$$s_{\lambda_1}(X(t_1), X(t_0)|X(t_0) = x(t_0)) = \sum_{i,j=1}^{n} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

...

$$s_{\lambda_{M-1}}(X(t_M), X(t_{M-1})|X(t_{M-1}) = x(t_{M-1})) = \sum_{i,j=1}^{n} \left| X_{ij}(t_M) - X_{ij}(t_{M-1}) \right|$$

- Statistics for the objective function parameters:

$$\sum_{m=1}^{M-1} s_{mk}(X(t_m)|X(t_{m-1}) = x(t_{m-1})) = \sum_{m=1}^{M-1} s_{mk}(X(t_m))$$

2. Setting the moment equations

The MoM estimator for θ is defined as the solution of the system of $M+K-1$ equations

$$\begin{cases} E_\theta\left[s_{\lambda_m}(X(t_m),X(t_{m-1})|X(t_{m-1})=x(t_{m-1}))\right]=s_{\lambda_m}(x(t_m),x(t_{m-1})) \\ E_\theta\left[\sum_{m=1}^{M-1} s_{mk}(X(t_m)|X(t_{m-1})=x(t_{m-1}))\right]=\sum_{m=1}^{M-1} s_{mk}(x(t_m)) \end{cases}$$

with $m=1,\ldots,M-1$ and $k=1,\cdots,K$

2. Setting the moment equations

Example

Let us assume to have observed a network at $M = 2$ time points

t_0:

	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	0	0	-	1
4	0	0	1	-

t_1:

	1	2	3	4
1	-	1	0	1
2	0	-	0	0
3	0	0	-	1
4	1	1	1	-

We want to model the network evolution according to the outdegree, the reciprocity and the transitivity effects

2. Setting the moment equations

Example

Let us assume to have observed a network at $M = 2$ time points

t_0:

	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	0	0	-	1
4	0	0	1	-

t_1:

	1	2	3	4
1	-	1	0	1
2	0	-	0	0
3	0	0	-	1
4	1	1	1	-

We want to model the network evolution according to the outdegree, the reciprocity and the transitivity effects

$$\theta = (\lambda, \beta_{out}, \beta_{rec}, \beta_{trans})$$

2. Setting the moment equations

Example

Statistics:

$$s_{\lambda}(X(t_1), X(t_0)|X(t_0) = x(t_0)) = \sum_{i,j=1}^{4} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

$$s_{out}\,(X(t_1)|X(t_0) = x(t_0)) = \sum_{i,j=1}^{4} X_{ij}(t_1)$$

$$s_{rec}\,(X(t_1)|X(t_0) = x(t_0)) = \sum_{i,j=1}^{4} X_{ij}(t_1) X_{ji}(t_1)$$

$$s_{trans}\,(X(t_1)|X(t_0) = x(t_0)) = \sum_{i,j,h=1}^{4} X_{ij}(t_1) X_{ih}(t_1) X_{jh}(t_1)$$

2. Setting the moment equations

Example

t_0:

	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	0	0	-	1
4	0	0	1	-

t_1:

	1	2	3	4
1	-	1	0	1
2	0	-	0	0
3	0	0	-	1
4	1	1	1	-

Observed values of the statistics:

$$s_\lambda = 5$$

$$s_{out} = 6 \quad s_{rec} = 4 \quad s_{trans} = 2$$

2. Setting the moment equations

Example

We look for the value of θ that satisfies the system:

$$\begin{cases} E_\theta\left[s_\lambda(X(t_1), X(t_0)|X(t_0)=x(t_0))\right]=5 \\ E_\theta\left[s_{out}\left(X(t_1)|X(t_0)=x(t_0)\right)\right]=6 \\ E_\theta\left[s_{rec}\left(X(t_1)|X(t_0)=x(t_0)\right)\right]=4 \\ E_\theta\left[s_{trans}\left(X(t_1)|X(t_0)=x(t_0)\right)\right]=2 \end{cases}$$

3. Solving the moment equations

Simplified notation:

- S: $(M-1+K)$-dimensional vector of statistics
- s: $(M-1+K)$-dimensional vector of the observed values of the statistics

Consequently, the system of moment equations can be written as

$$E_\theta[S] = s$$

or equivalently as

$$E_\theta[S - s] = 0$$

3. Solving the moment equations

Simplified notation:

- S: $(M-1+K)$-dimensional vector of statistics
- s: $(M-1+K)$-dimensional vector of the observed values of the statistics

Consequently, the system of moment equations can be written as

$$E_\theta[S] = s$$

or equivalently as

$$E_\theta[S-s] = 0$$

Problem:
analytical procedures cannot be applied to solve this system

3. Solving the moment equations

Definition

Stochastic approximation methods are a family of iterative stochastic algorithms that attempt to find zeros of functions which cannot be analytically computed.

The Robbins-Monro (RM) algorithm: iterative algorithm to find the solution to

$$E_\theta[X] = \alpha$$

The value of θ is iteratively updated according to:

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i \left(E_{\widehat{\theta}_i}[X] - \alpha \right)$$

where

$$\lim_{i \to \infty} a_i = 0 \qquad \sum_{i=1}^{\infty} a_i = \infty \qquad \sum_{i=1}^{\infty} a_i^2 < \infty$$

and $E_{\widehat{\theta}_i}[X]$ is an approximation of $E_\theta[X]$ based on θ_i.

3. Solving the moment equations

Adapting the RM step for the SAOM:

The MoM equation is:

$$E_\theta[S] = s$$

Let

$$\overline{S}_i \approx E_\theta[S]$$

according to $\widehat{\theta}_i$

3. Solving the moment equations

Adapting the RM step for the SAOM:

The MoM equation is:

$$E_\theta[S] = s$$

Let

$$\overline{S}_i \approx E_\theta[S]$$

according to $\widehat{\theta}_i$

The value of θ is iteratively updated according to:

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i \left(\overline{S}_i - s\right)$$

Updating the value of θ

Example

Let us consider the *"Teenage Friends and Lifestyle Study"* data set.

We model the network evolution according to the following parameter

$$\theta = (\lambda_1, \lambda_2, \beta_{out}, \beta_{rec}, \beta_{trans})$$

The MoM equations are:

$$\begin{cases} E_\theta\left[s_{\lambda_1}(X(t_1), X(t_0)|X(t_0) = x(t_0))\right] = 477 \\ E_\theta\left[s_{\lambda_2}(X(t_2), X(t_1)|X(t_1) = x(t_1))\right] = 437 \\ E_\theta\left[s_{out}(X(t_1)|X(t_0) = x(t_0))\right] = 909 \\ E_\theta\left[s_{rec}(X(t_1)|X(t_0) = x(t_0))\right] = 548 \\ E_\theta\left[s_{trans}(X(t_1)|X(t_0) = x(t_0))\right] = 1146 \end{cases}$$

Updating the value of θ

Example

- Guess $\theta_0 = (7.45, 6.83, -1.61, 0, 0)$
- Simulate the network evolution 1000 times according to $\widehat{\theta_0}$
- Approximation of the expected values

$\overline{S}_{\lambda_1} = 605.745$ $\overline{S}_{\lambda_2} = 573.715$

$\overline{S}_{\beta_{out}} = 1151.886$ $\overline{S}_{\beta_{rec}} = 141.406$ $\overline{S}_{\beta_{trans}} = 270.118$

Updating the value of θ

Example

- Guess $\theta_0 = (7.45, 6.83, -1.61, 0, 0)$

- Simulate the network evolution 1000 times according to $\widehat{\theta_0}$

- Approximation of the expected values

  $\overline{S}_{\lambda_1} = 605.745$ $\quad\overline{S}_{\lambda_2} = 573.715$

  $\overline{S}_{\beta_{out}} = 1151.886$ $\quad\overline{S}_{\beta_{rec}} = 141.406$ $\quad\overline{S}_{\beta_{trans}} = 270.118$

- Approximation of the moment equation

  $\overline{S}_{\lambda_1} - 477 = 128.745$ $\quad\overline{S}_{\lambda_2} - 437 = 136.715$

  $\overline{S}_{\beta_{out}} - 909 = 242.886$ $\quad\overline{S}_{\beta_{rec}} - 548 = -406.594$ $\quad\overline{S}_{\beta_{trans}} - 1146 = -875.882$

Updating the value of θ

Example

- Guess $\theta_1 = (7.1, 6.75, -1.70, 1.20, 0.25)$

- Simulate the network evolution 1000 times according to $\widehat{\theta_1}$

- Approximation of the expected values

  $\overline{S}_{\lambda_1} = 549.787$ $\quad$ $\overline{S}_{\lambda_2} = 532.551$

  $\overline{S}_{\beta_{out}} = 1478.988$ $\quad$ $\overline{S}_{\beta_{rec}} = 517.450$ $\quad$ $\overline{S}_{\beta_{trans}} = 1062.537$

Updating the value of θ

Example

- Guess $\theta_1 = (7.1, 6.75, -1.70, 1.20, 0.25)$

- Simulate the network evolution 1000 times according to $\widehat{\theta_1}$

- Approximation of the expected values

  $\overline{S}_{\lambda_1} = 549.787$ $\quad$ $\overline{S}_{\lambda_2} = 532.551$

  $\overline{S}_{\beta_{out}} = 1478.988$ $\quad$ $\overline{S}_{\beta_{rec}} = 517.450$ $\quad$ $\overline{S}_{\beta_{trans}} = 1062.537$

- Approximation of the moment equation

  $\overline{S}_{\lambda_1} - 477 = 72.787$ $\quad$ $\overline{S}_{\lambda_2} - 437 = 95.551$

  $\overline{S}_{\beta_{out}} - 909 = 569.988$ $\quad$ $\overline{S}_{\beta_{rec}} - 548 = -30.550$ $\quad$ $\overline{S}_{\beta_{trans}} - 1146 = -83.463$

Updating the value of θ

Example

- Guess $\theta_2 = (7.10, 6.75, -2.20, 1.40, 0.35)$

- Simulate the network evolution 1000 times according to $\widehat{\theta}_2$

- Approximation of the expected values

  $\overline{S}_{\lambda_1} = 446.853$ $\quad$ $\overline{S}_{\lambda_2} = 437.166$

  $\overline{S}_{\beta_{out}} = 1025.729$ $\quad$ $\overline{S}_{\beta_{rec}} = 414.484$ $\quad$ $\overline{S}_{\beta_{trans}} = 698.734$

Updating the value of θ

Example

- Guess $\theta_2 = (7.10, 6.75, -2.20, 1.40, 0.35)$

- Simulate the network evolution 1000 times according to $\widehat{\theta_2}$

- Approximation of the expected values

  $\overline{S}_{\lambda_1} = 446.853$ $\qquad \overline{S}_{\lambda_2} = 437.166$

  $\overline{S}_{\beta_{out}} = 1025.729$ $\qquad \overline{S}_{\beta_{rec}} = 414.484$ $\qquad \overline{S}_{\beta_{trans}} = 698.734$

- Approximation of the moment equation

  $\overline{S}_{\lambda_1} - 477 = -30.147$ $\qquad \overline{S}_{\lambda_2} - 437 = 0.166$

  $\overline{S}_{\beta_{out}} - 909 = 116.729$ $\qquad \overline{S}_{\beta_{rec}} - 548 = -133.516$ $\qquad \overline{S}_{\beta_{trans}} - 1146 = -447.266$

and so on...

Updating the value of θ

Example

- Guess $\theta_i = (10.71, 8.79, -2.63, 2.16, 0.46)$

- Simulate the network evolution 1000 times according to $\widehat{\theta}_i$

- Approximation of the expected values

  $\overline{S}_{\lambda_1} = 476.022$ $\overline{S}_{\lambda_2} = 436.983$

  $\overline{S}_{\beta_{out}} = 906.809$ $\overline{S}_{\beta_{rec}} = 545.578$ $\overline{S}_{\beta_{trans}} = 1147.795$

Updating the value of θ

Example

- Guess $\theta_i = (10.71, 8.79, -2.63, 2.16, 0.46)$
- Simulate the network evolution 1000 times according to $\widehat{\theta}_i$
- Approximation of the expected values

  $\overline{S}_{\lambda_1} = 476.022$ $\quad$ $\overline{S}_{\lambda_2} = 436.983$

  $\overline{S}_{\beta_{out}} = 906.809$ $\quad$ $\overline{S}_{\beta_{rec}} = 545.578$ $\quad$ $\overline{S}_{\beta_{trans}} = 1147.795$

- Approximation of the moment equation

  $\overline{S}_{\lambda_1} - 477 = -0.978$ $\quad$ $\overline{S}_{\lambda_2} - 437 = -0.017$

  $\overline{S}_{\beta_{out}} - 909 = -2.191$ $\quad$ $\overline{S}_{\beta_{rec}} - 548 = -2.422$ $\quad$ $\overline{S}_{\beta_{trans}} - 1146 = 1.795$

3.Solving the moment equations

The Robbins-Monro (RM) algorithm - remarks

1. Convergence:

$$\lim_{i\to\infty} \widehat{\theta}_i = \theta$$

2. Modified RM step: to improve the convergence of the algorithm

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i D^{-1}\left(E_{\widehat{\theta}_i}[X] - \alpha\right)$$

where D is a diagonal matrix with elements

$$D = \frac{\partial}{\partial \widehat{\theta}_i} E_{\widehat{\theta}_i}[X]$$

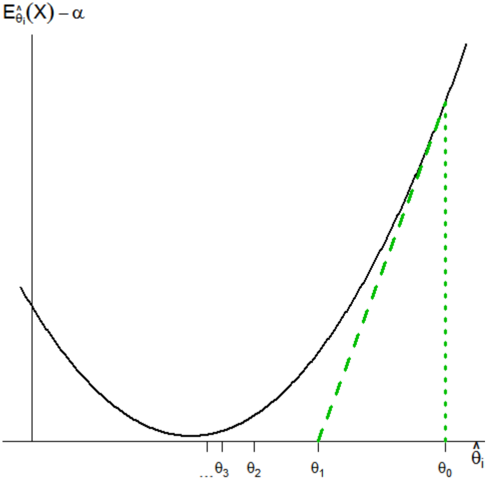
and estimate θ with:

$$\widehat{\theta} = \frac{1}{I}\sum_{i_1}^{I} \widehat{\theta}_i, \qquad I \; number \;\; of \;\; steps$$

Updating the value of θ

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i D^{-1} \left(E_{\widehat{\theta}_i}[X] - \alpha \right)$$

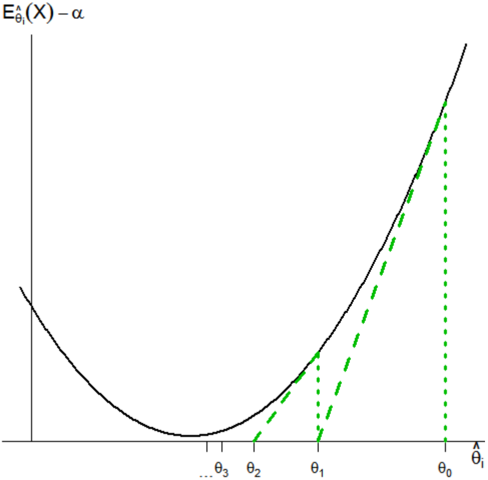
Intuitively:



Updating the value of θ

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i D^{-1} \left(E_{\widehat{\theta}_i}[X] - \alpha \right)$$

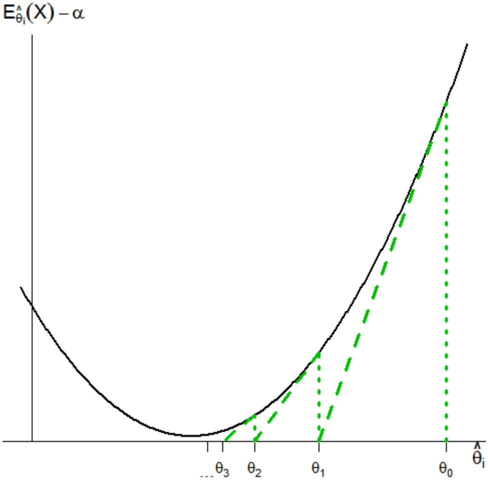
Intuitively:



Updating the value of θ

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i D^{-1} \left(E_{\widehat{\theta}_i}[X] - \alpha \right)$$

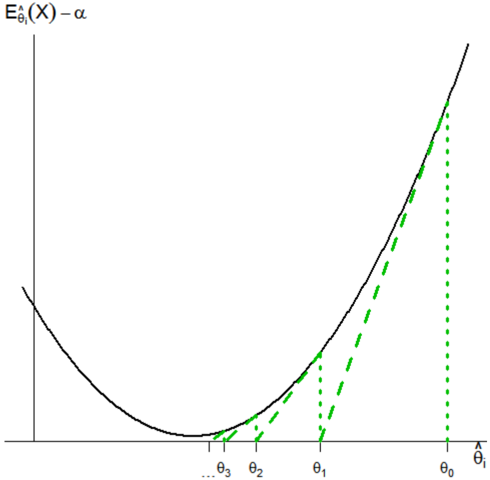
Intuitively:



Updating the value of θ

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i D^{-1} \left(E_{\widehat{\theta}_i}[X] - \alpha \right)$$

Intuitively:



3. Solving the moment equations

We want to solve:

$$E_{\theta}[S-s]=0$$

using the RM step:

$$\widehat{\theta}_{i+1}=\widehat{\theta}_i-a_i D^{-1}\left(E_{\theta_i}[S]-s\right)$$

but we cannot write $\left(E_{\theta_i}[S]-s\right)$ and D in a close form...

3. Solving the moment equations

We want to solve:

$$E_\theta[S - s] = 0$$

using the RM step:

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i D^{-1}\left(E_{\theta_i}[S] - s\right)$$

but we cannot write $\left(E_{\theta_i}[S] - s\right)$ and D in a close form...

Approximate unknown quantities via Monte Carlo (MC) methods
$\Rightarrow$ **stochastic**

Background: Monte Carlo method

Let X be a random variable with distribution function $f_X(x)$.
We want to estimate the expected value $E[s(X)]$.

Definition

The Monte Carlo method consists in:

1. generating a sample $(x_1, \cdots, x_q)$ from the distribution function $f_X(x)$

Background: Monte Carlo method

Let X be a random variable with distribution function $f_X(x)$.
We want to estimate the expected value $E[s(X)]$.

Definition

The Monte Carlo method consists in:

1. generating a sample $(x_1, \cdots, x_q)$ from the distribution function $f_X(x)$
2. computing $s(x_l),\ l = 1, \ldots,\ q$

Background: Monte Carlo method

Let X be a random variable with distribution function $f_X(x)$.
We want to estimate the expected value $E[s(X)]$.

Definition

The Monte Carlo method consists in:

1. generating a sample $(x_1, \cdots, x_q)$ from the distribution function $f_X(x)$
2. computing $s(x_l),\ l = 1, \ldots,\ q$
3. approximating the expected value with the empirical average, i.e.:

$$\overline{S} = \frac{1}{q} \sum_{l=1}^{q} s(x_l)$$

Background: Monte Carlo method

Let X be a random variable with distribution function $f_X(x)$.
We want to estimate the expected value $E[s(X)]$.

Definition

The Monte Carlo method consists in:

1. generating a sample $(x_1, \cdots, x_q)$ from the distribution function $f_X(x)$
2. computing $s(x_l)$, $l = 1, \ldots, q$
3. approximating the expected value with the empirical average, i.e.:

$$\overline{S} = \frac{1}{q} \sum_{l=1}^{q} s(x_l)$$

Reason

It can be proved that

$$\overline{S} \to E[s(X)]$$

as $q \to \infty$

Background: Monte Carlo method - empirical proof

Example

Let us consider the $\mathcal{G}(30, 0.08)$ model.

We are interested in estimating the expected number of edges.

Simulations from the $\mathcal{G}(30, 0.08)$ model for different values of q:

q	10	50	100	1000	10000
$\overline{S}$	32.90	36.18	35.17	34.60	34.81

Background: Monte Carlo method - empirical proof

Example

Let us consider the $\mathcal{G}(30, 0.08)$ model.

We are interested in estimating the expected number of edges.

Simulations from the $\mathcal{G}(30, 0.08)$ model for different values of q:

q	10	50	100	1000	10000
$\overline{S}$	32.90	36.18	35.17	34.60	34.81

Since the number of edges Y follows the binomial distribution,

$$E[Y] = Np = 34.8$$

Background: Monte Carlo method - empirical proof

Example

Let us consider the $\mathcal{G}(30, 0.08)$ model.

We are interested in estimating the expected number of edges.

Simulations from the $\mathcal{G}(30, 0.08)$ model for different values of q:

q	10	50	100	1000	10000
$\overline{S}$	32.90	36.18	35.17	34.60	34.81

Since the number of edges Y follows the binomial distribution,

$$E[Y] = Np = 34.8$$

Our results show that

$$\overline{S} \to E[Y]$$

as $q \to \infty$

The MC Method for the statistics in the SAOM

1. Given $x(t_0)$ and θ

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\cdots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ \ x^{(q)}(t_M)$$

The MC Method for the statistics in the SAOM

1. Given $x(t_0)$ and θ

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\ldots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ \ x^{(q)}(t_M)$$

2. For each sequence compute the value $S^{(l)}$ taken by S

The MC Method for the statistics in the SAOM

1. Given $x(t_0)$ and θ

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\ldots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ \ x^{(q)}(t_M)$$

2. For each sequence compute the value $S^{(l)}$ taken by S

3. Approximate the expected value by

$$\overline{S} = \frac{1}{q}\sum_{l=1}^{q} S^{(l)}$$

The MC Method for the statistics in the SAOM

1. Given $x(t_0)$ and θ

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\cdots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ x^{(q)}(t_M)$$

2. For each sequence compute the value $S^{(l)}$ taken by S

3. Approximate the expected value by

$$\overline{S} = \frac{1}{q}\sum_{l=1}^{q} S^{(l)} \to E_\theta[S]$$

The MC Method for the expected values of the statistics

Example

Approximating $E_\theta[s_{out}(X(t_1)|X(t_0) = x(t_0))]$ for the *"Teenage Friends and Lifestyle Study"* data set

1. Given:

   - $x(t_0)$
   - $\theta = (\lambda_1 = 10.69, \lambda_2 = 8.82, \beta_{out} = -2.63, \beta_{rec} = 2.17, \beta_{trans} = 0.46)$

   simulate the network evolution $q = 1000$ times

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\ldots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ \ x^{(q)}(t_M)$$

The MC Method for the expected values of the statistics

Example

Approximating $E_\theta[s_{out}(X(t_1)|X(t_0) = x(t_0))]$ for the *"Teenage Friends and Lifestyle Study"* data set

1. Given:

   - $x(t_0)$
   - $\theta = (\lambda_1 = 10.69, \lambda_2 = 8.82, \beta_{out} = -2.63, \beta_{rec} = 2.17, \beta_{trans} = 0.46)$

   simulate the network evolution $q = 1000$ times

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\ldots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ \ x^{(q)}(t_M)$$

```
myeff$initialValue[myeff$include] <- c(10.69,8.82,-2.63,2.17,0.46)
sim_model <- sienaModelCreate(projname = 'sim_model', cond = FALSE,
                              useStdInits = FALSE, nsub = 0 , n3=1000)
sim_ans <- siena07(sim_model, data = mydata, effects = myeff,returnDeps=TRUE)
```

The MC Method for the expected values of the statistics

Example

2. Compute the value assumed by S_{out} for each sequence of networks

$$S_{out}^{(l)} = \sum_{m=1}^{M-1} \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij}^{(l)}(t_m)$$

sim	1	2	3	4	5	6	7	8	...
Nr. Edges	942	874	1047	881	865	866	999	948	...

The MC Method for the expected values of the statistics

Example

2. Compute the value assumed by S_{out} for each sequence of networks

$$S_{out}^{(l)} = \sum_{m=1}^{M-1} \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij}^{(l)}(t_m)$$

sim	1	2	3	4	5	6	7	8	...
Nr. Edges	942	874	1047	881	865	866	999	948	...

```
stats <- t(t(sim_ans$sf) + sim_ans$targets)
stats
```

```
     [,1] [,2] [,3] [,4] [,5]
[1,]  480  447  942  570 1267
[2,]  469  450  874  524 1213
[3,]  514  458 1047  632 1567
[4,]  475  441  881  532 1118
[5,]  470  434  865  490  771
[6,]  457  422  866  510 1014
[7,]  509  469  999  632 2078
[8,]  483  428  948  564 1357
[9,]  525  474  962  586 1676
```

The MC Method for the expected values of the statistics

Example

3. Approximate the expected value by

$$\overline{S}_{out} = \frac{1}{q}\sum_{i=1}^{q} S_{out}^{(l)}$$

$$\overline{S}_{out} = \frac{942 + 874 + 1047 + 881 + 865 + 866 + 999 + 948 + \ldots}{1000} \approx 912$$

```
statsMC <- apply(stats,2,mean)
statsMC
```

475.492 438.963 911.737 550.602 1155.026

3. Solving the moment equations

We want to solve:

$$E_\theta[S - s] = 0$$

and using the RM:

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i D^{-1} \left(E_{\theta_i}[S] - s \right)$$

but we cannot write $\left(E_{\theta_i}[S] - s \right)$ and D in a close form...

3. Solving the moment equations

We want to solve:

$$E_\theta[S - s] = 0$$

and using the RM:

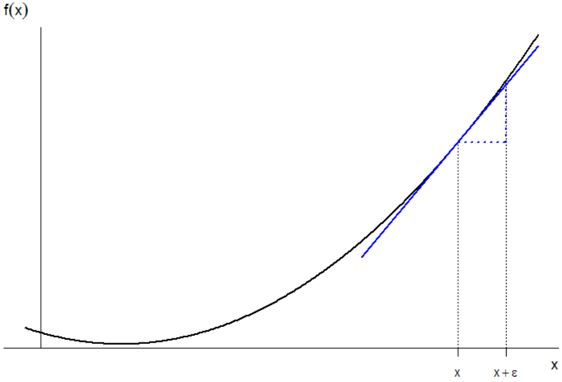
$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i D^{-1} \left(E_{\theta_i}[S] - s \right)$$

but we cannot write $\left(E_{\theta_i}[S] - s \right)$ and D in a close form...

Approximate unknown quantities via Monte Carlo (MC) methods
$\Rightarrow$ **stochastic**

Background: the MC Method for the approximation of derivatives

Univariate function:

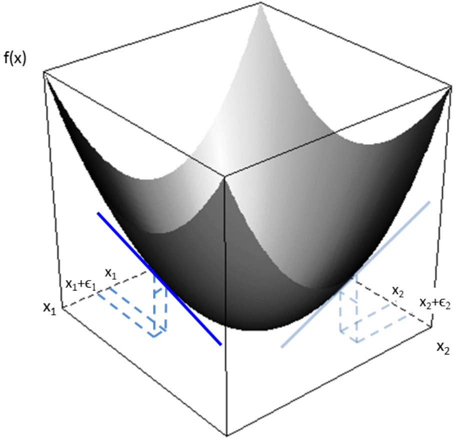


$$\frac{\partial}{\partial x} f(x) = \lim_{\epsilon \to 0} \frac{f(x+\epsilon) - f(x)}{\epsilon}$$

(Finite difference method)

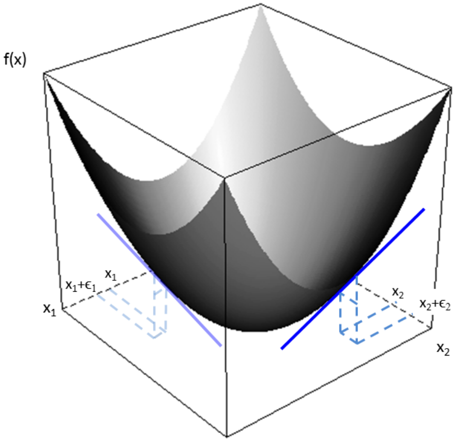
Background: the MC Method for the approximation of derivatives

Multivariate function $x = (x_1, \ldots, x_J)$*:*



Background: the MC Method for the approximation of derivatives

Multivariate function $x = (x_1, \ldots, x_J)$*:*

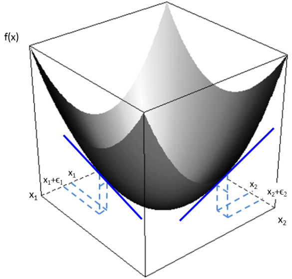


The MC Method for the derivatives of the statistics

Multivariate function $x = (x_1, \ldots, x_J)$*:*

J derivatives, one with respect to each variable

$$D = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x) & & & \\ & \frac{\partial}{\partial x_2} f(x) & & \\ & & \ddots & \\ & & & \frac{\partial}{\partial x_J} f(x) \end{bmatrix}$$

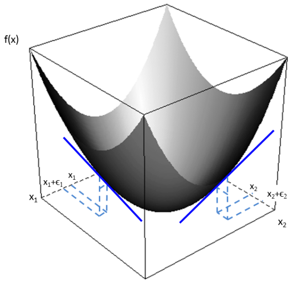


The MC Method for the derivatives of the statistics

Multivariate function $x = (x_1, \ldots, x_J)$:

J derivatives, one with respect to each variable

$$D = \begin{bmatrix} \frac{\partial}{\partial x_1} f(x) & & & \\ & \frac{\partial}{\partial x_2} f(x) & & \\ & & \ddots & \\ & & & \frac{\partial}{\partial x_J} f(x) \end{bmatrix}$$



The computation is done incrementing each variable at a time:

$e_j = (0, \ldots, 0, 1, 0, \ldots, 0)$ j-th unit vector

$\epsilon = (\epsilon_1, \ldots, \epsilon_J)$ vector of increments

$$\frac{\partial}{\partial x_j} f(x) = \lim_{\epsilon_j \to 0} \frac{f(x + \epsilon e_j) - f(x)}{\epsilon_j}$$

The MC Method for the derivatives of the statistics

For the SAOM D is a $M-1+K$ squared diagonal matrix:

$$D = \begin{bmatrix} \frac{\partial}{\partial \lambda_1} E_\theta[S-s] & & & & & \\ & \ddots & & & & \\ & & \frac{\partial}{\partial \lambda_{M-1}} E_\theta[S-s] & & & \\ & & & \frac{\partial}{\partial \beta_1} E_\theta[S-s] & & \\ & & & & \ddots & \\ & & & & & \frac{\partial}{\partial \beta_K} E_\theta[S-s] \end{bmatrix}$$

The diagonal element are computed increasing one parameter at a time by a "small" value ϵ_j

$$\frac{\partial}{\partial \theta_j} E[S-s] = \lim_{\epsilon_j \to 0} \frac{E_{\theta+\epsilon e_j}[S-s] - E_\theta[S-s]}{\epsilon_j} \approx \frac{\overline{S}_{\theta+\epsilon e_j} - \overline{S}_\theta}{\epsilon_j}$$

The MC Method for the expected values of the statistics

Example

Approximating $\frac{\partial}{\partial \beta_{out}} E_\theta[s_{out}(X)]$ for the *"Teenage Friends and Lifestyle Study"* data set, considering a model with outdegree, reciprocity and transitivity

1. Given:

   - $x(t_0)$
   - $\theta = (10.69, 8.82, -2.63, 2.17, 0.46)$

   simulate the network evolution $q = 1000$ times w.r.t. θ

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \dots,\ x^{(1)}(t_M)$$

$$\dots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \dots,\ x^{(q)}(t_M)$$

The MC Method for the expected values of the statistics

Example

Approximating $\frac{\partial}{\partial \beta_{out}} E_\theta[s_{out}(X)]$ for the *"Teenage Friends and Lifestyle Study"* data set, considering a model with outdegree, reciprocity and transitivity

1. Given:

    - $x(t_0)$
    - $\theta = (10.69, 8.82, -2.63, 2.17, 0.46)$

    simulate the network evolution $q = 1000$ times w.r.t. θ

$$x^{(1)}(t_1),\ x^{(1)}(t_2),\ \ldots,\ x^{(1)}(t_M)$$

$$\ldots$$

$$x^{(q)}(t_1),\ x^{(q)}(t_2),\ \ldots,\ \ x^{(q)}(t_M)$$

```
myeff$initialValue[myeff$include] <- c(10.69,8.82,-2.63,2.17,0.46)
sim_model <- sienaModelCreate(projname = 'sim_model', cond = FALSE,
                              useStdInits = FALSE, nsub = 0 , n3=1000)
sim_ans <- siena07(sim_model, data = mydata, effects = myeff,returnDeps=TRUE)
```

The MC Method for the expected values of the statistics

Example

2. Given:

   - $e_3 = (0,0,1,0,0)$ unit vector
   - $\epsilon = (0,0,0.1,0,0)$ vector of increments
   - $\theta + \epsilon e_3 = (10.69, 8.82, -2.53, 2.17, 0.46)$

   simulate the network evolution $q = 1000$ times w.r.t. $\theta + \epsilon e_3$

$$x^{*(1)}(t_1),\ x^{*(1)}(t_2),\ \ldots,\ x^{*(1)}(t_M)$$

$$\ldots$$

$$x^{*(q)}(t_1),\ x^{*(q)}(t_2),\ \ldots,\ x^{*(q)}(t_M)$$

```
myeff$initialValue[myeff$include] <- c(10.69,8.82,-2.63,2.17,0.46)+c(0,0,0.1,0,0)
sim_model2 <- sienaModelCreate(projname = 'sim_model', cond = FALSE,
                               useStdInits = FALSE, nsub = 0 , n3=1000)
sim_ans2 <- siena07(sim_model, data = mydata, effects = myeff,returnDeps=TRUE)
```

The MC Method for the expected values of the statistics

Example

3. Approximate the expected value using the MC method

$$\overline{S}_{out} = \frac{1}{q}\sum_{i=1}^{q} S_{out}^{(l)} = 911.737 \qquad \overline{S}^{*}_{out} = \frac{1}{q}\sum_{i=1}^{q} S_{out}^{*(l)} = 928.749$$

The MC Method for the expected values of the statistics

Example

3. Approximate the expected value using the MC method

$$\overline{S}_{out} = \frac{1}{q}\sum_{i=1}^{q} S_{out}^{(l)} = 911.737 \qquad \overline{S}^{*}_{out} = \frac{1}{q}\sum_{i=1}^{q} S_{out}^{*(l)} = 928.749$$

4. Approximate the first order derivative by

$$\frac{\partial}{\partial_{\beta_{out}}} E[S-s] = \lim_{\epsilon_3 \to 0} \frac{E_{\theta+\epsilon e_3}[S-s] - E_{\theta}[S-s]}{\epsilon_{\lambda_1}} \approx$$

$$\approx \frac{\overline{S}_{\theta+\epsilon e_3} - \overline{S}_{\theta}}{\epsilon_3} = \frac{928.749 - 911.737}{0.1} = 1701.2$$

The MC Method for the expected values of the statistics

Example

3. Approximate the expected value using the MC method

$$\overline{S}_{out} = \frac{1}{q}\sum_{i=1}^{q} S_{out}^{(l)} = 911.737 \qquad \overline{S}^{*}_{out} = \frac{1}{q}\sum_{i=1}^{q} S_{out}^{*(l)} = 928.749$$

4. Approximate the first order derivative by

$$\frac{\partial}{\partial_{\beta_{out}}} E[S-s] = \lim_{\epsilon_3 \to 0} \frac{E_{\theta+\epsilon e_3}[S-s] - E_{\theta}[S-s]}{\epsilon_{\lambda_1}} \approx$$

$$\approx \frac{\overline{S}_{\theta+\epsilon e_3} - \overline{S}_{\theta}}{\epsilon_3} = \frac{928.749 - 911.737}{0.1} = 1701.2$$

```
statsMC <- apply(t(t(sim_ans$sf) + sim_ans$targets),2,mean)
                                   475.492   438.963   911.737   550.602   1155.026
statsMC2 <- apply(t(t(sim_ans2$sf) + sim_ans2$targets),2,mean)
                                   478.205   439.442   928.749   562.564   1202.127
deriv <- (statsMC2[3]-statsMC[3])/0.01
                                                     1701.2
```

The Robbins-Monro algorithm

Phase 1:

- estimation of D
- first update of θ

Phase 2:

- estimation of θ through the RM step

Phase 3:

- estimation of the standard error of θ
- checking the convergence of the algorithm

The Robbins-Monro algorithm - Phase1

Algorithm 2: Robbins-Monro algorithm - Phase 1

Input: θ_0, s, q_1,ϵ
Output: $\widehat{\theta}_{q_1}$, $\widehat{D}$

$i \leftarrow 0$; $d \leftarrow 0$; $S_0 \leftarrow 0$
while $i < q_1$ **do**
  $i \leftarrow i + 1$
  $S_{i0} \sim \theta_0$
  $S_0 \leftarrow S_0 + S_{i0}$
  for *j=1, ... (M+K-1)* **do**
    $S_{ij} \sim \theta_1 + \epsilon e_j$
    $d_{ij} \leftarrow \epsilon_j^{-1}(S_{ij} - S_{i0})$
    $d \leftarrow d + d_{ij} e_j$
$\overline{S} \leftarrow \frac{1}{q_1} S_0$
$\widehat{d} \leftarrow \frac{1}{q_1} d$
$\widehat{D} \leftarrow diag(\widehat{d})$
$\widehat{\theta}_{q_1} \leftarrow \theta_0 - \widehat{D}^{-1}(\overline{S} - s)$
return $\widehat{\theta}_{q_1}$, $\widehat{D}$

The Robbins-Monro algorithm - Phase 2

Algorithm 3: Robbins-Monro algorithm - Phase2

Input: $\widehat{\theta}_{q_1}$, $\widehat{D}$, s
Output: $\widehat{\theta}$

$h \leftarrow 1$
$\widehat{\theta}_1 \leftarrow \widehat{\theta}_{q_1}$
while $h \leq c$ **do**
    $i \leftarrow 0$
    $\theta_0 \leftarrow \widehat{\theta}_h$
    if $i >= q_h^+$ OR $(i > q_h^-$ AND $(S_{ih} - s)(S_{(i-1)h} - s) < 0)$ **then**
        $i \leftarrow i + 1$
        $S_i \sim \widehat{\theta}_i$
        $\widehat{\theta}_{i+1} \leftarrow \widehat{\theta}_i - a_h \widehat{D}^{-1}(\overline{S}_i - s)$
        $\theta \leftarrow \widehat{\theta}_{i+1} + \theta$
    $\widehat{\theta}_h \leftarrow \frac{1}{i}\theta$
    $a_{h+1} \leftarrow a_h/2$
    $h \leftarrow h + 1;$
$\widehat{\theta} \leftarrow \theta_c$
return $\widehat{\theta}$

The Robbins-Monro algorithm - Phase 3

Algorithm 4: Robbins-Monro algorithm - Phase 3

Input: $\widehat{\theta}$, s, q_3,ϵ
Output: $\widehat{\Sigma}_\theta$

$i \leftarrow 0$; $d \leftarrow 0$; $S_0 \leftarrow 0$
while $i < q_3$ **do**
  $i \leftarrow i + 1$
  $S_{i0} \sim \widehat{\theta}$
  $S_0 \leftarrow S_0 + S_{i0}$
  for j=*1, . . . M+K-1* **do**
    $S_{ij} \sim \widehat{\theta} + \epsilon e_j$
    $d_{ij} \leftarrow \epsilon_j^{-1}(S_{ij} - S_{i0})$
    $d \leftarrow d + d_{ij} e_j$

$\overline{S} = \frac{1}{q_3} S_0$
$\widehat{d} = \frac{1}{q_3} d$
$\widehat{D} = diag(D)$
$\widehat{\Sigma}_\theta = \widehat{D}^{-1} \left[\frac{1}{q_3} (S_{i0} - \overline{S})(S_{i0} - \overline{S}) \right] \widehat{D}^{-1}$
return $\widehat{\Sigma}_\theta$

Recap: estimating the parameter of the SAOM

Issue

Given

$$x(t_0),\ x(t_1), \ldots,\ x(t_M)$$

and a parametrization of the SAOM

$$\theta = (\lambda_1,\ \ldots, \lambda_{M-1}, \beta_1,\ \ldots, \beta_K)$$

we want to estimate θ in a plausible way.

Recap: estimating the parameter of the SAOM

Issue

Given

$$x(t_0),\ x(t_1), \ldots,\ x(t_M)$$

and a parametrization of the SAOM

$$\theta = (\lambda_1,\ \ldots, \lambda_{M-1}, \beta_1,\ \ldots, \beta_K)$$

we want to estimate θ in a plausible way.

Different estimation methods are available:

1. Method of Moments:
   an estimation for θ is the value $\widehat{\theta}$ that solves:

$$E_\theta[S - s] = 0$$

Recap: estimating the parameter of the SAOM

Given an initial guess θ_0 for the parameter θ, the procedure can be roughly depicted as follows:

$$
\begin{array}{ccccc}
\theta_0 & \xrightarrow{\textit{simulation}} & E_{\theta_0}[S-s] & \xrightarrow{\textit{rule}} & \theta_1 \\
\theta_1 & \xrightarrow{\textit{simulation}} & E_{\theta_1}[S-s] & \xrightarrow{\textit{rule}} & \theta_2 \\
\dots & \xrightarrow{\textit{simulation}} & \dots & \xrightarrow{\textit{rule}} & \dots \\
\theta_{i-1} & \xrightarrow{\textit{simulation}} & E_{\theta_{i-1}}[S-s] & \xrightarrow{\textit{rule}} & \theta_i \\
\dots & \xrightarrow{\textit{simulation}} & \dots & \xrightarrow{\textit{rule}} & \dots
\end{array}
$$

until a certain criterion is satisfied

Recap: estimating the parameter of the SAOM

Issue

Given

$$x(t_0),\ x(t_1), \ldots,\ x(t_M)$$

and a parametrization of the SAOM

$$\theta = (\lambda_1,\ \ldots, \lambda_{M-1}, \beta_1,\ \ldots, \beta_K)$$

we want to estimate θ in a plausible way.

Different estimation methods are available:

1. Method of Moments:
   an estimation for θ is the value $\widehat{\theta}$ that solves:

   $$E_\theta[S - s] = 0$$

2. Maximum Likelihood Estimation:
   what is the most likely value of θ that could have generated the observed data?

The Maximum-likelihood estimation (MLE)

Definition

Let

- $$\mathcal{F} = \{F(\theta), \theta \in \Theta \subseteq \mathbb{R}^k\}$$

  be a collection of SAOMs parametrized by $\theta \in \Theta \subseteq \mathbb{R}^k$

- $x(t_0), \ldots, x(t_M)$ be the observed data

The Maximum-likelihood estimation (MLE)

Definition

Let

-
$$\mathcal{F} = \{F(\theta), \theta \in \Theta \subseteq \mathbb{R}^k\}$$
  be a collection of SAOMs parametrized by $\theta \in \Theta \subseteq \mathbb{R}^k$

- $x(t_0), \ldots, x(t_M)$ be the observed data

The likelihood function associated with the observed data is:

$$L : \Theta \to \mathbb{R}; \theta \longmapsto P_\theta(x(t_0), \ldots, x(t_M))$$

The Maximum-likelihood estimation (MLE)

Definition

Let

-
$$\mathcal{F} = \{F(\theta), \theta \in \Theta \subseteq \mathbb{R}^k\}$$
be a collection of SAOMs parametrized by $\theta \in \Theta \subseteq \mathbb{R}^k$

- $x(t_0), \ldots, x(t_M)$ be the observed data

The likelihood function associated with the observed data is:

$$L : \Theta \to \mathbb{R}; \theta \longmapsto P_\theta(x(t_0), \ldots, x(t_M))$$

A parameter vector $\widehat{\theta}$ maximizing L:

$$\widehat{\theta} = arg \max_{\theta \in \Theta} L(\theta)$$

is called a *maximum likelihood estimate* for θ

Computing the Likelihood function

For semplicity let us consider only two observations $x(t_0)$ and $x(t_1)$

The model assumptions allow to decompose the process in a series of micro-steps:

$$\{(T_r, i_r, j_r), r = 1, \ldots, R\}$$

where

- T_r is the time point for an opportunity for change
- i_r denotes the actor who has the opportunity to change
- j_r is the actor towards whom the tie is changed

Let R be the total number of micro-steps between t_0 and t_1.

We assume that the time point T_r are ordered increasingly:

$$t_0 = T_0 < T_1 < \ldots < T_R < t_1$$

Computing the Likelihood function

Definition

Given the sequence $\{(T_r, i_r, j_r), r = 1, \dots, R\}$

$$L(\theta) = \prod_{r=1}^{R} P_\theta((T_r, i_r, j_r)) \propto \frac{(n\lambda)^R}{R!} e^{-n\lambda} \prod_{r=1}^{R} \frac{1}{n} p_{i_r j_r}(\beta, x(T_r))$$

Then, the estimate for θ is

$$\widehat{\theta} = arg \max_{\theta \in \Theta} L(\theta)$$

or equivalently

$$\widehat{\theta} = arg \max_{\theta \in \Theta} log(L(\theta))$$

where $log(L(\theta))$ is called *the log-likelihood function*

The Maximum-likelihood estimation (MLE)

In practice finding

$$\widehat{\theta} = arg \max_{\theta \in \Theta} log(L(\theta))$$

means determining $\widehat{\theta}$ such that:

$$\frac{\partial}{\partial \theta} log(L(\theta)) = 0$$

where $\frac{\partial}{\partial \theta} log(L(\theta))$ is called *score function*.

The Maximum-likelihood estimation (MLE)

Example

Let us consider the *"Teenage Friends and Lifestyle Study"* data set.

We model the network evolution according to the following parameter

$$\theta = (\lambda_1, \lambda_2, \beta_{out}, \beta_{rec}, \beta_{trans})$$

We look for $\widehat{\theta}$ such that:

$$\begin{cases} \frac{\partial}{\partial \lambda_1} log(L(\theta)) = 0 \\ \frac{\partial}{\partial \lambda_2} log(L(\theta)) = 0 \\ \frac{\partial}{\partial \beta_{out}} log(L(\theta)) = 0 \\ \frac{\partial}{\partial \beta_{rec}} log(L(\theta)) = 0 \\ \frac{\partial}{\partial \beta_{trans}} log(L(\theta)) = 0 \end{cases}$$

The Maximum-likelihood estimation (MLE)

Problem:
we cannot observe the complete data, i.e., the complete series of micro-steps that lead from $x(t_0)$ to $x(t_1)$, from $x(t_1)$ to $x(t_2)$, ...

$$\Downarrow$$

we cannot compute the L of the observed data

$$\Downarrow$$

a stochastic approximation method must be applied.

The Maximum-likelihood estimation (MLE)

Given an initial guess θ_0 for the parameter θ, the procedure can be roughly depicted as follows:

$$
\begin{array}{ccccc}
\theta_0 & \xrightarrow{\textit{simulation}} & \frac{\partial}{\partial\theta} log(L(\theta_0)) & \xrightarrow{\textit{rule}} & \theta_1 \\
\theta_1 & \xrightarrow{\textit{simulation}} & \frac{\partial}{\partial\theta} log(L(\theta_1)) & \xrightarrow{\textit{rule}} & \theta_2 \\
\ldots & \xrightarrow{\textit{simulation}} & \ldots & \xrightarrow{\textit{rule}} & \ldots \\
\theta_{i-1} & \xrightarrow{\textit{simulation}} & \frac{\partial}{\partial\theta} log(L(\theta_{i-1})) & \xrightarrow{\textit{rule}} & \theta_i \\
\ldots & \xrightarrow{\textit{simulation}} & \ldots & \xrightarrow{\textit{rule}} & \ldots
\end{array}
$$

until a certain criterion is satisfied

Augmented data

To approximate the likelihood we use the *augmented data method*

Definition

The *augmented data* (or *sample path*) consist of the sequence of tie changes that brings the network from $x(t_0)$ to $x(t_1)$

$$(i_1, j_1), \ldots, (i_R, j_R)$$

Formally:

$$\underline{v} = \{(i_1, j_1), \ldots, (i_R, j_R)\} \in \mathcal{V}$$

where $\mathcal{V}$ is the set of all sample paths connecting $x(t_0)$ and $x(t_1)$.

Augmented data

To approximate the likelihood we use the *augmented data method*

Definition

The *augmented data* (or *sample path*) consist of the sequence of tie changes that brings the network from $x(t_0)$ to $x(t_1)$

$$(i_1, j_1), \ldots, (i_R, j_R)$$

Formally:

$$\underline{v} = \{(i_1, j_1), \ldots, (i_R, j_R)\} \in \mathcal{V}$$

where $\mathcal{V}$ is the set of all sample paths connecting $x(t_0)$ and $x(t_1)$.

We can approximate the likelihood function (and then the score function) of the observed data using the probability of $\underline{v}$

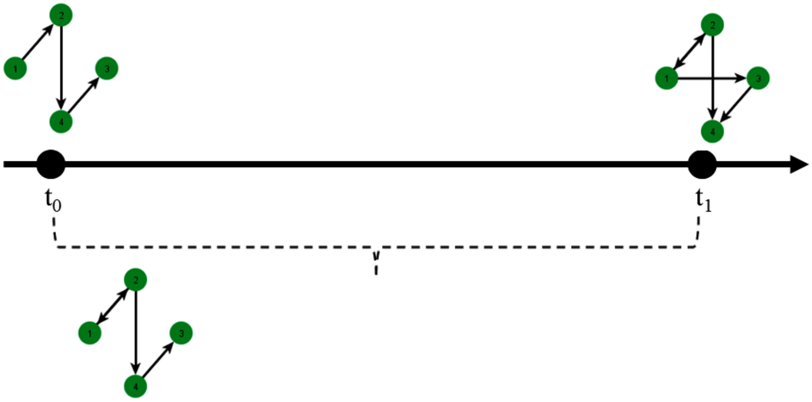
$$P(\underline{v} | x(t_0), x(t_1)) \propto \frac{(n\lambda)^R}{R!} e^{-n\lambda} \prod_{r=1}^{R} \frac{1}{n} p_{i_r j_r}(\beta, x(T_r))$$

Approximating the Likelihood function

Example

Approximating the Likelihood function

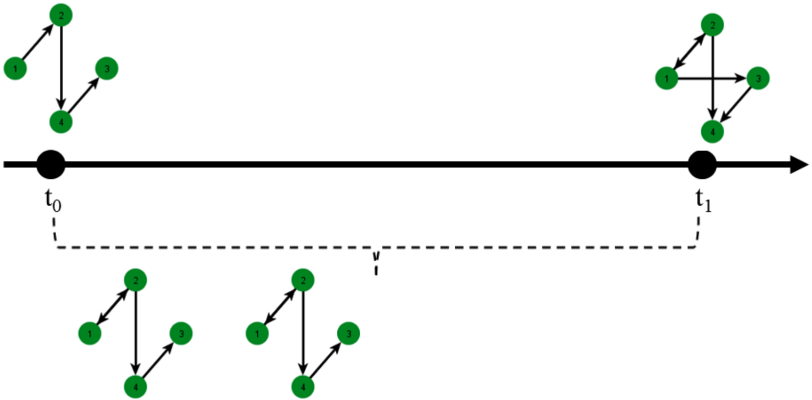
Example



$\underline{v} = \{(2,1) \qquad\qquad\qquad\qquad \}$

Approximating the Likelihood function

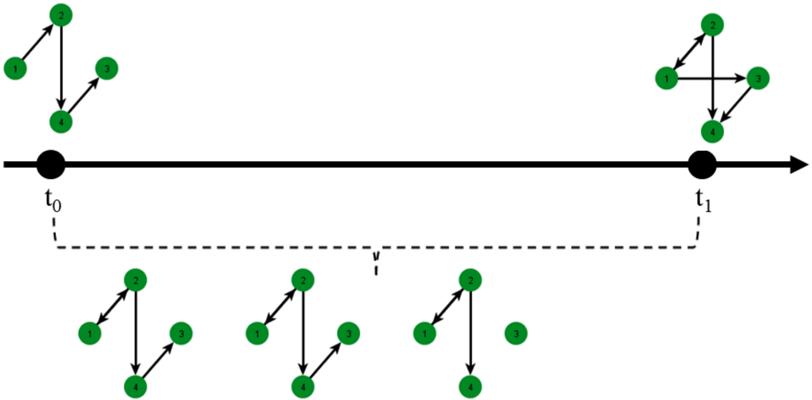
Example



$\underline{v} = \{(2,1),(1,1) \qquad\qquad\qquad \}$

Approximating the Likelihood function

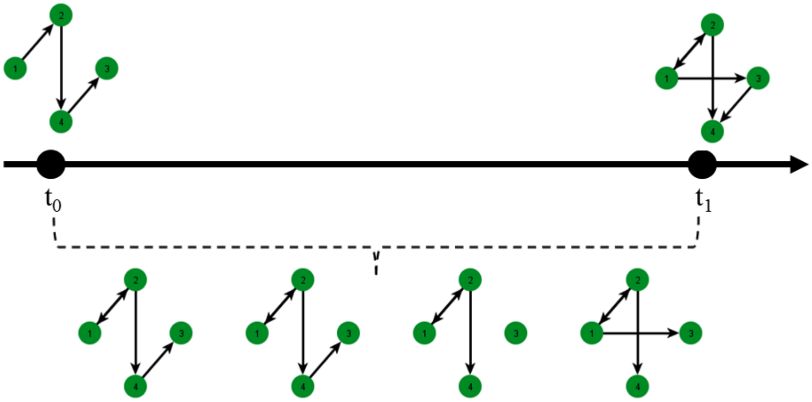
Example



$\underline{v} = \{(2,1),(1,1),(4,3) \qquad\qquad\qquad \}$

Approximating the Likelihood function

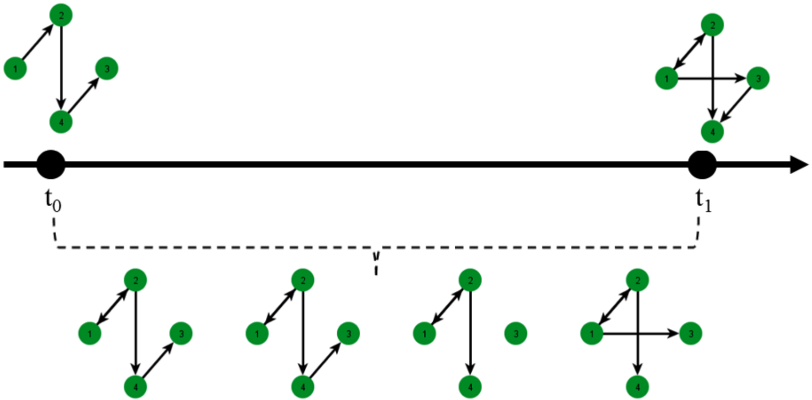
Example



$\underline{v} = \{(2,1),(1,1),(4,3),(1,3) \qquad \}$

Approximating the Likelihood function

Example



$$\underline{v} = \{(2,1),(1,1),(4,3),(1,3),(3,4)\}$$

Augmented data

Example

$$\mathcal{V} = \{\{(2,1),(1,1),(4,3),(1,3),(3,4)\},$$

$$\{(1,2),(3,4),(1,3),(4,3),(2,1),(1,2)\},$$

$$\{(3,3),(4,4),(2,3),(4,3),(2,1),(2,3),(3,4),(1,1)\},$$

$$\ldots\}$$

Sampling the augmented data

How to sample the augmented data from the distribution:

$$P(\underline{v}|x(t_0), x(t_1)) \propto \frac{(n\lambda)^R}{R!} e^{-n\lambda} \prod_{r=1}^{R} \frac{1}{n} p_{i_r j_r}(\beta, x(T_r))$$

given a certain value of the parameter θ?

The augmented data are sampled through the *Metropolis-Hastings algorithm*

Sampling the augmented data

The *Metropolis-Hastings algorithm* is defined by the following steps:

1. given $\underline{v}_i = \underline{v}$, generate $\underline{\widetilde{v}}$ from a proposal distribution $u(\underline{\widetilde{v}}|\underline{v}_i)$

Sampling the augmented data

The *Metropolis-Hastings algorithm* is defined by the following steps:

1. given $\underline{v}_i = \underline{v}$, generate $\widetilde{\underline{v}}$ from a proposal distribution $u(\widetilde{\underline{v}}\,|\,\underline{v}_i\,)$

   The proposal distribution $u(\widetilde{\underline{v}}\,|\,\underline{v}_i\,) : \mathcal{V} \to [0,1]$ assigns non-zero probabilities only to the following 5 cases:

   a. *Pairwise deletions*

   b. *Pairwise insertions*

   c. *Single deletion*

   d. *Single insertion*

   e. *Permutation*

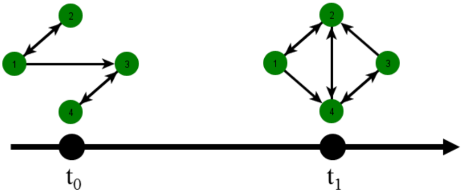
Sampling the augmented data

a. *Pairwise deletions*: r_1 and r_2 such that $(i_{r_1}, j_{r_1}) = (i_{r_2}, j_{r_2})$ is selected and the pairs (i_{r_1}, j_{r_1}) and (i_{r_2}, j_{r_2}) are deleted from $\underline{v}$

Sampling the augmented data

a. *Pairwise deletions*: r_1 and r_2 such that $(i_{r_1}, j_{r_1}) = (i_{r_2}, j_{r_2})$ is selected and the pairs (i_{r_1}, j_{r_1}) and (i_{r_2}, j_{r_2}) are deleted from $\underline{v}$

Example



$\underline{v} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$

- Select at random (r_1, r_2) in $\{(1,7),(1,10),(7,10),(2,8)\}$, e.g. $(r_1, r_2) = (1,7)$
- Delete the elements $(2,4)$

$$\widetilde{\underline{v}} = (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

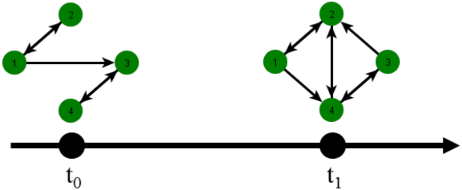
Sampling the augmented data

b. *Pairwise insertions*: $(i,j) \in \mathcal{N}^2$ and r_1 and r_2 are randomly chosen. The element (i,j) is inserted in $\underline{v}$ immediately before r_1 and r_2

Sampling the augmented data

b. *Pairwise insertions*: $(i,j) \in \mathcal{N}^2$ and r_1 and r_2 are randomly chosen. The element (i,j) is inserted in $\underline{v}$ immediately before r_1 and r_2

Example



$$\underline{v} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

- Select at random (i,j) and (r_1, r_2), e.g. $i = 4$, $j = 1$, $r_1 = 5$, $r_2 = 7$
- Insert the elements $(4,1)$ before $r_1 = 5$ and $r_2 = 7$

$$\widetilde{\underline{v}} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ \mathbf{(4,1)}\ (3,2)\ (1,4)\ \mathbf{(4,1)}\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

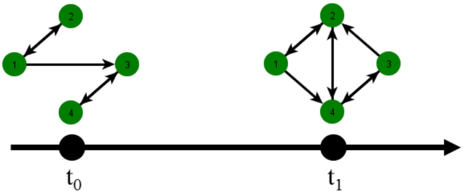
Sampling the augmented data

c. *Single deletion*: one pair (i_r, j_r) satisfying $i_r = j_r$ is randomly selected and deleted from $\underline{v}$

Sampling the augmented data

c. *Single deletion*: one pair (i_r, j_r) satisfying $i_r = j_r$ is randomly selected and deleted from $\underline{v}$

Example



$\underline{v} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$

- Select at random r in $\{3, 11\}$, e.g. $r = 11$
- Delete the elements $(3,3)$

$\widetilde{\underline{v}} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)$

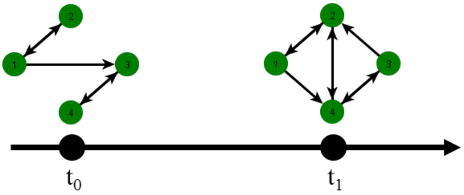
Sampling the augmented data

d. *Single insertion*: one actor $i \in \mathcal{N}$ and an index r are selected. The element (i, i) is inserted immediately before r

Sampling the augmented data

d. *Single insertion*: one actor $i \in \mathcal{N}$ and an index r are selected. The element (i, i) is inserted immediately before r

Example



$$\underline{v} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

- Select at random $i \in \mathcal{N}$ and r, e.g. $i = 4$ $r = 6$
- Insert the elements $(4, 4)$ before $r = 6$

$$\widetilde{\underline{v}} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ \mathbf{(4,4)}\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

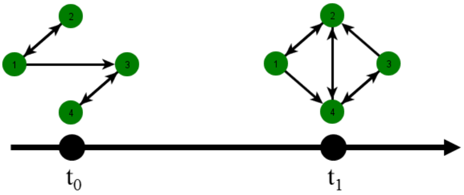
Sampling the augmented data

e. *Permutations*: for randomly chosen indices $r_1 < r_2$, the sequence $(i_{r_1}, j_{r_1}), \ldots, ((i_{r_2}, j_{r_2}))$ is randomly permuted

Sampling the augmented data

e. *Permutations*: for randomly chosen indices $r_1 < r_2$, the sequence $(i_{r_1}, j_{r_1}), \ldots, ((i_{r_2}, j_{r_2}))$ is randomly permuted

Example



$$\underline{v} = (2,4)\ (2,3)\ (1,1)\ (4,2)\ (3,2)\ (1,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

- Select at random (r_1, r_2) and r, e.g. $r_1 = 2, r_2 = 5$
- Permute the sequence $(i_2, j_2), \ldots, (i_5, j_5)$

$$\underline{v} = (2,4)\ \mathbf{(1,1)}\ \mathbf{(2,3)}\ \mathbf{(3,2)}\ \mathbf{(4,2)}\ (1,4)\ (4,4)\ (2,4)\ (2,3)\ (1,3)\ (2,4)\ (3,3)$$

Sampling the augmented data

The *Metropolis-Hastings algorithm* is defined by the following steps:

1. given $\underline{v}_i = \underline{v}$, generate $\widetilde{\underline{v}}$ from the proposal distribution $u(\widetilde{\underline{v}}|\underline{v}_i)$

Sampling the augmented data

The *Metropolis-Hastings algorithm* is defined by the following steps:

1. given $\underline{v}_i = \underline{v}$, generate $\widetilde{\underline{v}}$ from the proposal distribution $u(\widetilde{\underline{v}}|\underline{v}_i)$

2. take

$$\underline{v}_{i+1} = \begin{cases} \widetilde{\underline{v}} & \textit{with probability} \quad \rho(\widetilde{\underline{v}}, \underline{v}) \\ \underline{v} & \textit{with probability} \quad 1 - \rho(\widetilde{\underline{v}}, \underline{v}) \end{cases}$$

   where

$$\rho(\widetilde{\underline{v}}, \underline{v}) = \min\left\{ \frac{P(\widetilde{\underline{v}})u(\underline{v}|\widetilde{\underline{v}})}{P(\underline{v})u(\widetilde{\underline{v}}|\underline{v})}, 1 \right\}$$

The transition probabilities of the chain generate by the Metropolis-Hastings algorithm are given by $\rho(\widetilde{\underline{v}}, \underline{v})u(\widetilde{\underline{v}}|\underline{v})$

Sampling the augmented data

Theorem

The Metropolis-Hastings algorithm leads to an irreducible, aperiodic and reversible Markov-chain with stationary distribution:

$$P(\underline{v}|x(t_0),x(t_1)) \propto \frac{(n\lambda)^R}{R!} e^{-n\lambda} \prod_{r=1}^{R} \frac{1}{n} p_{i_r j_r}(\beta, x(T_r))$$

Proof

- *The Markov chain is irreducible*.
  Pairwise deletions and insertions and single deletion and insertion are sufficient for all $\underline{v} \in \mathcal{V}$ to communicate.

- *The Markov chain is aperiodic*.
  The graph associated to the resulting Markov-chain contains all the loops and thus the greatest common divisor of all cycles is one.

Sampling the augmented data

- *The Markov chain is reversible*. The detailed balance condition:

$$\rho(\underline{\widetilde{v}}, \underline{v}) u(\underline{\widetilde{v}} | \underline{v}) P(\underline{v}) = \rho(\underline{v}, \underline{\widetilde{v}}) u(\underline{v} | \underline{\widetilde{v}}) P(\underline{\widetilde{v}})$$

is satisfied.

$$\rho(\underline{\widetilde{v}}, \underline{v}) u(\underline{\widetilde{v}} | \underline{v}) P(\underline{v}) = \min\left\{ \frac{P(\underline{\widetilde{v}}) u(\underline{v} | \underline{\widetilde{v}})}{P(\underline{v}) u(\underline{\widetilde{v}} | \underline{v})}, 1 \right\} u(\underline{\widetilde{v}} | \underline{v}) P(\underline{v}) =$$

$$= \min\left\{ \frac{P(\underline{\widetilde{v}}) u(\underline{v} | \underline{\widetilde{v}})}{u(\underline{\widetilde{v}} | \underline{v})}, P(\underline{v}) \right\} u(\underline{\widetilde{v}} | \underline{v}) =$$

$$= \min\left\{ \frac{u(\underline{v} | \underline{\widetilde{v}})}{u(\underline{\widetilde{v}} | \underline{v})}, \frac{P(\underline{v})}{P(\underline{\widetilde{v}})} \right\} u(\underline{\widetilde{v}} | \underline{v}) P(\underline{\widetilde{v}}) =$$

$$= \min\left\{ 1, \frac{P(\underline{v}) u(\underline{\widetilde{v}} | \underline{v})}{P(\underline{\widetilde{v}}) u(\underline{v} | \underline{\widetilde{v}})} \right\} u(\underline{v} | \underline{\widetilde{v}}) P(\underline{\widetilde{v}}) =$$

$$= \rho(\underline{v}, \underline{\widetilde{v}}) u(\underline{v} | \underline{\widetilde{v}}) P(\underline{\widetilde{v}})$$

Sampling the augmented data

The ML estimation algorithm can be sketched in the following way:

1. For each $m = 1, \dots, M-1$ makes a large number of Metropolis-Hastings steps yielding $v^{(i)} = (v_1^{(i)}, \dots, v_{M-1}^{(i)})$

2. Compute the score function:

$$\frac{\partial}{\partial\theta} log(L(\widehat{\theta}_i; x; v_m^{(i)}))$$

3. Update the parameter estimate using the Robbins-Monro step

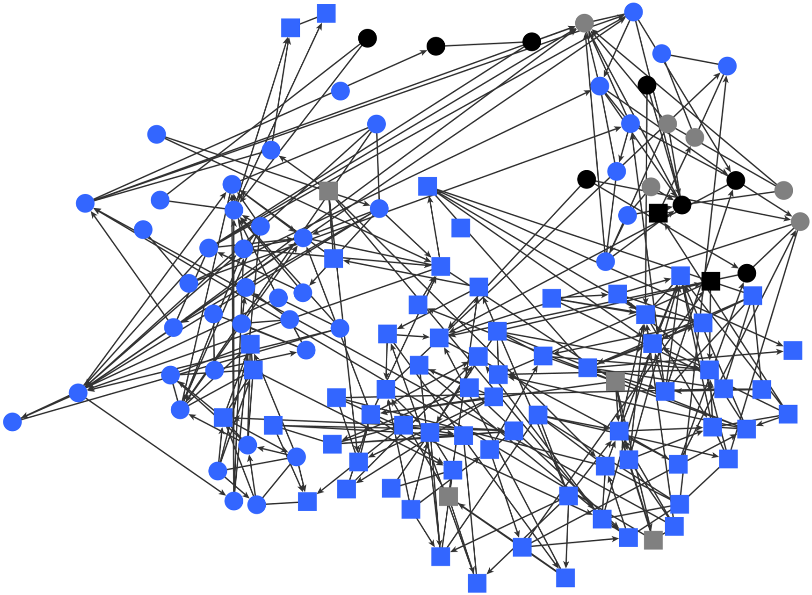
$$\theta_{i+1} = \theta_i + a_i D^{-1} U(L(\widehat{\theta}_i; x; v_m^{(i)}))$$

The estimate $\widehat{\theta}$ is calculated as the average of the θ_{i+1} values generated by this algorithm.

Outline

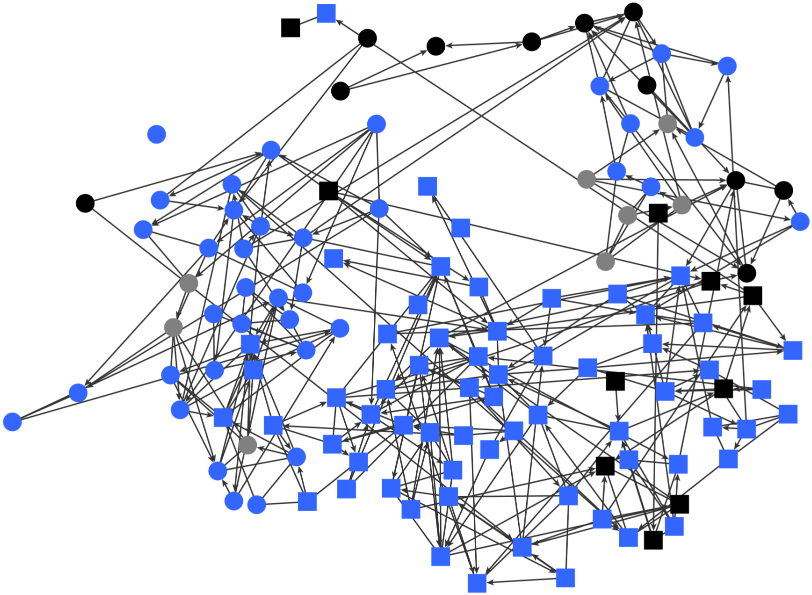
Networks are dynamic by nature: a real example

Ties and actors'characteristics can change over time.



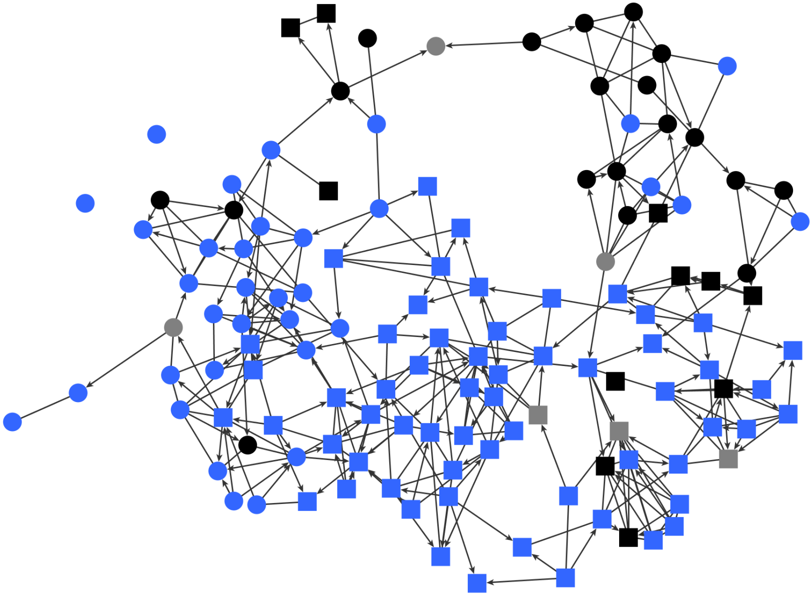
Networks are dynamic by nature: a real example

Ties and actors'characteristics can change over time.



Networks are dynamic by nature: a real example

Ties and actors'characteristics can change over time.



Motivation

1. Social network dynamics can depend on actors'characteristics.

Selection process: relationship *partners* are selected according to their characteristics

Motivation

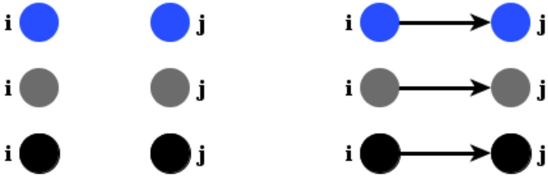
1. Social network dynamics can depend on actors'characteristics.

Selection process: relationship *partners* are selected according to their characteristics

Example

Homophily: the formation of relations based on the similarity of two actors

E.g. smoking behavior



Motivation

2. Changeable actors'characteristics can depend on the social network

E.g.: opinions, attitudes, intentions, etc. - we use the word behavior for all of these!

Influence process: actors adjust their characteristics according to the characteristics of other actors to whom they are tied

Motivation

2. Changeable actors'characteristics can depend on the social network

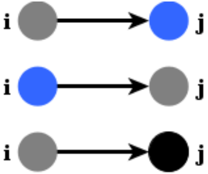
E.g.: opinions, attitudes, intentions, etc. - we use the word behavior for all of these!

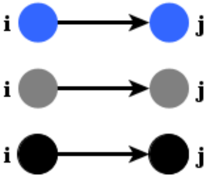
Influence process: actors adjust their characteristics according to the characteristics of other actors to whom they are tied

Example

Assimilation/contagion: connected actors become increasingly similar over time

E.g. smoking behavior





Competing explanatory stories

Homophily and assimilation give rise to the same outcome (similarity of connected individuals)

$$\Downarrow$$

study of influence requires the consideration of selection and vice versa.

Fundamental question: is this similarity caused mainly by influence or mainly by selection?

Competing explanatory stories

Homophily and assimilation give rise to the same outcome (similarity of connected individuals)

$$\Downarrow$$

study of influence requires the consideration of selection and vice versa.

Fundamental question: is this similarity caused mainly by influence or mainly by selection?

Extending the SAOM for the co-evolution of networks and behaviors

Competing explanatory stories

Example

Similarity in smoking:

Selection: "a smoker may tend to have smoking friends because, once somebody is a smoker, he or she is likely to meet other smokers in smoking areas and thus has more opportunities to form friendship ties with them"

Competing explanatory stories

Example

Similarity in smoking:

Selection: "a smoker may tend to have smoking friends because, once somebody is a smoker, he or she is likely to meet other smokers in smoking areas and thus has more opportunities to form friendship ties with them"

Influence: "the friendship with a smoker may have made an actor smoking in the first place"

Longitudinal network-behavior panel data

1. a network x represented by its adjacency matrix

2. a series of actors' attributes:
   - H constant covariates $V_1, \cdots, V_H$
   - L behavior covariates $Z_1(t), \cdots, Z_L(t)$
     Behavior variables are ordinal categorical variables.

Longitudinal network-behavior panel data

1. a network x represented by its adjacency matrix

2. a series of actors' attributes:
   - H constant covariates $V_1, \cdots, V_H$
   - L behavior covariates $Z_1(t), \cdots, Z_L(t)$
     Behavior variables are ordinal categorical variables.

Longitudinal network-behavior panel data: networks and behaviors observed at $M \geq 2$ time points $t_1, \cdots, t_M$

$$(x,z)(t_0),\ (x,z)(t_1), \cdots,\ (x,z)(t_M)$$

and the constant covariates $V_1, \cdots, V_H$.

Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space* $\mathcal{C}$*:* all the possible configurations arising from the combination of network and behaviors

     $$|C| = 2^{n(n-1)} \times B^n$$

     where B is the number of categories for the behavior variable.

Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

     $$|C| = 2^{n(n-1)} \times B^n$$

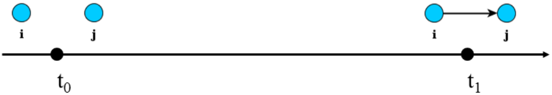
     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network

Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

     $$|C| = 2^{n(n-1)} \times B^n$$

     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network
   - *Continuous-time:*

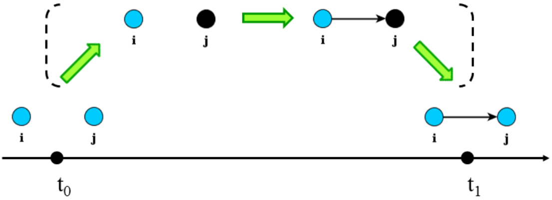


Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

     $$|C| = 2^{n(n-1)} \times B^n$$

     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network
   - *Continuous-time:*

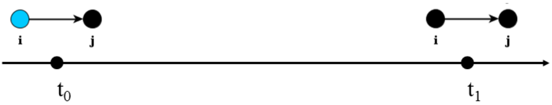


Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space* $\mathcal{C}$*:* all the possible configurations arising from the combination of network and behaviors

     $$|C| = 2^{n(n-1)} \times B^n$$

     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network and behavior
   - *Continuous-time:*

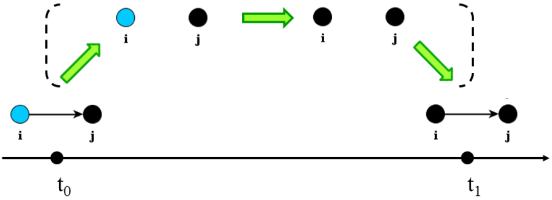


Assumptions

1. *Distribution of the process.*
   Changes between observational time points are modeled according to a continuous-time Markov chain.
   - *State space $\mathcal{C}$:* all the possible configurations arising from the combination of network and behaviors

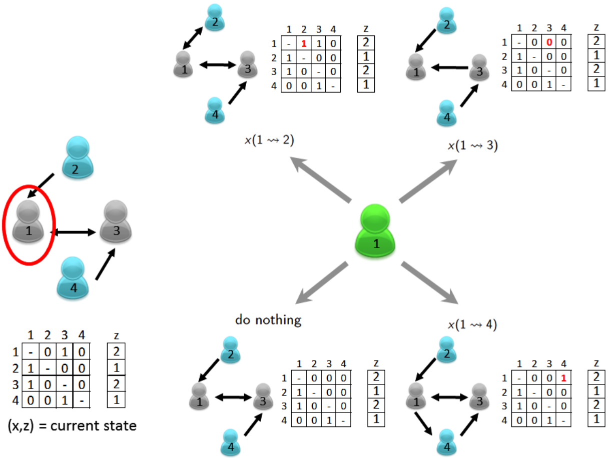
     $$|C| = 2^{n(n-1)} \times B^n$$

     where B is the number of categories for the behavior variable.
   - *Markovian assumption:* changes actors make are assumed to depend only on the current state of the network and behavior
   - *Continuous-time:*



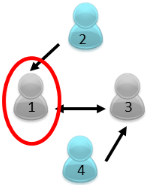
Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing ties or his behavior.



Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing ties or his behavior.

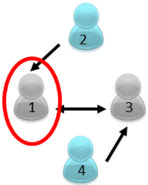


	1	2	3	4	z
1	-	0	1	0	2
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

(x,z) = current state

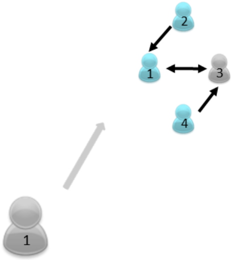
Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing ties or his behavior.



	1	2	3	4	z
1	-	0	1	0	2
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

(x,z) = current state

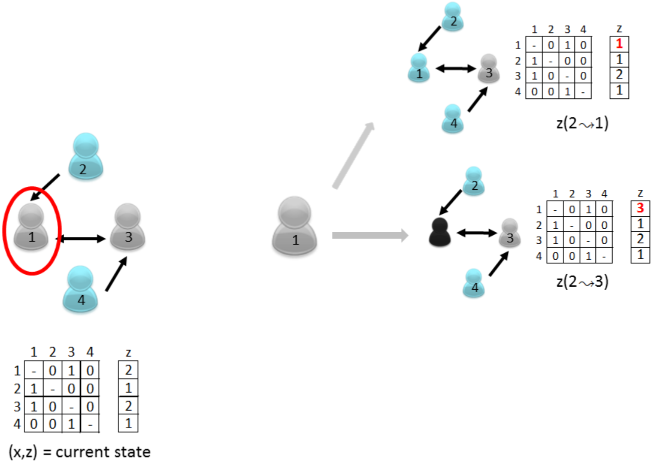


	1	2	3	4	z
1	-	0	1	0	1
2	1	-	0	0	1
3	1	0	-	0	2
4	0	0	1	-	1

z(2⇝1)

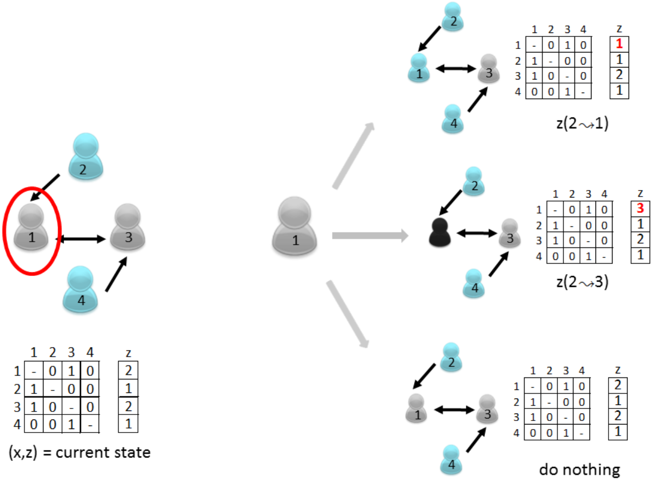
Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing ties or his behavior.



Assumptions

2. *Opportunity to change.*
   At any given moment one probabilistically selected actor has the opportunity to change one of his outgoing ties or his behavior.



Assumptions

3. *Absence of co-occurrence.*
   At each instant t, only one actor has the opportunity to change (one of his outgoing ties or his behavior)

Assumptions

3. *Absence of co-occurrence.*
   At each instant t, only one actor has the opportunity to change (one of his outgoing ties or his behavior)

4. *Actor-oriented perspective.*
   Actors control their outgoing ties as well as their own behavior.

   - the actor decide to change one of his outgoing ties or his behavior trying to maximize *a utility function*
   - two distinct objective functions: one for the network and one for the behavior change
   - actors have complete knowledge about the network and the behaviors of all the the other actors
   - the maximization is based on immediate returns (myopic actors)

Model definition

The co-evolution process is decomposed into a series of micro-steps:

- the opportunity of changing one network tie and the corresponding tie changed

- the opportunity of changing a behavior and the corresponding unit changed in behavior

Model definition

The co-evolution process is decomposed into a series of micro-steps:

- the opportunity of changing one network tie and the corresponding tie changed
- the opportunity of changing a behavior and the corresponding unit changed in behavior

$$\Downarrow$$

every micro-step requires the identification of a focal actor who gets the opportunity to make a change and the identification of the change outcome

Model definition

The co-evolution process is decomposed into a series of micro-steps:

- the opportunity of changing one network tie and the corresponding tie changed
- the opportunity of changing a behavior and the corresponding unit changed in behavior

$$\Downarrow$$

every micro-step requires the identification of a focal actor who gets the opportunity to make a change and the identification of the change outcome

	Occurrence	*Preference*
Network changes	Network rate function	Network objective function
Behavioral changes	Behavioral rate function	Behavioral objective function

The rate functions

The frequency by which actors have the opportunity to make a change is modeled by the *rate functions*, one for each type of change.

Why must we specify two different rate functions?

The rate functions

The frequency by which actors have the opportunity to make a change is modeled by the *rate functions*, one for each type of change.

Why must we specify two different rate functions?

Practically always, one type of decision will be made more frequently than the other

The rate functions

The frequency by which actors have the opportunity to make a change is modeled by the *rate functions*, one for each type of change.

Why must we specify two different rate functions?

Practically always, one type of decision will be made more frequently than the other

Example

In the joint study of friendship and smoking behavior at high school, we would expect more frequent changes in the network than in behavior

The rate functions

Network rate function
T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda_i^{net})$$

Behavior rate function
T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda_i^{beh})$$

The rate functions

Network rate function
T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda_i^{net})$$

Behavior rate function
T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda_i^{beh})$$

Waiting time for a new micro-step
$T_i^{net \vee beh}$ = the waiting time until i gets the opportunity to make any change

$$T_i^{net \vee beh} \sim Exp(\lambda_{tot})$$

where

$$\lambda_{tot} = \sum_i (\lambda_i^{net} + \lambda_i^{beh})$$

The rate functions (simplest specification)

Network rate function

T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda^{net})$$

Behavior rate function

T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda^{beh})$$

The rate functions (simplest specification)

Network rate function
T_i^{net} = the waiting time until i gets the opportunity to make a network change

$$T_i^{net} \sim Exp(\lambda^{net})$$

Behavior rate function
T_i^{beh} = the waiting time until i gets the opportunity to make a behavior change

$$T_i^{beh} \sim Exp(\lambda^{beh})$$

Waiting time for a new micro-step
$T_i^{net \vee beh}$ = the waiting time until i gets the opportunity to make any change

$$T_i^{net \vee beh} \sim Exp(\lambda_{tot})$$

where

$$\lambda_{tot} = n(\lambda^{net} + \lambda^{beh})$$

The rate functions (simplest specification)

The rate functions (simplest specification)

Probabilities for an actor to make a micro-step

$$P(i\ can\ make\ a\ network\ micro-step) = \frac{\lambda^{net}}{\lambda_{tot}}$$

$$P(i\ can\ make\ a\ behavioral\ micro-step) = \frac{\lambda^{beh}}{\lambda_{tot}}$$

The rate functions (simplest specification)

Probabilities for an actor to make a micro-step

$$P(i\ \ can\ \ make\ \ a\ \ network\ \ micro-step) = \frac{\lambda^{net}}{\lambda_{tot}}$$

$$P(i\ \ can\ \ make\ \ a\ \ behavioral\ \ micro-step) = \frac{\lambda^{beh}}{\lambda_{tot}}$$

Probabilities for a micro-step

$$P(network\ \ micro-step) = \frac{n\lambda^{net}}{\lambda_{tot}} = \frac{\lambda^{net}}{\lambda^{net} + \lambda^{beh}}$$

$$P(behavioral\ \ micro-step) = \frac{n\lambda^{beh}}{\lambda_{tot}} = \frac{\lambda^{beh}}{\lambda^{net} + \lambda^{beh}}$$

The objective functions

Why must we specify two different objective functions?

The objective functions

Why must we specify two different objective functions?

- The network objective function represents how likely it is for *i* to change one of his outgoing ties
- The behavioral objective function represents how likely it is for the actor *i* the current level of his behavior

The objective functions

Why must we specify two different objective functions?

- The network objective function represents how likely it is for i to change one of his outgoing ties
- The behavioral objective function represents how likely it is for the actor i the current level of his behavior

Network utility function

$$\begin{aligned} u_i^{net}(\beta, x(i \rightsquigarrow j), z, v) &= f_i^{net}(\beta, x(i \rightsquigarrow j), z, v) + \epsilon_i(t, x, j) \\ &= \sum_{k=1}^{K} \beta_k s_{ik}^{net}(x, z, v) + \epsilon_i(t, x, j) \end{aligned}$$

The objective functions

Behavioral utility function

$$\begin{aligned} u_i^{beh}(\gamma, z(l \rightsquigarrow l'), x, v) &= f_i^{beh}(\gamma, z(l \rightsquigarrow l'), x, v) + \epsilon_i(t, z, l, l') \\ &= \sum_{w=1}^{W} \gamma_w s_{iw}^{beh}(x, z(l \rightsquigarrow l'), v) + \epsilon_i(t, z, l, l') \end{aligned}$$

where

- $s_w^{beh}(x, z, v)$ are effects
- γ_w are statistical parameters
- $\epsilon_i(t, z, l, l')$ is a random term

The objective functions

Behavioral utility function

$$
\begin{aligned}
u_i^{beh}(\gamma, z(l \rightsquigarrow l'), x, v) &= f_i^{beh}(\gamma, z(l \rightsquigarrow l'), x, v) + \epsilon_i(t, z, l, l') \\
&= \sum_{w=1}^{W} \gamma_w s_{iw}^{beh}(x, z(l \rightsquigarrow l'), v) + \epsilon_i(t, z, l, l')
\end{aligned}
$$

where

- $s_w^{beh}(x, z, v)$ are effects
- γ_w are statistical parameters
- $\epsilon_i(t, z, l, l')$ is a random term

The probability that an actor *i* changes his own behavior by one unit is:

$$
p_{ll'}(\gamma, z(l \rightsquigarrow l'), x, v) = \frac{exp\left(f_i^{beh}(\gamma, z(l \rightsquigarrow l'), x, v)\right)}{\sum_{l'' \in \{l+1, l-1, l\}} exp\left(f_i^{beh}(\gamma, z(l \rightsquigarrow l''), x, v)\right)}
$$

p_{ll} is the probability that *i* does not change his behavior

The objective functions

The specification of the behavioral objective function

- *Basic shape effects*

$$s^{beh}_{i\_linear}(x,z,v) = z_i \qquad\qquad s^{beh}_{i\_quadratic}(x,z,v) = z_i^2$$

The basic shape effects must be always included in the model specification

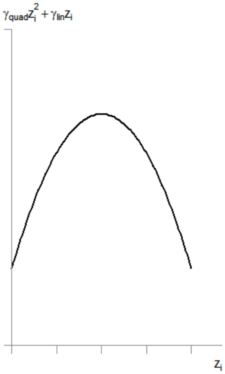
The objective functions

The specification of the behavioral objective function

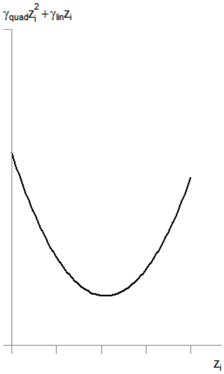
- *Basic shape effects*

$$s^{beh}_{i\_linear}(x,z,v) = z_i \qquad\qquad s^{beh}_{i\_quadratic}(x,z,v) = z_i^2$$

The basic shape effects must be always included in the model specification



negative γ_{quad}



positive γ_{quad}

The objective functions

The specification of the behavioral objective function

- *Classical influence effects*

1. The *average similarity effect* $s^{beh}_{i\_avsim}(x, z, v)$

$$s^{beh}_{i\_avsim}(x, z, v) = \frac{1}{x_{i+}} \sum_{j=1}^{n} x_{ij}(sim_z(ij) - \widehat{sim}_z)$$

where

$$sim_z(ij) = 1 - \frac{|z_i - z_j|}{R_z}$$

R_z is the range of the behavior z and $\widehat{sim}_z$ is the mean similarity value

The objective functions

The specification of the behavioral objective function

- *Classical influence effects*

1. The *average similarity effect* $s^{beh}_{i\_avsim}(x,z,v)$

$$s^{beh}_{i\_avsim}(x,z,v) = \frac{1}{x_{i+}} \sum_{j=1}^{n} x_{ij}(sim_z(ij) - \overline{sim_z})$$

where

$$sim_z(ij) = 1 - \frac{|z_i - z_j|}{R_z}$$

R_z is the range of the behavior z and $\overline{sim_z}$ is the mean similarity value

2. The *total similarity effect* $s^{beh}_{i\_totsim}(x,z,v)$

$$s^{beh}_{i\_totsim}(x,z,v) = \sum_{j=1}^{n} x_{ij}(sim_z(ij) - \overline{sim_z})$$

The objective functions

The specification of the behavioral objective function

- *Position-dependent influence effects*

  Network position could also have an effect on the behavior of dynamics

  1. *outdegree effect*

  $$s^{beh}_{i\_out}(x, z, v) = z_i \sum_{j=1}^{n} x_{ij}$$

  2. *indegree effect*

  $$s^{beh}_{i\_ind}(x, z, v) = z_i \sum_{j=1}^{n} x_{ji}$$

- *Effects of other actor variables.*
  For each actor's attribute a main effect on the behavior can be included in the model

Example

Example data: excerpt from the “Teenage Friends and Lifestyle Study” data set

We will use the SAOM for the co-evolution of networks and behaviors to distinguish influence from selection.

1. Do pupils select friends based on similar smoking behavior?
2. Are pupils influenced by friends to adjust to their smoking behavior?

Dependent variables: friendship networks and smoking behavior

Covariate: gender

Precondition of the analysis

To find out whether it makes sense to analyze the data with a co-evolution model one should check whether:

1. the data are sufficiently informative to allow for identification of effects

$$J = \frac{N_{11}}{N_{11} + N_{01} + N_{10}} > 0.3 \qquad \textit{Jaccard index}$$

```
Tie changes between subsequent observations:
 periods        0 =>  0   0 =>  1   1 =>  0   1 =>  1   Distance Jaccard   Missing
  1 ==>  2      15827       237       240       208      477     0.304    0 (0%)
  2 ==>  3      15839       228       209       236      437     0.351    0 (0%)
```

Precondition of the analysis

2. there is interdependence between network and behavioral variables

$$I = \frac{n \sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\left(\sum_{ij} x_{ij}\right)\left(\sum_{i}(z_i - \overline{z})^2\right)} \qquad \textit{Moran index}$$

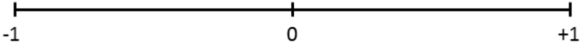
where $\overline{z}$ is the mean of z over all the periods

Precondition of the analysis

2. there is interdependence between network and behavioral variables

$$I = \frac{n \sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\left(\sum_{ij} x_{ij}\right)\left(\sum_{i}(z_i - \overline{z})^2\right)} \qquad \textit{Moran index}$$

where $\overline{z}$ is the mean of z over all the periods

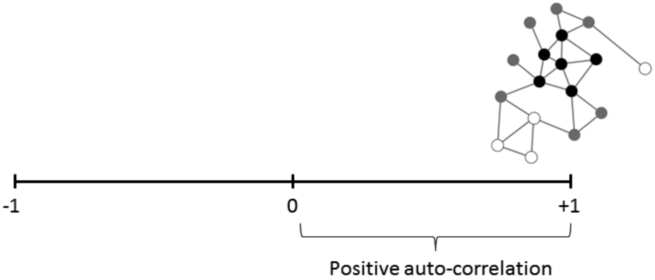


Precondition of the analysis

2. there is interdependence between network and behavioral variables

$$I = \frac{n \sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\left(\sum_{ij} x_{ij}\right)\left(\sum_{i}(z_i - \overline{z})^2\right)}$$ *Moran index*

where $\overline{z}$ is the mean of z over all the periods

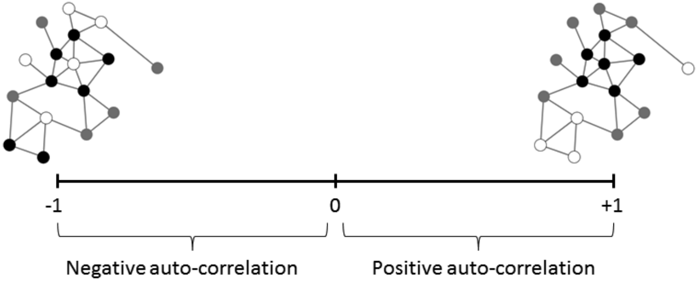


Precondition of the analysis

2. there is interdependence between network and behavioral variables

$$I = \frac{n \sum_{ij} x_{ij}(z_i - \bar{z})(z_j - \bar{z})}{\left(\sum_{ij} x_{ij}\right)\left(\sum_{i}(z_i - \bar{z})^2\right)} \qquad \textit{Moran index}$$

where $\bar{z}$ is the mean of z over all the periods

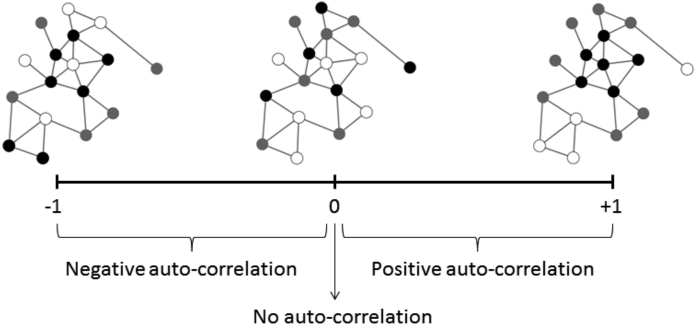


Precondition of the analysis

2. there is interdependence between network and behavioral variables

$$I = \frac{n \sum_{ij} x_{ij}(z_i - \overline{z})(z_j - \overline{z})}{\left(\sum_{ij} x_{ij}\right)\left(\sum_{i}(z_i - \overline{z})^2\right)} \qquad \textit{Moran index}$$

where $\overline{z}$ is the mean of z over all the periods



Precondition of the analysis

The computation of the index *I* for the data leads to

0.244	0.258	0.341

Conclusion:
there is considerable dependence between networks and behaviors and it is reasonable to apply the SAOM

```
moran1 <- nacf(net1,tobacco[,1],lag.max=1,neighborhood.type = "out",
          type="moran",mode="digraph")
moran2 <- nacf(net2,tobacco[,2],,lag.max=1,neighborhood.type = "out",
          type="moran",mode="digraph")
moran3 <- nacf(net3,tobacco[,3],,lag.max=1,neighborhood.type = "out",
          type="moran",mode="digraph")

moranInd <- c(moran1[2],moran2[2],moran3[2])
```

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Network rate parameters:

- about 9 opportunities for a network change in the first period
- about 7 opportunities for a network change in the second period

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Network objective function parameters:

- outdegree parameter: the observed networks have low density
- reciprocity parameter: strong tendency towards reciprocated ties

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Behavioral rate parameters:

- about 4 opportunities for a behavioral change in the first period
- about 4 opportunities for a behavioral change in the second period

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Parameter interpretation: a baseline model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	8.6287	(0.6666)	
constant friendship rate (period 2)	7.2489	(0.5466)	
outdegree (density)	-2.4084	(0.0407)	-59.1268
reciprocity	2.7024	(0.0823)	32.8337
Behavior Dynamics			
rate smokebeh (period 1)	3.8922	(1.9689)	
rate smokebeh (period 2)	4.4813	(2.3679)	
behavior smokebeh linear shap	-3.5464	(0.4394)	-8.0712
behavior smokebeh quadratic shape	2.8464	(0.3628)	7.8447

Behavioral objective function parameters:

attractiveness of different behavioral levels based on the current structure of the network and the behavior of the others

Parameter interpretation: a baseline model

- Smoking behavior: coded with 1 for “no”, 2 for “occasional”, and 3 for “regular” smokers.

Parameter interpretation: a baseline model

- Smoking behavior: coded with 1 for "no", 2 for "occasional", and 3 for "regular" smokers.

- The smoking covariate is centered: $\overline{z} = 1.377$ is the mean of the covariate

$$z_i - \overline{z} = \begin{cases} -0.377 & \textit{for no smokers} \\ 0.623 & \textit{for occasional smokers} \\ 1.623 & \textit{for regular smokers} \end{cases}$$

Parameter interpretation: a baseline model

- Smoking behavior: coded with 1 for "no", 2 for "occasional", and 3 for "regular" smokers.

- The smoking covariate is centered: $\bar{z} = 1.377$ is the mean of the covariate

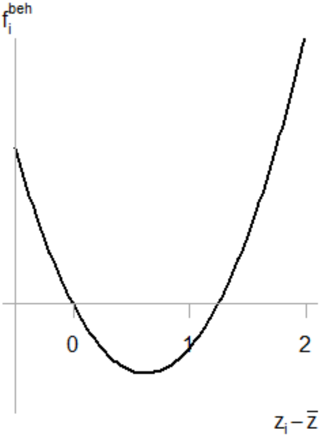
$$z_i - \bar{z} = \begin{cases} -0.377 & \textit{for no smokers} \\ 0.623 & \textit{for occasional smokers} \\ 1.623 & \textit{for regular smokers} \end{cases}$$

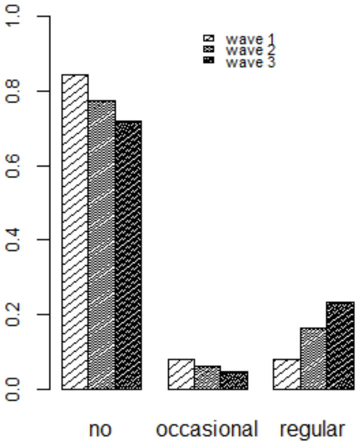
- The contribution to the behavioral objective function is

$$\gamma_{linear}(z_i - \bar{z}) + \gamma_{quadratic}(z_i - \bar{z})^2 =$$

$$= -3.5464(z_i - \bar{z}) + 2.8464(z_i - \bar{z})^2$$

Parameter interpretation: a baseline model





U-shaped changes in the behavior are drawn to the extreme of the range

A more complex model

The baseline model does not provide any information about selection and influence processes:

- the network dynamics are explained by the preference towards creating and reciprocating ties
- the behavior dynamics are described only by the distribution of the behavior in the population

A more complex model

The baseline model does not provide any information about selection and influence processes:

- the network dynamics are explained by the preference towards creating and reciprocating ties
- the behavior dynamics are described only by the distribution of the behavior in the population

If we want to distinguish selection from influence we should include in the objective functions specification:

- the effects that capture the dependence of social network dynamics on actor's characteristic
- the effects that capture the dependence of behavior dynamics on social network

A more complex model

Effects for the dependence of network dynamics on actor's characteristic

- pupils prefer to establish friendship relations with others that are similar to themselves $\rightarrow$ covariate similarity

A more complex model

Effects for the dependence of network dynamics on actor's characteristic

- pupils prefer to establish friendship relations with others that are similar to themselves $\rightarrow$ covariate similarity

This effect must be controlled for the sender and receiver effects of the covariate.

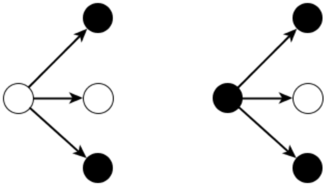
- Covariate ego effect

- Covariate alter effect

A more complex model

Effects for the dependence of behavior dynamics on network

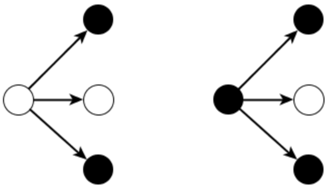
- pupils tend to adjust their smoking behavior according to the behaviors of their friends $\rightarrow$ average similarity effect



A more complex model

Effects for the dependence of behavior dynamics on network

- pupils tend to adjust their smoking behavior according to the behaviors of their friends $\rightarrow$ average similarity effect



This effect must be controlled for the indegree and the outdegree effects

- Indegree effect

- Outdegree effect

A more complex model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	10.7166	(1.4036)	
constant friendship rate (period 2)	9.0005	(0.7709)	
outdegree (density)	-2.8435	(0.0572)	-49.6776
reciprocity	1.9683	(0.0933)	21.1077
transitive triplets	0.4447	(0.0322)	13.7964
sex ego	0.1612	(0.1206)	1.3368
sex alter	-0.1476	(0.1064)	-1.3871
sex similarity	0.9104	(0.0882)	10.3244
smoke ego	0.0665	(0.0846)	0.7857
smoke alter	0.1121	(0.0761)	1.4719
smokebeh similarity	0.5114	(0.1735)	2.9479

A more complex model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	10.7166	(1.4036)	
constant friendship rate (period 2)	9.0005	(0.7709)	
outdegree (density)	-2.8435	(0.0572)	-49.6776
reciprocity	1.9683	(0.0933)	21.1077
transitive triplets	0.4447	(0.0322)	13.7964
sex ego	0.1612	(0.1206)	1.3368
sex alter	-0.1476	(0.1064)	-1.3871
sex similarity	0.9104	(0.0882)	10.3244
smoke ego	0.0665	(0.0846)	0.7857
smoke alter	0.1121	(0.0761)	1.4719
smokebeh similarity	0.5114	(0.1735)	2.9479

Network objective function parameters:

tendency towards reciprocity, transitivity and homophily with respect to gender

A more complex model

	Estimates	s.e.	t-score
Network Dynamics			
constant friendship rate (period 1)	10.7166	(1.4036)	
constant friendship rate (period 2)	9.0005	(0.7709)	
outdegree (density)	-2.8435	(0.0572)	-49.6776
reciprocity	1.9683	(0.0933)	21.1077
transitive triplets	0.4447	(0.0322)	13.7964
sex ego	0.1612	(0.1206)	1.3368
sex alter	-0.1476	(0.1064)	-1.3871
sex similarity	0.9104	(0.0882)	10.3244
smoke ego	0.0665	(0.0846)	0.7857
smoke alter	0.1121	(0.0761)	1.4719
smokebeh similarity	0.5114	(0.1735)	2.9479

Network objective function parameters:

pupils selected others with similar smoking behavior as friends

→ evidence for selection process

A more complex model

The contribution to the network objective function is given by:

$$\beta_{ego}(z_i - \overline{z}) + \beta_{alter}(z_j - \overline{z}) + \beta_{same}\left(1 - \frac{|z_i - z_j|}{R_z} - sim_z\right) =$$

$$= 0.0665(z_i - 1.377) + 0.1121(z_j - 1.377) + 0.5114(1 - \frac{|z_i - z_j|}{R_z} - 0.7415)$$

z_i/z_j	no	occasional	regular
no	0.0648	-0.0787	-0.2223
occasional	-0.1243	0.2435	0.0999
regular	-0.3135	0.0543	0.4221

- preference for similar alters
- this tendency is strongest for high values on smoking behavior

A more complex model

	Estimates	s.e.	t-score
Behavior Dynamics			
rate smokebeh (period 1)	3.9041	(1.7402)	
rate smokebeh (period 2)	3.8059	(1.4323)	
behavior smokebeh linear shape	-3.3573	(0.5678)	-5.9129
behavior smokebeh quadratic shape	2.8406	(0.4125)	6.8864
behavior smokebeh indegree	0.1711	(0.1812)	0.9444
behavior smokebeh outdegree	0.0128	(0.1926)	0.0662
behavior smokebeh average similarity	3.4361	(1.4170)	2.4250

Behavioral objective function parameters:

U-shaped distribution of the smoking behavior

A more complex model

	Estimates	s.e.	t-score
Behavior Dynamics			
rate smokebeh (period 1)	3.9041	(1.7402)	
rate smokebeh (period 2)	3.8059	(1.4323)	
behavior smokebeh linear shape	-3.3573	(0.5678)	-5.9129
behavior smokebeh quadratic shape	2.8406	(0.4125)	6.8864
behavior smokebeh indegree	0.1711	(0.1812)	0.9444
behavior smokebeh outdegree	0.0128	(0.1926)	0.0662
behavior smokebeh average similarity	3.4361	(1.4170)	2.4250

Behavioral objective function parameters:

indegree and outdegree effects are not significant

A more complex model

	Estimates	s.e.	t-score
Behavior Dynamics			
rate smokebeh (period 1)	3.9041	(1.7402)	
rate smokebeh (period 2)	3.8059	(1.4323)	
behavior smokebeh linear shape	-3.3573	(0.5678)	-5.9129
behavior smokebeh quadratic shape	2.8406	(0.4125)	6.8864
behavior smokebeh indegree	0.1711	(0.1812)	0.9444
behavior smokebeh outdegree	0.0128	(0.1926)	0.0662
behavior smokebeh average similarity	3.4361	(1.4170)	2.4250

Behavioral objective function parameters:

pupils are influenced by the smoking behavior of the others

$\rightarrow$ evidence for influence process

A more complex model

The contribution to the behavioral objective function is given by:

$$\gamma_{linear}(z_i - \overline{z}) + \gamma_{quadratic}(z_i - \overline{z})^2 + \gamma_{avsim}\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - sim_z) =$$

$$= -3.3573(z_i - \overline{z}) + 2.8406(z_i - \overline{z})^2 + 3.4361\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - 0.7415)$$

where $sim_z(ij) = 1 - \frac{|z_i - z_j|}{R_Z} = 1$

Example

a) i adjusts his behavior to "no-smoker" when all of his friends are no-smokers

$$sim_z(ij) = 1 - \frac{|1-1|}{2} = 1$$

$$-3.3573(1 - 1.377) + 2.8406(1 - 1.377)^2 + 3.4361(1 - 0.7415) = 2.56$$

A more complex model

The contribution to the behavioral objective function is given by:

$$\gamma_{linear}(z_i - \overline{z}) + \gamma_{quadratic}(z_i - \overline{z})^2 + \gamma_{avsim}\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - sim_z) =$$

$$= -3.3573(z_i - \overline{z}) + 2.8406(z_i - \overline{z})^2 + 3.4361\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - 0.7415)$$

where $sim_z(ij) = 1 - \frac{|z_i - z_j|}{R_Z} = 1$

Example

b) *i* adjusts his behavior to "no-smoker" when all of his friends are occasional smokers

$$sim_z(ij) = 1 - \frac{|1-2|}{2} = 0.5$$

$$-3.3573(1 - 1.377) + 2.8406(1 - 1.377)^2 + 3.4361(0.5 - 0.7415) = 0.84$$

A more complex model

The contribution to the behavioral objective function is given by:

$$\gamma_{linear}(z_i - \overline{z}) + \gamma_{quadratic}(z_i - \overline{z})^2 + \gamma_{avsim}\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - sim_z) =$$

$$= -3.3573(z_i - \overline{z}) + 2.8406(z_i - \overline{z})^2 + 3.4361\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - 0.7415)$$

where $sim_z(ij) = 1 - \frac{|z_i - z_j|}{R_Z} = 1$

Example

b) i adjusts his behavior to "no-smoker" when all of his friends are regular smokers

$$sim_z(ij) = 1 - \frac{|1-3|}{2} = 0$$

$$-3.3573(1 - 1.377) + 2.8406(1 - 1.377)^2 + 3.4361(0 - 0.7415) = -0.88$$

A more complex model

The contribution to the behavioral objective function is given by:

$$\gamma_{linear}(z_i - \overline{z}) + \gamma_{quadratic}(z_i - \overline{z})^2 + \gamma_{avsim}\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - sim_z) =$$

$$= -3.3573_{linear}(z_i - \overline{z}) + 2.8406_{quadratic}(z_i - \overline{z})^2 + 3.4361\frac{1}{x_{i+}}\sum_{j=1}^{n} x_{ij}(sim_z(ij) - 0.7415)$$

z_j / z_i	no	occasional	regular
no	2.56	-1.82	-0.51
occasional	0.84	-0.10	1.20
regular	-0.88	-1.82	2.92

- the focal actor prefers to have the same behavior as all these friends (except for the occasional smokers)

- friends do not smoke at all: the preference toward imitating their behavior is less strong

Simulating the co-evolution of networks and behavior

Algorithm 5: Co-evolution of networks and behavior

Input: $x(t_0)$, $z(t_0)$, λ^{net}, λ^{beh}, β, γ, n
Output: $x^{sim}(t_1)$, $z^{sim}(t_1)$
$t \leftarrow 0$; $x \leftarrow x(t_0)$; $z \leftarrow z(t_0)$
while *condition=TRUE* **do**
    $T^{net} \sim Exp(\lambda^{net}), \quad T^{beh} \sim Exp(\lambda^{beh})$
    if $\min\{T^{net}, T^{beh}\} = T^{net}$ **then**
        $i \sim Uniform(1, \ldots, n), \quad j \sim p_{ij}$
        if $i \neq j$ **then**
            $x \leftarrow x(i \rightsquigarrow j)$
        else
            $x \leftarrow x$
        $t \leftarrow t + T^{net}$
    else
        $i \sim Uniform(1, \ldots, n), \quad l' \sim p_{ll'}$
        if $l \neq l'$ **then**
            $z \leftarrow z(l \rightsquigarrow l')$
        else
            $z \leftarrow z$
        $t \leftarrow t + T^{beh}$
$x^{sim}(t_1) \leftarrow x$
$z^{sim}(t_1) \leftarrow z$
return $x^{sim}(t_1)$, $z^{sim}(t_1)$

Simulating the co-evolution of networks and behavior

1. *Unconditional* simulation:

   simulation carries on until a predetermined time length has elapsed (usually until $t = 1$).

Simulating the co-evolution of networks and behavior

1. *Unconditional* simulation:

   simulation carries on until a predetermined time length has elapsed (usually until $t = 1$).

2. *Conditional* simulation on the observed number of changes:

   - simulation runs on until

$$\sum_{\substack{i,j=1 \\ i\neq j}}^{n} \left| X_{ij}^{obs}(t_1) - X_{ij}(t_0) \right| = \sum_{i,j=1}^{n} \left| X_{ij}^{sim}(t_1) - X_{ij}(t_0) \right|$$

Simulating the co-evolution of networks and behavior

1. *Unconditional* simulation:

   simulation carries on until a predetermined time length has elapsed (usually until $t = 1$).

2. *Conditional* simulation on the observed number of changes:

   - simulation runs on until

$$\sum_{\substack{i,j=1 \\ i \neq j}}^{n} \left| X_{ij}^{obs}(t_1) - X_{ij}(t_0) \right| = \sum_{i,j=1}^{n} \left| X_{ij}^{sim}(t_1) - X_{ij}(t_0) \right|$$

   - simulation runs on until

$$\sum_{i=1}^{n} \left| z_i^{obs}(t_1) - z_i(t_0) \right| = \sum_{i=1}^{n} \left| z_i^{sim}(t_1) - z_i(t_0) \right|$$

The parameter estimation (MoM)

Aim: estimate the parameter θ for the co-evolution model

- $M-1$ rate parameters for the network rate function

$$\lambda_1^{net},\ \ldots,\ \lambda_{M-1}^{net}$$

- $M-1$ rate parameters for the behavior rate function

$$\lambda_1^{beh},\ \ldots,\ \lambda_{M-1}^{beh}$$

- K parameters for the network objective function

$$\sum_{k=1}^{K} \beta_k s_{ik}^{net}(x, z, v)$$

- W parameters for the behavior objective function

$$\sum_{w=1}^{W} \gamma_w s_{iw}^{beh}(x, z(l \rightsquigarrow l'), v)$$

The parameter estimation (MoM)

Aim: estimate the $2(M-1)+K+W$-dimensional parameter θ using the MoM

The parameter estimation (MoM)

Aim: estimate the $2(M-1)+K+W$-dimensional parameter θ using the MoM

In practice:

1. find $2(M-1)+K+W$ statistics
2. set the theoretical expected value of each statistic equal to its sample counterpart
3. solve the resulting system of equations

   $$E_\theta[S-s]=0$$

   with respect to θ

The parameter estimation (MoM)

Statistics:

- Network rate parameters for the period m

$$s^{net}_{\lambda_m}(X(t_m), X(t_{m-1}) | X(t_{m-1}) = x(t_{m-1})) = \sum_{i,j=1}^{n} \left| X_{ij}(t_m) - X_{ij}(t_{m-1}) \right|$$

- Behavior rate parameters for the period m

$$s^{beh}_{\lambda_m}(Z(t_m), Z(t_{m-1}) | Z(t_{m-1}) = z(t_{m-1})) = \sum_{i=1}^{n} |Z_i(t_m) - Z_i(t_{m-1})|$$

The parameter estimation (MoM)

Statistics:

- Network objective function effects

$$\sum_{m=1}^{M-1} s_{mk}^{net}((X,Z,V)(t_m)|(x,z,v)(t_{m-1})) = \sum_{m=1}^{M-1} s_{mk}^{net}((X,Z,V)(t_m),(X,Z,V)(t_{m-1}))$$

- Behavior objective function effects

$$\sum_{m=1}^{M-1} s_{mw}^{beh}((X,Z,V)(t_m)|(x,z,v)(t_{m-1})) = \sum_{m=1}^{M-1} s_{mw}^{beh}((X,Z,V)(t_m),(X,Z,V)(t_{m-1}))$$

The parameter estimation (MoM)

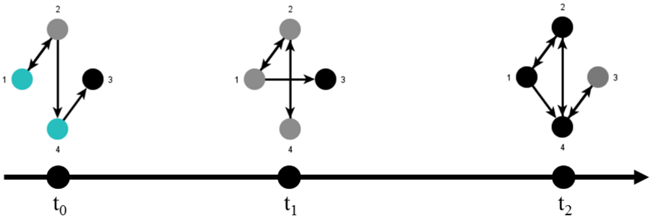
Consequently the MoM estimator for θ is provided by the solution of the system of equations:

$$
\begin{cases}
E_\theta \left[s_{\lambda_m} (X(t_m), X(t_{m-1}) | X(t_{m-1}) = x(t_{m-1})) \right] = s_{\lambda_m}(x(t_m), x(t_{m-1})) \\
\\
E_\theta \left[s_{\lambda_m} (Z(t_m), Z(t_{m-1}) | Z(t_{m-1}) = z(t_{m-1})) \right] = s_{\lambda_m}(z(t_m), z(t_{m-1})) \\
\\
E_\theta \left[\sum_{m=1}^{M-1} s_{mk}^{net} ((X, Z, V)(t_m) | (x, z, v)(t_{m-1})) \right] = \sum_{m=1}^{M-1} s_{mk}^{net}((x, z, v)(t_m), (x, z, v)(t_{m-1})) \\
\\
E_\theta \left[\sum_{m=1}^{M-1} s_{mw}^{beh} ((X, Z, V)(t_m) | (x, z, v)(t_{m-1})) \right] = \sum_{m=1}^{M-1} s_{mw}^{beh}((x, z, v)(t_m), (x, z, v)(t_{m-1}))
\end{cases}
$$

The parameter estimation (MoM)

Example

Let us assume to have observed a network at $M = 3$ time points



t_0:

	1	2	3	4	beh
1	-	1	0	0	0
2	1	-	0	1	1
3	0	0	-	0	2
4	0	0	1	-	0

t_1:

	1	2	3	4	beh
1	-	1	1	0	1
2	1	-	0	1	1
3	0	0	-	0	2
4	0	1	0	-	1

t_2:

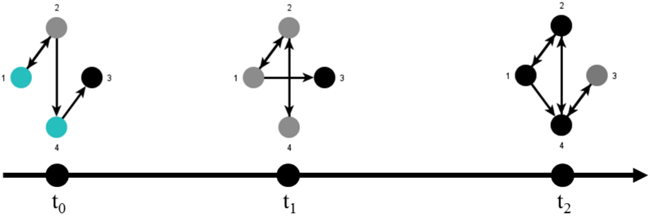
	1	2	3	4	beh
1	-	1	0	1	2
2	1	-	0	1	2
3	0	0	-	1	1
4	0	1	1	-	2

We want to model the network evolution according to the outdegree, the reciprocity, the linear shape and the quadratic shape effects

The parameter estimation (MoM)

Example

Let us assume to have observed a network at $M = 3$ time points



t_0

	1	2	3	4	beh
1	-	1	0	0	0
2	1	-	0	1	1
3	0	0	-	0	2
4	0	0	1	-	0

t_1

	1	2	3	4	beh
1	-	1	1	0	1
2	1	-	0	1	1
3	0	0	-	0	2
4	0	1	0	-	1

t_2

	1	2	3	4	beh
1	-	1	0	1	2
2	1	-	0	1	2
3	0	0	-	1	1
4	0	1	1	-	2

We want to model the network evolution according to the outdegree, the reciprocity, the linear shape and the quadratic shape effects

$$\theta = (\lambda_1^{net}, \lambda_2^{net}, \lambda_1^{beh}, \lambda_2^{beh}, \beta_{out}, \beta_{rec}, \gamma_{linear}, \gamma_{quadratic})$$

The parameter estimation (MoM)

Example

Statistics for the network evolution:

$$s_{\lambda_1^{net}}(X(t_1), X(t_0) | X(t_0) = x(t_0)) = \sum_{i,j=1}^{4} \left| X_{ij}(t_1) - X_{ij}(t_0) \right|$$

$$s_{\lambda_2^{net}}(X(t_2), X(t_1) | X(t_1) = x(t_1)) = \sum_{i,j=1}^{4} \left| X_{ij}(t_2) - X_{ij}(t_1) \right|$$

$$\sum_{m=1}^{M-1} s_{out}\left(X(t_m) | X(t_{m-1}) = x(t_{m-1})\right) = \sum_{m=1}^{2} \sum_{i,j=1}^{4} X_{ij}(t_m)$$

$$\sum_{m=1}^{M-1} s_{rec}\left(X(t_m) | X(t_{m-1}) = x(t_{m-1})\right) = \sum_{m=1}^{2} \sum_{i,j=1}^{4} X_{ij}(t_m) X_{ji}(t_m)$$

The parameter estimation (MoM)

Example

Statistics for the behavior evolution:

$$s_{\lambda_1^{beh}}(Z(t_1), Z(t_0)|Z(t_0) = z(t_0)) = \sum_{i=1}^{4} |Z_i(t_1) - Z_i(t_0)|$$

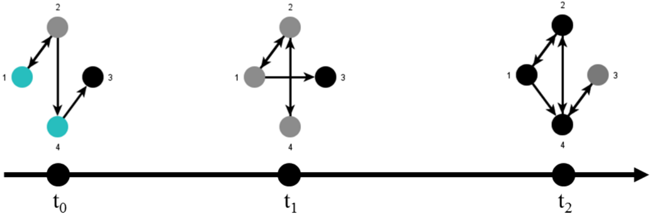
$$s_{\lambda_2^{beh}}(Z(t_2), Z(t_1)|Z(t_1) = z(t_1)) = \sum_{i=1}^{4} |Z_i(t_2) - Z_i(t_1)|$$

$$\sum_{m=1}^{M-1} s_{linear}\,(Z(t_m)|Z(t_{m-1}) = z(t_{m-1})) = \sum_{m=1}^{2} \sum_{i=1}^{4} z_i(t_m)$$

$$\sum_{m=1}^{M-1} s_{quadratic}\,(Z(t_m)|Z(t_{m-1}) = z(t_{m-1})) = \sum_{m=1}^{2} \sum_{i=1}^{4} z_i^2(t_m)$$

The parameter estimation (MoM)

Example



t_0

	1	2	3	4	beh
1	-	1	0	0	0
2	1	-	0	1	1
3	0	0	-	0	2
4	0	0	1	-	0

t_1

	1	2	3	4	beh
1	-	1	1	0	1
2	1	-	0	1	1
3	0	0	-	0	2
4	0	1	0	-	1

t_2

	1	2	3	4	beh
1	-	1	0	1	2
2	1	-	0	1	2
3	0	0	-	1	1
4	0	1	1	-	2

$s_{\lambda_1^{net}} = 3$ $\quad$ $s_{\lambda_2^{net}} = 4$

$s_{\lambda_1^{beh}} = 2$ $\quad$ $s_{\lambda_2^{beh}} = 4$

$s_{out} = 5 + 7 = 12$ $\quad$ $s_{rec} = 4 + 6 = 10$

$s_{linear} = 5 + 7 = 12$ $\quad$ $s_{quadratic} = 7 + 13 = 20$

The parameter estimation (MoM)

Example

We look for the value of θ that satisfies the system:

$$\begin{cases} E_\theta\left[S_{\lambda_1^{net}}\right] = 3 \\ E_\theta\left[S_{\lambda_2^{net}}\right] = 4 \\ E_\theta\left[S_{\lambda_1^{beh}}\right] = 2 \\ E_\theta\left[S_{\lambda_2^{beh}}\right] = 4 \\ E_\theta[S_{out}] = 12 \\ E_\theta[S_{rec}] = 10 \\ E_\theta[S_{linear}] = 12 \\ E_\theta[S_{quadratic}] = 20 \end{cases}$$

The parameter estimation (MoM)

In a more compact notation, we look for the value of θ that satisfies the system:

$$E_\theta[S - s] = 0$$

but we know that we cannot solve it analytically.
We can use the Robbins-Monro algorithm:

Phase 1: provide the initial value for θ and for D

Phase 2: updating the value of θ via the RM step:

$$\widehat{\theta}_{i+1} = \widehat{\theta}_i - a_i D^{-1} \left(E_{\theta_i}[S] - s \right)$$

Phase 3: estimate the s.e. of the estimate

Outline

Creating and deleting ties

Terminating a tie is not just the opposite of creating a tie

Example

- the loss in terminating a tie is greater than the reward in creating one
- transitivity plays an important role especially in creating ties

This is modeled by adding to the objective function one of the two components:

1. the *creation function*
2. the *endowment function*

The creation function

Models the gain in satisfaction incurred when a network tie is created:

$$c_i(\delta, x) = \sum_k \delta_k s_{ik}(x)$$

where

- δ_k are statistical parameters
- $s_{ik}(x)$ are the effects

The utility function for an actor *i* when he creates a new tie is provided by:

$$u_i(x) = f_i(\beta, x) + c_i(\delta, x) + \epsilon_i(t, x, j)$$

The creation function is zero for the dissolution of ties

The endowment function

Models the loss in satisfaction incurred when a network tie is deleted

$$e_i(\eta, x) = \sum_k \eta_k s_{ik}(x)$$

where

- η_k are statistical parameters
- $s_{ik}(x)$ are the effects

The utility function for an actor i when he deletes a tie is provided by:

$$u_i(x) = f_i(\beta, x) + e_i(\eta, x) + \epsilon_i(t, x, j)$$

The endowment function is zero for the creation of ties

Creating and deleting ties - Remarks

- creation and deletion functions must not be included when ties mainly are created or terminated

- it could also happen that increasing a behavior is not the same as decreasing a behavior. Thus, there are also:

  1. the *creation behavior function*
  2. the *endowment behavior function*

  but their usage is still under investigation

Creating and deleting ties

Example

Example data: excerpt from the "Teenage Friends and Lifestyle Study" data set

We estimate the SAOM for investing the evolution of friendship networks according to:

- outdegree
- reciprocity
- transitivity
- reciprocity for the endowment function

```
myeff <- includeEffects(myeff,transTrip)
myeff <- includeEffects(myeff,recip,type="endow")
myeff
mymodel <- sienaModelCreate(useStdInits = FALSE, projname = 'tfls')
model1 <- siena07(mymodel, data = mydata, effects=myeff,useCluster=TRUE,
nbrNodes=2, initC=TRUE,clusterString=rep("localhost", 2))
```

Creating and deleting ties

Example

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	6.70	0.73	
Rate parameter period 2	5.81	0.58	
Other parameters:			
outdegree	-2.58	0.05	-51.62
reciprocity	3.23	0.29	11.15
reciprocity (endow)	-2.23	0.58	-3.85
transitive triplets	0.44	0.03	14.55

The utility function for an actor i when he deletes a tie is provided by:

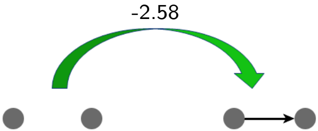
$$\begin{aligned}
u_i(x) &= f_i(\beta, x) + e_i(\eta, x) + \epsilon_i(t, x, j) = \\
&= \beta_{out} s_{i_out}(x) + \beta_{rec} s_{i_rec}(x) + \beta_{trans} s_{i_trans}(x) + \eta_{rec} s_{i_rec}(x) \\
&= -2.58 s_{i_out}(x) + 3.23 s_{i_rec}(x) + 0.44 s_{i_trans}(x) - 2.23 s_{i_rec}(x)
\end{aligned}$$

Creating and deleting ties

Example

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.44	0.73	
Rate parameter period 2	7.09	0.58	
Other parameters:			
outdegree	-2.58	0.05	-51.62
reciprocity	3.23	0.29	11.15
reciprocity (endow)	-2.23	0.58	-3.85
transitive triplets	0.44	0.03	14.55

Ties formation/deletion

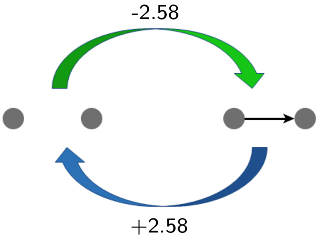


Creating and deleting ties

Example

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.44	0.73	
Rate parameter period 2	7.09	0.58	
Other parameters:			
outdegree	-2.58	0.05	-51.62
reciprocity	3.23	0.29	11.15
reciprocity (endow)	-2.23	0.58	-3.85
transitive triplets	0.44	0.03	14.55

Ties formation/deletion

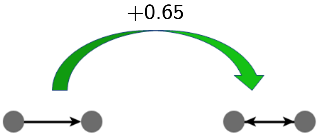


Creating and deleting ties

Example

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.44	0.73	
Rate parameter period 2	7.09	0.58	
Other parameters:			
outdegree	-2.58	0.05	-51.62
reciprocity	3.23	0.29	11.15
reciprocity (endow)	-2.23	0.58	-3.85
transitive triplets	0.44	0.03	14.55

Reciprocation/ending reciprocation

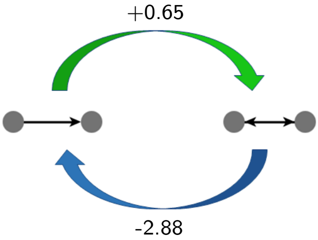


Creating and deleting ties

Example

	Estimates	s.e.	t-score
Rate parameters:			
Rate parameter period 1	8.44	0.73	
Rate parameter period 2	7.09	0.58	
Other parameters:			
outdegree	-2.58	0.05	-51.62
reciprocity	3.23	0.29	11.15
reciprocity (endow)	-2.23	0.58	-3.85
transitive triplets	0.44	0.03	14.55

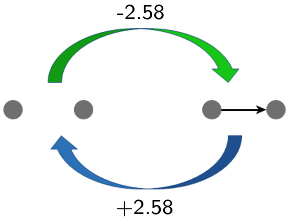
Reciprocation/ending reciprocation



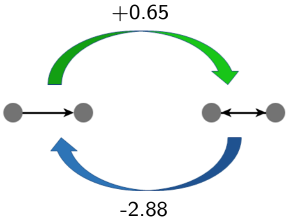
Creating and deleting ties

Example

Ties formation/deletion



Reciprocation/ending reciprocation



Conclusions:

1. formation of reciprocal ties is more rewarding than the formation of a non-reciprocal tie

2. dissolution of reciprocal ties is more costly than the dissolution of a non-reciprocal tie and the creation of a reciprocal tie

Non-directed relations

For directed relation we assumed that:

1. an actor gets the opportunity to make a change
2. he decided for the change that assures him the highest payoff

Are this assumptions still reliable when we consider undirected relations such as: collaboration, trade, strategic alliance?

Non-directed relations

For directed relation we assumed that:

1. an actor gets the opportunity to make a change
2. he decided for the change that assures him the highest payoff

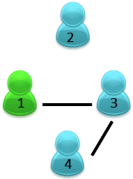
Are this assumptions still reliable when we consider undirected relations such as: collaboration, trade, strategic alliance?

- Yes, if one actor (*dictator*) can impose a decision about a tie to another
- No, if there is coordination or negotiation about a tie

Non-directed relations

1. Dictatorial choice: *i* chooses his action and imposes his decision to *j*

Actor 1 gets the opportunity to change



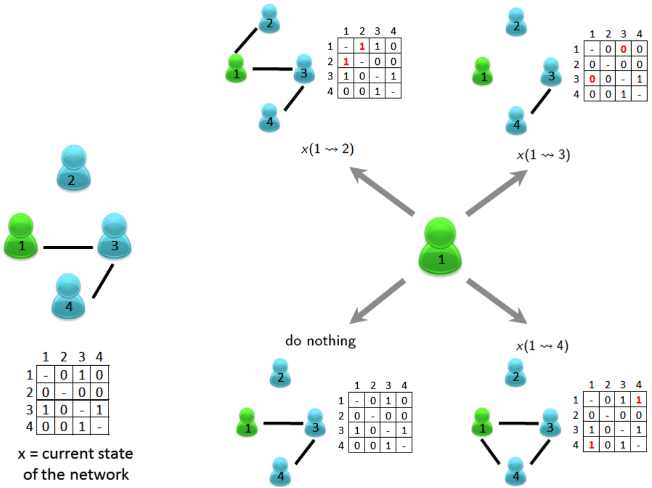
	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

x = current state of the network

Non-directed relations

1. Dictatorial choice: *i* chooses his action and imposes his decision to *j*

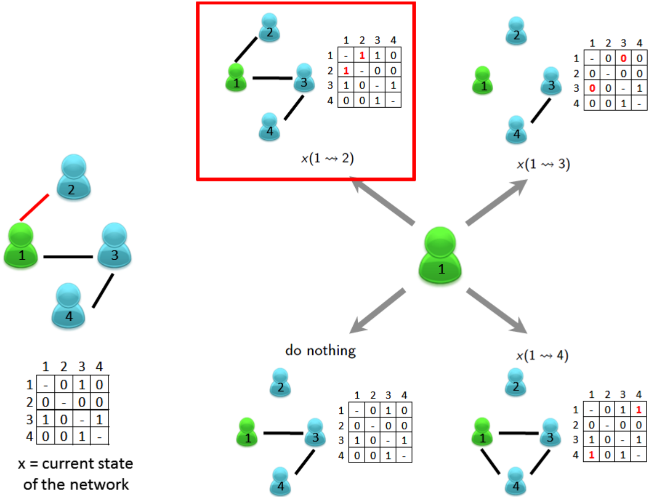
Actor 1 evaluates the alternatives and the corresponding objective functions



Non-directed relations

1. Dictatorial choice: i chooses his action and imposes his decision to j

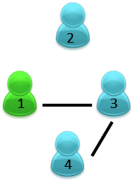
E.g. actor 1 imposes his choice to actor 1



Non-directed relations

2. Mutual agreement: both actors must agree

Actor 1 gets the opportunity to change



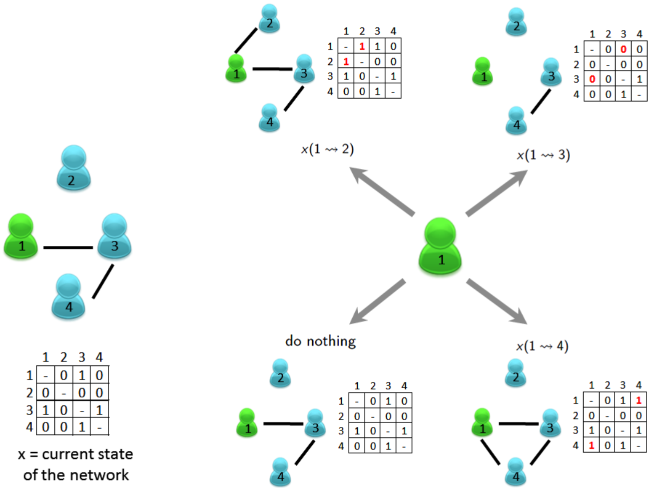
	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

x = current state of the network

Non-directed relations

2. Mutual agreement: both actors must agree

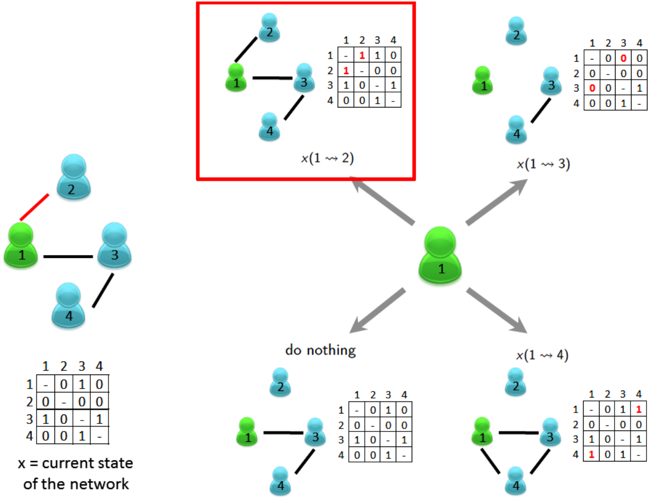
Actor 1 evaluates the alternatives and the corresponding objective functions



Non-directed relations

2. Mutual agreement: both actors must agree

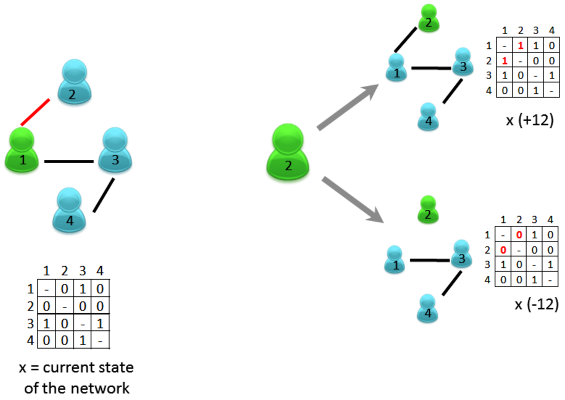
Actor 1 suggests to modify the tie towards actor 2



Non-directed relations

2. Mutual agreement: both actors must agree

Actor 2 evaluates the proposal of actor 1



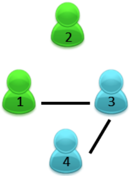
and accepts it with probability

$$P(2\ accepts\ tie\ proposal) = \frac{exp(f_2(x^{+12}))}{exp(f_2(x^{+12})) + exp(f_2(x^{-12}))}$$

Non-directed relations - Tie-based approach

A couple (i,j) of actors is selected with rate λ_{ij} and gets the opportunity to revise the tie among them

1. Dictatorial choice: one actor can impose the decision (e.g. actor 1)



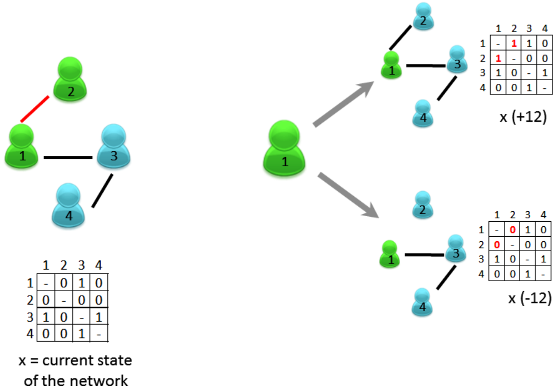
	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

x = current state of the network

Non-directed relations - Tie-based approach

A couple (i, j) of actors is selected with rate λ_{ij} and gets the opportunity to revise the tie among them

1. Dictatorial choice: one actor can impose the decision (e.g. actor 1)



	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

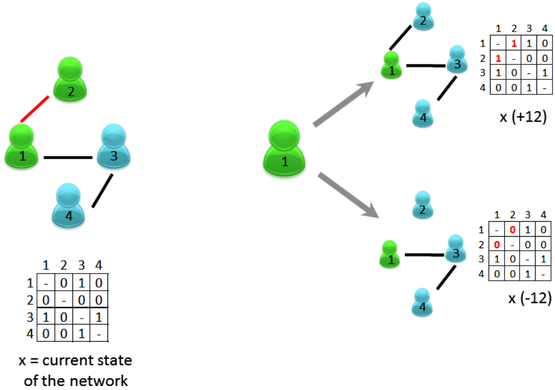
	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	1	0	-	1
4	0	0	1	-

	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

Non-directed relations - Tie-based approach

A couple (i,j) of actors is selected with rate λ_{ij} and gets the opportunity to revise the tie among them

1. Dictatorial choice: one actor can impose the decision (e.g. actor 1)



	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

x = current state of the network

	1	2	3	4
1	-	1	1	0
2	1	-	0	0
3	1	0	-	1
4	0	0	1	-

x (+12)

	1	2	3	4
1	-	0	1	0
2	0	-	0	0
3	1	0	-	1
4	0	0	1	-

x (-12)

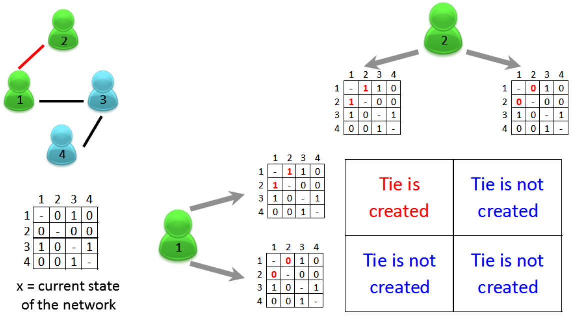
Actor 1 chooses his action with probability

$$P(1 \ imposes \ a \ tie \ on \ 2) = \frac{exp(f_1(x^{+12}))}{exp(f_1(x^{+12})) + exp(f_1(x^{-12}))}$$

Non-directed relations - Tie-based approach

A couple (i,j) of actors is selected with rate λ_{ij} and gets the opportunity to revise the tie among them

2. Mutual agreement: both actors propose a tie



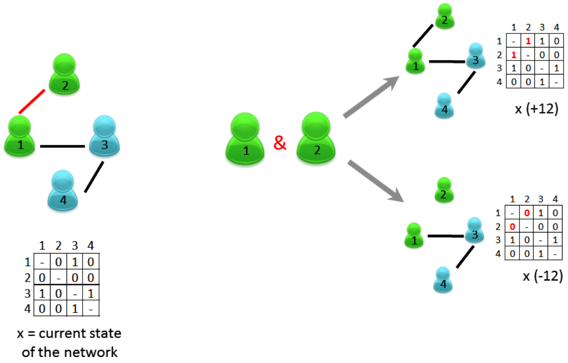
Actor 1 and 2 created a tie with probability

$$P(+12) = \frac{exp(f_1(x^{+12}))}{exp(f_1(x^{+12})) + exp(f_1(x^{-12}))} \frac{exp(f_2(x^{+12}))}{exp(f_2(x^{+12})) + exp(f_2(x^{-12}))}$$

Non-directed relations - Tie-based approach

A couple (i,j) of actors is selected with rate λ_{ij} and gets the opportunity to revise the tie among them

3. Compensatory: the decision is made on the combined interest



Actor 1 and 2 choose their action with probability

$$P(+12) = \frac{exp(f_1(x^{+12}) + f_2(x^{+12}))}{exp(f_1(x^{+12}) + f_2(x^{+12})) + exp(f_1(x^{-12}) + f_2(x^{-12}))}$$

And others...

- Improving the estimation procedures (MLE)
- New estimation procedures (bayesian estimation)
- Goodness of fit of the model
- Model selection
- Time-heterogeneity tests
- Missing data
- Analysis of multiple relations
- ...