Reading Group on Stochastic Modelling

https://matteoiacopini.github.io/stochmodgroup/index.html

Meetings, Discussants and Material

Discussed paper:

A Structural Model of Dense Network Formation

Mele (2017)

Econometrica

Introduction

Literature

Related works

- similar paper [7]
- extensions [14], [3], [11]

Model

- potential games [13]
- network formation models [1]
- ERGM theory [8], [18]

Asymptotics

- many networks asymptotics [17]
- large network asymptotics [9], [15], [12]
  - graph limits [2]
  - large deviations [6]
  - variational methods for the exponential family [20]

Estimation

- ERGM estimation [18], [5]

Introduction

Strategic models of network formation provide a framework to interpret the observed network as the equilibrium of a (potential) game.

Estimation and identification of strategic models is challenging

1) multiple equilibria ⇒ links generate externalities not fully accounted for by agents
2) curse dimensionality ⇒ # network configs grows exponentially with # agents
3) data on single graph ⇒ only one network snapshot is observable

Proposed model of network formation

- combines features from the strategic and random network formation literature
- players' utilities depend on payoffs from direct links and link externalities (e.g., reciprocity, indirect friends, popularity, . . .)
- network formation is dynamic: each period, a player meets another one and decides whether to form a new link, keep an existing link, or do nothing
- process generates a sequence of directed dense graphs

Model of Network Formation

Model of Network Formation

Setup

- n agents, with characteristics $X_i \in \mathbb{R}^A, \ \forall i \in \mathcal{I} := \{1, \ldots, n\}$
- discrete time $t \in \mathbb{N}$
- directed, binary network $G \in \mathcal{G}$, realisations each time g^t

Definition 1 (Individual utility function).

Let $u_{ij}^{\theta_u} = u(X_i, X_j|\theta_u)$, $m_{ij}^{\theta_m} = m(X_i, X_j|\theta_m)$, $v_{ij}^{\theta_v} = v(X_i, X_j|\theta_v)$, $w_{ij}^{\theta_w} = w(X_i, X_j|\theta_w)$ where $\boldsymbol{\theta} = (\theta_u, \theta_m, \theta_v, \theta_w)' \in \mathbb{R}^4$ are parameters.

The **utility of agent** i from network g is given by the sum of four components

$$U_i(g, X|\boldsymbol{\theta}) = \underbrace{\sum_{j=1}^{n} g_{ij} u_{ij}^{\theta_u}}_{\text{direct links}} + \underbrace{\sum_{j=1}^{n} g_{ij} g_{ji} m_{ij}^{\theta_m}}_{\text{mutual links}} + \underbrace{\sum_{j=1}^{n} g_{ij} \sum_{\substack{k=1 \\ k \neq i,j}}^{n} g_{jk} v_{ik}^{\theta_v}}_{\text{indirect links}} + \underbrace{\sum_{j=1}^{n} g_{ij} \sum_{\substack{k=1 \\ k \neq i,j}}^{n} g_{ki} w_{kj}^{\theta_w}}_{\text{popularity}} .$$

▶ "Markovian" only indirect links are valuable and are perfect substitutes (no utility from two-links-away contacts)

Potential Game

Definition 2 (Potential Game).

A game is said to be a **Potential Game** if the incentive of all players to change their strategy (here: link formation choice) can be expressed using a single global function called the **potential function** $\mathcal{Q} : \mathcal{G} \times \mathcal{X} \to \mathbb{R}$ such that:

$$\mathcal{Q}(g_{ij}, g_{-ij}, X) - \mathcal{Q}(g'_{ij}, g_{-ij}, X) = U_i(g_{ij}, g_{-ij}, X) - U_i(g'_{ij}, g_{-ij}, X), \quad \forall i, j \; \forall g_{-ij}$$

Remark 1.

The **potential function** is useful for:

- analyse equilibrium properties of games,
  set of pure Nash equilibria corresponds to the local optima of potential function; existence of profitable deviations performed using the potential, instead of checking each player's possible deviation
- study convergence and finite-time convergence of iterated game towards a Nash equilibrium

Assumption 1. (Preferences)

$$m_{ij}^{\theta_m} = m(X_i, X_j|\theta_m) = m(X_j, X_i|\theta_m) = m_{ji}^{\theta_m} \quad \forall i, j \in \mathcal{I} \times \mathcal{I}$$
$$w_{kj}^{\theta_v} = w(X_k, X_j|\theta_v) = v(X_k, X_j|\theta_v) = v_{kj}^{\theta_v} \quad \forall k, j \in \mathcal{I} \times \mathcal{I}$$

- first is necessary for identification of the utility from indirect links and popularity;
- second makes another agent i internalise the externality she creates.

Proposition 1 (Existence Potential Function).

Under Assumption 1, the deterministic components of the incentives of any player in any state of the network are summarized by a **potential function** $\mathcal{Q} : \mathcal{G} \times \mathcal{X} \to \mathbb{R}$ and the network game is a Potential Game

$$\mathcal{Q}(g, X|\boldsymbol{\theta}) = \sum_{i=1}^{n}\sum_{j=1}^{n} g_{ij} u_{ij}^{\theta_u} + \sum_{i=1}^{n}\sum_{j>i}^{n} g_{ij} g_{ji} m_{ij}^{\theta_m} + \sum_{i=1}^{n}\sum_{\substack{j=1 \\ j\neq i}}^{n}\sum_{\substack{k=1 \\ k\neq i,j}}^{n} g_{ij} g_{jk} v_{ik}^{\theta_v}$$

$\mathcal{Q}$ is an aggregate function summarising: (i) state of network; (ii) deterministic incentives of players in each state.

Reading Group on Stochastic Modelling

A brief overview of:

Incomplete Simultaneous Discrete Response Model with Multiple Equilibria

Tamer (2003)

Review of Economic Studies

In relation to:

A Structural Model of Dense Network Formation

Mele (2017)

Econometrica

Mele (2017) footnote n.17 p.830

> *"The* **second part of the assumption 1 [see below]** *is an identification restriction, that guarantees the model's* **coherency** *in the sense of Tamer (2003)."*

Individual i **values his popularity** effect as much as k **values the indirect link to** j through any "bridging" individual:

$$\underbrace{w_{kj}^{\theta_v} = w(X_k, X_j|\theta_v)}_{\text{indirect link}} = \underbrace{v(X_k, X_j|\theta_v) = v_{kj}^{\theta_v}}_{\text{popularity}} \quad \forall k, j \in \mathcal{I} \times \mathcal{I}$$

indirect link:
utility for $k \neq i,j$
for indirectly linking with j
through any $i \neq k,j$
ⓚ → ⓘ → ⓙ

popularity:
utility for any $i \neq k,j$
for bridging k to j
ⓚ → ⓘ → ⓙ

Under assumption 1.2 any individual $i \in \mathcal{I}$ **internalizes all the externalities generated by his links**:

- The **popularity** component of $U_i(g, \mathbf{X}|\Theta)$ is **equal to** the sum, over all $k \in \mathcal{I} - i$, of the utility of **indirect links of** k **passing through** i, which are the indirect links that can be **influenced** by i's link-formation **decisions**;

Summary of Tamer (2003)

Modelling framework - **2x2 entry-game with perfect information**

two players ($i \in \{-1, 1\}$), action set ($y_i \in \{0, 1\}$) and externalities δ_i. The payoff π_i of player i is defined as:

$$\pi_i := y_i(x_i\beta_i + y_{-i}\delta_i + u_i)$$

Where:

- $\mathbf{y} = (y_{-1}, y_1)$ are response variables;
- $\mathbf{x} = (x_{-1}, x_1) \in \mathcal{R}^d$ are observable exogenous variables;
- $\mathbf{u} = (u_{-1}, u_1)$ are random variables unobservable to the econometricians;
- $\beta = (\beta_{-1}, \beta_1, \delta_{-1}, \delta_1)$ are parameters of interest;

Distinction of model identification issues - **Incoherency Vs Incompleteness**

1− **incoherent model**: hasn't a well-defined reduced form, or, is logically inconsistent. For example:
if externalities δ_{-1} and δ_1 are both negative, the above model gives
$Pr[(0,0)|x] + Pr[(0,1)|x] + Pr[(1,0)|x] + Pr[(1,1)|x] > 1$

2− **incomplete model**: the relationship from input variables (x_is and u_is) to responses (y_is) is a **correspondence** and not a function. For example:
if δ_is are both negative, $\exists$ a non-empty region of **u**'s *support for which the model predicts a non-unique outcome* $(1,0)$ OR $(0,1)$

Contribution and findings of Tamer (2003)

- **Identifies sufficient conditions for parameter point identification (when externalities have same sign);**
- **Develops a technique for semi-parametric maximum (quasi)likelihood (SML) estimation:** by "replacing" $Pr[(y_{-1}, y_1)|x]$ for outcomes (0,1) and/or (1,0), with local approximations of the the empirical relative frequencies of these outcomes as a function of exogenous variables;

Why Assumption 1.2 relevant for identification in Mele (2017)?

1. **externalities are "paired"**: each **indirect-link** effect has a **corresponding popularity effect** with same sign, value and parameter;
2. **number of parameters of the model is reduced**: from 4 $(\theta_u, \theta_m, \theta_w, \theta_v)$ to 3 $(\theta_u, \theta_m, \theta_v)$. Condition necessary for model completeness (?);
3. guarantees that the system of **conditional linking probabilities** implied by the model **generate a proper joint distribution** of the network matrix;
4. can **use** the potential function $\mathcal{Q}$ **to construct the network as a best-response potential game**. Via **sequential** link-formation **decisions** the game converges through an **improvement path** to a **Pure Strategy Nash Equilibrium** network with $Pr = 1$;

Network formation process

Stochastic best-response dynamics, generating a Markov chain of graphs:

- for each t, randomly chosen player i meets j according to meeting technology
- *meeting process* is a stochastic sequence $\mathbf{m} = \{m^1, ..., m^t\}_t$ supported on $\mathcal{I} \times \mathcal{I}$, with realisations $m^t = ij = \{i, j\}$ whose probability is

$$\mathbb{P}(m^t = ij | g^{t-1}, X) = \rho(g^{t-1}, X_i, X_j)$$

Assumption 2. (Meeting process)

The meeting probability between i, j does not depend on the existence of a link between them, and each meeting has a positive probability of occurring, that is

$$\rho(g^{t-1}, X_i, X_j) = \rho(g^{t-1}_{-ij}, X_i, X_j) > 0 \qquad \forall ij$$

▶ guarantees any equilibrium network can be reached with positive probability

Idea: Assumption 2 makes the Markov chain irreducible

▶ identification: if ρ depends on link $g_{ij} \Rightarrow$ prevents closed form likelihood

Players' rules

- conditional on meeting $m^t = ij$, player i updates link g_{ij} to maximise her utility
- existing network g_{-ij}^{t-1} is taken as given
- complete information: everybody known each others' attributes and whole network
- myopia: agents not account for effects of their linking strategy on future evolution of network

Assumption 3. (Idiosyncratic shocks)

Idiosyncratic shock on individual preferences: $\varepsilon_{ij,t} \overset{iid}{\sim} EV_1(\varepsilon_{ij,t})$ Type I extreme value distribution, *iid* among links and across time

$\Rightarrow\ > 0$ proba moving out from any state $\rightarrow$ eliminates absorbing states

- link established if and only if

$$U_i(g_{ij}^t = 1, g_{-ij}^{t-1}, X|\boldsymbol{\theta}) + \varepsilon_{1t} > U_i(g_{ij}^t = 0, g_{-ij}^{t-1}, X|\boldsymbol{\theta}) + \varepsilon_{0t}$$

Process generates a Markov chain of networks:

✓ transition proba determined by: (i) meeting process, (ii) agents' linking choices

✓ irreducible (from Assumption 2), aperiodic (from Assumption 3)

Equilibrium

Remark 2.

Any change in utility for any agent is equivalent to change in potential $\mathcal{Q}$. So, any deviation from Nash (equilibrium) network must decrease the potential.

Thus, the Nash network is a local maximizer of the potential function over the set of networks that differ from the current network for at most one link.

Theorem 1 (Uniqueness and Characterisation of Stationary Equilibrium).

The network formation game, under Assumptions 1–3, converges to a **unique** stationary distribution

$$\pi(g, X|\boldsymbol{\theta}) = \frac{\exp\{\mathcal{Q}(g, X|\boldsymbol{\theta})\}}{\sum_{\omega \in \mathcal{G}} \exp\{\mathcal{Q}(\omega, X|\boldsymbol{\theta})\}} \tag{1}$$

Comments

- existence and uniqueness come from irreducibility and aperiodicity of Markov chain
- closed form stationary $\pi(g, X|\boldsymbol{\theta})$ corresponds to the likelihood of observing a specific network configuration in the long run
- estimation: uniqueness avoids multiple equilibria
  $\Rightarrow$ unique stationary = unique likelihood

$\Rightarrow$ can estimate $\boldsymbol{\theta}$ with only one network, assumed drawn from stationary equilibrium

- $\pi(g, X|\boldsymbol{\theta})$ coincides with likelihood of ERGM (Exponential Random Graph Model), where probability observing a network is proportional to exponential of linear combination of network statistics

Corollary 2.1.

Let Assumptions 1–3 hold. If the utility functions are linear in parameters, the stationary distribution $\pi(g, X|\boldsymbol{\theta})$ describes an ERGM, with $\mathbf{t}(g, X)$ a vector of canonical statistics

$$\pi(g, X|\boldsymbol{\theta}) = \frac{\exp\{\boldsymbol{\theta}'\mathbf{t}(g, X)\}}{\sum_{\omega \in \mathcal{G}} \exp\{\boldsymbol{\theta}'\mathbf{t}(\omega, X)\}} \tag{2}$$

Model without shocks

Proposition 2 (Model Without Shocks: Equilibria and Long Run).

Consider the model without idiosyncratic preference shocks. Under Assumptions 1-2:

(i) there exists at least one pure-strategy Nash equilibrium network.

(ii) the set $\mathcal{NE}(\mathcal{G}, X, U)$ of all pure-strategy Nash equilibria of the network formation game is completely characterized by the local maxima of the potential function:

$$\mathcal{NE}(\mathcal{G}, X, U) = \left\{ g^* : g^* = \arg \max_{g \in \mathcal{N}(g^*)} Q(g, X) \right\}.$$

(iii) any pure-strategy Nash equilibrium is an absorbing state.

(iv) as $t \to \infty$, the network converges to one of the Nash networks with probability 1.

Extensions

Utility functions
☞ possible to include additional utility components, as long as possible to find restrictions on payoffs that guarantee the existence of a potential function

Undirected networks
☞ possible to extend existence results, characterisation of equilibrium, relation with ERGM and asymptotic results to undirected networks

Sparsity
☞ model with negative linking externalities is compatible with a certain degree of sparsity

Estimation and Identification

Estimation and Identification

Likelihood function

$$L(g, X|\boldsymbol{\theta}) = \pi(g, X|\boldsymbol{\theta}) = \frac{\mathcal{Q}(g, X|\boldsymbol{\theta})}{\sum_{\omega \in \mathcal{G}} \mathcal{Q}(\omega, X|\boldsymbol{\theta})} = \frac{\mathcal{Q}(g, X|\boldsymbol{\theta})}{c(\mathcal{G}, X, \boldsymbol{\theta})}$$

whose normalizing constant $c(\mathcal{G}, X, \boldsymbol{\theta})$ is intractable since it sums $2^{n(n-1)}$ terms.

✗ standard ML infeasible
✗ MCMC with standard MH step infeasible (ratio of normalizing constants)

Estimation Algorithm

ERGM literature $\Rightarrow$ approximate $c(\mathcal{G}, X, \boldsymbol{\theta})$ via MCMC (for fixed $\boldsymbol{\theta}_0$)

Algorithm 1 Metropolis-Hastings for Network Simulations

procedure $\text{MH\_NETSIM}(\boldsymbol{\theta}_0, g_0, R)$
  for $r = 1, \ldots, R$ **do**
    1) propose network $g' \sim q_g(g'|g^{(r)})$
    2) accept network g' with probability

$$\alpha(g^{(r)}, g') = \min\left\{1, \frac{\exp\{\mathcal{Q}(g', X|\boldsymbol{\theta}_0)\}}{\exp\{\mathcal{Q}(g^{(r)}, X|\boldsymbol{\theta}_0)\}} \frac{q_g(g^{(r)}|g')}{q_g(g'|g^{(r)})}\right\}$$

  end for
  return sequence of R networks $\{g^{(r)}\}_r$
end procedure

✓ not requires $c(\mathcal{G}, X, \boldsymbol{\theta})$

✗ *slow convergence*
✗ *local sampler* at each iteration, update link g_{ij} according to $\alpha(\cdot, \cdot)$
✗ *degeneracy* problem: large probability mass on few networks

Estimation Algorithm

ERGM literature $\Rightarrow$ approximate $c(\mathcal{G}, X, \boldsymbol{\theta})$ via MCMC (for fixed $\boldsymbol{\theta}_0$)

Algorithm 1 Metropolis-Hastings for Network Simulations

1: **procedure** MH_NETSIM($\boldsymbol{\theta}_0, g_0, R$)
2: **for** $r = 1, \ldots, R$ **do**
3: 1) propose network $g' \sim q_g(g'|g^{(r)})$
4: 2) accept network g' with probability

$$\alpha(g^{(r)}, g') = \min\left\{1, \frac{\exp\{\mathcal{Q}(g', X|\boldsymbol{\theta}_0)\}}{\exp\{\mathcal{Q}(g^{(r)}, X|\boldsymbol{\theta}_0)\}} \frac{q_g(g^{(r)}|g')}{q_g(g'|g^{(r)})}\right\}$$

5: **end for**
6: **return** sequence of R networks $\{g^{(r)}\}_r$
7: **end procedure**

▶ how to choose $q_g(\cdot|g^{(r)})$?

Asymptotic behaviour

Classes of asymptotics for networks

1) *many networks* $\Rightarrow$ same players, growing number of networks
2) *large networks* $\Rightarrow$ growing number of players, same network

- Hp: homogeneous players (i.e. $X_i = X_j, \forall i, j$)
- potential function re-scaled by $n^{\nu(H)}$, with $\nu(H)$ $\#$ players in each utility term
- example **re-scaled likelihood**, with $\mathcal{T}(g) = \alpha t(H_1, g) + \beta t(H_2, g)$ re-scaled potential and $\psi_n = n^{-2} \log(c(\alpha, \beta, \mathcal{G}_n))$ log-normalising constant

$$\pi_n(g|\alpha,\beta) = \frac{\exp\left\{ n^2 \left[\alpha \frac{\sum_{i=1}^n \sum_{j=1}^n g_{ij}}{n^2} + \beta \frac{\sum_{i=1}^n \sum_{j=1}^n \sum_{k\neq i}^n g_{ij} g_{jk}}{n^3} \right] \right\}}{c(\alpha, \beta, \mathcal{G}_n)}$$

$$= \exp\left\{ n^2 \left[\mathcal{T}(g) - \psi_n \right] \right\} \tag{3}$$

Asymptotic behaviour

Theorem 2 (Nonnegative Link Externalities).

Model (3) with nonnegative link externalities $\beta \geq 0$ exhibits the following behaviour

1) asymptotic normalizing constant ψ solves

$$\psi = \lim_{n\to\infty} \psi_n = \max_{\mu\in[0,1]} \left\{ \alpha\mu + \beta\mu^2 - \mu \log(\mu) - (1-\mu)\log(1-\mu) \right\} \qquad (4)$$

2) networks generated by the model are indistinguishable from directed Erdös–Rényi graph with linking probability μ^*, defined as follows:

   (a) if the maximization (4) has a unique solution, then μ^* satisfies $2\beta\mu(1-mu) < 1$ for almost all $\alpha \in \mathbb{R}$ and $\beta \geq 0$, and solves

$$\mu = \frac{\exp\{\alpha + 2\beta\mu\}}{1 + \exp\{\alpha + 2\beta\mu\}} \qquad (5)$$

   (b) if the maximization (4) has two solutions, then μ^* picked randomly from same proba distribution over μ_1^* and μ_2^*, such that $\mu_1^* < 0.5 < \mu_2^*$, and both solve (5) and satisfy $2\beta\mu(1-\mu) < 1$.

Comments on Theorem 2

- consistent estimator of log-normalising constant – analogue of variational representation of the discrete exponential family
- $\beta \geq 0 \Rightarrow$ realisations using (α, β) indistinguishable from those using $(\alpha', 0) = (\log(\mu^*/(1-\mu^*)), 0)$, that is from Erdös-Rényi model

Corollary 2.2.

When $\beta \geq 0$, the **externality cannot be identified**.

Corollary 2.3.

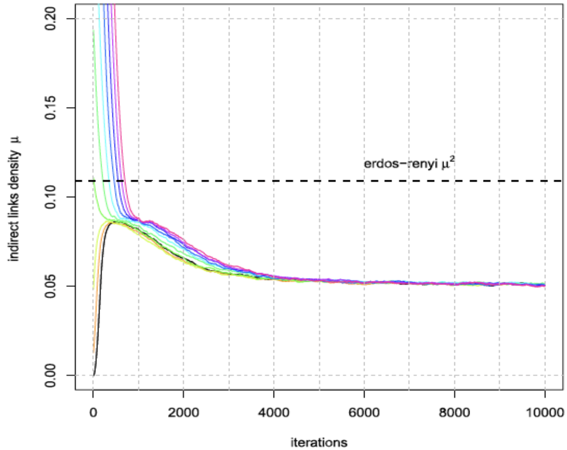
When $\beta \geq 0$, Algorithm 1 is not necessary since Erdös-Rényi graphs can be simulated using Bernoulli draws.

Asymptotic behaviour

Theorem 3 (Negative Link Externalities).

If $\beta < 0$ and sufficiently large in magnitude, model (3) is asymptotically different from a directed Erdös-Rényi model.

✓ sparser graphs than Erdös-Rényi



Asymptotic behaviour

Theorem 3 (Negative Link Externalities).

If $\beta < 0$ and sufficiently large in magnitude, model (3) is asymptotically different from a directed Erdös-Rényi model.

- how much "sufficiently large" magnitude?

- how to know it, since we must estimate β?

Asymptotic behaviour

Consider an additional utility component (cyclic triangles):

$$T(g) = \alpha t(H_1, g) + \beta t(H_2, g) + \gamma t(H_3, g), \qquad t(H_3, g) = \sum_{i=1}^{n} \sum_{j=1}^{n} \sum_{k \neq i}^{n} g_{ij} g_{jk} g_{ki} \quad (6)$$

Theorem 4.

Consider model (6) as $n \to \infty$

1) If $\beta \geq 0$ and $\gamma \geq 0$, then the asymptotic normalising constant ψ solves

$$\psi = \lim_{n \to \infty} \psi_n = \max_{\mu \in [0,1]} \left\{ \alpha\mu + \beta\mu^2 + \gamma\mu^3 - \mu \log(\mu) - (1-\mu)\log(1-\mu) \right\} \quad (7)$$

and model is asymptotically indistinguishable from directed Erdös-Rényi graph, with μ^* maximising (7). If the maximisation problem has multiple solutions, then μ^* picked randomly from some distribution on maximisers.

2) If at least **one** externality is negative (i.e. $\beta < 0$ or $\gamma < 0$) and sufficiently large, then model (6) not converge asymptotically to directed Erdös-Rényi graph and externalities can be identified.

Summary of Asymptotics

Remark 3.

Homogeneous players ($X_i = X_j$, $\forall i, j$):

(a) <u>positive externalities</u>

- asymptotically indistinguishable from Erdös-Rényi graph
- externalities not identified
- can approximate likelihood of model via likelihood of Erdös-Rényi graph

(b) at least <u>one externality negative and large</u>

- asymptotically sparser than Erdös-Rényi graph
- externalities identified

Heterogeneous players ($\exists i, j$ such that $X_i \neq X_j$):

- no results
- preliminary study in Mele & Zhu (2017) - working paper

Sampler Convergence

Theorem 5 (Convergence of Local Sampler with Nonnegative Externalities).

Model (6), with probability of meeting $\rho_{ij} = 1/(n(n-1))$. Fix $\gamma \geq 0$. Then, in the case of nonnegative externalities $\beta \geq 0$, there exists a V-shaped region of the parameter space delimited by functions $S_\gamma(\phi_1(\alpha)), S_\gamma(\phi_2(\alpha))$ such that

1) if (α, β) belongs to the V-shaped region, then model converges to stationarity in e^{Cn^2} steps, $C > 0$. This results holds for any local sampler.
2) otherwise, model converges in $Cn^2 \log(n)$ steps, $C > 0$.

Intuition:

(1a) in the V-shaped region problem (7) has 2 *local maxima*, the sampler spend exponential time at one of them (i.e. probability e^{-Cn^2} to escape from local max)

(1b) increasing $\gamma \implies$ increase area of exponentially slow convergence

(2a) when convergence is quadratic $\implies$ sampler feasible for $n < 500$

(2b) this happens when model is indistinguishable from directed Erdös-Rényi graph

Simulation and Estimation

Simulation and Estimation in Finite Networks

Posterior inference via approx version of exchange algorithm of MGM06 [16]

- double Metropolis-Hastings step to avoid computing $c(\mathcal{G}, X, \boldsymbol{\theta})$
- data augmentation via auxiliary network g'
- higher $R \implies$ better approximation of posterior, but higher rejection rate

Algorithm 2 Approximate Exchange Algorithm

procedure $\mathrm{AEA}(\boldsymbol{\theta}, g, M, R)$
  for $m = 1, \ldots, M$ **do**
    1) propose parameter $\boldsymbol{\theta}' \sim q_{\boldsymbol{\theta}}(\cdot|\boldsymbol{\theta})$
    2) run Algorithm 1 for R iterations using $\boldsymbol{\theta}'$. Keep last simulated network g'
    3) accept parameter $\boldsymbol{\theta}'$ with probability

$$\alpha(\boldsymbol{\theta}, \boldsymbol{\theta}', g', g) = \min\left\{1, \frac{\exp\{\mathcal{Q}(g', X|\boldsymbol{\theta})\}}{\exp\{\mathcal{Q}(g, X|\boldsymbol{\theta})\}} \frac{p(\boldsymbol{\theta}')}{p(\boldsymbol{\theta})} \frac{q_{\boldsymbol{\theta}}(\boldsymbol{\theta}|\boldsymbol{\theta}')}{q_{\boldsymbol{\theta}}(\boldsymbol{\theta}'|\boldsymbol{\theta})} \frac{\exp\left\{\mathcal{Q}(g, X|\boldsymbol{\theta}')\right\}}{\exp\left\{\mathcal{Q}(g', X|\boldsymbol{\theta}')\right\}}\right\}$$

  end for
  return sequence of M parameters $\{\boldsymbol{\theta}^{(m)}\}_m$
end procedure

Simulation and Estimation in Finite Networks

Algorithm 2 Approximate Exchange Algorithm

procedure $\mathrm{AEA}(\boldsymbol{\theta}, g, M, R)$
  for $m = 1, \ldots, M$ **do**
    1) propose parameter $\boldsymbol{\theta}' \sim q_{\boldsymbol{\theta}}(\cdot|\boldsymbol{\theta})$
    2) run Algorithm 1 for R iterations using $\boldsymbol{\theta}'$. Keep last simulated network g'
    3) accept parameter $\boldsymbol{\theta}'$ with probability

$$\alpha(\boldsymbol{\theta}, \boldsymbol{\theta}', g', g) = \min\left\{1, \frac{\exp\{\mathcal{Q}(g', X|\boldsymbol{\theta})\}}{\exp\{\mathcal{Q}(g, X|\boldsymbol{\theta})\}} \frac{p(\boldsymbol{\theta}')}{p(\boldsymbol{\theta})} \frac{q_{\boldsymbol{\theta}}(\boldsymbol{\theta}|\boldsymbol{\theta}')}{q_{\boldsymbol{\theta}}(\boldsymbol{\theta}'|\boldsymbol{\theta})} \frac{\exp\left\{\mathcal{Q}(g, X|\boldsymbol{\theta}')\right\}}{\exp\left\{\mathcal{Q}(g', X|\boldsymbol{\theta}')\right\}}\right\}$$

  end for
  return sequence of M parameters $\{\boldsymbol{\theta}^{(m)}\}_m$
end procedure

▶ what prior distribution $p(\boldsymbol{\theta})$?

▶ what proposal distribution $q_{\boldsymbol{\theta}}(\cdot|\boldsymbol{\theta})$?

Simulation results

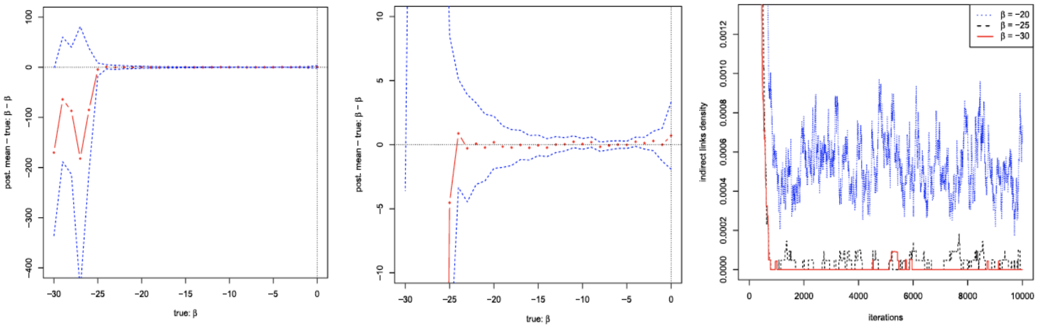


Figure: *Left*: Estimates of $\beta < 0$, with 95% credibility intervals (*middle*: zoom-in). *Right*: indirect links density.

- $\beta \geq 0 \implies$ Erdös-Rényi case, not identified
- $\beta < 0 \implies$ identified
- $\beta \ll 0 \implies$ estimation impossible: $\#$ indirect links close to 0

Simulation results

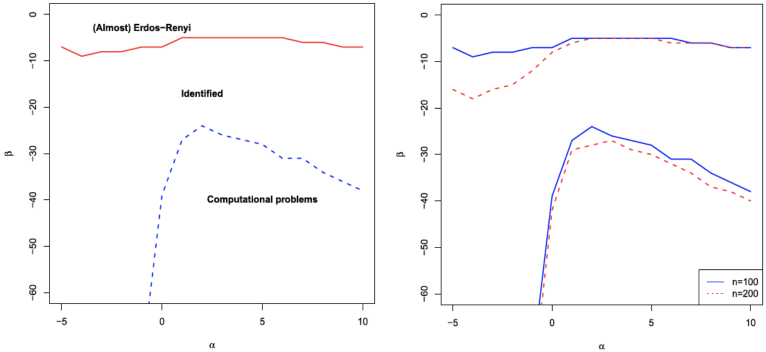


Figure: *Left*: approximate regions of identified parameters, for $n = 100$. *Right*: comparison of regions for $n = 100$, $n = 200$.

Remark 4.

Regions of identified parameters (α, β) vary with n, the number of players.

Simulation results

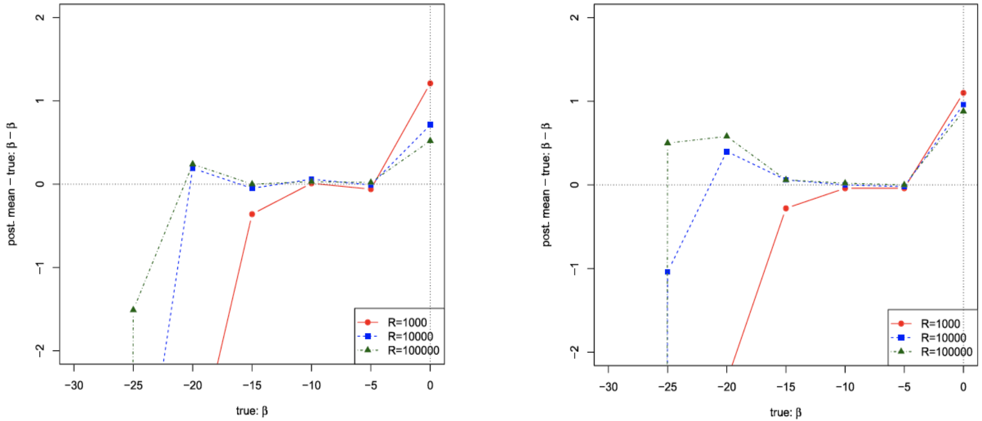


Figure: Difference between posterior estimates and true, for varying number of network simulations R: $n = 100$ (*left*) and $n = 200$ (*right*).

Simulation results

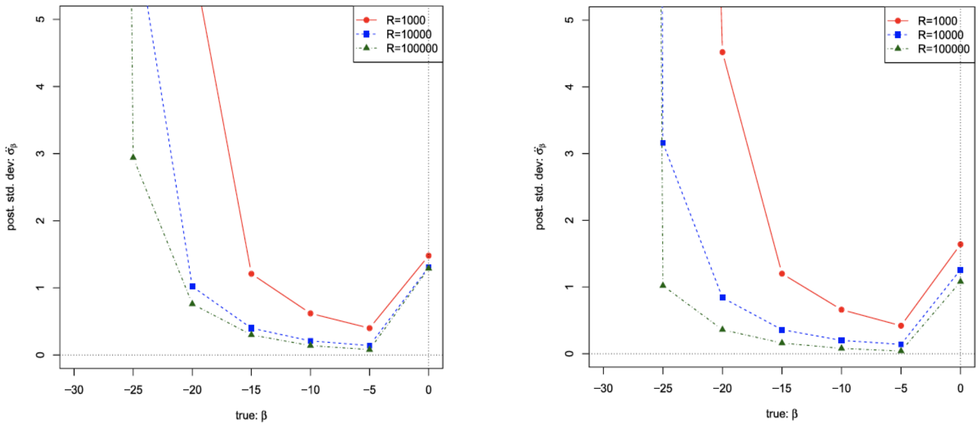


Figure: Posterior standard deviation for varying number of network simulations R: $n = 100$ (*left*) and $n = 200$ (*right*).

Simulation results

- $R = 1000 \implies$ imprecise estimates
- no significant difference between $R = 10,000$ and $R = 100,000 \implies$ suggest rule-of-thumb $R = 10,000$
- cost of increasing network simulations $\implies$ almost linear $\mathcal{O}(R)$
- results suggest convergence is almost quadratic $\mathcal{O}(n^2)$ in this area of parameter space

Simulation results

- $R = 1000 \implies$ imprecise estimates
- no significant difference between $R = 10,000$ and $R = 100,000 \implies$ suggest rule-of-thumb $R = 10,000$
- cost of increasing network simulations $\implies$ almost linear $\mathcal{O}(R)$
- results suggest convergence is almost quadratic $\mathcal{O}(n^2)$ in this area of parameter space

- what was the computing time?

- what about real data applications?

Conclusions

Conclusions

The paper in a nutshell:

- network formation model, combining strategic and random networks features
- payoffs depend on links: direct + indirect (externalities)
- homogeneous players meet sequentially at random, myopically updating links
- network formation process is a potential game and converges to ERGM, generating directed dense networks
- identification: only if at least 1 externality negative and sufficiently large
- standard estimation for ERGMs exponentially slow $\Rightarrow$ Bayesian MCMC (almost quadratic time)

Conclusions

Unclear points and questions:

- theoretical quantification of "sufficiently large" (negative) magnitude of β?
- choice of prior for parameters $p(\boldsymbol{\theta})$?
- choice of proposal for network $q_g(\cdot|g)$?
- choice of proposal for parameters $q_{\boldsymbol{\theta}}(\cdot|\boldsymbol{\theta})$?
- duration of computing time in simulations?
- real data applications?

Thanks for your attention!

Any question?

References

References i

- Jackson, M. O. (2008), Social and Economics Networks. Princeton
- Lovasz, L. (2012), Large Networks and Graph Limits. American Mathematical Society Colloquium Publications, Vol. 60. American Mathematical Society

- Badev, A. (2017), "Discrete Games in Endogenous Networks: Equilibria and Policy", Workin Paper. Available at arXiv https://arxiv.org/abs/1705.03137v1
- Besag, J. (1974), "Spatial Interaction and the Statistical Analysis of Lattice Systems", Journal of the Royal Statistical Society. Series B (Methodological), 36(2):192–236.
- Caimo, A., & N. Friel (2011), "Bayesian Inference for Exponential Random Graph Models", Social Networks, 33(1):41–55
- Chatteriee, S., & S. R. S. Varadhan (2011), "The Large Deviation Principle for the Erdös-Rényi Random Graph" European Journal of Combinatorics, 32(7):1000–1017. Homomorphisms and Limits

References ii

- Christakis, N., J. Fowler, G. W. Imbens, & K. Kalyanaraman (2010), "An Empirical Model for Strategic Network Formation" Harvard University
- Frank, O., & D. Strauss (1986), "Markov Graphs", Journal of the American Statistical Association, 81:832–842
- Graham, B. (2017), "An Econometric Model of Network FormationWith Degree Heterogeneity", Econometrica, 85(4):1033-1063
- Gouriéroux, J., Lafont, J. J. and Monfort, A. (1981), "Coherency Conditions in Simultaneous Linear Equation Models with Endogenous Switching Regimes", Econometrica, 48(3):675–696.
- Hsieh, C.-S., & L.-F. Lee (2016), "A Social Interactions Model With Endogenous Friendship Formation and Selectivity", Journal of Applied Econometrics, 31(2):301–319
- Leung, M. (2018), "A weak law for moments of pairwise-stable networks", Working Paper. Available at SSRN: https://ssrn.com/abstract=2663685

References iii

- Monderer, D., & L. Shapley (1996), "Potential Games", Games and Economic Behavior, 14, 124–143.
- Mele, A., & L. Zhu (2017), "Approximate Variational Estimation for a Model of Network Formation", Working Paper. Available at arXiv: https://arxiv.org/abs/1702.00308
- Menzel, K. (2017), "Strategic Network Formation With Many Agents", Working Papers, NYU
- Murray, I. A., Z. Ghahramani, & D. J. C. MacKay (2006), "MCMC for Doubly-Intractable Distributions", Uncertainty in Artificial Intelligence
- Miyauchi, Y. (2016), "Structural Estimation of a Pairwise Stable Network Formation With Nonnegative Externality", Journal of Econometrics, 195(2):224–235
- Snijders, T. A. B. (2002), "Markov Chain Monte Carlo Estimation of Exponential Random Graph Models", Journal of Social Structure, 3(2)

References iv

- Tamer, E. (2003), "Incomplete Simultaneous Discrete Response Model with Multiple Equilibria", The Review of Economic Studies, 70:147–165.
- Wainwright, M. J., & M. Jordan (2008), "Graphical Models, Exponential Families, and Variational Inference", Foundations and Trends® in Machine Learning, 1(1–2):1–305

Meetings, Discussants and Material

Links: Mele (2017) paper; Mele (2017) supplements;

1^{st} session: 11/10/2018 - Matteo Iacopini - Main slides on Mele (2017);

2^{nd} session: 18/10/2018 - Carlo Santagiustina - Sup. slides on Tamer (2003);