Social network analysis: social learning

Donglei Du
(ddu@unb.edu)

Faculty of Business Administration, University of New Brunswick, NB Canada Fredericton
E3B 9Y2

October 20, 2016

Table of contents

Section 1

Introduction

Learning on network: social learning I

- This material is adopted from [Jackson, 2010]-Chapter 8.
- Question to be answered: How the structure of social networks affects learning and the diffusion of information?
  - whether individuals in a society come to hold a common belief or remain divided in opinions (Peron-Frobenius Theorem);
  - which individuals have the most influence over the beliefs in a society (centrality);
  - how quickly individuals learn (fast mixing of Markov chain);
  - whether initially diverse information scattered throughout the society can be aggregated in an accurate manner (wisdom of crowd).

Section 2

The DeGroot Model [DeGroot, 1974]

Subsection 1

The Model

The DeGroot Model [DeGroot, 1974] I

- This is a network interaction model of information transmission, opinion formation, and consensus formation.
- It is a very simple and quite natural starting point for a theory that will allow us to more fully understand how the structure of a network influences the spread of information and opinion formation.
- It describes a mechanism on how individuals in a group can reach a consensus.
- It pools the opinions of a group into a consensus.
- It forms a common probability distribution of a group consensus.
- It provides an explicit basis for some of eigenvector-based measures of centrality and influence.
- The DeGroot model and its variants are tractable and powerful as tools for studying a variety of issues associated with the diffusion of information and learning.

The DeGroot Model [DeGroot, 1974] II

- As the influence matrix is something that can be examined empirically, it holds substantial promise as a tool for empirical research.

The model I

- Consider a society with n agents V where everybody has an opinion on a subject, represented by a stochastic column vector

$$p(0) = \begin{pmatrix} p_1(0) \\ p_2(0) \\ \vdots \\ p_n(0) \end{pmatrix}$$

- Agents interact via a weighted, directed network $G = (V, A, w)$.
  - The weight w_{ij} on arc $(i, j) \in A$ represents the weight or trust that agent i places on the current belief of agent j in forming his or her belief for the next period.
  - The weight matrix $W = (w_{ij})$ is row stochastic (i.e., $W \geq 0$ and row sum $We_i = 1, \ \forall i$).
  - Agents cannot update their interaction patterns, but they can communicate with their neighbors to update their belief.

The model II

- Belief are updated over time according to the following dynamic:

$$\boxed{p(t) = Wp(t-1) = W^t p(0)}$$

Subsection 2

Convergence

Convergence

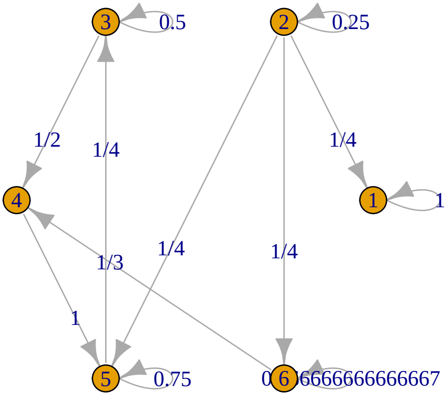
- An absorbing set of nodes is a $C \subset V$ such that there is no directed link from an agent in C to an agent outside of C (that is, there is no pair $j \in C$ and $i \notin C$ such that $w_{ij} > 0$).
- A social influence matrix W is convergent if the following limit exists for all initial vectors of beliefs $p(0)$:

$$\lim_{t \to \infty} p(t) = p(\infty)$$

$$\Updownarrow$$

Every strongly connected absorbing set of nodes is aperiodic.

An example I



- There are two strongly connected absorbing sets: $\{1\}$ and $\{3,4,5\}$

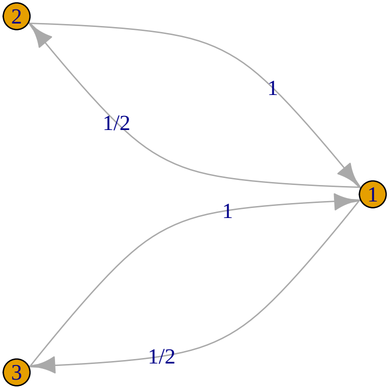
An example II

- There are quite a few absorbing node sets: $\{1\}$ $\{3, 4, 5\}$, $\{1, 3, 4, 5\}$, $\{3, 4, 5, 6\}$, $\{1, 3, 4, 5, 6\}$, $\{1, 2, 3, 4, 5, 6\}$.

An example of convergence, but not consensus: W^∞ exists, but rows are not identical

```
##           1 2         3         4         5            6
## 1 1.0000000 0 0.0000000 0.0000000 0.0000000  0.000000e+00
## 2 0.3333333 0 0.1904762 0.0952381 0.3809524 4.862865e-177
## 3 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 4 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 5 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 6 0.0000000 0 0.2857143 0.1428571 0.5714286 8.104775e-177
##           1 2         3         4         5            6
## 1 1.0000000 0 0.0000000 0.0000000 0.0000000  0.000000e+00
## 2 0.3333333 0 0.1904762 0.0952381 0.3809524 3.241910e-177
## 3 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 4 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 5 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 6 0.0000000 0 0.2857143 0.1428571 0.5714286 5.403183e-177
```

An example of non-convergence



R output: W^∞ does not exist.

```
##      [,1] [,2] [,3]
## [1,]    1  0.0  0.0
## [2,]    0  0.5  0.5
## [3,]    0  0.5  0.5
##      [,1] [,2] [,3]
## [1,]    0  0.5  0.5
## [2,]    1  0.0  0.0
## [3,]    1  0.0  0.0
```

Social influence of the agents on the limiting belief I

- Consider a social influence matrix $W = (w_{ij})$ (row stochastic) and its associated weighted directed network $G(V, A, w)$.
- Let $V = B_1 \cup \ldots \cup B_\ell \cup R$ be a partition of nodes such that
  - B_k is the maximal strongly connected absorbing set.
  - R are the remaining agents (not in any strongly connected absorbing set).
- A (row) stochastic matrix W is convergent if and only if there is a nonnegative row vector $s \in \mathbb{R}^n_+$ such that
  - $s_{B_k} > 0$ is a stochastic row vector for each B_k and it is the left leading eigenvector of W restricted to B_k;;
  - $s_i = 0$ iff $i \in R$;
  - For any vector $p \in \mathbb{R}^n$ and B_k, $k \in \{1, \ldots, \ell\}$

$$\left(\lim_{t\to\infty} W^t_{B_k} p\right)_{B_k} = s_{B_k} p_{B_k};$$

Social influence of the agents on the limiting belief II

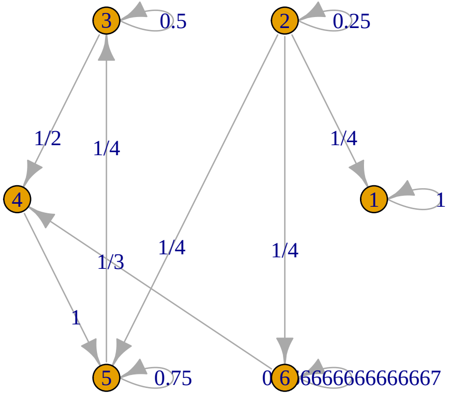
- For any vector $p \in \mathbb{R}^n$ and each agent $j \in R$, there exists a stochastic row vector $w^j_{B_k} \geq 0$ such that

$$\left(\lim_{t\to\infty} W^t p\right)_j = \sum_k w^j_{B_k} s_{B_k} p_{B_k}.$$

Intepretation

- The previous result states that (provided a society converges) each strongly connected absorbing set of nodes converges to a consensus belief that is determined by the social influence vector for that group times their initial beliefs.
- Agents outside of the closed and strongly connected sets then converge to some weighted average of the strongly connected absorbing groups limiting beliefs.

An example of the limiting belief



R output

```
##           1 2         3         4         5            6
## 1 1.0000000 0 0.0000000 0.0000000 0.0000000  0.000000e+00
## 2 0.3333333 0 0.1904762 0.0952381 0.3809524 4.862865e-177
## 3 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 4 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 5 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 6 0.0000000 0 0.2857143 0.1428571 0.5714286 8.104775e-177
##           1 2         3         4         5            6
## 1 1.0000000 0 0.0000000 0.0000000 0.0000000  0.000000e+00
## 2 0.3333333 0 0.1904762 0.0952381 0.3809524 3.241910e-177
## 3 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 4 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 5 0.0000000 0 0.2857143 0.1428571 0.5714286  0.000000e+00
## 6 0.0000000 0 0.2857143 0.1428571 0.5714286 5.403183e-177
```

Social influence weights: an example I

- Evidently $s_1 = 1$ from the first row on Slide 22.
- $\{3, 4, 5\}$ will reach a consensus as if they were an isolated society. The left leading eignevector of the submatirx

$$W_{\{3,4,5\}} = \begin{pmatrix} \frac{1}{2} & \frac{1}{2} & 0 \\ 0 & 0 & 1 \\ \frac{1}{4} & 0 & \frac{1}{2} \end{pmatrix}$$

is

$$(s_3, s_4, s_5) = \left(\frac{2}{7}, \frac{1}{7}, \frac{4}{7}\right) \underbrace{\approx}_{\text{Slide 22}} (0.2857143, 0.1428571, 0.5714286).$$

Social influence weights: an example II

- 6 only pays attention to 6 and 4. Given that 4 eventually converges, 6's belief will converge to 4's belief regardless of 6's initial belief. Therefore, 6 has a weight of 1 on $\{3, 4, 5\}$, and so

$$
\begin{aligned}
w^6_{\{1\}} &= 1 \\
w^6_{\{3,4,5\}} &= 0
\end{aligned}
$$

- Agent 2 is paying equal attention to 1, 2, 5, and 6. As 1, 5 and 6's beliefs converge to various limits, 2's initial belief will not matter. Given that 6 will converge to the same belief as 5, then effectively 2 has twice as much weight on the limiting belief of 5 (2/4) compared

Social influence weights: an example III

to that of 1 (1/4). We can also see this by simply noting that 2's limiting beliefs have to satisfy the following:

$$\begin{aligned} p_2(\infty) &= \frac{1}{4}(p_1(\infty) + p_2(\infty) + p_5(\infty) + p_6(\infty)) \\ &\underbrace{=}_{p_5(\infty)=p_6(\infty)} \frac{1}{4}(p_1(\infty) + p_2(\infty) + 2p_5(\infty)) \end{aligned}$$

$$\Downarrow$$

$$p_2(\infty) = \frac{1}{3}p_1(\infty) + \frac{2}{3}p_2(\infty)$$

Social influence weights: an example IV

- Check the output from Slide 22:

$$
\begin{aligned}
& (0.3333333, 0, 0.1904762, 0.0952381, 0.3809524, 0) \\
\approx\; & \frac{1}{3}(1, 0, 0, 0, 0, 0) \\
+\; & \frac{2}{3}(0, 0, 0.2857143, 0.1428571, 0.5714286, 0)
\end{aligned}
$$

- So

$$
\begin{aligned}
w^2_{\{1\}} &= \frac{1}{3} \\
w^2_{\{3,4,5\}} &= \frac{2}{3}
\end{aligned}
$$

Subsection 3

Consensus

Consensus

- A social influence matrix W is reaching a consensus if

$$p_1(\infty) = \ldots = p_n(\infty) = \pi p(0)$$

$$\Updownarrow$$

There is only one strongly connected absorbing aperiodic set

$$\Updownarrow$$

$\exists t$ such that some column of T^t has all positive entries.

- Here π is the left leading (row) eigenvector vector of W.
    - since starting with $p(0)$ or $p(1) = Wp(0)$ yields the same limit, it must be that $sp(1) = sp(0)$. Therefore,

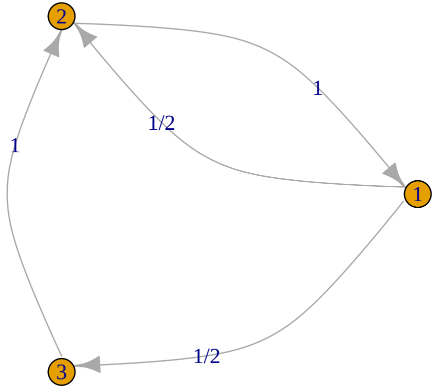
$$\begin{aligned} s(Wp(0)) &= sp(0),\ \forall p(0) \\ &\Downarrow \\ sW &= s \end{aligned}$$

Two methods to find the consensus

- Direct method: one can solve $\pi W = \pi$ directly in cases where n is not too large.
- Power method: or approximate π by the power method:

$$W^t \to W^\infty = \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix}$$

An example to find the consensus



Power method

```
##      [,1] [,2] [,3]
## [1,]  0.4  0.4  0.2
## [2,]  0.4  0.4  0.2
## [3,]  0.4  0.4  0.2
##      [,1] [,2] [,3]
## [1,]  0.4  0.4  0.2
## [2,]  0.4  0.4  0.2
## [3,]  0.4  0.4  0.2
```

Direct method

```
##       [,1] [,2] [,3]
## [1,]  0.0    1    0
## [2,]  0.5    0    1
## [3,]  0.5    0    0
## [1] 1+0i
## [1] 0.4+0i 0.4+0i 0.2+0i
```

Section 3

Appendix: Markov chain theory [Luenberger, 1979]-Chapter 7 or [Lovász, 1993] on time-reversible Markov chain

Markov chain (MC) theory

- A Markov chain is a stochastic process $(X_1, X_2, \ldots)$ defined on a countable states $\{1, \ldots, n\}$, satisfying the Markov property:

$$\begin{aligned} & \mathbb{P}(X_{n+1} = j \mid X_1 = i_1, X_2 = i_2, \ldots, X_2 = i_{n-1}, X_n = i) \\ = \ & \mathbb{P}(X_{n+1} = j \mid X_n = i). \end{aligned}$$

- Therefore a Markov chain is completely specified by a row stochastic matrix (state transition matrix)

$$P = \begin{pmatrix} p_{11} & p_{12} & \cdots & p_{1n} \\ p_{21} & p_{22} & \cdots & p_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ p_{n1} & p_{n2} & \cdots & p_{nn} \end{pmatrix}.$$

representing the conditional probability

$$\mathbb{P}(X_{k+1} = j \mid X_k = i) = p_{ij}$$

Graph representation of MC

- The transition matrix can be represented by a weighted and directed network $(G = (V, A, P)$ (state diagram), where $V = S = \{1, \ldots, n\}$ is the state space; and there is an arc $(i, j) \in A$ with weight p_{ij} whenever $p_{ij} > 0$.
- The transition matrix P (and hence the corresponding Markov chain) is irreducible iff G is strongly connected.
- The transition matrix P (and hence the corresponding Markov chain) is regular iff G is strongly connected and aperiodic.

Algebraic characterization of irreducibility and regularity

- **Irreducible** transition matrices: for any pair of i, j:

$$\boxed{\begin{array}{r} (P^{k_{ij}})_{ij} > 0, \text{ for some } k_{ij} \geq 1 \\ \Updownarrow \\ \forall \text{ permutation matrix } Q : Q^T P Q \neq \begin{bmatrix} X & Y \\ 0 & Z \end{bmatrix} \end{array}}$$

- **Regular** transition matrices (a.k.a. primitive transition matrices):

$$\boxed{P^k > 0, \text{ for some } k \geq 1}$$

The dynamics on Markov chain

- Let $\pi(k)$ be an n-dimensional row stochastic vector specifying the knowledge/belief on the states at time k.
- Successive knowledge are generated by the linear time-invariant system:

$$\pi(k+1) = \pi(k)P$$

- Note that the $\pi(k)$ is not really the state of the Markov process.
  - At each step the state is one of the n distinct states $V = S = \{1, \ldots, n\}$.
  - The vector $\pi(k)$ gives the probabilities that the Markov process takes on specific values.
  - Thus, the sequence of $\pi(k)$ does not record a sequence of actual states, rather it is a projection of our probabilistic knowledge.

Limiting distribution of Markov chain

- Let P be the transition matrix (row stochastic) of a regular Markov chain. Then:
  1. there exists a unique row stochastic vector π satisfying
  $$\pi = \pi P$$
  2. The limit of the power matrix
  $$\lim_{k\to\infty} P^k = \begin{pmatrix} \pi \\ \vdots \\ \pi \end{pmatrix}$$

Classification of states

- Communicating class C: strongly connected
    -
- Absorbing class (closed class): no possible transitions from C to any state outside.
- Transient class C: possible transitions from C to some state outside

Classification of states

- Canonical form for MC

$$\begin{bmatrix} C_{r\times r} & 0 \\ R_{(n-r)\times r} & T_{(n-r)\times(n-r)} \end{bmatrix}$$

- Assume there are r states in absorbing classes and $n-r$ in transient classes:
  - C is an $r \times r$ (row) stochastic matrix representing the transition probabilities within the closed classes;
  - T is an $(n-r)\times(n-r)$ sub-stochastic matrix (at least one row sum is less than 1) representing the transition probabilities among the transients states;
  - R is an $(n-r)\times r$ matrix representing the transition probabilities from transient states to states within a closed class.

Fundamental matrix of Markov chain

- Given a Canonical form for MC, fundamental matrix is

$$M = (I - T)^{-1} = I + T + T^2 + \ldots + T^k + \ldots$$

- M exists and positive ((Peron-Frobenius Theorem)).

The mean number of times the process is in transient state j when initiated in transient state i

- The element m_{ij} of the fundamental matrix M of a Markov chain with transient states is equal to the mean number of times the process is in transient state j if it is initiated in transient state i.

The mean number of steps before entering a closed class when the process is initiated in transient state i

- In a Markov chain with transient states, the ith component of the vector $M\mathbb{1}$ equal to the mean number of steps before entering a closed class when the process is initiated in transient state i.

The probability that if a Markov chain is started in transient state i it will first enter a closed class by visiting state j

- Let b_{ij} be the probability that if a Markov chain is started in transient state i it will first enter a closed class by visiting state j. Let B be the $(n-r)\times r$ matrix with entries b_{ij}. Then

$$B = MR$$

References I

DeGroot, M. H. (1974).
Reaching a consensus.
Journal of the American Statistical Association, 69(345):118–121.

Jackson, M. O. (2010).
Social and economic networks.
Princeton University Press.

Lovász, L. (1993).
Random walks on graphs: A survey.
Combinatorics, Paul erdos is eighty, 2(1):1–46.

Luenberger, D. (1979).
Introduction to dynamic systems: theory, models, and applications.