How homophily affects the speed of learning and best-responding dynamics_Session 1

The Quarterly Journal of Economics (2012)

Benjamin Golub & Matthew O.Jackson

The Road Map

- Introduction of the Models
  - Multi-type random networks
  - The average-based updating processes
  - Convergence and the Measure of the speed
- How homophily affects the learning
  - Define spectral homophily
  - Show how spectral homophily affects the convergence time of average-based updating processes
- A special case: The islands model
  - Two measures of homophily coincides

Multi-type radom networks

Set-up

- a set of n nodes, $N = \{1, \ldots, n\}$
- a network (undirected) represented via a n-by-n adjacency matrix A, $A_{ij} \in \{0, 1\}$
- the degree of node i, $d_i(A) = \sum_{j=1}^{n} A_{ij}$
- the sum of all degrees, $D(A) = \sum_i d_i(A)$ which is twice of the total number of links
- there are m different types of nodes, $N_k \subset N$ denote the set of nodes of type k ,the size of group k denotes by $n_k = |N_k|$
- $\mathbf{n} = (n_1, n_2, \ldots, n_m)$ be the corresponding vector of cardinalities, n denotes the total number of agents
- P_{kl} represents the probability that an agent of type k links to an agent of type l and P is a m-by-m symmetric matrix
- $A_{ij} = A_{ji}$ and $A_{ii} = 0$ for each i

A multi-type random network is defined by the cardinality vector $\mathbf{n}$ together with a symmetric m-by-m matrix P : $A(P, n)$

The average based updating process

- Let $T(A)$ be defined by $T_{ij}(A) = \frac{A_{ij}}{d_i(A)}$ (uniformization)
- the initial vector of beliefs $b(0) \in [0,1]^n$

Agent i's choice at date t is

$$b_i(t) = \underbrace{\sum_j T_{ij}(A) b_j(t-1)}_{\text{Agent } i \text{ only updates the average of his or her neighbors' last period choices, because } T_{ii}=0}$$

In matrix form:

$$b(t) = T(A)b(t-1) \text{ for } t \geq 1$$

then

$$b(t) = T(A)^t b(0)$$

Convergence and the measure of the speed

Lemma 1

If A is $\underbrace{\text{connected}}_{\text{irreducible}}$ and has $\underbrace{\text{at least one cycle of odd length}}_{\text{aperiodicity}}$, then $T(A)^t$ converges to a limit $T(A)^\infty$ such that $(T(A)^\infty)_{ij} = \frac{d_j(A)}{D(A)}$

It follows from standard results on Markov chains and implies that for any given initial vector of beliefs $b(0)$, all agents' behaviors or beliefs converge to an equilibrium in which consensus obtains. That is:

$$lim_{t\to\infty} b(t) = T(A)^\infty b(0) = (b, \dots, b) \text{ where } \quad b = \sum_j b_j(0) \frac{d_j(A)}{D(A)}$$

Golub and Jackson (2010) propose **aperiodicity** as a condition ensuring convergence in strongly connected stochastic matrices.

Definition 1

The matrix T is aperiodic if the greatest common divisor of the lengths of its simple cycles is 1.

Convergence and the measure of the speed

- At each moment, message through one link twice
- For the network A and starting belief b, we denote the consensus distance at time t by $CD(t; A, b)$
- The distance of beliefs at time t from consensus is $CD(t; A, b) = |T(A)^t b - T(A)^\infty b|_{s(A)}$ where $s(A)$ is denoted by $s(A) = (\frac{d_1(A)}{D(A)}, \ldots, \frac{d_n(A)}{D(A)})$
- We examine how many periods are needed for the vector describing agents' belief to get within some distance ε of its limit.

Definition 2

The consensus time to $\varepsilon > 0$ of a connected network A is

$$CT(\varepsilon; A) = sup_{b \in [0,1]^n} min\{t : CD(t; A, b) < \varepsilon\}$$

Four characters of multi-type random networks

Definition 3

- *A sequence of multi-type random networks is* **sufficiently dense** *if the ratio of the minimum expected degree to* $\log^2 n$ *tends to infinity:*

$$\frac{\min_k d_k[Q(P,n)]}{\log^2 n} \to \infty$$

- *A sequence of multi-type random networks has no* **vanishing groups** *if* $\liminf_n \frac{n_k}{n} > 0$
- *A sequence of multi-type random networks has* **interior homophily** *if*

$$0 < \liminf_n h^{spec}(P,n) \leq \limsup_n h^{spec}(P,n) < 1$$

- *Let* $\underline{P}$ *denote the smallest nonzero entry of* P *and* $\overline{P}$ *denote the largest nonzero entry. A sequence of multi-type random networks has* **comparable densities** *if:*

$$0 < \liminf_n \frac{\overline{P}}{\underline{P}} \leq \limsup_n \frac{\overline{P}}{\underline{P}} < \infty$$

Formalization of "Relevance"

Definition 4

Given two sequences of random variables $x(n)$ *and* $y(n)$*, we write* $x(n) \approx y(n)$ *to denote that for any* $\varepsilon > 0$*, if* n *is large enough, then the probability that*

$$\frac{(1-\varepsilon)y(n)}{2} \leq x(n) \leq 2(1+\varepsilon)y(n)$$

is at least $1 - \varepsilon$

In other words, $x(n) \approx y(n)$ indicates that the two random expressions $x(n)$ and $y(n)$ are within a factor of 2 (with a vanishingly small amount of slack) for large enough n with a probability going to 1.

A general measure of homophily: Spectral homophily

- $Q_{kl}(P, n) = n_k n_l P_{kl}$ represents the expected total contribution to the degrees of N_k from N_l (when $k \neq l$).
- let $d_k[Q(P, n)] = \sum_l Q_{kl}(P, n)$ be the expected total degree of nodes of type k.
- F_{kl} represents the expected fraction of the links that N_k will have with N_l (take $\frac{0}{0} = 0$):

$$F_{kl}(P, n) = \frac{Q_{kl}(P, n)}{d_k[Q(P, n)]}$$

Definition 5

The spectral homophily of a multi-type random network (P, n) is the second-largest eigenvalue of $F(P, n)$. We denote it as $h^{spec}(P, n)$.

Main results about homophily affect speeds

Theorem 1

Consider a sequence of multi-type random networks satisfying the conditions in Def 3. Then, for any $\gamma > 0$:

$$CT(\frac{\gamma}{n}; A(P, n)) \approx \frac{log(n)}{log(\frac{1}{|h^{spec}(P,n)|})}$$

Explanation of the theorem:

- The speed of the process essentially depends only on population size and homophily
- The approximation for consensus time on the right-hand side is always within a factor of 2 of the true consensus time
- Properties of the network other than spectral homophily can change the consensus time by at most a factor of 2 relative to the prediction made based on spectral homophily alone

A special case: The islands model

In the multi-type random network notation, we say the multi-type random network $(P, \mathbf{n})$ is an **islands network** with parameters (m, p_s, p_d) if:

- m islands of equal size;
- $P_{kk} = p_s$ for all k;
- $P_{kl} = p_d$ for all $k \neq l$, where $p_d \leq p_s$ and $p_s > 0$.

Definition of the **homophily in islands model**:

The natural measure of homophily is to compare the difference between same and different linking probabilities to the average linking probability.

Let $p = \frac{p_s + (m-1)p_d}{m}$ be the average linking probability, we define:

$$h^{\text{islands}}(m, p_s, p_d) = \frac{p_s - p_d}{m^* p}$$

It coincides with Coleman's(1958) homophily index:

$$\frac{\frac{p_s}{mp} - \frac{1}{m}}{1 - \frac{1}{m}}$$

- $h^{\text{islands}}(m, p_s, p_d) \in [0, 1]$

A simple case: The islands model

Proposition 1

If (P,n) is an islands network with parameters (m, p_s, p_d), then:

$$h^{islands}(m, p_s, p_d) = h^{spec}(P, n)$$

Proof:

$h^{\text{islands}}(m, p_s, p_d) = h^{\text{spec}}(P, n) \longleftrightarrow \frac{p_s - p_d}{mp}$ is the second largest eigenvalue of $F(P, n)$

Proof of the proposition 1

$$F_{kl}(P,n) = \frac{Q_{kl}(P,n)}{d_k(Q(P,n))} \text{ where } Q_{kl}(P,n) = \underbrace{n_k}_{n}\,\underbrace{n_l}_{n}\quad \overbrace{P_{kl}}^{\begin{cases} p_s & k=l \\ p_d & k \neq l \end{cases}}$$

$$d_k(Q(P,n)) = \sum_{l=1}^{m} Q_{kl}(P,n) = Q_{kk}(P,n) + \sum_{l \neq k}^{m} Q_{kl}(P,n) = n^2 p_s + (m-1)n^2 p_d$$

So $F_{kl}(P,n) = \frac{Q_{kl}(P,n)}{d_k(Q(P,n))} = \begin{cases} \frac{n^2 p_s}{n^2 p_s + (m-1)n^2 p_d} & k = l \\ \frac{n^2 p_d}{n^2 p_s + (m-1)n^2 p_d} & k \neq l \end{cases} = \begin{cases} \frac{p_s}{p_s + (m-1)p_d} & k = l \\ \frac{p_d}{p_s + (m-1)p_d} & k \neq l \end{cases} =$

$$\begin{cases} \frac{p_d}{p_s + (m-1)p_d} + \frac{p_s - p_d}{p_s + (m-1)p_d} & k = l \\ \frac{p_d}{p_s + (m-1)p_d} & k \neq l \end{cases}$$

Proof of the proposition 1

Take E_m as m-by-m matrix of 1 and I_m as m-by-m identity matrix, then:

$$F(P, n) = \frac{p_d}{p_s + (m-1)p_d} E_m + \frac{p_s - p_d}{p_s + (m-1)p_d} I_m$$

- Notice that λ is a eigenvalue of E_m if and only if $\frac{p_d}{p_s+(m-1)p_d}\lambda$ is a eigenvalue of $\frac{p_d}{p_s+(m-1)p_d}E_m$
- Adding kI_m is just shifts all the eigenvalues(and 0) by adding to them the k multiplying the identity.

As:

$$|\lambda I_m - E_m| = 0 \leftrightarrow \lambda = 0 \text{ or } \lambda = m$$

So:

$$|\lambda I_m - E_m - \frac{p_s - p_d}{p_s + (m-1)p_d} I_m| = 0 \leftrightarrow \lambda = \frac{p_s - p_d}{p_s + (m-1)p_d} \text{ or } \lambda = m + \frac{p_s - p_d}{p_s + (m-1)p_d}$$

As $m > 0$, then the second largest eigenvalue is $\frac{p_s - p_d}{p_s+(m-1)p_d}$ which equivalent to $\frac{p_s - p_d}{mp}$

Thank you ☺