

ランダムグラフと関連するネットワークモデル

システム情報数理学 II 橋 秀治

2007 年 2 月 15 日

1 現実世界に見られるネットワーク

例えば, インターネット, 食物連鎖, 友達関係, 電話回線網などがある. 以下では, ネットワークをグラフで捉えていく.

2 グラフ

定義 2.1 空でない集合 V と $E \subset \{\{x, y\}; x, y \in V\}$ の組 $G = (V, E)$ をグラフという. $\{x, y\} \in E$ を $x \sim y$ と書く. また, G の辺数を $e(G)$ と書く. $d(x) = \#\{y \in V; y \sim x\}$ を x の次数という.

定義 2.2 $x_0, x_1, x_2, \dots, x_{n-1}, x_n \in V$ に対して, $x_0x_1, x_1x_2, \dots, x_{n-1}x_n$ の形をした辺の有限列を n -step の歩道という. 歩道を $x_0 \sim x_1 \sim x_2 \sim \dots \sim x_{n-1} \sim x_n$ と書く.

定義 2.3 歩道 $x_0 \sim x_1 \sim \dots \sim x_{n-1} \sim x_0$ が $n \geq 3$ で, $x_0, x_1, \dots, x_{n-1}$ がすべて異なるとき, 閉路という.

定義 2.4 $G = (V, E)$ が連結とは, 異なる任意の 2 頂点 $x, y \in V$ が歩道で結ばれるときをいう.

定義 2.5 連結で閉路のないグラフを木という.

3 ネットワークを特徴付ける指標

定義 3.1 $G = (V, E)$ に対して,

$$p(k) = \frac{\#\{v \in V; d(v) = k\}}{n}, \quad k = 0, 1, 2, \dots$$

をグラフ G の次数分布という.

定義 3.2 グラフ $G = (V, E)$, $V = \{v_1, v_2, \dots, v_n\}$ に対して,

$$L = \binom{n}{2}^{-1} \sum_{1 \leq i < j \leq n} d(v_i, v_j)$$

をグラフ G の平均頂点間距離という. 但し, $d(v_i, v_j)$ は, 考えられるすべての v_i から v_j への歩道の step 数のうち, 最小のものとする.

定義 3.3 頂点 $v_i \in V$ に隣接している頂点の集合を $\Gamma(v_i) = \{v_1^{(i)}, v_2^{(i)}, \dots, v_{k_i}^{(i)}\}$ とするとき,

$$C_i = \binom{k_i}{2}^{-1} \#\{v_s^{(i)} v_t^{(i)} \in E; 1 \leq s < t \leq k_i\}$$

を頂点 v_i のクラスター係数という. これをもとに, グラフのクラスター係数 C を $C = (\sum_{i=1}^n C_i) / n$ で定義する.

現実世界のネットワークは, $p(k) \propto k^{-\gamma}$ (べき則), $L = O(\log n)$, C は 1 に近いことが特徴である.

4 ランダムグラフ

Ω を $V = \{1, 2, \dots, n\}$ を頂点集合とするグラフ全体とし, $0 < p < 1$ とする. Ω 上の確率 $\mathbf{P}_{n,p}$ を

$$\mathbf{P}_{n,p}(\{G\}) = p^{e(G)}(1-p)^{N-e(G)}, \quad N = \binom{n}{2}$$

によって定義する. このとき, 確率空間 $\mathcal{G}(n, p) = (\Omega, \mathbf{P}_{n,p})$ が考察するランダムグラフである. $\mathcal{G}(n, p)$ に属するグラフは, 異なる頂点 x と y の間に辺を引くかどうかを確率 p で独立に定めたものと捉えられる.

定理 4.1 $\mathcal{G}(n, p)$ の次数分布の平均 $\tilde{p}(k)$ は二項分布 $B(n-1, p)$ に従い,

$$\tilde{p}(k) = \binom{n-1}{k} p^k (1-p)^{n-1-k}$$

となる.

$n \rightarrow \infty$ において, 平均次数 $(n-1)p = \lambda$ (定数) が保たれている状況を考える. 当然, $p \rightarrow 0$ であり, $np \sim (n-1)p \rightarrow \lambda$ が成り立つ. このとき,

$$\tilde{p}(k) \rightarrow \frac{e^{-\lambda} \lambda^k}{k!}$$

となることが容易に判る (これは, ポワソンの少数の法則と同じ). ポワソン分布はべき則とは異なるので, エルデシュ・レニーのランダムグラフは, 現実世界に見られる複雑なネットワークのモデルとして不十分である.

定義 4.2 (WS モデル) $V = \{0, 1, \dots, n-1\}$ を番号順に輪状に並べる. k を偶数とし, 各頂点を両隣 $k/2$ 個の頂点と辺で結ぶ. 次に, 辺 $i(i+1)$ ($i \in V$) を頂点 i は固定したまま, 確率 $0 < p < 1$ で頂点 $i+1$ から引き離し, 端点を頂点 $i+1$ 以外の頂点から等確率で選び結ぶ. 但し, 多重辺やループができないようにする. 辺 $i(i+2)$ に対しても, 頂点 i は固定したまま, 確率 p で頂点 $i+2$ から引き離し, 端点を頂点 $i+2$ 以外の頂点から等確率で選び結ぶ. 同様のことを, 辺 $i(i+k/2)$ まで続ける.

5 スペクトル解析

定義 5.1 $G = (V, E)$ に対して,

$$A_{xy} = \begin{cases} 1 & (x \sim y) \\ 0 & (x \not\sim y) \end{cases}$$

で定義される行列 $A = (A_{xy})_{x,y \in V}$ を G の隣接行列という.

定義 5.2 A の固有値を $\lambda_1 < \lambda_2 < \dots < \lambda_s$, それらに対応する重複度を $m_1, m_2, \dots, m_s$ として得られる表

λ_1	λ_2	$\dots$	λ_s
m_1	m_2	$\dots$	m_s

を G のスペクトルという.

$\#V(G) = n$ とし, 次のスペクトル分布 (固有値分布) を考える.

$$\mu_A = \frac{1}{n} \sum_{k=1}^s m_k \delta_{\lambda_k}.$$

簡略化された WS モデルの構成方法

頂点集合 V を

$$V = \{0, 1, \dots, n-1\}$$

とおき, $k \geq 2$ を偶数の定数とする. $i, j \in V$ ($i \neq j$) の順序対に対して, ベルヌイ型確率変数 $X_{i \rightarrow j}$ を次のように定義する: $1 \leq t \leq k/2$ のとき,

$$\mathbf{P}_{n,p}(X_{i \rightarrow i+t} = 1) = 1 - p, \quad \mathbf{P}_{n,p}(X_{i \rightarrow i+t} = 0) = p.$$

$k/2 + 1 \leq t \leq n-1$ のとき,

$$\begin{aligned} \mathbf{P}_{n,p}(X_{i \rightarrow i+t} = 1) &= \frac{p}{n - (k/2 + 1)}, \\ \mathbf{P}_{n,p}(X_{i \rightarrow i+t} = 0) &= 1 - \frac{p}{n - (k/2 + 1)}. \end{aligned}$$

さらに, 確率変数の族 $\{X_{i \rightarrow j}; i \neq j\}$ は独立であるとする. この $X_{i \rightarrow j}$ を用いて,

$$A_{ij} = \frac{1}{2}(X_{i \rightarrow j} + X_{j \rightarrow i}), \quad A_{ii} = 0$$

とすると, ランダム行列 $A = (A_{ij})$ が得られる. A は対角成分が 0 となる対称行列であるが, 非対角成分の値は 0, 1 に加えて $1/2$ もとる.

行列 A のスペクトル分布 μ_A はランダム測度になる. その平均測度 $\mu_{n,p} = \mathbf{E}(\mu_A)$ を行列 A の平均スペクトル分布という. 以下, モデルのパラメータを $k = 2$ とする. このモデルの平均次数は 1 (n, p によらない).

補題 5.3 $\mu_{n,p}$ の m 次モーメントは, $0 \neq j_1 \neq j_2 \neq \dots \neq j_{m-1} \neq 0$ を V に属する頂点列とすると,

$$M_m(\mu_{n,p}) = \sum_{0 \neq j_1 \neq j_2 \neq \dots \neq j_{m-1} \neq 0} \mathbf{E}_{n,p}(A_{0j_1} A_{j_1 j_2} \dots A_{j_{m-1} 0})$$

と表せる.

補題 5.4 任意の $s \geq 1$ に対して,

$$\begin{aligned} \mathbf{E}_{n,p}(A_{ij}^s) &= \frac{1}{2^s} \{ \mathbf{E}_{n,p}(X_{i \rightarrow j}) + \mathbf{E}_{n,p}(X_{j \rightarrow i}) \\ &\quad + (2^s - 2) \mathbf{E}_{n,p}(X_{i \rightarrow j}) \mathbf{E}_{n,p}(X_{j \rightarrow i}) \} \end{aligned}$$

が成り立つ.

$1 \leq e(T) \leq m/2$ を満たすすべての木 T を次の写像

$$\Phi: E(T) \longrightarrow \{\alpha, \beta\} \times \{2, 4, 6, \dots\}, \quad \Phi(e) = (\gamma(e), k(e))$$

によって, ラベル付きグラフにする. 但し, $\sum_{e \in E(T)} k(e) = m$ を満たしかつ, 任意の $x \in V(T)$ に対して,

$$\#\{e \in E(T); e \text{ の端点が } x \text{ であり, } e \text{ のラベルが } \alpha\} \leq 2$$

を満たすようにする.

定義 5.5 $T = (V, E)$ が $T' = (V', E')$ とラベル付き同形であるとは, ラベルを保存するグラフ同形が存在するときをいう.

$\sum_{e \in E(T)} k(e) = m$ を満たすラベル付き木のラベル付き同形類の集合を $\mathcal{T}_m$ とおく.

また, $h(T)$ を出発した頂点に戻るような, ラベル $\{2, 4, 6, \dots\}$ に適合した m -step 歩道の数とする. 但し, $h(T)$ を考えるとき, ラベル付き自己同形を除く. つまり,

$$x \sim y \iff \varphi(x) \sim \varphi(y), \quad \Phi(xy) = \Phi(\varphi(x)\varphi(y))$$

を満たす全単射 $\varphi: V(T) \rightarrow V(T)$ が存在する木 T に対し, m -step 歩道 $v_0 \sim v_1 \sim \dots \sim v_{m-1} \sim v_0$ と $\varphi(v_0) \sim \varphi(v_1) \sim \dots \sim \varphi(v_{m-1}) \sim \varphi(v_0)$ が等しいときは, 歩道の数に 1 つとして数えることにする.

定理 5.6 T からラベル β の付いた辺を除き, さらに孤立点を除いてできるグラフの連結成分の個数を $p(T)$ とする. また, T のラベル α の辺の個数を $s(T)$, ラベル β の辺の個数を $t(T)$ とすると, m が奇数ならば,

$$\lim_{n \rightarrow \infty} M_m(\mu_{n,p}) = 0.$$

m が偶数ならば,

$$\lim_{n \rightarrow \infty} M_m(\mu_{n,p}) = \sum_{T \in \mathcal{T}_m} 2^{p(T)+t(T)-m} (1-p)^{s(T)} p^{t(T)} h(T)$$

が成り立つ.

参考文献

- [1] B. Bollobás: “Modern Graph Theory,” Springer, 2002.
- [2] S. Liang, N. Obata and S. Takahashi: Asymptotic spectral analysis of generalized Erdős-Rényi random graphs, preprint.
- [3] 増田直紀・今野紀雄: “複雑ネットワークの科学,” 産業図書, 2005.