

複雑ネットワークの視点から見た地域間交流データの 分析手法

Methodology for Inter-regional Flow Data from point of view of
Complex Networks

中 村 有 一 *
Yuichi NAKAMURA

Keywords : complex network, inter-regional flow data, scale-free, small world,
cluster

1. はじめに

前回の研究ノートにおいては、グラフ理論的な視点から、地域間交流データを分析した。この研究を踏まえて、今回はさらに複雑ネットワークの視点を追加して、研究を進めていきたい。まず前回までの研究ノートの要点を整理しておこう。^{(1), (2)}

(1) べき乗則が成立

各地域から発信されるデータの合計について、各地域の大きさ（サイズ）と各地域の順位（ランク）には、ランクサイズルールが成り立つ。つまり両対数プロットするとおおむね直線になる。これは一般にはべき乗則の一種であり、複雑ネットワークでいわれるある種の秩序を持っていることが推測される。

(2) 対地データを使った木構造の抽出

OD表の中で、横方向に最大となる地域を第1対地、2番目を第2対地（以下同様）という。この対地のデータをゼロイチのデータとして、グラフ理論的な意味でのグラフに表すことにより全体像を視覚化する。ここでは、データの大きな値のみを抽出しているので、大局的な構造をとらえているに過ぎない。つまり階層的な構造と圏域を求めることはできたが、階層内の構造はフラットで、それ以上分解することはできなかった。

(3) 最大木による分析

これはもとのデータを重み付きグラフとしてとらえ、最大木（重みの和が最大となる木）を

* 多摩大学経営情報学部 School of Management and Information Sciences, Tama University

求めることにより、主要な構造を抽出するものである。結果としては、第1対地による分析とほぼ同じで、計算の複雑さから見てもあまりメリットがないことが分かった。

なお今回の分析の元データについては、前回の研究ノートと同じなので、説明については省略する。^{(1), (3)}

2. 複雑ネットワークの効果的な視覚化

複雑ネットワークは、現実のネットワークをモデル化するのに有効な手法であることが認められてきているが、それを視覚的に表現することはできないだろうか。複雑ネットワークの特性を3つの視点から考察してみよう。それが実際の地域間交流データについて当てはまるのかを考えてみたい。

(1) スケールフリー性

スケールフリー (scale-free) とは、「特徴的な大きさを持たない」という意味である。別のいいかたをすれば、フラクタル的 (自己相似的) と呼んでもよい。この様子を視覚的に表現するには、情報が集中する中心性の高い地域を結節点として、階層的に枝分かれするようなグラフに表すことが可能である。

ただし前回の研究では、データの一部 (ただし重要な部分) を取り上げているに過ぎなかった。今回の研究では、圏域の中をフラットな状態からさらに階層に分解することにより、データを有効利用したい。また地域間交流データはべき乗則に従うことから、スケールフリーであることが示されたが、これを有効に視覚化することも目標となる。

(2) スモールワールド性

スモールワールド (small world) とは、文字通り「小さな世界」という意味で、現実のネットワークで連結された世界は、思ったより小さく、少ないリンクをたどって連結されているということを指す。⁽⁴⁾

OD データでは、すべてのリンクに小さいながらも重みがあるので、リンクを通れるか、通れないかという、ゼロイチの判断はされないが、圏域間の成分を評価すれば、スモールワールドの特徴が、視覚化できるものと考えられる。情報をバイパスするようなリンクが各所に埋め込まれているといってもよい。完全な階層構造を持つ会社のような組織でも、非公式なリンクによって効率が保たれ、秩序ができていると考えられる。

この方面の研究においては、何らかの指標によって表現できるかどうか、また有効な視覚化の方法があるかどうか課題である。⁽⁵⁾

(3) クラスタ性

クラスタとは、ここでは圏域のことである。各地域が集まって圏域を形作るが、情報は均等に分布するのではなく、情報の集積は、さらなる情報をひきつけ、拡大していこうとする特性を持っている。これは凝集性といってもよい。インターネットのサービスにおいては、大規模なところほど顧客を集め、さらに拡大していく。⁽⁵⁾ このようなことを検証するには、特徴的なパラメータを設定して、動的なシミュレーションにより実験することが多いが、今回は対象外

とした。

3. 圏域の分解

前回の研究ノートにおいては、第1対地によるグラフを求めている。このグラフでは中心となる地域が求められ、その圏域があることが示されている。しかし圏域の中はフラットであり、その内部構造はわからない。これをさらに細かく表現することはできないだろうか。

まず考えられるのは、膨大なデータの中でごく一部のデータしか使われていないことである。またデータを多くすると、リンクの線が多くなり、交差が煩雑となる。この点を考慮すると、圏域が特定されたことを前提として、その内部をさらに分解していくことが有効ではないかと考えられる。圏域が特定されれば、それを表すリンクは必要なくなり、線で囲むだけでわかるからである。

圏域を特定するには、グラフを見れば明らかであるが、同じ第1対地をもつ地域を集めてくればよい。たとえば東京の場合、東京が第1対地である地域をまとめて1つの圏域と考える。これをOD表の行と列の並べ替えによって、1つの領域にまとめる。この領域について第1対地のデータを作り、グラフの形に表現する。これにより第2階層の構造が見えてくるはずである。

図1、図2は、上記の方法で第1対地によるグラフ化を東京圏と大阪圏についてそれぞれ求めたものである。この方法では、対地のデータを圏域の外地域にも許している。これにより、圏域外とのつながりも表せているが、内部に限定するのであれば、外部との関係は無視して、対地として取り上げないようにすればよい。ここでは広い範囲で分析することを考えて、前者の方法を用いた。なおノードは円で表し、中にある数字は標準的な都道府県コードである。⁽¹⁾

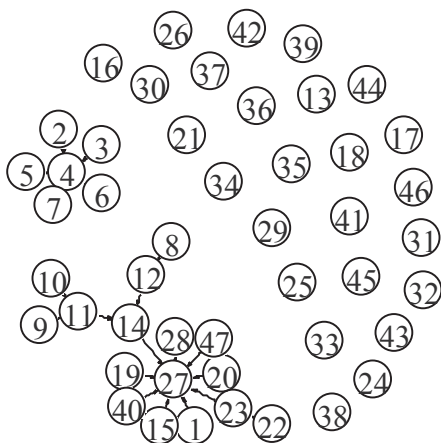


図1 東京圏内部の分解

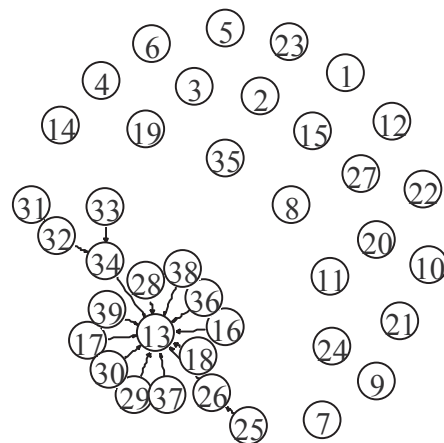


図2 大阪圏内部の分解

東京圏は21個の地域からなる。その中心となる東京は含めないことにした。グラフの右上には孤立したノードが並んでいるが、これは今回分析の対象ではないので無視してよい。大阪圏についても同様である。

東京圏の内部の構造については、まず宮城（コード＝4）を中心とした東北の圏域が見いだされる。次に圏域の外部につながる地域であるが、27,28,23,22という近畿・東海の地域が現れる。

これは圏域間を短絡的につなぐリンクで、スモールワールド性に関係するものと思われる。

大阪圏は15個の地域からなり、その中心となる大阪は含まれていない。ここにも東京、岡山、広島など圏域の外延部とのリンクが現れている。このような短絡的リンクをどのようにうまく視覚的に表現するかは、今後の課題である。

4. 圏域内のスケールフリー性

次に圏域に分解したとして、その内部もスケールフリーになっていることを確かめてみよう。図3は、東京圏を取り上げて地域の規模を集計し、地域の順位（ランク）と地域の規模（サイズ）を両対数プロットしたものである。

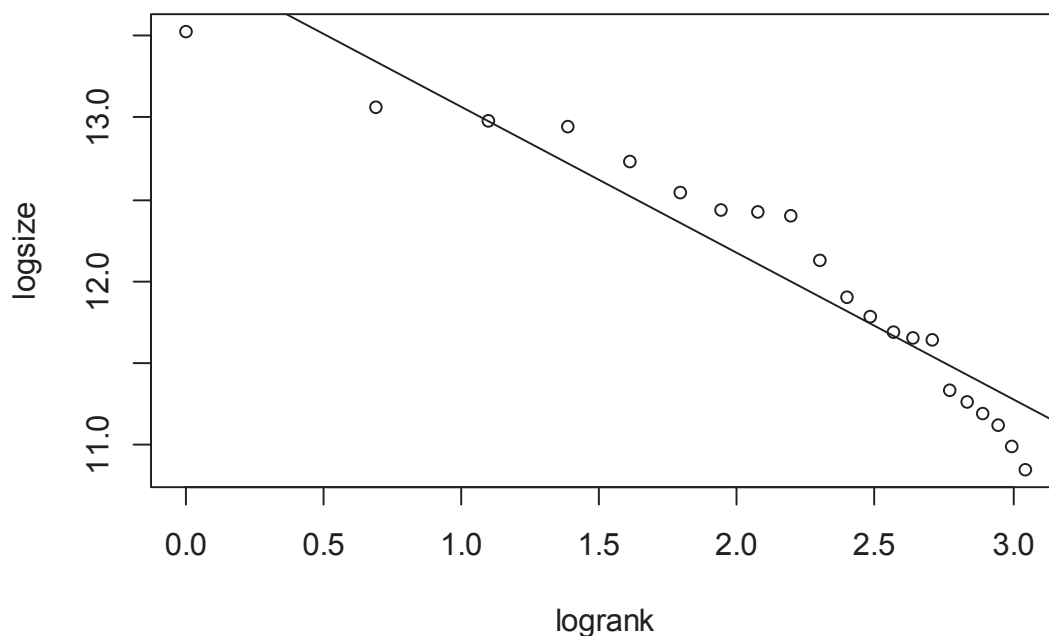


図3 東京圏の21地域のランクサイズプロット
(直線近似の決定係数 $R^2=0.9028$)

全体のランクサイズプロットと同様、ほぼ直線にのっており、圏域に分解してもスケールフリーという性質は変わらないことがわかる。圏域が小さくなりすぎると意味がなくなるが、再帰的に圏域を分解していても問題ないと思われる。

5. 今後の発展性

これまで複雑ネットワークの特性を前提として、地域間のつながりを視覚的に表現する手法について検討してきた。さらに今後の研究の可能性について以下で考えてみたい。

- (1) 今まで地域間交流データを主に使ってきたので、データの特性を無意識のうちに前提としてきたかもしれない。一般的なネットワーク分析においては、異質なデータを取り扱うこともあり、ネットワークの特性に応じて、手法の適用可能性が異なる場合も出てくると思われる。つまりデータの特性と手法の対応関係を整理しておきたい。

- (2) 時間の概念を入れて、変化の方向性を探ること、あるいはパラメータを与えて時系列のシミュレーションを行うことが考えられる。問題はかなり複雑になるが、興味深い視点が得られるだろう。
- (3) これまで使ってきた手法のうち、経験的に有効な手法であるが理論的な根拠がはっきりしないものについては、その根拠をデータの特性和との関係で検討したい。たぶんデータの特性和としては、暗黙のうちに様々な前提が置かれていたのではないだろうか。
- (4) 複雑ネットワークの世界では、これまで中心性を表す指標など、様々な指標が提案されてきた。⁽⁵⁾ これらの指標を検討することにより、ネットワークの特性和をより広範な視点から検討していきたい。
- (5) 地域の構造を再帰的に分解していくことを考えると、手法は複雑化してくる。ある程度手法として確立したら、ライブラリの形にまとめて簡単に使えるようにしたい。

以上思いつくままに書いてきたが、複雑ネットワークの世界では、理論的にすっきりした形では、問題は把握できないことが多い。手探りで統計的な手法を適用して初めてネットワークの特性和が明らかになってくる。このため直観と地道な実験の繰り返しが重要であることが実感された。また複雑系の世界では、単純な原理の積み重ねが複雑な結果を生み出す。その結果から、元の単純な原理を求めることは、一般に難しい。このような特性和のため理解を進めることが難しいという点を十分意識しておく必要がある。

6. おわりに

複雑ネットワークの分野、特にグラフ理論的なアプローチは、近年急速に発展してきた分野である。今まで意識しないで使ってきた手法にも、違う角度から光を当てると、また違った景色が見えることもある。また違った分野の研究成果が思わぬところで生かされることもある。複雑な問題をさらに複雑にするのではなく、複雑な構造をいかに単純な原理でシンプルに説明できるかということを心がけて研究を進めていきたい。

参考文献

- (1) 中村有一「地域間フローデータの階層的視覚化」経営・情報研究（多摩大学研究紀要）研究ノート 2017, No.21, pp.125-131 (2017)
- (2) 中村有一「地域間フローデータの視覚について」経営・情報研究（多摩大学研究紀要）研究ノート 2013, No.17, pp.105-112 (2013)
- (3) 電気通信事業者協会「テレコムデータブック 2015」第2章情報通信サービス利用状況 pp.21-25
<http://www.tca.or.jp/databook/> (2015)
- (4) ダンカン・ワッツ「スモールワールド」東京電機大学出版局 (2006)
- (5) 増田直紀、今野紀雄「複雑ネットワーク」近代科学社 (2010)