

## 地域間フローデータの視覚化について

### Visualization of Inter-Regional Flow Data

中村 有一\* Yuichi Nakamura

**Keywords:** visualization, inter-regional flow data, information quantity, scale-free

#### 1. はじめに

これまで通話や郵便など地域間の情報の流れを定量的に分析し、モデル化する手法について研究してきたが<sup>6)</sup>、本研究ノートでは、さらに視覚的にわかりやすい表現方法で要約する手法の開発をめざして、モデルを発展させていきたい。

分析の対象となるデータは、従来と同様、地域間 OD データであるが、今回はあまりこれにはこだわらず、もう少し広い視野でネットワークの構造を特定するにはどうすべきかについて検討していきたい。実際にデータを当てはめて検証するのは次の段階とし、今回は大まかな方向性をさぐることに重点を置くものである。

#### 2. モデル化の方法

ネットワークの結合構造を解析するには、何らかの単純な原理に基づきモデル化を行うが、本研究では以下の 3 種類のモデルに分けて、それぞれの特徴を比較検討することから始めた。ここで出てくる 3 つのモデルは、すべて広い意味では数学的モデルといえるが、その発想の原点を比較する意味で、3 つに分けて検討することにした。したがって一般的に使われている用語の意味とは異なる点に注意しておく必要がある。

##### (1) 数学的モデル

これは現実のデータを単純な数式で近似するものであり、その数式は自由に設定でき、その理論的根拠は必ずしも必要ない。適合の度合いが高ければそれでいいことになる。この種のモデルは自由度が非常に高く、最初のモデルは単純なものであっても、段々と改良が重ねられていくにしたがって、パラメータの数が多くなり、適合度は高まるが恣意的なモデルになっていく傾向がある。このようにして複雑化したモデルは、別のデータに適用した場合に、大きくその挙動が変化し、適合度も下がり不安定な結果をもたらすことがし

---

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

(原稿受理日 2012.11.17)

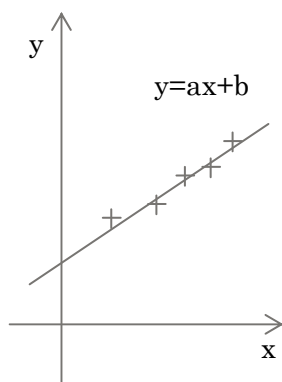
ばしば起こる。

図1の(a)は、数学的モデルのイメージを示したもので、ここでは1次関数でデータを近似する例を取り上げている。

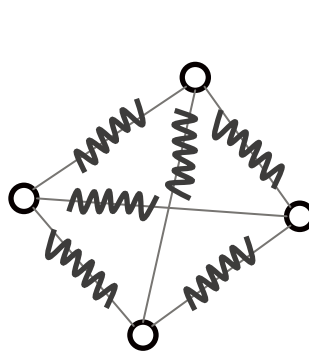
### (2) 物理学的モデル

ここでいう「物理学的モデル」においては、現実の世界を支配している何らかの物理法則のアナロジーとして数式を限定し、モデル化した結果に違和感が持たれないようにしたものである。たとえばエネルギー保存の法則に反するような世界は、現実的ではないので排除したり、統計力学の法則などを考慮したりする。このモデルは、宇宙空間に重力が隅々にまで及んでいるかのように、物体同士が力を及ぼしあっているというイメージでとらえることができる。このような世界では、お互いに及ぼしあっている力が大きいか小さいかだけの違いしかなく、全体の構造は単純で平等なものとなる。ただしさまざまな物理法則が絡み合っただけで複雑な関係が生ずると、太陽系や銀河系のような階層的な構造が形成されることがある。

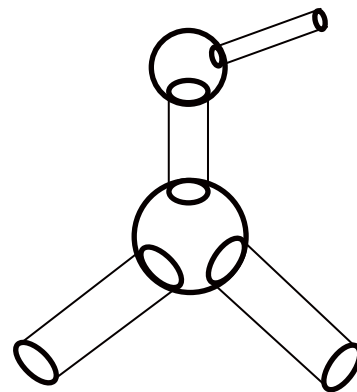
図1の(b)は、物体同士がばねで結ばれて力を及ぼしあっているというイメージを表したものである。



(a) 数学的モデル



(b) 物理学的モデル



(c) 生物学的モデル

図1 3種類のモデルのイメージ

### (3) 生物学的モデル

ここでいう「生物学的モデル」というのは、単なる物理法則の式に当てはめるのではなく、お互いの作用が構造を作り、その構造を前提として、また相互作用が深まっていくようなモデルのことである。この場合は、実体同士はネットワークでつながっているということであり、全く関係のないものもあるが、強くつながっているものもある。ネットワークは時間的な経過に伴って成長していき、ある一定の環境では定常状態に達すると考える。

いったん安定的な構造ができあがれば、大きな環境変化が起きないかぎりそれが維持されるが、大きく環境が変われば、次の安定状態に移行することもありうる。

平野に雨が降ってそれが川になり地形が形作られるように、最初はちょっとした偶然から地形が作られる。これが安定状態に到達すると、それを前提として人間の活動が展開される。人間の活動も、完全に設計して作れる範囲は限られている。大都市の成長パターンの研究成果によると、空間的にも時間的にもスケールの大きな現象では、人間の意志の影響は小さいことが知られている。<sup>2)</sup>

以上のような生物と類似した構造を持つモデルをここでは「生物学的モデル」と呼ぶことにする。このモデルでは、ネットワークと階層的構造が大きな意味を持つてくる。面ではなく線が重要であり、体積よりもむしろ表面積が重要となる場合がある。ネットワークが前提となっているので、物理学で扱うような一様な空間はあまり意味を持たない。したがって距離の概念もユークリッド的な距離ではなく、ネットワーク上での距離に意味がある。

実際にものが運ばれるためにはそのためのネットワークが必要となる。それは植物の場合、茎の中を通る維管束であり、人間の場合には血管網や神経系であったりする。これが地域であれば、道路網、通信網などのネットワークに相当する。動物の血管網の構造は、いろいろな制約のもとで、血液という液体を最小のエネルギーで運べるように最適化されている。物理法則が生物の形に影響した結果であり、地域間の物流ネットワークなどにもこれと同様の力が働いていると想像できる。

生物学的モデルでは、地域を単なる大きさだけで語ることはできない。個性の異なる地域同士の連携によって、その規模以上の影響力を持つことも可能である。これは生物の世界で「共生」と呼ばれる現象である。このような共生関係を扱うモデルは、重要ではあるが、本質的に質的に異なるもの同士の関係を扱う必要があるため、従来の定量的な手法では困難であることが予想される。

図1の(c)は、血管網のように液体を運ぶパイプで結ばれたネットワークをイメージして表したものである。

### 3. グラフ理論的モデル

本研究の目標は、「生物学的モデル」の視点に立って、地域間を流れる情報などのデータを分析して、わかりやすく視覚的に表現する手法を開発することである。そのためには、自然な形でその特徴を浮かび上がらせることが重要である。この場合、細かい例外的な挙動を表すデータは、むしろ全体の大きな構造を明らかにするうえでは、ノイズとして取り去るべきものと考えられる。どのような特徴を残せば、人間から見て自然な構造が見えてくるのか、この点が本質的な課題となる。

生物学的モデルとしては、いろいろな可能性があると考えられるが、その特徴を持ったままもっとも単純化して考えると、「グラフ理論的モデル」になる。これはグラフ理論で扱

うグラフのような単純化されたモデルである。関係があるかないかをゼロとイチであらわし、関係があるもの同士を線で結ぶとグラフができあがる。グラフ理論では、システムの要素となるものを「ノード」と呼び、それらの間の関係を表す線を「辺」と呼ぶ。<sup>1)</sup>

グラフ理論では、すべてのもの同士が関係のあるようなグラフのことを「完全グラフ」といい、すべての関係がないグラフ、つまり辺が一つもないグラフのことを「空グラフ」と呼んでいる。現実的なグラフは、この間にあることになる。ノードの数を  $n$  とすると、完全グラフでは、ノードを結ぶ辺の本数は、

$${}_nC_2 = \frac{1}{2}n(n-1)$$

となり、空グラフではゼロとなる。現実的なグラフでは、この間の本数となるが、この中で一番組み合わせの数が多いものが、ちょうど中間の本数の場合である。辺がランダムにとられると仮定すれば、この場合エントロピーは最大となり、表現できる情報量も最大となる。

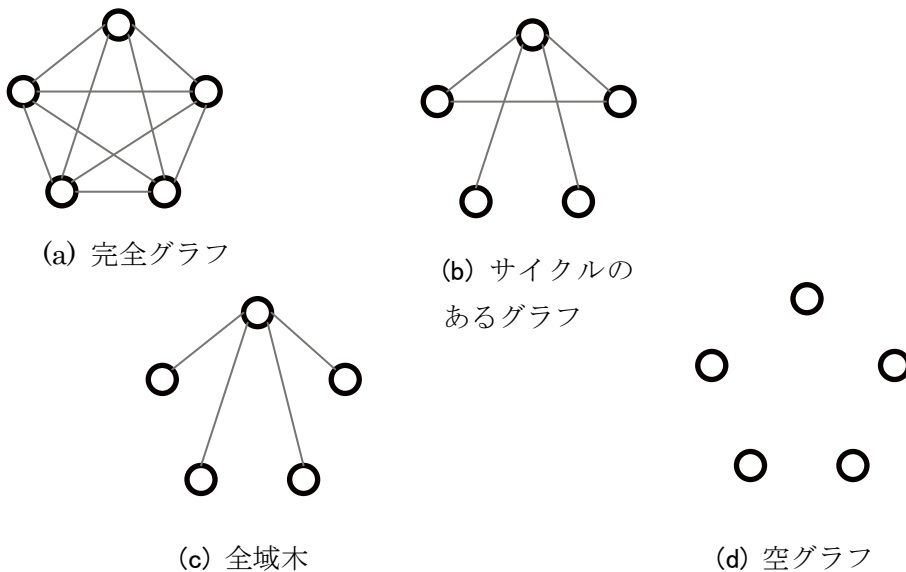


図2 完全グラフ、サイクルのあるグラフ、全域木、空グラフの例

グラフ理論的モデルの例として、以下の3つのモデルを考えよう。

#### (1) 全域木モデル

ネットワークの構造を、連結性の視点から考えてみよう。グラフが「連結」であるとは、任意の2つのノードを取ったとき行き来できる道が存在することである。完全グラフから辺を1本ずつ取り除いていくと、あるところでグラフは連結でなくなる。この連結性が失

われる直前のグラフのことを「木」と呼ぶ。「サイクル」とは、辺をぐるっと回ってもとに戻って来られる道のことであるが、木にはサイクルがない。

ネットワークを木の形に集約するメリットとしては、任意に 2 つのノードを選ぶと、その間の道が一意的に決まるという性質がある。サイクルがあると一意的に決まらなくなる。これは情報を表現する上で非常に便利な性質である。一方で、木にするデメリットは、1 点にフローが集中してしまいボトルネックとなる可能性が高いという点である。

グラフに含まれるすべてのノードをつなぐ木のことを特に「全域木」と呼んでいる。グラフのノードの数を  $n$  とすると、全域木の辺の数は  $(n-1)$  であることが知られている。

全域木で、各辺に距離に相当する重みを割り当てる場合には、距離の合計を最小化する（すなわちネットワークを張るコストを最小化する）という問題がよく取り上げられる。これは「最小全域木(minimum spanning tree)」を求める問題である。この問題の解法については、クラスカルやプリムのアルゴリズムなどいくつかの方法がある。<sup>1)</sup> 一方、上と同等の問題であるが、各辺に流量を重みとして割り当てる場合には、流量を最大化する問題として定式化できる。

木には中心となるノードがあるだろうか。各辺に距離の重みが与えられていれば、計算できる。他のノードからの距離の和を最小にするノードを求めれば、それが中心となるはずである。この中心付近は、フローが集中する重要な地点と考えられるが、逆にボトルネックともなる可能性がある。

木にはもともとノード間の上下関係はない。ただ「根(root)」を決めるとノード間の順位を決めることができるようになる。これは根となるノードからいくつの辺を経て結び付いているかによって決まる。2 つのノードの間には一意的な道があり、通った辺の数を距離とすれば、一意に根からの距離が求められるからである。根となるノードは、単純に考えれば中心となる要素、すなわち最大規模の要素をとればよい。

グラフ理論的モデルで、階層構造があるモデルでは、対象間に上下関係があることになり、これは重力が働いている空間のイメージで考えることができる。これに対し、単につながっているかどうかしか問題にならない場合、無重量状態のように、ふわふわと漂っているイメージとしてとらえられる。このように階層構造は、グラフ理論的モデルに秩序を与える重要な手段と考えられる。視覚的表現においても、階層構造を仮定するのは有効な方法だと思われる。

図 3 は、根をもった全域木で階層構造を表した例である。

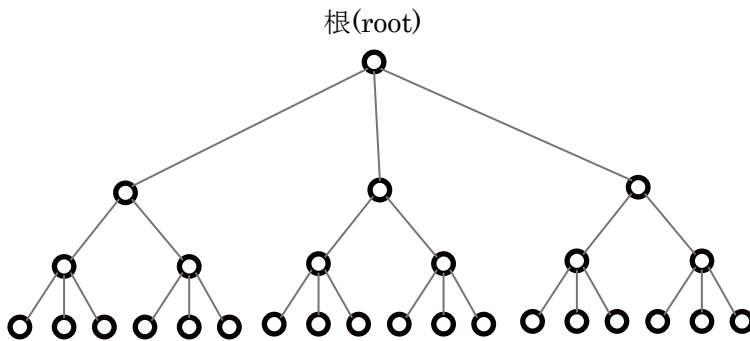


図3 階層的全域木モデル

(2) DAG モデル

次にグラフが木でない場合について考えてみよう。つまりサイクルが存在する場合である。グラフ上のノードに順位をつけられるためには、グラフが **DAG(Directed Acyclic Graph=有向無サイクルグラフ)**であることが必要である。ノードに順位をつけるとは、ノードに大小関係を付けて離散的な重みを割り振ることに相当する。これにより階層構造ができていく。

階層構造を仮定することにより対象間を結ぶ辺の数を減らすことが可能となる。たとえば、 $A>B$ ,  $B>C$  が成り立てば、当然  $A>C$  であるから、これに相当する辺は省略することができる。これは大小関係においては推移律が成り立つからである。以上のような原理で、グラフを単純化し、階層構造を明らかにするモデルとしては、**ISM(Interpretive Structural Modeling)**と呼ばれる手法が有名である。<sup>3)</sup>

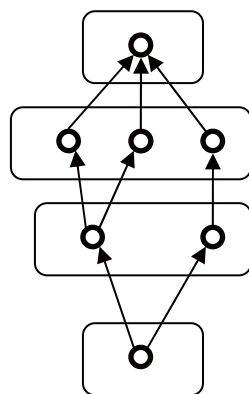


図4 DAG(Directed Acyclic Graph)の例

図4は、DAGの例を示したものである。図の中の枠は階層構造を表しており、この場合、4段階の階層になっている。

### (3) スモールワールドモデル

現在、ネットワークのモデルとして研究が進んでいるものとして、スモールワールドネットワークがある。これは、グラフ理論の手法で、人間同士の知り合い関係をグラフに表したとき、見かけ上遠い人に、予想以上に近い距離で到達できるという現象（スモールワールド性）を説明できるモデルとして注目された。1990年代後半、ワッツとストロガッツによりこの疑問に明確な答えが与えられた。<sup>4)</sup> 彼らはネットワークを格子構造からランダム構造に徐々に変形していくモデルを考え、その中間段階において平均距離は小さいが、クラスター性が大きいという性質を持つネットワークが存在することを示し、このような性質を有するネットワークを「スモールワールドネットワーク」と呼ぶことにした。ただランダムネットワークで作られたモデルは人工的なものであり、現実の世界に当てはまるかどうかは疑問であった。

1999年にバラバシらの研究グループは、WWWのリンク関係をグラフとして解析することにより、ページに含まれるリンク数とそのページ数の間には、べき乗則が成り立っていることを示した。<sup>5)</sup> べき乗則で構造が決まるシステムは、特徴的な大きさを持たないという意味で「スケールフリー(scale-free)」と呼ばれている。つまり現実の世界は「スケールフリーネットワーク」であり、単純なランダムネットワークではないということが明らかになったのである。このスケールフリーネットワークの世界では、多くのページからリンクされているページは有名なページであり、ますます多くのリンクを引き付ける。つまり有名なページは、ますます有名になるという結果になる。情報が多く集まる場所には、さらに情報が集まり、逆に情報の少ないところはますますさびれていく、と言ってもいいかもしれない。現実の世界では、特に規制がない自然な状態では、このような状況にしばしば遭遇する。

スケールフリーネットワークの特徴は、ところどころに巨大なクラスタが現れるという点である。これは結節点といってもよく、グラフの辺が密集している。ランダムネットワークを仮定すると、もっと一様に分布したグラフとなる。

## 4. 自然なネットワーク構造の視覚化

ネットワークのモデルは、つながっているかどうかだけを見る0-1の「デジタル型モデル」と、重み付きグラフを使った「アナログ型モデル」に分けて考えることができる。またこの中間に位置するものとして、離散的な段階の重みを持つモデルも考えることができる。

単純化してモデルを考える場合は、もともとのデータがアナログデータであっても、何らかの変換によって、デジタル型モデルにあてはめることは可能である。このときのデジタル変換は、単純に閾値をもうけて切り捨てるだけでは、地域間の構造をうまく表現でき

ない可能性があるので、いろいろな手法を試してみる必要がある。本来アナログ的なデータの場合には、そのまま分析し、視覚化できる手法が望ましいと考えられる。この場合、情報量的な視点からデータを要約して、大まかな構造を抽出することが求められる。このための手法についても研究していく必要があるだろう。

現実の世界に現れるスケールフリー性を、ネットワークの問題として地域間データに当てはめて、データの骨組みにあたる構造を抽出するにはどうすればよいか。ある程度、統計的な処理が必要と考えられるので、大規模なネットワークで分析する必要がありそうである。

## 5. おわりに

生物界にみられる現象は、いろいろ興味深いものがあり、人間社会の活動にも応用できる可能性がある。単にアナロジーとして当てはめるだけでなく、その根拠となる物理法則や制約条件を理解してモデル化できたら、応用範囲は大きく広がると思われる。大都市の成長過程や、通信ネットワークの普及過程など、極めて大きな時間的・空間的なスケールで起こる現象にも、人間の個人の意思を超えた最適化の力が働いていると考えられる。

## 参考文献

- 1) R.J.ウィルソン, J.J.ワトキンス (大石泰彦 訳)「グラフ理論へのアプローチ」日本評論社 (1997)
- 2) フィリップ・ボール(桃井緑美子 訳)「枝分かれ — 自然が創り出す美しいパターン」早川書房 (2012)
- 3) 樫木義一・河村和彦「参加型システムズ・アプローチ — 手法と応用」日刊工業新聞社 (1981)
- 4) ダンカン・ワッツ (栗原聡・佐藤進也・福田健介 訳)「スモールワールド — ネットワークの構造とダイナミクス」東京電機大学出版局 (2006)
- 5) バラバシ (青木薫 訳)「新ネットワーク思考 — 世界の仕組みを読み解く」NHK出版 (2002)
- 6) 中村有一「地域間情報交流の分析手法について」経営・情報研究 (多摩大学研究紀要) 研究ノート 2010, No.14, pp.41-54 (2010)
- 7) 中村有一「大都市圏における情報交流の分析手法」経営・情報研究 (多摩大学研究紀要) 研究論文 2007, No.11, pp.15-25 (2007)
- 8) 中村有一・高嶋裕一「通話トラヒックデータに基づく単料金区域の統合化手法に関する考察」TIMIS Journal No.32 (1993)