

大学入学人口移動空間の視覚化

古藤 浩

大学入学による都道府県単位の人口移動データから、視覚化技術を用いて地域構造を示す方法を提案し、実際にその構造を議論する。本研究では、分析の枠組みとして冪型ハフモデルを導入し、2種類の未知数の与え方で推定を行い段階的に分析する。最初の段階ではハフモデルで重要となる「地域の魅力」を検討し、次の段階で時間地図等の作成時に使われる考え方を応用して、大学入学による人口移動空間を仮想地図（変形地図）によって視覚化する。そして、仮想地図の時系列的な変遷を考察する。なお、ここでは内々移動のデータの扱いも重視して考察を進めるところにも特徴がある。

キーワード：ハフモデル、大学入学、視覚化、最尤法

1. はじめに

本研究では、大学入学による都道府県（以下では省略して「県」と記述する）単位の人口移動データから、その地域構造を明らかにし、それを視覚化する新しい方法を提案し、その構造を議論する。

少子化傾向のため大学入試を廻る募集環境は大学側にとって年々厳しくなる一方である。文部科学省等[1]によれば、2007年には大学及び短大入学希望者数が入学定員数を下回ると予測されている。このような新しい時代に対応するため、多くの大学では、大学の教育の質の向上・募集活動の充実にむけた努力を進めている。募集活動での地域的な視点では、活動の重点をどこにどのような順序で置くかのウェイト設定は重要である。そのための補助資料、新しい交通機関の整備による進学先配分の変化の予測などに役立つような視覚化結果を得ることが本研究の目的である。

2. 研究の概要

本研究では、大学入学者数の分析の枠組みとして冪型ハフモデルを導入し、異なる2種類の未知数の与え方で推定を行い分析する。最初の未知数の与え方ではハフモデルで重要となる「地域の魅力」を検討し、次に、時間地図や通話地図（例えば、古藤・長谷川[2]）等の作成時に使われる考え方を応用して、大学入学時における人口移動を最もよく表すような仮想地図を作成できるように未知数を設定する。そして、仮想地図

の時系列的な変遷を視覚的に考察する。

大学入学による人口移動のデータには学校基本調査[3]を用いた。大学入学時の人口移動人数は距離に応じて減少し、それは一般の人口移動と同様である。一方、大学入試の問題を考えると入学者の総数をある程度正しく推測・予測することも重要である。総数を考える場合、人数が最も多いのは同県内進学者数なので、移動しない、県内々の移動（入学）人数にも注目する必要がある。そこで、本研究では内々移動のデータの扱いも重視して考えていく。

なお、本研究では大学と短期大学を合わせ、まとめて「大学」として分析する。

3. 大学進学による人口移動の概況

まず、学校基本調査[3]で公表されている大学進学による都道府県間の人口移動の状況を2005年度入試の場合を中心に簡潔に説明する。大学入学者数は1993年度入試の81万人から、2005年度入試には71万5千人に減少した。都道府県間の出入を散布図で表すと図1になる。ここで点線は転出と転入が等しい場合を示す。出身者の多い順に東京都、大阪府、神奈川県、愛知県、埼玉県、兵庫県、千葉県の順になる。

全体的傾向として大学入学によって、地域から大都市圏に移動する傾向がわかる。流入傾向が特に顕著なのは東京都・京都府である。また、大都市の周辺は大学志願者が多いが、大学も多いため入学者数も多く、兵庫県や千葉県など、出入人数がほぼ等しい傾向が読み取れる。

データは47都道府県間の行列になるのでデータ数は2,209(=47²)となる。都道府県間の移動人数の

ことう ひろし

東北芸術工科大学 デザイン工学部
〒990-9530 山形市上桜田200番地

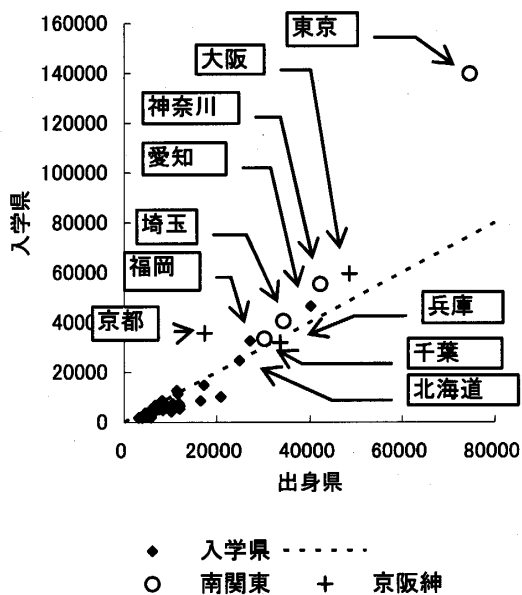


図1 高校生の大学入学による人口移動：出入人数

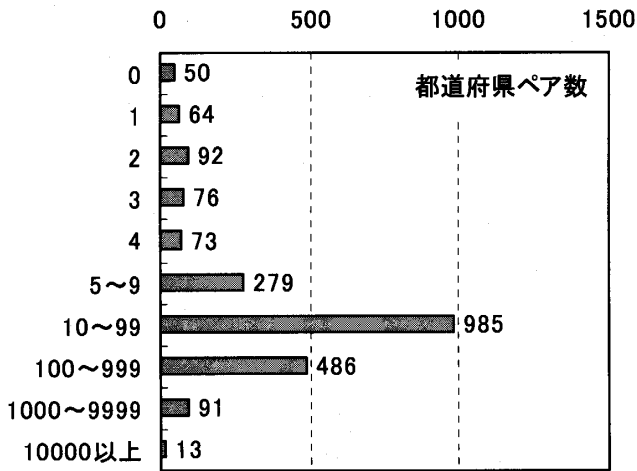


図2 都道府県間ペアの移動人数のヒストグラム

ヒストグラムを作成すると図2になる。71%の都道府県ペアで10人以上、84%の都道府県ペアで5人以上の移動があり、全国規模の移動とわかる。なお、2005年度入試での大検合格者の入学者数は2万人弱（入学者全体の2.8%）いるが、この人数は本分析では使わなかった。

4. モデル

4.1 モデルの概形

ハフモデルとは、客が店舗を選択する確率を、店舗の魅力と客と店舗間の距離によって説明する立地分析のための数理モデルである[4]。そのアナロジーをここでは大学入学に応用する。まず、各県（進学候補地）の大学の総和としての魅力を考える。次に志願者出身県と進学先候補地との距離と候補地の魅力によ

て決まる県間の引力を定義する。A県からB県への大学入学者数は、全県のA県への引力の合計を分母、B県のA県への引力を分子とする確率によって与える。

まず、記号を次のように定義する。

m : 県の数 (=47)

S_j : 県 j の魅力 (>0)

d_{ij} : 県 i から県 j への距離

f_{ij} : 県 i の志願者を県 j の大学が引く力

n_{ij} : 県 i 出身者の、県 j の大学への入学者数

α : 距離に関する減衰パラメータ (>0)

本研究では冪型のハフモデルを基本とする。すなわち県 i の志願者を県 j の大学が入学させようとする引力を

$$f_{ij} = S_j d_{ij}^{-\alpha} \quad (1)$$

とし、県 i の志願者が県 j の大学に入学する確率 \hat{p}_{ij} を

$$\hat{p}_{ij} = \frac{f_{ij}}{\sum_{k=1}^m f_{ik}} \quad (2)$$

と与えるモデルである。

県 i の志願者数を o_i とするならば、県 i から県 j の大学への入学者数推定値 \hat{n}_{ij} は以下に与えられる。

$$\hat{n}_{ij} = o_i \hat{p}_{ij}$$

以上によるモデルでの未知数の推定は、ロジットモデルの枠組みを使い、同時確率の最大化、すなわち最尤推定法を用いて行う[5]。すなわち尤度関数 L ならびに対数尤度 $\ln L$ を以下のように定義し、対数尤度関数 $\ln L$ の最大化を行う。尤度関数 L および、目的関数 F は

$$\begin{aligned} L &= \prod_{i=1}^m \prod_{j=1}^m \hat{p}_{ij}^{n_{ij}} \\ \max F &= \ln L = \sum_i \sum_j n_{ij} \ln \hat{p}_{ij} \\ &= \sum_i \sum_j n_{ij} \ln \frac{S_j d_{ij}^{-\alpha}}{\sum_{k=1}^m S_k d_{ik}^{-\alpha}} \quad (3) \end{aligned}$$

と書ける。実例での計算は、逐次二次計画法を用いた。

4.2 距離の与え方について

一部の例外を除き、大学入学者の人数で一番多いのは自県出身者なので、大学入学の分析では、県内々移動による入学者数の推計が重要である。また、近県への移動を含め、実人数をなるべく正確に推定できるようモデルを考えたい。そのため、内々距離や近県間の距離の与え方を工夫しなくてはならない。内々距離を決める要因には、県の形状や人口分布、大学の分布状

況など様々に考えられるが、ここでは領域間平均距離の概念[6]を活用する。

それは、各県が県庁所在地を中心とする（真の面積と）等面積の円であると仮定し、その場合の面対面の領域間平均距離（一次近似）を使う方法である。すなわち大学志願者と大学は県内にまんべんなく分布していると考えた距離を用いた。距離を求める県 i, j の面積を z_i, z_j 、県庁所在地間の直線距離を d_{ij} とすれば、以下のように与えられる。

$$\text{内々} : \frac{128\sqrt{z_i}}{45\pi^{1.5}}, \quad \text{県間} : d_{ij} + \frac{z_i + z_j}{8\pi d_{ij}} \quad (4)$$

県間距離の第二項は直線距離 d_{ij} が大きくなるのに反比例して小さくなる。言い換えれば近隣県間ほど第二項の効果が強く出る。また、面対点（県庁所在地）の平均距離なども比較して最終的に(4)式の利用を決めた。

4.3 何を未知数とするか

次に、式(3)に対するデータの当てはめで、何を県の魅力とし、何を未知数とするかを検討する。また、大学入学による空間構造を視覚化したいので、最終的に県の位置や面積を未知数として考えていきたい。そこで、本研究では次のような手順で分析を進めた。

- 手順1：県庁所在地の真の位置と面積を所与とし、距離減衰パラメータ・各県の魅力を未知数として式(3)を計算する
- 手順2：手順1による魅力を外生的に与えるとしたら何によって最もよく近似できるかを考察する。
- 手順3：手順2の結果を用い、距離減衰パラメータ・県の魅力を所与として、県の位置・面積を未知数として式(3)を計算する。

手順3によって、大学入学での人口移動から見た人口移動の地域構造が視覚化される。時系列データでの比較のため、距離減衰パラメータは所与とし、固定して計算した。

5. 分析結果

5.1 地理的關係を固定したハフモデルによる分析

手順1の結果、 $\alpha=1.916$ 、真の値と推定値との決定係数は0.977となり、その当てはまりの様子は図3に示される。かなりの説明力に見えるが、都道府県個別に調べると、中小の県ではかなりの誤差があることがわかった。人口規模が東京都の10分の1未満の県が出身地ベースで19県、入学県ベースで36県と、規模の分散が大きいことがその理由と考えられる。なお、推定された魅力は東京都、北海道、神奈川県、千葉県

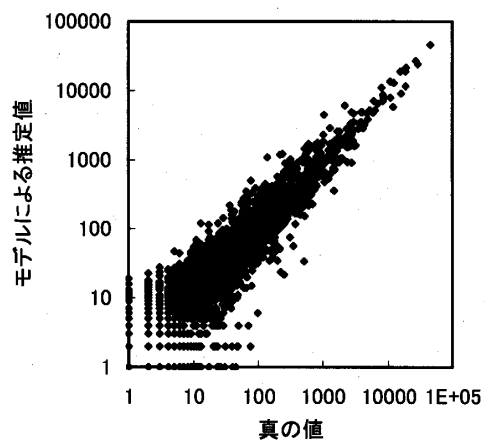


図3 ハフモデルの適用結果（対数尺）

の順で大きな値となった。値が大きい方から、関東、愛知、近畿、その他の順となり、大都市圏の魅力値が大きくなった。

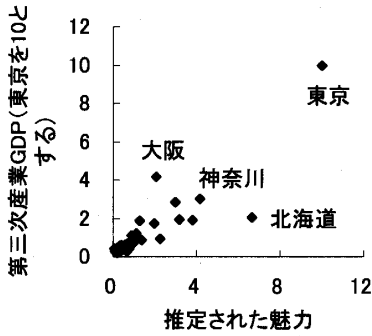
手順2では手順1で推定された各県の魅力を表す社会指標を考える。候補として人口、高等学校卒業者数、第三次産業GDP、全産業GDP、小売業販売額、会社数、図書館数などをとりあげて調べた。これは二つの目的があり、一つは大学入学人口移動の地域構造を視覚化するときになるべく安定した結果を得るため、もう一つは本モデルを活用して将来の大学入学人口移動を推定する場合に、何らかの形でそのときの各県の魅力を予測して与える必要があるためである。

分析の結果から、県の魅力を第三次産業による県内総生産額（入試年度の2年前の値）と決めた。手順1でのモデルから導かれた県の魅力と、上記指標の相関係数は0.883となった（2005年度入試）。その関係の様子は図4に示される。これより大きな相関係数を得る指標も候補の中にはあったが、中小県の魅力を評価するためのデータに対数をとった場合の比較（図4(b)）や、第三段階のモデルでの魅力を様々に与えて分析した結果、県の魅力として最も適切なのは第三次産業による県内総生産額と判断した。

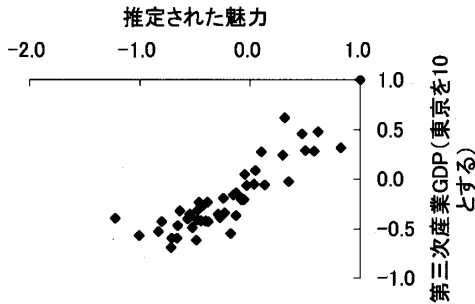
5.2 仮想地図作成の目的関数

手順3では、仮想地図の正則性を保つため、東京を原点とし、沖縄のy座標を0と固定した。これによって地図の比較もたやすくなる。距離減衰パラメータ α は1.8に固定した。推定での目的関数は、式(5)のように書くことができる。

$$\begin{aligned} \max F(\mathbf{z}, \mathbf{x}, \mathbf{y}) &= \ln L = \sum_i \sum_j n_{ij} \ln \hat{p}_{ij} \\ &= \sum_i \sum_j n_{ij} \ln \frac{S_j d_{ij}^\alpha}{\sum_{k=1}^m S_k d_{ik}^\alpha} \end{aligned}$$



(a) 実数尺で見た散布傾向



(b) 対数尺上での散布傾向

図4 手順1のモデルでの魅力と第三次産業GDP

$$\begin{aligned} \mathbf{z} &= \{z_1, z_2, \dots, z_{47}\}, \\ \mathbf{x} &= \{x_1, x_2, \dots, x_{46}\}, \\ \mathbf{y} &= \{y_1, y_2, \dots, y_{45}\} \end{aligned} \quad (5)$$

$$d_{ii} = \frac{128\sqrt{z_i}}{45\pi^{1.5}},$$

$$d_{ij} = \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} + \frac{z_i + z_j}{8\pi\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}}, \quad (i \neq j)$$

未知数の数は138と多く、大域的な最適解という保証はない。そのため、本研究での計算例では3種類の初期値を与えて計算し、その中での最適な結果を選んだ。日本の物理的な値、別な目的関数で日本の仮想地図を作ったときの値、一定の範囲内の乱数の3種類の初期値で計算したが、多くの場合で、物理的な日本の配置を初期値とした場合の結果が最も良かった。

距離減衰パラメータ α を1.8に決めるに当たっては、パラメータ α を様々に変えて目的関数の値を調べた。1993年度入試と2005年度入試の2時点についてその結果を図5に示す。図5では各年度の結果について最良の値に対する目的関数の悪化の具合を比率で示している。1993年度入試よりも2005年度入試で、パラメータ α が少し大きいように見えるが、目的関数の値は1.7~1.9の範囲でそれほど変化しない。以上から、年度にかかわらず $\alpha=1.8$ と定めた。

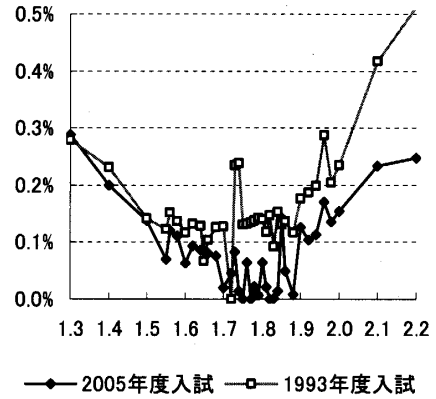


図5 パラメータ α と目的関数値：最良の結果からの乖離の比率

どの年度でも推定結果と実際の値の決定係数は0.999以上だった。ただし、未知数が多いので当然とも言えるだろう。

5.3 仮想地図の大きさについて

式(5)による仮想地図の全体的なサイズは一般的な時間地図のような、地域の近接の程度を意味せず、近隣県との関係の強さを表す。すなわち近隣県との関係が強ければ仮想地図は大きくなり、近隣県との関係が弱ければ仮想地図は小さくなる。これは次のような数式の展開によって示される。式(5)で地図の大きさ(縦方向・横方向とも)を q 倍して、目的関数値がどう変化するかを考察する。そのとき目的関数 F の値は次の式(6)のように整理できる。

$$\begin{aligned} & \sum_i \sum_j n_{ij} \ln \bar{f}_{ij} \\ &= \sum_{i,j;i \neq j} n_{ij} \ln \left[\frac{S_j \left(\sqrt{(qx_i - qx_j)^2 + (qy_i - qy_j)^2} + \frac{z_i + z_j}{8\pi\sqrt{(qx_i - qx_j)^2 + (qy_i - qy_j)^2}} \right)^{-\alpha}}{\sum_{k=1}^m S_k \left(\sqrt{(qx_i - qx_k)^2 + (qy_i - qy_k)^2} + \frac{z_i + z_j}{8\pi\sqrt{(qx_i - qx_k)^2 + (qy_i - qy_k)^2}} \right)^{-\alpha}} \right] \\ &+ \sum_i n_{ii} \ln \frac{S_j \left(\frac{128\sqrt{z_i}}{45\pi^{1.5}} \right)^{-\alpha}}{\sum_{k=1}^m S_k \left(\frac{128\sqrt{z_i}}{45\pi^{1.5}} \right)^{-\alpha}} \\ &= \sum_{i,j;i \neq j} n_{ij} \ln \left[\frac{S_j \left(\sqrt{(x_i - x_j)^2 + (y_i - y_j)^2} + \frac{z_i + z_j}{8\pi q^2 \sqrt{(x_i - x_j)^2 + (y_i - y_j)^2}} \right)^{-\alpha}}{\sum_{k=1}^m S_k \left(\sqrt{(x_i - x_k)^2 + (y_i - y_k)^2} + \frac{z_i + z_j}{8\pi q^2 \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2}} \right)^{-\alpha}} \right] \\ &+ \sum_i n_{ii} \ln \frac{S_j \left(\frac{128\sqrt{z_i}}{45\pi^{1.5}} \right)^{-\alpha}}{\sum_{k=1}^m S_k \left(\frac{128\sqrt{z_i}}{45\pi^{1.5}} \right)^{-\alpha}} \end{aligned} \quad (6)$$

式(6)の変形後を見ると、ピタゴラスの定理に等しい距離の式の部分では q は消え、 q が残っているのは県間距離での面積 z が関係する項の分母のみとわかる。

そこから、

- (1) もしも領域間平均距離ではなく都道府県庁所在地間の直線距離で計算するならば、仮想地図の大きさは推定精度に無関係、
- (2) 地図サイズを q 倍にすると目的関数の値が大きくなるならば、それは領域間平均距離での面積の影響を q^2 分の 1 にした方がよい推定結果を得る、

ことを意味する。つまり、仮想地図が大きい (= q が大きい) ほど隣接県など近県が領域間平均距離よりも近く、近県への入学が相対的に多い傾向を意味する。

5.4 大学入学人口移動から見る日本の構造

以上をふまえ、式(5)による 1993 年度、1999 年度、2005 年度入試での大学入学の仮想地図を図 6、図 7、図 8 に示す。ここで各県を示す円の面積は推定された面積と等しく、円の中心の黒丸(点)の大きさは県の魅力に比例した面積となっている。また、新幹線経路を線で示した。比較のためこの描き方による物理的な日本地図も図 9 に示す。ここから、次のように空間構造を読み取ることができる。

□一般的に考えられる性質

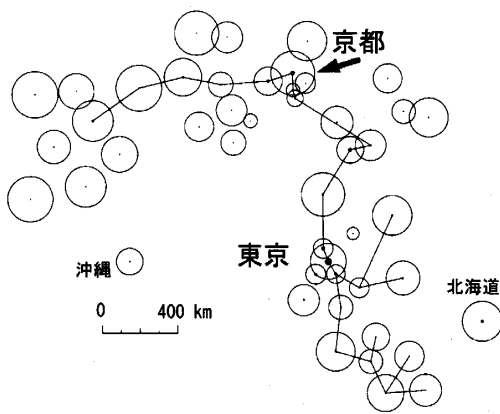


図 6 大学入学の仮想地図 (2005)

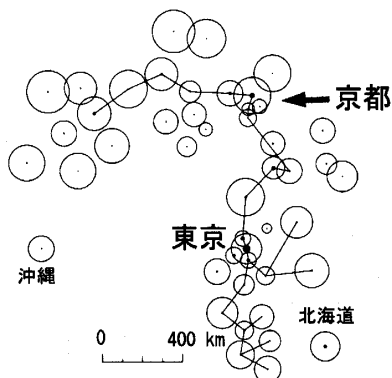


図 7 大学入学の仮想地図 (1999)

- (A) 大学入学する確率が高ければ、関係する県間距離は短くなる→交通の便がよければ物理的な距離よりも近接する。時間地図と似た配置になる可能性がある。ただし、金銭的な影響もあるだろう。
- (B) 一部地域に志願者人口が集中すれば、面積は小さくなる。
(B2) 出身地から外に出る傾向が強ければ、自県は相対的に遠いことになり、面積は広くなる。
- (C) 地図の全体サイズが大きいほど、近隣県への移動確率が高く、県外に出る場合に近県を選ぶ傾向がある。

□仮想地図からの知見

・年度に関わらない全体的傾向

- (i) 仮想地図は物理的な地図よりも大きくなった。県外に出る場合に近隣県に移動する傾向が強くなることになる。ただし、大学は都市部に立地する機会が多いため、面対面の近似式(式(4))ほどには距離が効かないことを意味しているのかもしれない。
- (ii) 多くの県では面積が大きくなった。自県が大きいということは、自県が相対的に遠いことを示すので県外に出る傾向がわかる。これは図 1 からも窺えた傾向である。
- (iii) 北海道など一部の県では面積が小さくなり、全体として県面積の分散が小さくなった。県の面積にかかわらず、県というバリアの影響がどの県でも同じようにあると見えそうである。

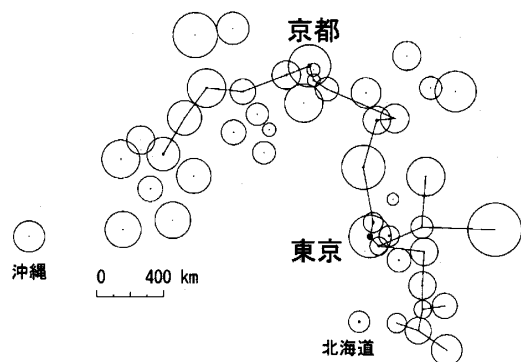


図 8 大学入学の仮想地図 (1993)

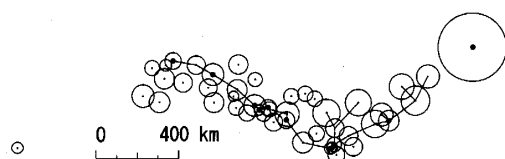


図 9 物理的な日本地図

- (iv) 東京都・京都府は特に広がった。県内に広く人口が分布していることや、自県内にそれほどは入学していない、県外に入学する確率が高いことを示している。
- (v) 日本全体が大きくS字型(2005年, 1999年), C型(1993)にくびれた形となった。全体としては遠距離がそれほど効かず、九州や北海道が関東に近い(多く関東の大学に入学し)傾向といえる。一方で四国は近畿に近くなった。このような形になる傾向は、通話など他の交通流の視覚化研究(例えば 古藤・長谷川(2004))でも見ることができる。

・経年的な変化

- (vi) 北海道の位置に注目すると、93年はC型の内側に位置したのが、99年、2005年では外側、特に西日本から離れた位置となった。
- (vii) 北海道の面積は、93年では物理的な面積の4分の1と小さかったのが、年を経るごとに物理的な大きさに近づきつつある。93年では道内での進学が非常に多かったが、2005年では道外に出る傾向が強まっているといえるだろう。
- (viii) 沖縄の位置は、北海道とは逆に徐々に関東に近づきつつあるように見える。
- (ix) 滋賀県は90年代以降いくつかの大学の移転・新設が多くあり、大学定員は93年度入試時の2,700人が2005年度入試時点では7,600人と大幅に増え、定員も満たして入学者を確保している。その影響が図に現れ、93年度では京都に含まれている位置にあったのが、2005年度では京都から離れ独立した位置になった。滋賀県出身で滋賀県内に進学する学生が増え、隣県である京都から離れる位置となったのだろう。

6. おわりに

以上によって大学入学による人口移動を対象にその構造を視覚化する方法と、その結果の吟味の結果の説

明を終わる。本研究はロジットモデルの枠組みを使って視覚化した。今回の同時確率の最大化という考えによる仮想地図の作成という方法は、一つの新しい試みであり有意義だろう。ただし、結果としての地図では地図が大きいことが「全国が遠い」ことではなく「相対的に近隣県が近い」ことを意味するという点で、直感的には少しわかりにくい仮想地図となってしまった。普通に重力モデルを用いて、そこから逆算して県の位置を求めることも検討し、比較して考察したい。

本研究での視覚化では内々距離や近県間の距離の扱いに力を入れた。本研究の方法・表示方法はその一つの例となると考えている。

本仮想地図の用途は、各大学で募集戦略を考えるとときに、どの地域に重点を置くべきか、新しい交通機関ができる時などに、その効果がどのように現れるかの推定・考察に活用できると考えている。そのための今後の展開として、大学・短大を分離しての分析もしていきたい。

参考文献

- [1] 文部科学省 中央教育審議会(2005): 我が国の高等教育の将来像(答申), http://www.mext.go.jp/b_menu/shingi/chukyo/chukyo0/toushin/05013101.htm
- [2] 古藤浩, 長谷川文雄(2004): 逆算距離を利用した通話地図による日本の構造分析. GIS—理論と応用, Vol. 12, No. 2, pp. 165-175.
- [3] 文部科学省: 学校基本調査(1993, 1999, 2005速報値), http://www.mext.go.jp/b_menu/toukei/001/index01.htm
- [4] Huff, D. L. (1964): Defining and Estimating a Trading Area, *Journal of Marketing*, Vol. 28, No. 3, pp. 34-38.
- [5] 土木学会編(1995): 非集計行動モデルの理論と実際, 土木学会.
- [6] 栗田治・腰塚武志(1988): 領域間平均距離の近似理論とその応用, 都市計画論文集, No. 23, pp. 43-48.