

年齢構成の共通性に着目した 地域社会の比較分析法

元井 初音¹・奥村 誠²・水谷 大二郎³

¹ 非会員 東北大学 大学院工学研究科土木工学専攻 (〒980-8579 仙台市青葉区荒巻字青葉 6-6)
E-mail: hatsune.motoi.t7@dc.tohoku.ac.jp

² 正会員 東北大学教授 災害科学国際研究所 (〒980-0845 仙台市青葉区荒巻字青葉 468-1)
E-mail: okumura@irides.tohoku.ac.jp

³ 正会員 東北大学助教 災害科学国際研究所 (〒980-0845 仙台市青葉区荒巻字青葉 468-1)
E-mail: mizutani@irides.tohoku.ac.jp

地域の社会・経済的な状況の変化や自然災害の影響を受けて、人口移動や残留の人数は経時的に変化する。これらの人口移動や残留は、進学、就職、住み替えなどの目的に応じて異なる年齢構成において発生し、その年齢構成パターンは多くの地域社会で共通している。年齢構成に基づいて、多くの都市圏に共通の都市間の移動、残留の要因を特定するとともに、個々の地域社会の固有性を把握できれば、各地域社会の政策検討のための有用な情報を得ることができると考えられる。本研究では、非負値行列因子分解により、地域社会ごとの転入、転出、残留の要因を年齢構成を用いて説明するとともに、それを他の地域と比較分析するための方法論を提案する。実際に最近の 3 回の国勢調査で得られた都道府県別の転入、転出、残留者数のデータに適用して、複数地域に共通する特性と個々の地域に特徴的な特性が抽出できることを示し、提案する方法論の有用性を検証する。

Key Words: *demography, regional community, non-negative matrix factorization, The Population Census*

1. はじめに

地域社会が何らかの原因で大きな影響を受け社会状況、経済状況が変化した場合、公共サービスや施設への新しいニーズが発生するほか、摩擦的に起こる問題の影響を和らげる政策等が必要となる。各地域の自治体は生じつつある変化を的確に把握し対応することが求められるが、特に未経験の変化だった場合には把握が難しい。このとき、そのような変化について、他地域や過去に観測された変化との類似性や共通性を部分的に把握できれば、今後の推移傾向や発生しうる社会的な問題とその対応策、今後重要となる公共サービスのニーズを考える際に有用となりうる。

地域の変化は居住者の転居に影響を与えることが多く、変化の原因によって異なる年齢階層に影響する。また、年齢階層ごとに人々の公共サービスに対するニーズも異なる。したがって、年齢構成比を通して地域社会の特徴を観察し、他地域や過去の年齢構成と比較の上で把握することでその地域で起こっている変動の原因を推測するとともに、今後発生する公共サービスのニーズを把握で

きる可能性がある。

多くの観測データの中に共通する変動パターンを、少数の要因の加算として把握する手法としてパターン認識が挙げられる。中でも、すべての要因が非負の要素を持ち、その加算の程度も非負の数値として把握する分析手法として、非負値行列因子分解 (Non-negative Matrix Factorization; NMF) がある。

本研究で着目する地域の転出、転入、残留者数は、測定誤差や記録時の誤差が含まれない限りは非負の値を取り、これらの数が負の値を取ることも許容した因子分析などの手法により地域間の転出入、残留を分析した場合には分析結果にバイアスが生じる可能性がある。そのため、本研究では NMF に基づいた地域特性の把握のための方法論を提案する。

2. 年齢構成データへの非負値行列因子分解の適用

NMF については数多くの既往研究や成書が存在するが²⁾土木計画学における利用事例は少ないため、本章で、

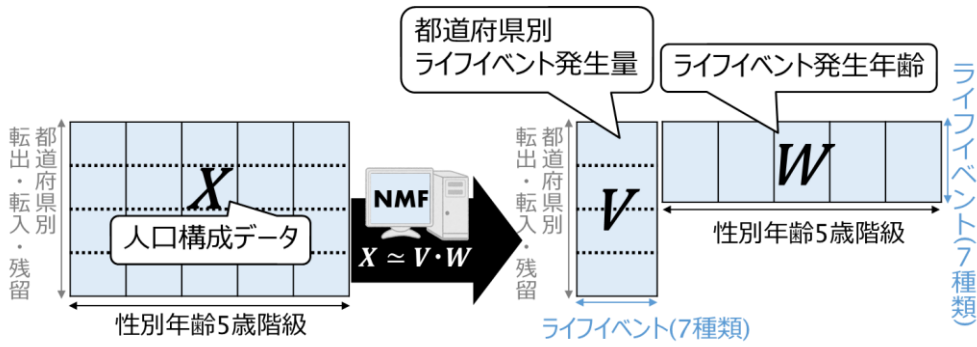


図-1 NMFの行列 X , V , W の関係

一般的な NMF の概要について説明するとともに、年齢構成データに適用する際の注意点を述べる。

(1) 非負値行列因子分解の概要

主成分分析や独立成分分析が所与のデータを複数の加法的な成分に分解する一方で、NMF は、非負値のデータを加法的な構成成分に分解することを目的とした多変量解析手法²⁾であり、得られる結果が全て非負値である点に最大の特徴がある。現実世界のデータの多くは非負値で表されるため、NMF によって音声や顔画像、文書内の単語の出現回数といった非負値データの構成成分を抽出し、音源分離や顔認証、文書の自動分類をする研究が進められている。

NMF は、サイズ $K \times N$ の観測行列 X を

$$V \cdot W \approx X \quad (1)$$

を満たすようなサイズ $K \times M$ の因子得点の行列 V とサイズ $M \times N$ の因子負荷量の行列 W に分解する図-1 のような操作である。なお、分解後の因子数 M は解析者が事前に決めておき³⁾、

$$M \leq \min(K, N) \quad (2)$$

を満たす必要がある。

(2) 分析対象データ

本研究において分析対象データは表-1 のように、性別各年齢階層（年齢 5 歳階級）の人口移動数を、各期間の各都道府県の転出、転入、残留ごとに並べた形式を採用する。ここで「転出」は 5 年前に当該都道府県に居住していた者のうち調査時点で他都道府県に居住地を移した人数である。また「転入」は調査時点で当該都道府県に居住する者のうち 5 年前に他都道府県に居住していた人数である。更に「残留」は調査時点と 5 年前で同一の都道府県に居住する者であり、住所を変えなかった者と

当該都道府県内で転居した者の両方が含まれる。なおこれらの定義より、調査時点で満 5 才に達していない者や当該 5 年間の間に死亡した者、国外との転出入は含まれない。都道府県を越えた転出入者と都道府県内残留を続けた者のどちらもそれぞれ年齢構成比を持っており、共通のライフイベントを持っている可能性があることから、個々の都道府県において、転出、転入、残留を別個に取り扱うこととした。

表-1 分析データの形式

[人]	5~10歳 男性	...	80~85歳 女性	85歳~ 女性
2010-2015年 北海道転出				
⋮				
2010-2015年 沖縄県転出				
2010-2015年 北海道転入				
⋮				
2010-2015年 沖縄県転入				
2010-2015年 北海道残留				
⋮				
2010-2015年 沖縄県残留				
2000-2005年 北海道転出				
⋮				
2000-2005年 沖縄県残留				
1995-2000年 北海道転出				
⋮				
1995-2000年 沖縄県残留				

(3) 因子の意味づけと因子数

本研究では、図-1 のように人口の変動をライフイベントの加算として把握することを目的とするため、得られた因子がライフイベントとして妥当であるかによって M を決定した。都道府県外へ転居するかどうかの選択は、基本的に入学・進学、職場の変化、結婚・離婚や住み替えといった重大なイベントごとに起こる⁹⁾。この人生の節目をライフイベントとする。分析による結果をライフイベントの数を変えながら比較し、最も妥当な解釈が可能であったライフイベントを7つとした場合の結果を採用し、 $M=7$ とした。

因子負荷量行列 V は、NMF の条件の一つである非負制約により疎になる傾向がある。すなわち、因子は互いに独立になるように求まる傾向にある²⁾。したがって、各ライフイベントがそれぞれ異なる年齢階層をカバーする傾向にあるため、分析結果の解釈はより容易になる。

(2) 非負値行列因子分解の目的関数

NMF は、生成行列 Y の要素 y_{kn} と観測行列 X の要素 x_{kn} の乖離度を非負制約の下で最小化することで、生成行列を観測行列に近似する。乖離度基準には、二乗誤差や I ダイバージェンス⁵⁾、板倉斎藤擬距離⁶⁾などがある。これらの乖離度基準を1つのパラメータ β で統一的に記述した β ダイバージェンスがあり²⁾、以下のように表される。

$$\beta \in \mathbb{R}, \beta \neq 0, \beta \neq 1$$

$$D_{\beta}(x|y) = x \frac{x^{\beta-1} - y^{\beta-1}}{\beta - 1} - \frac{x^{\beta} - y^{\beta}}{\beta} \quad (3)$$

この乖離度を用いることは、 $V \cdot W$ が Tweedie 分布に従うことに相当しており、そのパラメータ p を調整することにより、複合ポアソン分布に従う数値を扱うことができる。

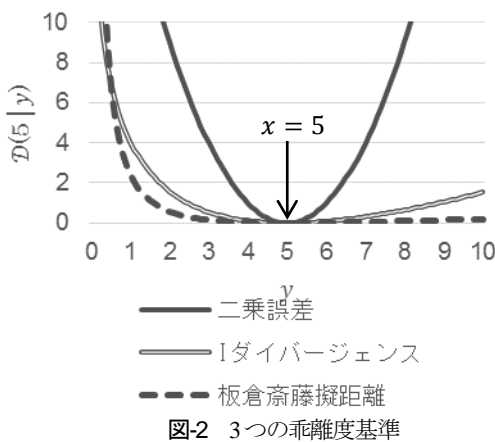


図-2 3つの乖離度基準

x_{kn} と y_{kn} の乖離度は $x_{kn}=5$ のとき、 y_{kn} の値によって図-2のように変化する。

転出、転入、残留の理由となるようなライフイベントはガンマ分布に従って起こり、一回あたりの転出、転入、残留の人数はポアソン分布に従う。したがって、転出、転入、残留は複合ポアソン分布にしたがって生成していると考えることができる。

3. 分析結果と考察

(1) 分析対象データの概要

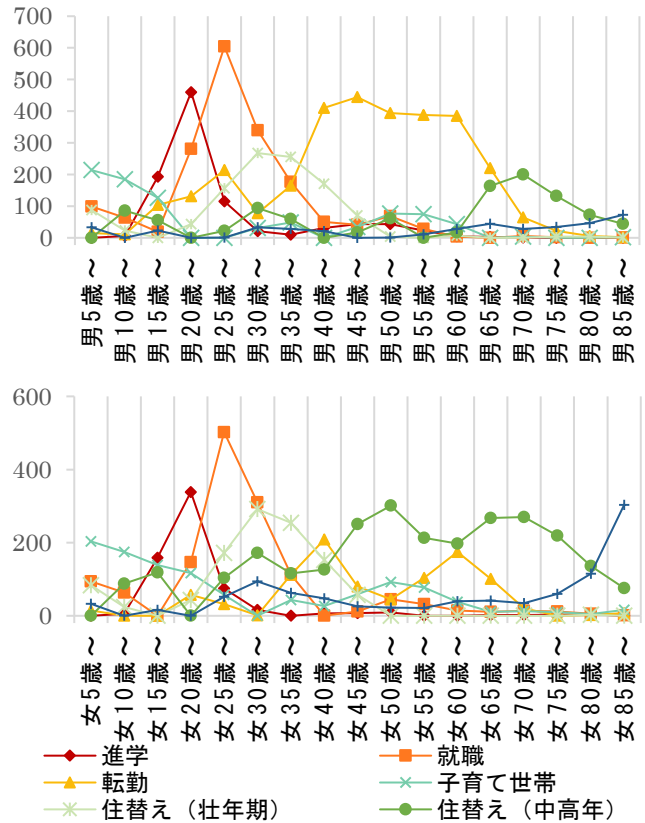


図-3 因子負荷量

表-3 各因子の解釈

因子名	ライフイベントとしての解釈
進学	高校卒業後の進学。
就職	高校、大学、短大、専門、大学院等からの卒業後の就職。
転勤	生産年齢人口で占められる。単身赴任は男性が多く上の世代では男性に偏っている。転職に伴う転居も含まれる。
子育て世帯	成人前後までの年齢の子供とその両親。
住替え(壮年期)	主に壮年期の男女。結婚、子供が増えるなどでより広い家へ引っ越し他、新社会人になり何年か経過し落ち着いた成人の住み替え、独立など。
住替え(中高年)	男性労働者の退職の時期に伴う住み替え。
後期高齢者	後期高齢者の介護を行う世帯、あるいは未亡人となった高齢女性。

日本において公的に調査され公開されている転出、転入、残留のデータとして毎年の住民基本台帳移動人口調査データと国勢調査のデータが挙げられる。しかし、住

民基本台帳は住民票の場所を記録したものであり住民票を移さない場合も多いため、より転出、転入、残留の実態を表している国勢調査を採用した。

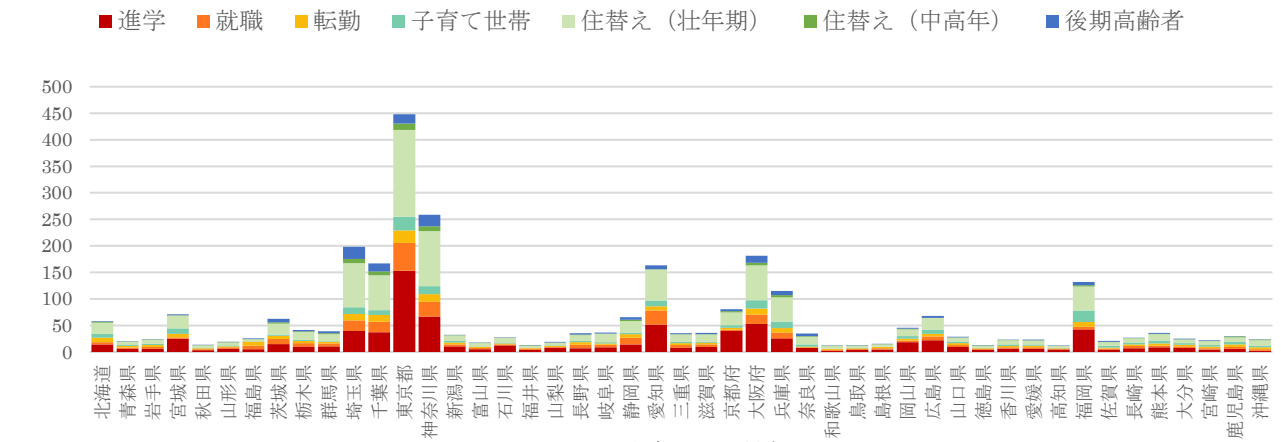
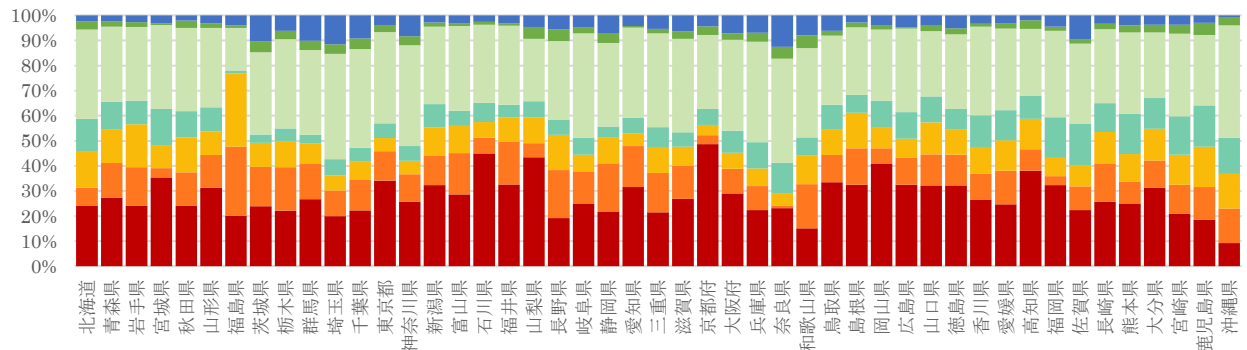
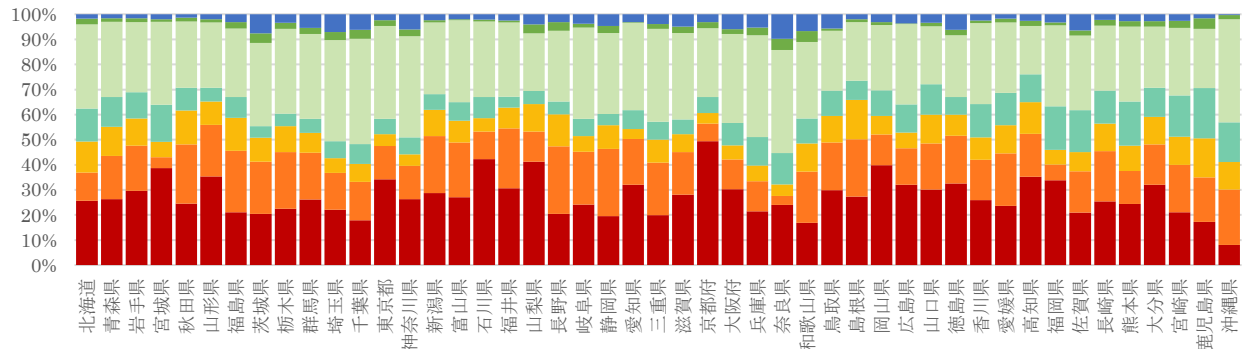


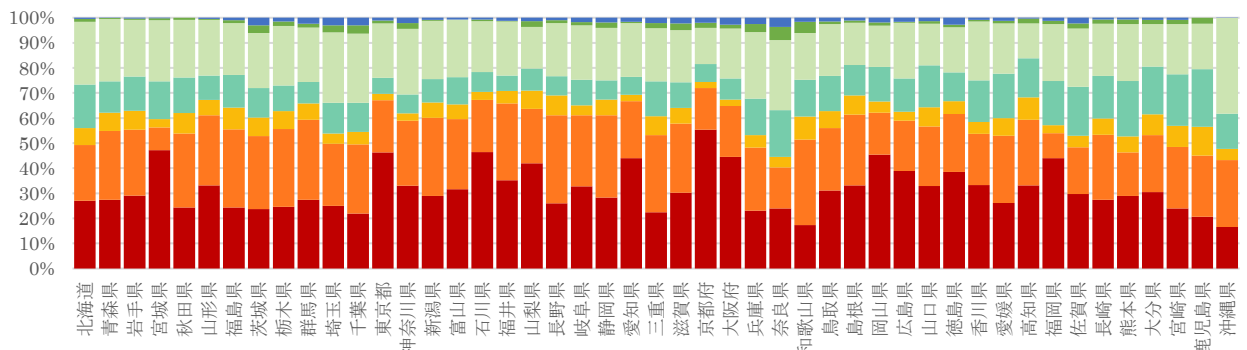
図-4 2010-2015年転入因子得点



2010-2015年転入



2005-2010年転入



1995-2000年転入

図-5 3期間転入因子得点 (100%積上げ)

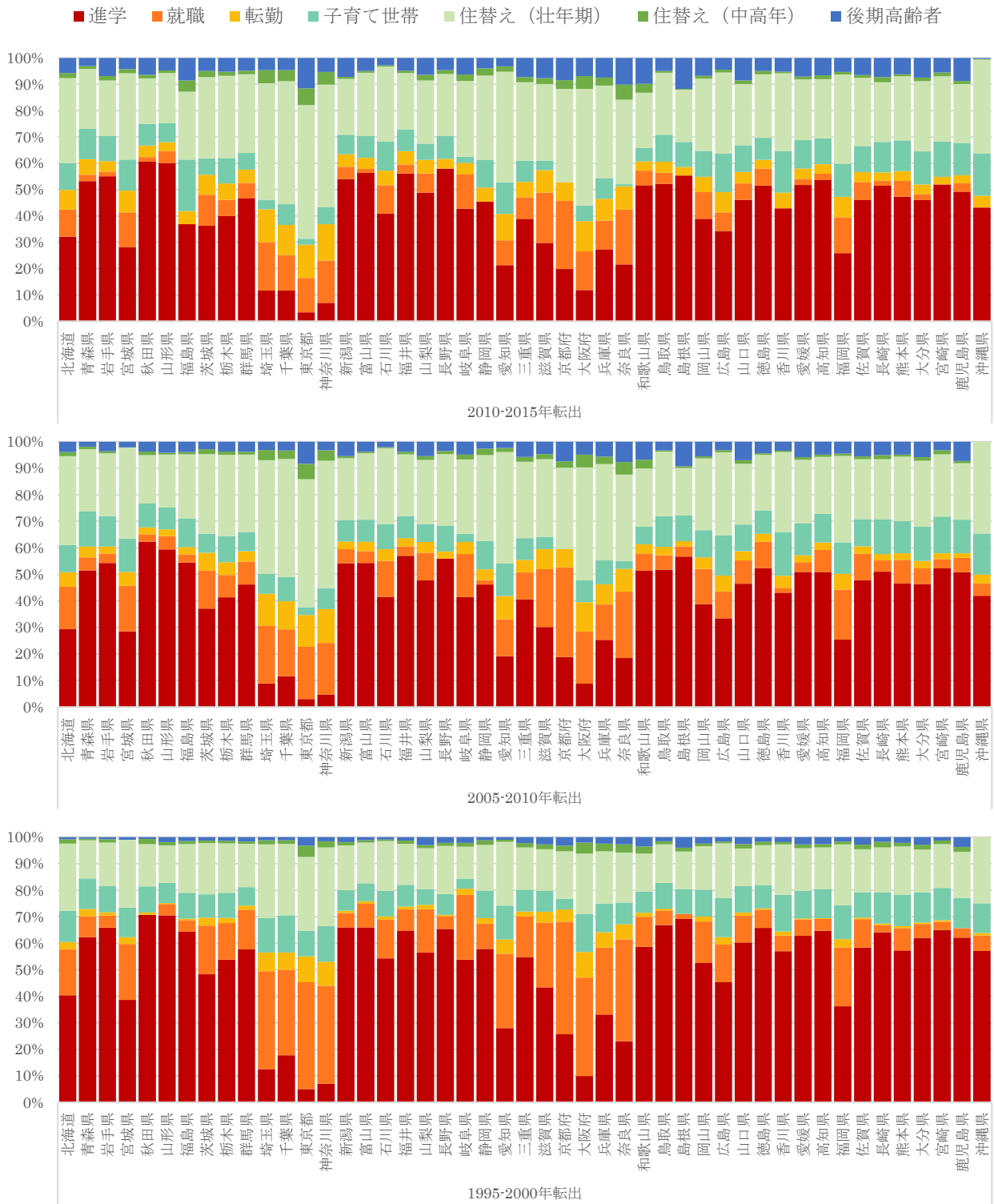


図-6 3期間転出因子得点 (100%積上げ)

国勢調査では西暦の末尾が0の年に大規模調査，末尾が5の年に簡易調査を行うが，2015年度調査では2011年に起こった東日本大震災の影響を把握するため従来大規模調査でのみ調査される「5年前の住居の所在地」項

目が加わっている¹²⁾。分析対象期間として，1995年から2015までの1995-2000年，2005-2010年，2010-2015年の3期間とした。

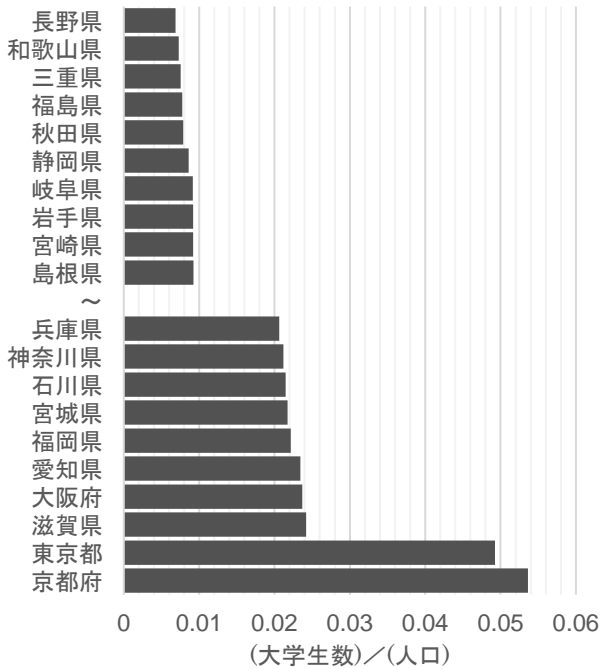


図-7 2010年度都道府県別大学生数対人口
(上位と下位各 10 都道府県)

(2) 分析結果と考察

各因子の因子負荷量は図-3 のように得られた。各因子は表-3 のようなライフイベントとして解釈し、それに基づいて各期間別、各都道府県別の転出、転入、残留の因子得点について以下で考察する。

a) 学都の都道府県 (一般的なライフイベント)

図-4 より「進学」する転入者が最も多いのは東京都である。他年度も同様であった。東京都は全国の「進学」というライフイベントを迎えた人々の進学先として最も選ばれる都道府県であることがわかる。これは、全国の大学生の 4 人に 1 人が東京都の大学生である¹³⁾¹³⁾ という実際の大学生数の都道府県分布とも一致しており、多くの若者が進学目的で東京都へ集まっていることが確認できる。一方で図-5 のように因子得点を 100% 積上げにし、割合に注目すると、東京都の転入者のうち「進学」の割合は目立つほど大きくない。最も「進学」の割合が大きいのは京都府であり、図-7 の通り大学生数の対人口比は東京都よりも大きい。

次に転出の因子得点を積み上げた図-6 に着目すると、埼玉県・千葉県・東京都・神奈川県 の 1 都 4 県は「進学」する転出者の割合が少ない。これは東京都へ進学する際に転居の必要がないため転出を行っていないことを表している。また、転出者の中の「就職」の割合を見ると、京都府、東京圏からの多さが特徴的である。大学在学中に居住した大学生がこれらの都道府県から就職とともに出ていく傾向にあることが分かる。

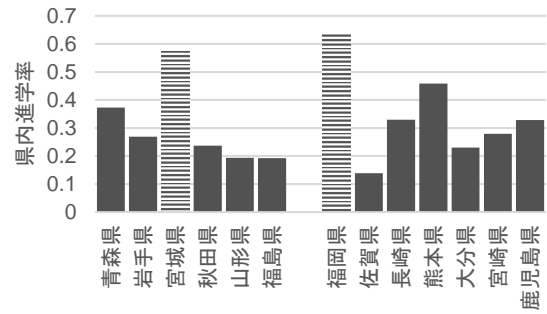


図-8 2015年度県内大学進学率 (東北・九州地方)

また宮城県と福岡県は、これらの都府県と同様に転入者の「進学」、転出者の「就職」が多い。さらにこの 2 県の周囲の県では転出者の「進学」、転入者の「就職」が多い。これらの傾向は、大学進学に際して地方中心都市に転居し、卒業後に地元に戻って就職、あるいは地方外の三大都市圏等へ就職する者が多いためであると考えられる。図-8 は、総務省のデータに基づき、出身高校が当該県の大学進学者のうち進学大学の所在県が同一の者の比率を表した県内大学進学率を示している。宮城県、福岡県が各地方内で最も高い値を示していることがわかる。

b) 震災後の福島県 (特殊なイベント)

2011 年に発生した東日本大震災の原発事故により、子どもの放射線リスクを回避したい母子が原発避難をする一方、父親は被災地の働き手として残留する母子避難が起こった¹⁵⁾。新潟県が 2013 年に行ったアンケートでは福島県から新潟県への原発避難世帯の半数以上が「震災の影響により離れて生活する家族がいる」と回答した¹⁵⁾。母子避難の他に、広い避難先が見つからないため分散して暮らさざるを得ない 3 世代以上の大規模家族も見られる。避難先は元の居住地とは異なる県であることも多く、福島県を除く 46 都道府県全てに及んだ¹⁶⁾。以上から、震災後の福島県では平常時や他の都道府県では起こらない世帯の分裂が起こったと考えられる。

被災地で急増した復旧・復興事業の多くが男性中心の仕事内容であったため男性労働者が他都道府県から流入した一方で多くの女性が職を失った¹⁶⁾ことから福島県への転入者の性別年齢階層構成比も通常と異なることが予想できる。

図-6 を見ると、福島県では転出における「子育て世帯」、転入における「転勤」の比率が大きい。急増した母子避難が「子育て世帯」に、単身男性労働者が「転勤」に含まれていることが分析結果に表れている。

4. おわりに

(1) 本研究のまとめ

本研究では地域の変化を地域の変化を非負のライフイベントの出現の加算として表現し、特殊な変化の他地域や過去事例との類似性や共通性を把握するために NMF を人口データへ適用した。その結果、「進学」、「就職」、「転勤」、「子育て世帯」、「住替え（壮年期）」、「住替え（中高年）」、「後期高齢者」と解釈可能な 7 つの因子を得た。これらで一般的なライフイベントを解釈可能であった。また、震災後の福島県で見られたような特殊な避難というイベントも一般的なライフイベントにより説明できることを確認した。

(2) 今後の課題

本研究では行列対象の NMF を用いたが、今後はテンソルとした年齢構成データへの適用を考えていくことでより立体的な人口構成分析ができると考えられる。年齢階層別 OD データ、また属性として移動者の職業などの付加が挙げられる。

また、時代によるライフイベントの種類の変化を観測するためにはコーホートに着目する方法が考えられる。コーホートの動きまで明らかにできるようになれば、時代の影響まで考慮できるため、より多い期間を分析することができ、より多くのライフイベントを見つけ出すことが可能になる。

更に、より疎な因子負荷量を持つ結果を得るためにより良い目的関数を見つけ、精度を上げていくことを今後の課題としたい。

参考文献

- 1) 奥村誠：国勢調査メッシュデータに基づく地区の将来人口構成予測手法，都市計画論文集，Vol.40-3，pp.193-198，2005.
- 2) 亀岡弘和：非負値行列因子分解，計測と制御，第 51 巻第 9 号，pp.835-844，2012.
- 3) 澤田宏：非負値行列因子分解 NMF の基礎とデータ／信号解析への応用，電気情報通信学会誌，Vol.95，

- No.9，pp.829-833，2012.
- 4) 清水昌人：近年の人口移動理由，人口問題研究，57-1，pp.8-24，2001.
- 5) Imre Csiszár: I-divergence geometry of probability distributions and minimization problems, The Annals of Probability, 3-1, pp.146-158, 1975.
- 6) 板倉文忠：統計的手法による音声分析合成系に関する研究，博士論文，名古屋大学大学院工学研究科，1972.
- 7) Bent Jørgensen: Exponential Dispersion Models, J. R. Statist. Soc. B, Vol.49, No.2, pp.127-162, 1987.
- 8) Daniel D. Lee and H. Sebastian Seung: Learning the parts of objects with non-negative matrix factorization, Nature, 401, pp.788-791, 1999.
- 9) Daniel D. Lee and H. Sebastian Seung: Algorithms for non-negative matrix factorization, Proceedings of the 13th International Conference on NIPS, pp.556-562, 2001.
- 10) Masahiro Nakano, Hirokazu Kameoka, Jonathan Le Roux, Yu Kitano, Nobutaka Ono, and Shigeki Sagayama: Convergence-guaranteed multiplicative algorithms for non-negative matrix factorization with β -divergence, MLSP 2010, pp.283-288, 2010.
- 11) Derry Fitzgerald, Matt Cranitch, and Eugene Coyle: On the use of the Beta Divergence for Musical Source Separation, ISSC 2008, Galway, 2009.
- 12) 総務省統計局：統計局ホームページ/国勢調査の基本に関する Q&A，2018/01/18 参照
<http://www.stat.go.jp/data/kokusei/qa-6.htm#f1>
- 13) 総務省：社会・人口統計体系，都道府県データ，基礎データ
- 14) 宝田惇史：福島第一原発事故に伴う避難・移住における交通関連の課題—避難者の二重生活と支援者の全国ネットワーク化を中心として—，交通権，No.33，pp.53-62，2016.
- 15) 新潟県県民生活・環境部広域支援対策課：「避難生活の状況に関する調査」結果について，2018/01/19 参照
http://www.reconstruction.go.jp/topics/main-cat2/20140228_Niigata01_hinansaki03.pdf
- 16) 日本学術会議東日本大震災復興支援委員会福島復興支援分科会：東京電力福島第一原子力発電所事故による長期避難者の暮らしと住まいの再建に関する提言，第 13 回東日本大震災復興支援委員会，2014.

(2018.?? 受付)

A COMPARATIVE COMMUNALITY ANALYSIS METHOD BASED ON THE AGE COMPOSITIONS

Hatsune MOTOI, Makoto OKUMURA and Daijiro MIZUTANI

Interregional migration and remaining are generated by several reasons such as enrollment into higher grade schools, getting a job, housing upgrading, and so on. Each of these reasons may have different frequency pattern in age composition. We can grasp valuable information of status and changes of local community, based on the calculated factor scores relating to the age composition of migration and remaining population. We propose a methodology of comparative community characteristic analysis based on NMF (Non-negative Matrices Factorization) over age composition of emigrants, immigrants and remaining people. We verify empirically that typical migration and remaining of Japanese 47 prefectures in

recent 20 years can be examined using only seven factors.