

非線形次元縮約に基づく地域定着意思に関する アンケート分析

井上 寛・柳 永珍

1. はじめに

本研究は、若年層が特定の地域に対して定着意思の有無を判断する際に、その地域にどのような要素を求めるとかを分析する。日本は人口減少問題を国家レベルの課題として認識し、多様な対策を講じてきた。特に、地域の人口減少問題は非常に重要な争点として取り上げられ、地域の人口減少を抑えるために様々な努力が続けられている。その中で、地域間人口移動や誘引に関する調査や研究が多様な分野で活発に行われている。

そこで本研究は、北九州市内の5つの大学を対象に、各大学の在學生に基づいた252個のアンケートデータ（19個の質問項目）を用い、若年層の地域定着意思に大事な要素を、非線形次元縮約法に基づき分析する。この研究を基に、北九州市内のみならず、その他の地域に対する定着意思の分析にも有効活用できることが期待できる。

2. データについて¹

本研究では若年層の主観的な認識から定着意思を把握するために、既存のインフラや地域環境の量的な現状ではなく、主観的な認識を把握できる質問

を構成する必要がある。そのためにまず北九州市に居住しかつ地域内の大学に通っている12名の学生を対象に、非構造化されたインタビューを用いたパイロット調査を実施し、最終的に「経済」・「インフラ」・「文化」・「包容」のカテゴリにほぼ均等に属する19個の質問項目を確定した。(表1)

調査方法について、本研究は探索的な研究であるため可能な限り多数のサンプルを確保することに重点を置いた。北九州市の大学生という基準だけを設定し、それ以外の条件については特に統制を行っていない。調査対象者のサンプリング方法も接近性が高い便宜表集 (Convenience Sampling) を採択した。ただし、可能な限り多様な大学が含まれるように北九州市内の5つの大学を対象に、各大学の在學生に基づいた割り当てを適用した。調査は配布調査とオンライン調査を併行した。対面授業及び地域活動に参加した学生の場合は紙媒体の質問紙を用いた配布調査を実施し、COVID-19の影響により配布調査が難しい場合はSNS、Formsを用いたオンライン調査を実施した。調査は2020年4月6日から7月14日まで100日間に渡って実施された。質問紙は総330部を配布し、271部が回収できた。回収したものの中で無回答が多くデータとして活用できない質問紙19部を除いた252部を最終分析に用いた。調査対象者の基本属性は(表2)の通りである。

また、質問項目の信頼度分析に関しては、使用される質問の信頼度を検定するために全体の質問(Q2~Q20)に対してCronbach's α 係数とMcDonald ω 係数を用いた検定を行った。実施結果は(表4)のとおりである。

2つの係数がすべて良好な信頼度の基準である0.6以上を満たしており、厳格な基準である0.7以上も満たしているので、信頼度のある質問と判断できる。削除したら信頼度が上昇する質問 (if item dropped) も存在しないことを確認した。以下の質問を用いて、Infra, Culture, Inclusive, Economyという4つの要素(表2)として定着意思に影響を与える原因を因子分析を通じて分類した。

¹ 第1節の「データについて」は、柳, 井上(2022)の未刊行論文の一部を引用している。

表1 最終質問項目

番号	質問（括弧の中は変数名）	測定尺度
Q 1	現在の地域に定着したいと思いますか。	1：定着したい 2：定着したくない
Q 2	物価は安定的だと感じる。(V 1)	1：全く同意しない 2：同意しない 3：どちらともいえない 4：同意する 5：完全に同意する (リッカート5点)
Q 3	働き口が多様だと思う。(V 2)	
Q 4	就職が容易である。(V 3)	
Q 5	この地域に就職したら十分な所得が期待できる。(V 4)	
Q 6	自然災害から安全だと思う。(V 5)	
Q 7	治安・防犯的に安全だと思う。(V 6)	
Q 8	技術的に先端である。(V 7)	
Q 9	交通環境は便利である。(V 8)	
Q 10	医療サービスは充実している。(V 9)	
Q 11	買い物の環境は利便性が良い。(V 10)	
Q 12	文化施設が豊富である。(V 11)	
Q 13	多様な店舗がある。(V 12)	
Q 14	景観が美しい。(V 13)	
Q 15	アメニティは充実している。(V 14)	
Q 16	楽しめるものが多い。(V 15)	
Q 17	誰でも歓迎されている。(V 16)	
Q 18	寛大な雰囲気である。(V 17)	
Q 19	お互いが配慮している。(V 18)	
Q 20	心安らかである。(V 19)	
Q 21	あなたの性別を選んでください。	1：男性 2：女性 3：その他
Q 22	あなたの年齢は何歳ですか。	開放型
Q 23	あなたの出身地域に該当するものを選んでください。	1：北九州市内 2：北九州市外の福岡県 3：九州内の他県 4：九州外の他県 5：その他

表 2 因子分析の結果

	因子				信頼度 上 = α 下 = ω
	1 (インフラ)	2 (文化)	3 (包容)	4 (経済)	
V 5	.892	.016	.036	-.084	.899 .902
V 6	.853	.038	-.077	-.037	
V 7	.836	.029	-.134	.015	
V 8	.821	.127	-.015	.098	
V 9	.699	.093	-.274	.344	
V 10	.622	.227	-.300	.348	
V 15	.088	.940	.200	-.017	.945 .952
V 14	.107	.936	.185	-.011	
V 12	.092	.914	.264	-.029	
V 13	.093	.782	.276	-.060	
V 17	-.210	.165	.894	.047	.918 .922
V 18	-.201	.210	.891	.076	
V 16	.063	.446	.787	-.064	
V 19	-.123	.314	.785	.032	
V 2	.071	-.053	.047	.981	.985 .985
V 3	.085	-.037	.032	.974	
V 1	.068	-.040	.032	.967	
eigenvalue	4.346	4.074	3.664	3.175	
分散の%	22.875	21.440	19.282	16.708	
累積%	22.875	44.315	63.598	80.306	

KMO = .845, Bartlett の球面性検定 = 5353.109 ($p < .001$)

表3 調査対象者の基本属性

	分類	頻度	有効%
性別	男	141	56.0
	女	111	44.0
年齢	18歳	15	6.0
	19歳	38	15.1
	20歳	49	19.4
	21歳	34	13.5
	22歳	55	21.8
	23歳	48	19.0
	24歳	13	5.2
出身地域	北九州市内	62	24.6
	北九州市外の福岡県	98	38.9
	九州内の他県	64	25.4
	九州外の他県	28	11.1
	合計	252	100

表4 信頼度検定

	Cronbach's α	McDonald ω
Scale	0.845	0.857

3. 非線形次元圧縮法（多様体学習：拡散マップ）

データの次元縮約（圧縮）は、機械学習における教師あり学習の中のパターン認識における特徴抽出などの前処理として行われる。また、教師無し学習や、データマイニングにおける可視化などに関連して様々な領域で用いられている操作である。しかし、通常よく使われる主成分分析などの線形の次元縮約手法は適用限界があり、近年、高次元のユークリッド空間の中で低次元多様体の周囲に集中して散布しているとみなせるデータに対して、その多様体構造を利用した非線形次元縮約法が、多様体学習という名の下に研究されている。^(1,2)

簡易的に述べると、多様体学習とは、 \mathbb{R}^n 内の d 次元部分多様体 M からサンプルされた N 個のデータ $X = (x_1, x_2, \dots, x_N), x_i \in \mathbb{R}^n$ が得られた時、それらの点を M 内の隣接関係をできるだけ保存するように低次元空間 $\mathbb{R}^d (d < n)$ にマッピングすることを指す。

本研究では、固有ベクトルを用いたスペクトル埋め込み型の多様体学習アルゴリズムである拡散マップ (diffusion Maps)^(3,4)を用い、地域定着に関するアンケートデータを解析する。この手法のアルゴリズムは、前処理としてデータの多様体に関する情報を近傍グラフに変換する。データの各点がグラフの頂点に対応しており、各頂点は近傍にふくまれる頂点とのみ辺で結ばれる。近傍については、多くは \in 近傍や k 近傍が用いられる。その後、以下のステップによって非線形次元縮約を行う。

◆近傍グラフの各辺 $x_i x_j$ に重み $W_{ij} \geq 0, W_{ij} = W_{ji}$ を割り当てる。これを正規化して $N \times N$ の推移確率行列 P を作る。

$$P_{ij} = p_t(x_i, x_j) = \frac{W_{ij}}{D_{ii}}, \quad (D_{ii} = \sum_{j=1}^n W_{ij}, \quad D_{ij} = 0 \quad (i \neq j))$$

$p_t(x_i, x_j)$ は P で表現されるグラフ上のランダムウォークによって x_i を出発して t ステップ後に x_j に到達する確率を表す。推移行列の性質から $p_t(x_i, x_j)$ は $t \rightarrow \infty$ で定常分布 $\varphi_0(x_i)$ に収束する。この時、点 x_i と点 x_j の拡散距離を

$$D_t(x_i, x_j)^2 = \frac{\sum_{k=1}^N (p_t(x_i, x_k) - p_t(x_j, x_k))^2}{\varphi_0(x_k)}$$

で定義している。

◆ P の固有値及び固有ベクトルを

$$P\psi_i = \lambda_i \psi_i, \quad 1 = |\lambda_0| \geq |\lambda_1| \geq \dots \geq |\lambda_{N-1}| \geq 0$$

とする。このとき、

$$D_t^2(x_i, x_j) = \sum_{k=1}^N \lambda_k^{2t} (\psi_k(x_i) - \psi_k(x_j))^2$$

が成り立つ。 $|\lambda_i| \leq 1$ なので、 N より小さな適当な次元 $d(t)$ まで固有ベクトルをとり、

$$y_i = \Psi_t(x_i) = (\lambda_1^t \phi_1(x_i), \dots, \lambda_{d(t)}^t \phi_{d(t)}(x_i))^T$$

と次元縮約する。この時、

$$\|y_i - y_j\|^2 \approx D_t^2(x_i, x_j)$$

が成り立つ。

本研究は、地域定着に関する252個のアンケートデータ(19個の質問項目)を用い、上記で紹介した拡散マップを適用する。非線形次元縮約したデータを改めて、 k -means法でクラスタリングし、実際に得られたアンケート調査結果と比較することにより、地域定着に大事な要素を考察していく。

4. 結 果

本研究で用いたアンケートデータに拡散マップを適用し、その後クラスタリングをした結果は以下の通りである。(図1：2次元に縮約した結果)(図2：3次元に縮約した結果)

図1 拡散マップ(2次元)

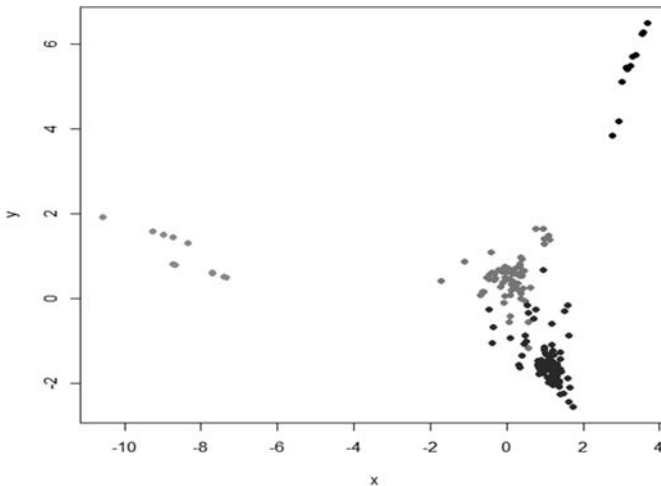
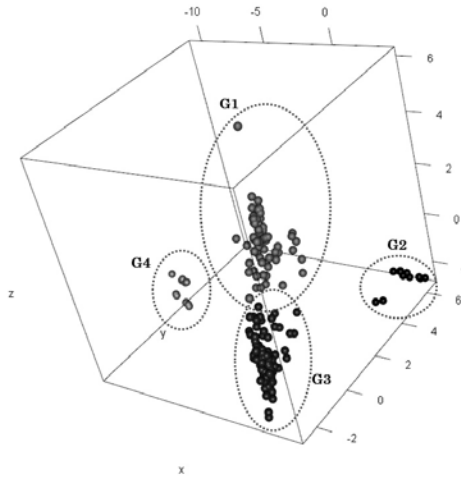
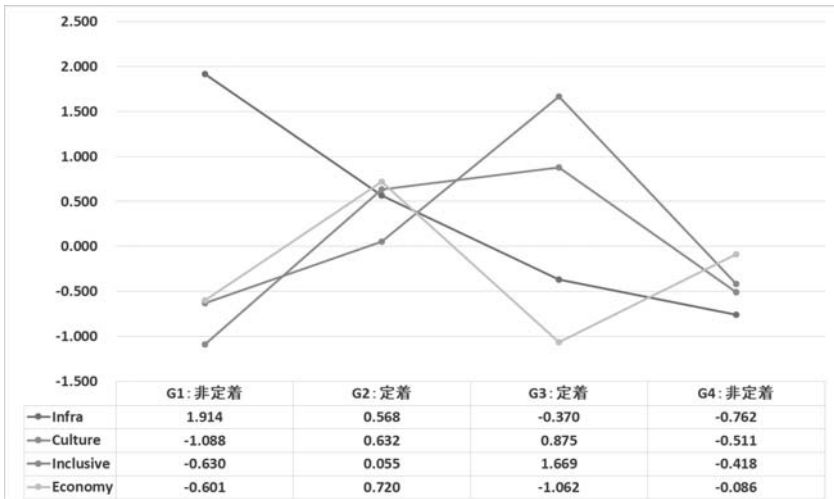


図2 拡散マップ（3次元）



この非線形縮約を行った後のクラスタリング結果とアンケートデータの内容を比較し、見える化したのが図3である。

図3 クラスタリング結果に基づいた分類に関する



5. 考 察

本研究は、北九州地域への定着に対して4つのグループに分類することができた。各グループの比較を通じて以下のような考察ができる。

第一、「G1」に該当する回答者は、地域の都市的な環境の整備について、他の3つのグループに比べて最も高い満足度を示しているが、定着の意思はないことが確認できる。もう一つ非定着の「G4」グループのように4つの要素の評価点が全てマイナスを示しているケースと比べて推察してみると、定着意思に Infra という都市的な環境は大きな意味がないと考えられる。言い換えれば、G1グループにとって地域の魅力は Infra ではなく、地域の文化的な楽しさや包容力をより大事だと認識している可能性が高い。もし、「G1」グループの定着意思を高めようとする場合は、政策的に文化や地域の多様性などを培うことが有意であると考えられる。

第二、「G4」グループは政策的には、優先順位が非常に低い対象であるといえる。該当のグループは前述したようにすべての因子において、低い評価点数を見せている。つまり、完全に地域の雰囲気全般が変化しないかぎり、地域への定着意思はないともいえる。これは地域が持っている根本的な環境やアイデンティティーそのものと不一致である。

第三、定着意思を示している「G2」と「G3」のグループにおいて、共通している要素は文化と包容であることがわかる。「G2」は全般的に地域に対して高い満足度を見せているが、「G3」は都市的な環境や経済面においては点数が相対的に低いことが把握できる。そして、その低い点数に対して文化と包容の数値が非常に高いことが把握できる。言い換えれば、都市的な環境や経済的な部分は、文化的な楽しさや包容と相互補完的な相関を持つ可能性がある」と推察できる。

以上の考察から、北九州地域の若者において、定着意思を高めるためには文化と包容という分野に集中する必要があるといえる。

【参考文献】

- 1) Roweis S.T, Saul L.K. Nonlinear Dimensionality Reduction by Locally Linear Embedding. *Science*. 2000; 290 (5500): 2323-2326.
- 2) Tenenbaum J.B, de Silva V, Langford J.C. A global geometric framework for nonlinear dimensionality reduction. *Science*. 2000; 290 (5500): 2319-2323.
- 3) Coifman R.R, Lafon S. Diffusion maps. *Applied and Computational Harmonic Analysis*. 2006; 21: 5-30.
- 4) Nadler B, Lafon S, Coifman R.R, Kevrekidis I.G. Diffusion maps, spectral clustering, and eigenfunctions of Fokker–Planck operators. *Neural Information Processing Systems*. 2005; 18: 8 pages.
- 5) 柳永珍, 井上寛. 若い世代の地域定着意思に関する探索的研究—北九州市の大学生を中心に—. 未刊行論文. 2022.