

鹿児島県における地域特性分析

—主成分分析法を用いて—

石塚 孔 信

1. はじめに

地域経済の分析を行う場合、たくさんの変量を少数の尺度で評価したり、あるいは、多くの変量を共通の要因で説明する単純なモデルを作ることが必要となることがある。たとえば、それぞれの地域は、人口、工業出荷額、生産額、道路舗装率、山林面積比率、犯罪発生件数などの経済的、社会的あるいは地理的条件を表すたくさんの指標を用いてそれぞれの視点から評価することが出来るけれども、一方、少数の尺度で総合的に地域の性格づけを行うほうが合理的な評価が出来ることがある。最も簡単な方法は、経済、社会、文化、人口などのいくつかの部門に分け、各部門ごとに地域の性格を端的に表すと考えられるいくつかの指標を選んで単純平均をとったり、分析目的にあわせて主観的に賦与したウェイトを用いて加重平均したりする方法であるが、単純平均はウェイトがすべての指標について同一であると考えているわけであるから、これらの方法はいずれもウェイトについての客観的妥当性が保証されない欠点を持っている。

また、地域別の行政に対する需要を予測するために、各地域の部門別の行政水準がどのような要素によってどのように規定されているかを定量的に分析したい場合がある。ところが行政施策に対する住民の需要が地域の経済社会の状態や他の部門の行政水準の影響を受けていることはわかっているけれども、需要を規定する因果関係が理論的に明らかになっていないこと、説明変数の数が多いこと、あるいは変数が相互依存的で原因と結果に区分しにくいことなどの理由で回帰分析の方法が利用できないことが多い。

このような二つの問題への接近法としては、主成分分析法や因子分析法がある。主成分分析では、変量がいくつかある場合に、それらの変量を要約して、新しい指標で代表させようとするものである。その結果できる新しい指標を主成分と呼ぶ。もっとも主要な主成分1個だけで、元の変量のあらかずものほとんどすべてを言い尽くすことが出来る場合もあれば、第2のあるいは第3の主成分を考慮しなければならない場合もある。主要なほうから第1主成分、第2主成分、・・・と命名するが、このうち、第1主成分は、もとのデータを総合化する傾向が大きいので、多くの変量からの総合的評価を得たい場合に有効である。したがって、主成分分析はもとのデータを要約し、表現し直すための手法であるが、同時に、主成分をグラフ上にプロットし、新しい視点で各個体を捉えなおし、個体をグルーピングするという目的にも適用される。一方、因子分析法では、分析対象となる全変数が共通因子とよばれる少数個の共通の変量と独自因子とよばれる各変数に固有の変量で説

明できるという仮定のもとに共通因子の係数と因子の量を計測する手法である。

本稿においては、主成分分析法を用いて鹿児島県の地域特性分析を試みる。分析の方法については、まず、人口、住宅形態、産業の各部門から変数を選び、鹿児島県内の全市町村についてそのデータを抽出し、それに主成分分析法を適用することにした。

2. 主成分分析法

主成分分析法は、経営分析や地域、商圏の総合指標を作成する場合にしばしば利用される。それは、多数の要因を少数の総合指標に集約するのに有用であるからである。

主成分分析の入力データは、 p 個の数量的な変数 x_{pn} に関するデータを並べた多変量データ行列である。次のようなケースを考える。

	X_1	X_2	...	X_p
1	X_{11}	X_{21}	...	X_{p1}
2	X_{12}	X_{22}	...	X_{p2}
⋮	⋮	⋮		⋮
⋮	⋮	⋮	...	⋮
⋮	⋮	⋮		⋮
n	X_{1n}	X_{2n}	...	X_{pn}

このデータから、次のような分散共分散行列 S を求める。

$$S = \begin{bmatrix} S_{11} & S_{12} & \cdots & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & \cdots & S_{2p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ S_{p1} & S_{p2} & \cdots & \cdots & S_{pp} \end{bmatrix}$$

次に、 S の固有値 λ を求める。

$$\begin{vmatrix} S_{11} - \lambda & S_{12} & \cdots & \cdots & S_{1p} \\ S_{21} & S_{22} - \lambda & \cdots & \cdots & S_{2p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ S_{p1} & S_{p2} & \cdots & \cdots & S_{pp} - \lambda \end{vmatrix} = 0$$

この固有値は正の実数であり、次のような関係が見られる。

$$\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_p \geq 0$$

さらに、 $a_{i1}^2 + a_{i2}^2 + \cdots + a_{ip}^2 = 1$ という条件のもとで固有値 λ_i の固有ベクトル $(a_{i1}, a_{i2}, \cdots, a_{ip})$ を求める。

$$\begin{bmatrix} S_{11} & S_{12} & \cdots & \cdots & S_{1p} \\ S_{21} & S_{22} & \cdots & \cdots & S_{2p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ S_{p1} & S_{p2} & \cdots & \cdots & S_{pp} \end{bmatrix} \begin{bmatrix} a_{i1} \\ a_{i2} \\ \cdots \\ a_{ip} \end{bmatrix} = \lambda_i \begin{bmatrix} a_{i1} \\ a_{i2} \\ \cdots \\ a_{ip} \end{bmatrix}$$

したがって、最大固有値 λ_1 の固有ベクトル $(a_{11}, a_{12}, \dots, a_{1p})$ に対して

$$\text{第1主成分 } Z_1 = a_{11}X_1 + a_{12}X_2 + \cdots + a_{1p}X_p$$

を得る。2番目に大きい固有値 λ_2 の固有ベクトル $(a_{21}, a_{22}, \dots, a_{2p})$ から

$$\text{第2主成分 } Z_2 = a_{21}X_1 + a_{22}X_2 + \cdots + a_{2p}X_p$$

.....

p番目の固有値 λ_p の固有ベクトル $(a_{p1}, a_{p2}, \dots, a_{pp})$ から

$$\text{第p主成分 } Z_p = a_{p1}X_1 + a_{p2}X_2 + \cdots + a_{pp}X_p$$

が得られる。

また、

$$\frac{\lambda_i}{\lambda_1 + \lambda_2 + \cdots + \lambda_p}$$

を第 i 主成分の寄与率と呼び、

$$\frac{\lambda_1 + \lambda_2 + \cdots + \lambda_i}{\lambda_1 + \lambda_2 + \cdots + \lambda_p}$$

を第 1 - 第 i 主成分の累積寄与率とよぶ。

p個の説明変数に対して、第1主成分から第p主成分まで主成分もp個定義されるが、出来るだけ少ない主成分の数でデータの情報を反映できることが望ましい。

主成分得点は、次のような関数で与えられる。第 i 主成分の主成分得点 $Z(x_1, x_2, \dots, x_p)$ は、

$$Z(x_1, x_2, \dots, x_p) = a_{i1}X_1 + a_{i2}X_2 + \cdots + a_{ip}X_p$$

となる。

なお、分散共分散行列を用いると説明変量の単位の変化の影響を受けてしまうが、それを避けるために用いられるのが相関行列である。したがって、分散共分散行列の代わりに相関行列 R

$$R = \begin{bmatrix} 1 & r_{12} & \cdots & \cdots & r_{1p} \\ r_{21} & 1 & \cdots & \cdots & r_{2p} \\ \cdots & \cdots & \cdots & \cdots & \cdots \\ r_{p1} & r_{p2} & \cdots & \cdots & 1 \end{bmatrix}$$

を用いて主成分分析を行うことが多い。

3. 鹿児島県における地域特性分析

3-1 分析方法

都市や地域を形成している要因となる変数は数多くある。そして、それらの変数間には独立な関係のものもあれば、相互依存的な関係のものも存在する。したがって、ある一つの変数だけで地域の特性を表現することは不可能である。そこで、多変量データを数個の主成分で代表させ、その簡略化された主成分に強くかかわる都市や地域を分類することの出来る主成分分析を用いることにする。その際、変数の選択については、人口、住宅形態、産業の各セクションから33の変数を選択した。^{注1)~注4)} その際、それぞれの変数について異常に大きいか、異常に小さいデータによって分析結果が歪められるという問題点を避けるためにここでのほとんどの変数を比率の形に加工して処理している。

まず、鹿児島県の全96市町村に対して選択した33個の変数についてのデータを抽出し、それをもとに33×33の相関行列を作成した。そして、その相関行列を用いて固有方程式を作成し、33個の固有値を導出した。その固有値の大きいほうから3つを選択し、それを元に第1主成分から第3主成分を取り上げることにした。(表-1) この場合の構造ベクトルが主成分負荷量として表-2に表示されている。さらにそれを直観的に理解するために模式化したのが図-1~図-3である。

3-2 分析結果の概要

図-1~図-3からそれぞれの主成分に対する各変数の関わりを考えるとそれぞれの主成分についての特徴は以下のように表現される。

<第1主成分について>

表-1から、第1主成分は全体の変動の27.65%を説明している。表-2、図-1より第1主成分が両極端に大きい変数に注目すると、金融保険業比、人口当たり歳入、卸売小売業比、不動産業比、商店当り販売額、民間借家比、人口密度、商店当り売場面積、集中地区人口比、世帯当り乗用車がプラスに特に強く作用しており、さらに生産人口比、幼年人口比、事業所当り出荷額、製造業比、世帯当り人口、サービス業比、運輸通信業比がプラスに強く作用している。一方、高齢人口比がマイナスに特に強く作用しており、さらに、建設業比、農業比、漁業比、持ち家比、公務比がマイナスに強く作用している。したがって、第1主成分は、都市化の水準をある程度説明しているものと判断できる。表-3より、第1主成分得点が大きな市町村は、鹿児島市、国分市、隼人町、始

表-1 固有値

主成分No.	固有値	寄与率(%)	累積(%)
1 都市化の水準	9.12	27.65	27.65
2 地域開発の推進力	5.70	17.27	44.92
3 農業の相対的地位	2.89	8.76	53.68

鹿児島県における地域特性分析

表-2 主成分負荷量

	主成分 1	主成分 2	主成分 3
持ち家比	-0.3878	-0.8177	-0.0568
世帯当り乗用車	0.5692	-0.6260	0.1364
集中地区人口比	0.6953	0.3662	-0.2119
人口密度	0.7310	0.1645	-0.1586
民営借家比	0.8008	0.3881	-0.1035
卸売小売業比	0.8555	0.0673	-0.2047
不動産業比	0.8498	0.1837	-0.1534
金融保険業比	0.8617	0.0683	-0.0785
生産人口比	0.4727	0.0343	0.1828
世帯当り人口	0.4326	-0.5583	0.2986
農業比	-0.4225	-0.4910	0.4801
高齢人口比	-0.8051	-0.1858	-0.4288
公営住宅比	-0.1280	0.2453	0.1941
公務比	-0.3713	0.4424	-0.2894
林業比	-0.2296	-0.1232	0.0531
幼年人口比	0.4583	0.2085	0.7096
建設業比	-0.4321	0.6526	0.1427
商店当り売り場面積	0.7197	0.0248	-0.0301
サービス業比	0.3888	0.5079	-0.4025
病院診療所当り人口	0.1728	-0.4515	0.3138
商店当り販売額	0.8205	0.0040	0.0430
製造業比	0.4429	-0.5160	0.0393
男性比	-0.1012	0.6440	0.6015
女性比	0.1012	-0.6440	-0.6015
給与住宅比	-0.0898	0.7352	-0.0681
事業所当り出荷額	0.4454	-0.0345	0.2505
間借り比	0.1713	0.4853	0.1091
運輸通信業比	0.3848	0.0679	-0.3342
事業所当り貨物車台数	-0.0909	0.0645	0.3541
漁業比	-0.4098	0.2512	-0.4485
電気ガス水道比	-0.2356	0.6859	0.1431
鉱業比	-0.0952	0.0727	0.0746
人口当り歳入	0.8602	-0.2663	0.0022

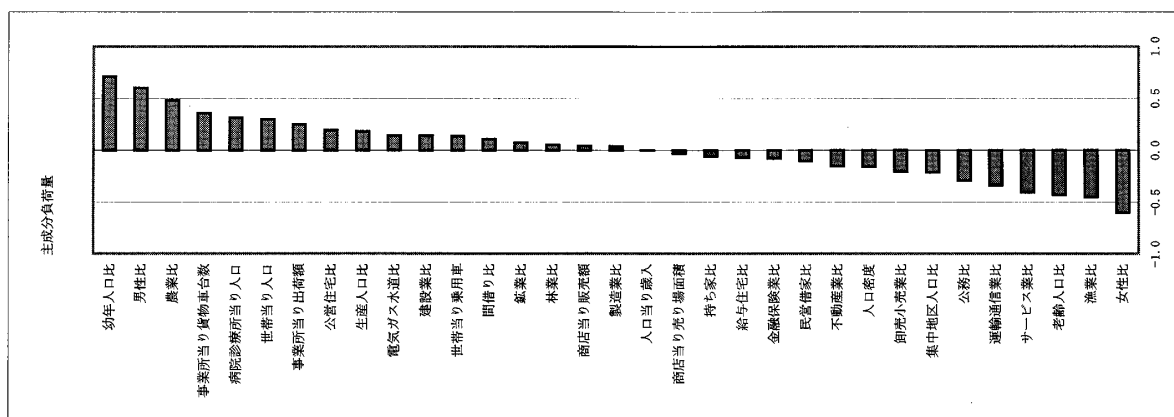


図-1 第1主成分の主成分負荷量

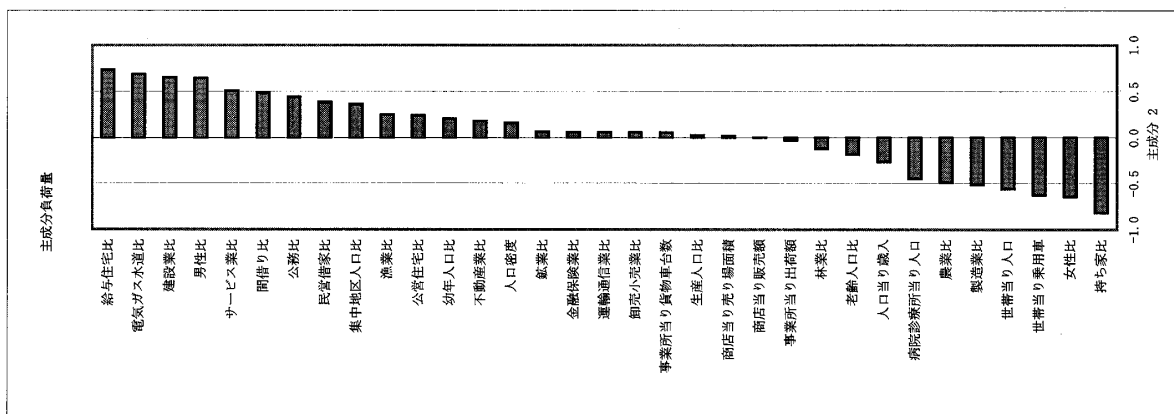


図-2 第2主成分の主成分負荷量

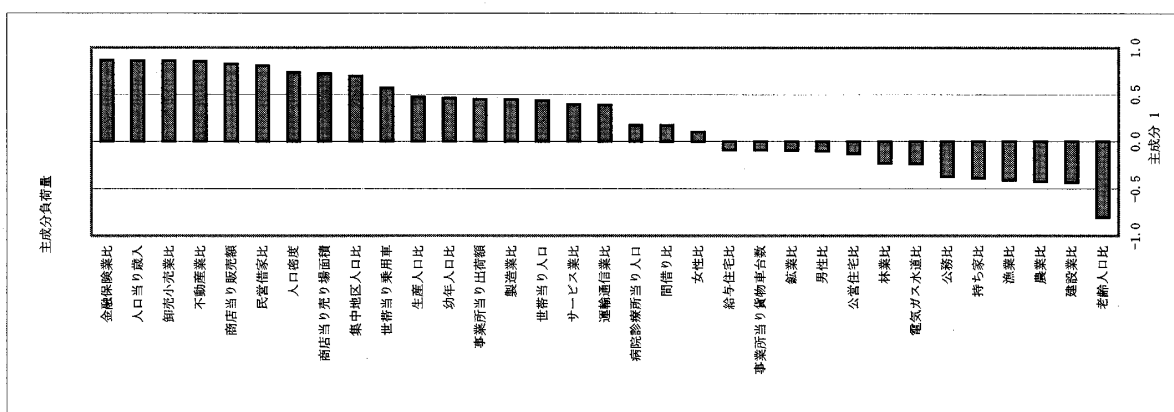
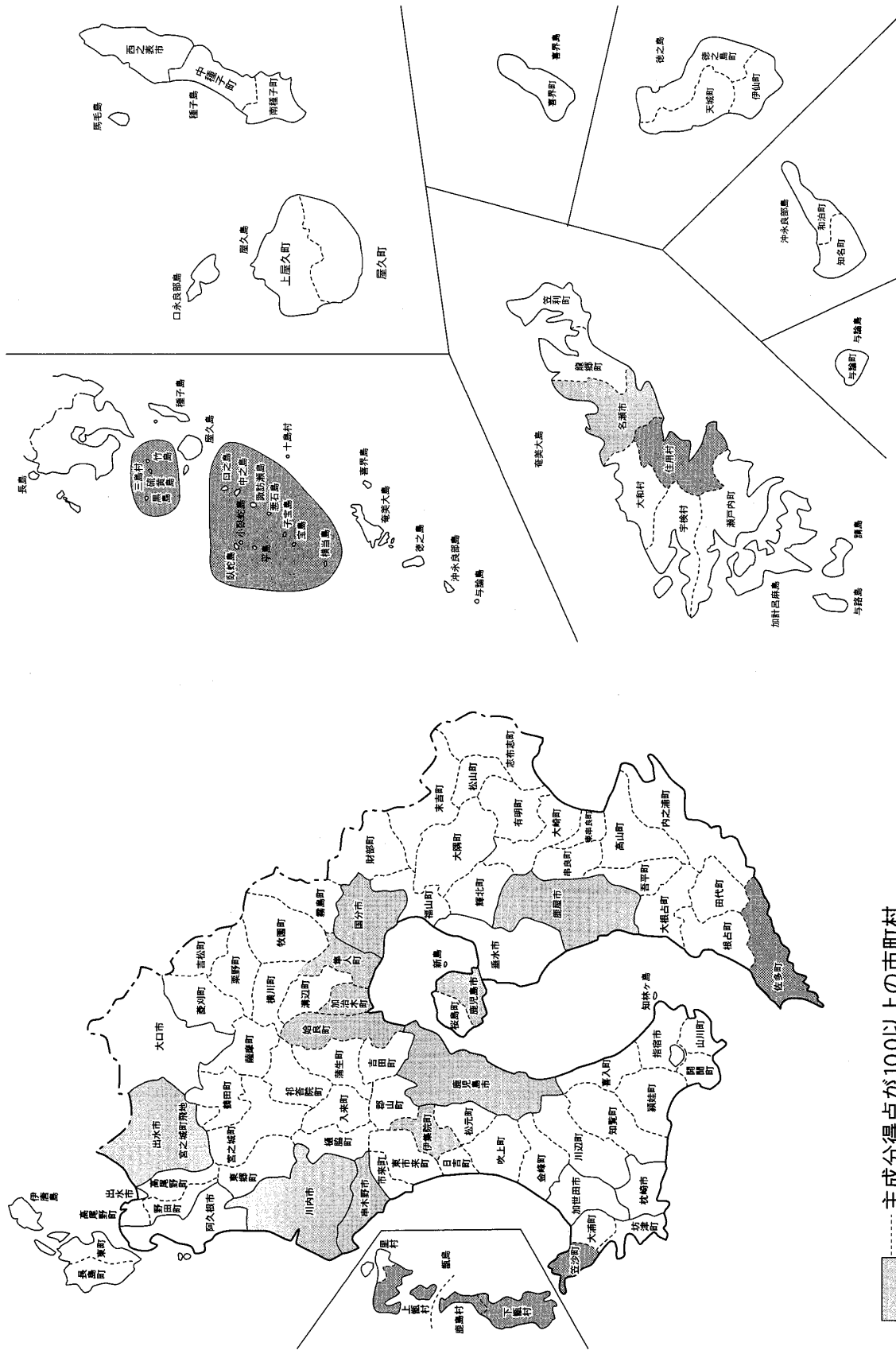
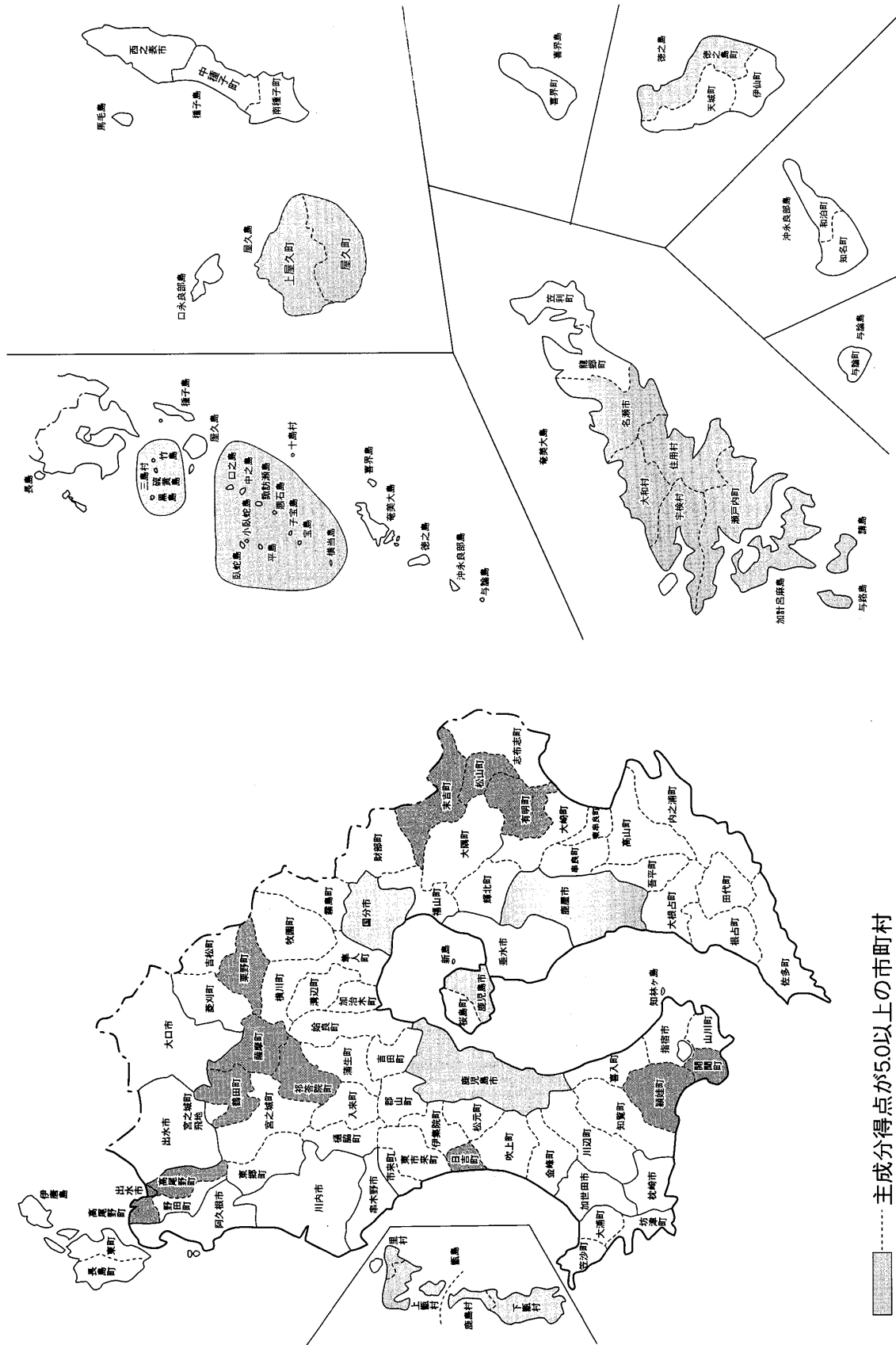


図-3 第3主成分の主成分負荷量

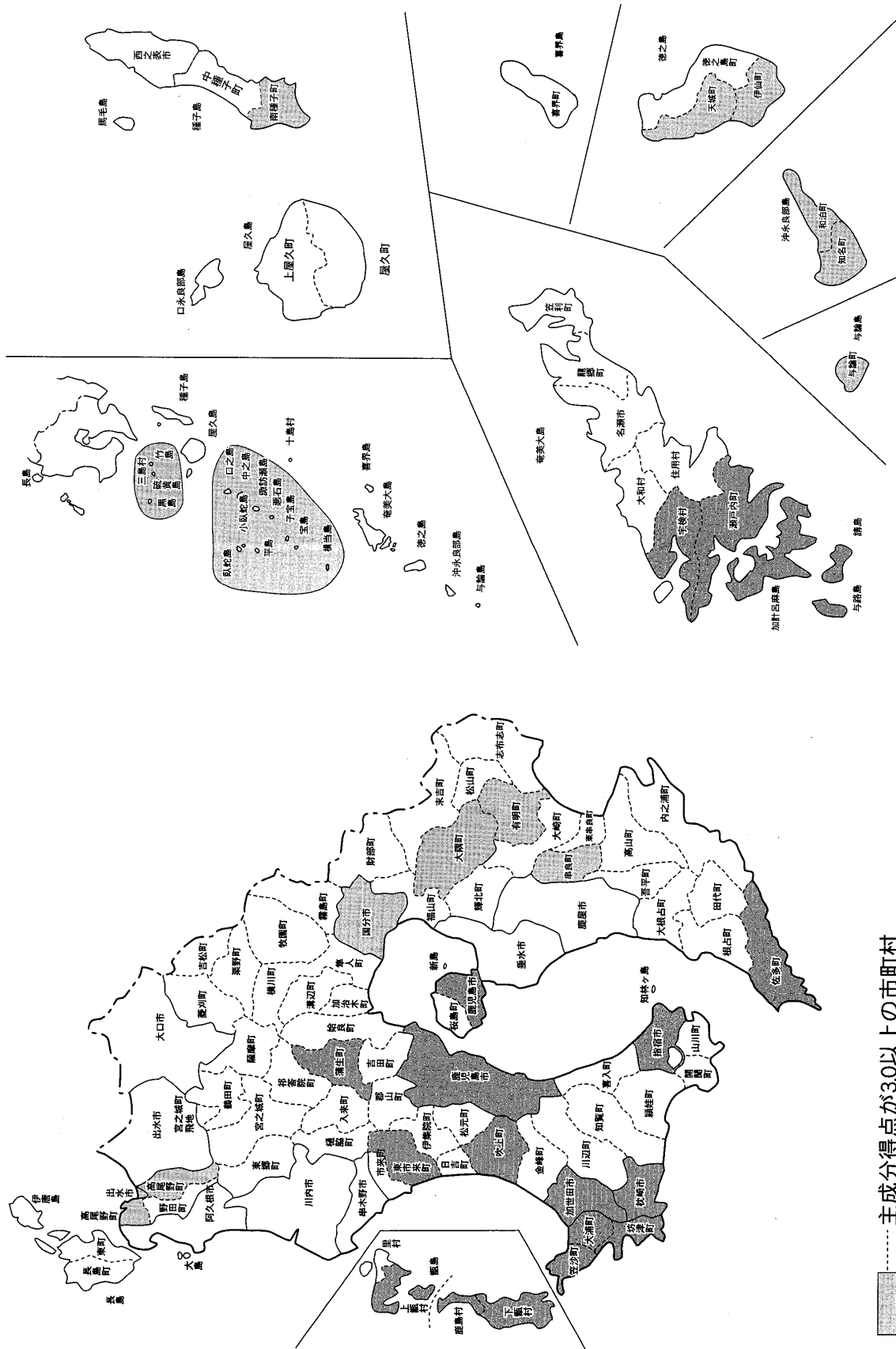


図一4 第1主成分の得点の分布

..... 主成分得点が10.0以上の市町村
 主成分得点が10.0以下の市町村



図一5 第2主成分の得点の分布



図一六 第3主成分の得点の分布

良町，加治木町，鹿屋市，名瀬市，伊集院町，川内市，串木野市，出水市（主成分得点が10.0以上）などがあり，とりわけ，錦江湾北岸の鹿児島市から国分市まで並んでいる市町村が高得点を保持している。逆に第1主成分得点が小さな市町村は，十島村，三島村，上甌村，鹿島村，佐多町，下甌村，住用村，笠沙町（主成分得点が負でその絶対値が10.0以上）などがあげられ，トカラ列島や甌列島の小村が並んでいる。（図－4）

<第2主成分について>

表－1から，第2主成分は全変動の17.27%を説明している。表－2，図－2より第2主成分得点が両極端に大きい変数に注目すると，給与住宅比，電気ガス水道比，建設業比，男性比，サービス業比がプラスに特に強く作用しており，間借り比，公務比，民営住宅比，集中地区人口比がプラスに強く作用している。一方，持ち家比，女性比，世帯当り人口，製造業比がマイナスに特に強く作用しており，農業比，病院診療所当り人口がマイナスに強く作用している。したがって，第2主成分は，地域開発の推進力を示しているものと思われる。その際，それが公共事業によるものか民間デベロッパーによるものかの2つのタイプがあると考えられる。表－4により，第2主成分得点が高い市町村は，三島村，十島村，下甌村，名瀬市，鹿児島市，瀬戸内町（主成分得点が10.0以上）などがあり，離島と鹿児島市が高得点をあげている。おそらく前者は公共投資主導型で後者は民間デベロッパー主導型ということではないだろうか。逆に，第2主成分得点が小さな市町村は，薩摩町，穎娃町，鶴田町，日吉町（主成分得点が負でその絶対値が6.0以上）などがあげられる。（図－5）

<第3主成分について>

表－1から，第3主成分は全変動の8.76%を説明している。表－2，図－3より第3主成分得点が両極端に大きい変数に注目すると，幼年人口比，男性比がプラスに特に強く作用しており，農業比，事業所当り貨物車台数，病院診療所当り人口がプラスに強く作用している。一方，女性比がマイナスに特に強く作用しており，漁業比，高齢人口比，サービス業比，運輸通信業比，公務比がマイナスに強く作用している。したがって，第3主成分は，農業の相対的地位を示しているものと考えることが出来る。表－5により，第3主成分得点が高い市町村は，三島村，南種子町，国分市（主成分得点が5.0以上）などであり，第3主成分得点が小さな市町村は，鹿島村，笠沙町，上甌村，坊津町，大浦町（主成分得点が負でその絶対値が5.0以上）などがあげられる。（図－6）

ここで，合成された第1～第3までの主成分は，もともと33の変量相互の相関性を考慮して統合された独立した新たな変量である。したがって，表－3～表－5に基づいて県内各市町村を主成分を軸とした新たな座標軸上に配置すると各市町村相互の類似性を見ることが出来る。図－7～図－9に（第1主成分と第2主成分），（第1主成分と第3主成分），（第2主成分と第3主成分）をそれぞれの座標軸とした各市町村の散布図を示しておく。

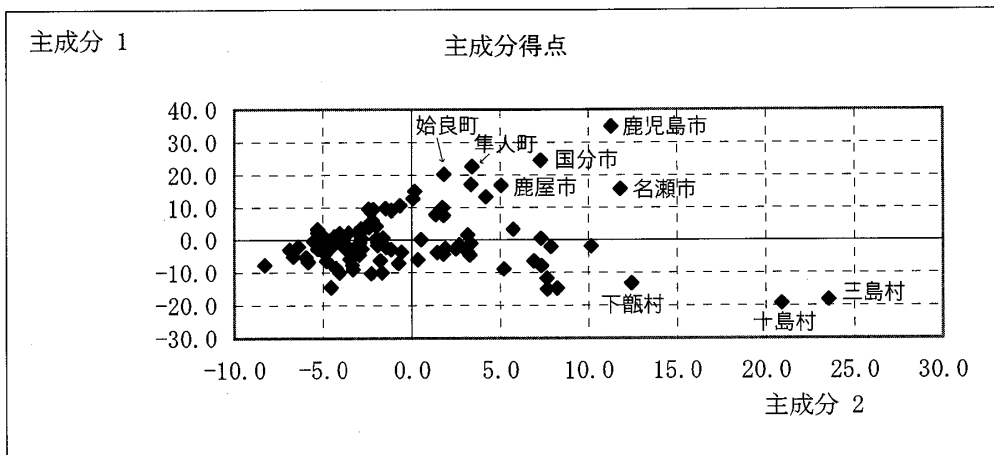


図-7 第1主成分-第2主成分による散布図

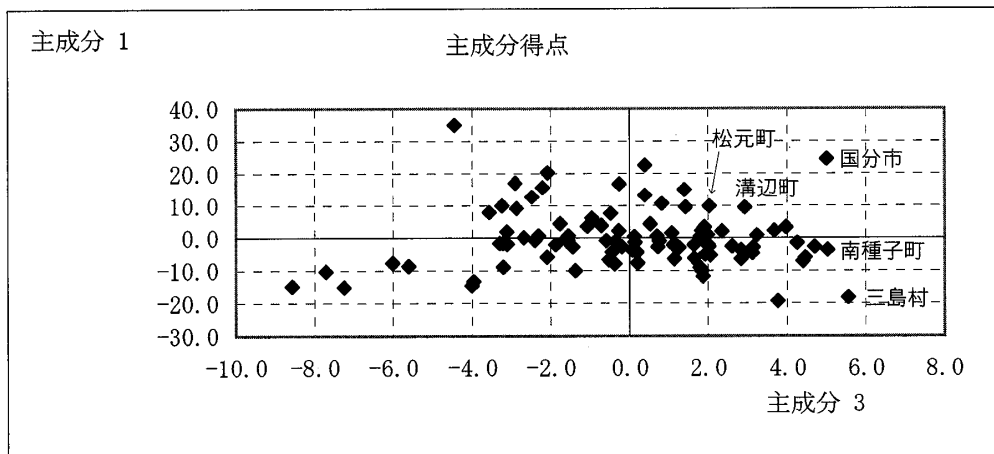


図-8 第1主成分と第3主成分による散布図

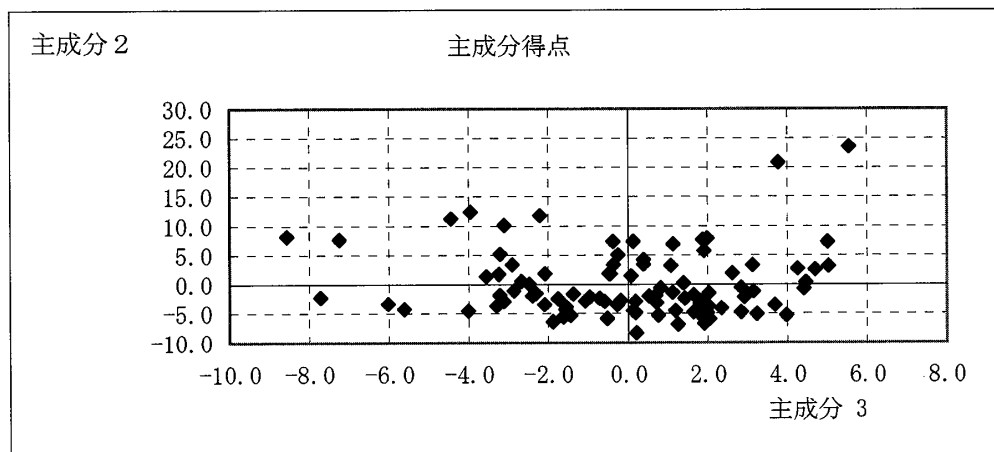


図-9 第2主成分と第3主成分による散布図

第1主成分が全体の27.65%とウェイトの大きいものであるので、その関与する図-7、図-8を検討してみる。これらの図に示す各主成分座標での標本に関しては、その絶対的な数値ではなく、各標本の位置、他の標本との近さの程度、類似の集団についてのグルーピングなどが重要である。

図-7より、第1象限に属する鹿児島市、国分市、鹿屋市、川内市、隼人町、加治木町、名瀬市などは、都市化も進み、かつ、地域開発も盛んな地域と判断することが出来る。一方、第4象限に属する三島村、十島村、下甕村、瀬戸内町、鹿島村などは、都市化はまったく進まないが、地域開発は盛んに行われている地域と判断することが出来る。このことから、同じ開発推進地域でも前者は、都市化の伴う開発型でおそらく民間デベロッパー主導型であるのに対して後者は、都市化の伴わない開発型で公共事業主導型であるのではないかと思われる。離島を多く抱える鹿児島県においては明瞭な形であらわれているが、他の過疎化の進んでいる地方の県でも同様な結果が見られるのではないかと推察できる。

次に、図-8を見ると、第1象限に属する国分市、溝辺町、松元町などは都市化も進み、かつ、農業の相対的地位も高い地域と判断することが出来る。一方、第4象限の三島村、南種子町、和泊町、天城町、伊仙町、与論町、十島村などは、都市化は進まず、かつ、農業の相対的地位の高い地域と判断できる。この場合、前者は近郊農業的な色彩の強い農業が、後者は、伝統的な農業が中心となっているのではないかと思われる。

以上のように、3つの主成分を使って、その地域特性を吟味し、それによって市町村をいくつかの類型にグルーピングすることが可能になるのである。

4. おわりに

本稿においては、鹿児島県の全96市町村について、経済的社会的観点から人口、住宅形態、産業の各セクションから33の変数を選択して主成分分析を行った。その結果、第1主成分から第3主成分までの累積寄与率が53.68%であるので、全変動の半分以上がこの3つの主成分で説明される。したがって、第1主成分（都市化の水準）、第2主成分（地域開発の推進力）、第3主成分（農業の相対的地位）によって、各市町村の地域特性を吟味し、地域分類を行った。その結果、離島が多く、過疎化地域を多く抱える鹿児島県においては、都市化の進む地域が鹿児島市を中心とした地域に集中し、また、地域開発のあり方が、過疎化地域への公共投資という形に偏っている傾向があることが明らかになった。ただし、選択した33の変数が鹿児島県のような離島を多く抱える地域を対象にする場合に適切であったのかどうか、また、3つの主成分の累積寄与率が53%強ということで全体を説明するには少し説明力不足が否めない等、さらに検討を要するが、それは今後の課題としたい。

[注]

注1) 変数の選択については、神頭〔4〕を参考にした。

注2) 商店販売額が掲載されていない市町村のデータについては、売場面積との関係が大きいと考えて回帰分析を用いて推定した値を用いた。

注3) 各産業の比は各市町村の全産業の就業者に対する比を示しており、各住宅形態の比は各市町村の全住宅形態に対する比を示している。

注4) 産業就業者数及び住宅形態についてのデータは、平成12年度の国勢調査を、他のデータについては、『2002年度版地域経済総覧』東洋経済を利用した。

参考文献

- [1] Armstrong, H. and J. Taylor Regional Economics & Policy, Prentice-Hall, 1993.
坂下昇監訳『地域経済学と地域政策』流通経済大学出版会, 1998年。
- [2] 有馬哲, 石村貞夫『多変量解析のはなし』東京図書, 1987年。
- [3] 江沢譲爾, 金子敬生編『地域経済の計量分析』勁草書房, 1973年。
- [4] 神頭広好『計量分析にもとづく「まちづくり」－三好町を対象として－』愛知大学経営総合科学研究所, 1998年。
- [5] 神頭広好『都市と地域の立地論－立地モデルの理論と応用－』古今書院, 2000年。
- [6] 杉原敏夫, 藤田渉『多変量解析』牧野書店, 1998年。