

Title	主成分分析の地理学的応用：都市化による地域区分
Sub Title	Geographical application of the principal component analysis
Author	高橋, 潤二郎
Publisher	慶應義塾経済学会
Publication year	1971
Jtitle	三田学会雑誌 (Keio journal of economics). Vol.64, No.4 (1971. 4) ,p.161(25)- 181(45)
JaLC DOI	10.14991/001.19710401-0025
Abstract	
Notes	論説
Genre	Journal Article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00234610-19710401-0025

慶應義塾大学学術情報リポジトリ(KOARA)に掲載されているコンテンツの著作権は、それぞれの著作者、学会または出版社/発行者に帰属し、その権利は著作権法によって保護されています。引用にあたっては、著作権法を遵守してご利用ください。

The copyrights of content available on the KeiO Associated Repository of Academic resources (KOARA) belong to the respective authors, academic societies, or publishers/issuers, and these rights are protected by the Japanese Copyright Act. When quoting the content, please follow the Japanese copyright act.

えた一般組合の運動とそのイデオロギーが、何故に、その主体的勢力となりえなかったのか、イギリス労働運動史におけるひとつの興味ある問題が提起されるであろう(未完)。

主成分分析の地理学的応用

— 都市化による地域区分 —

高橋潤二郎

1. はじめに

最近の地域開発政策の進展、地域経済分析方法の発展とともに、中央官庁、地方自治体さらに大学をはじめとするさまざまな調査研究諸機関がそれぞれ独自の地域区分を設定し、それにもとづいて作業をすすめるようになったが、これら地域区分に共通にみられる一つの特色は、それがいわゆる「形式地域」、ブードヴィユによって「計画地域」planning region として規定された性格を多分にもっていることである。たとえば、新全国総合開発計画は、日本全土にわたって、七つのブロックと三つの地帯を設定しているが、これらが、たとえ機能性ないし同質性原理を考慮して、区割されたものであるにせよ、本質的に開発を目的として設定された計画地域であることには変わりはない。これと同様に、経済企画庁をはじめ、いくつかの研究機関が試みている我国経済に関する地域計量モデルにおける地域区分も、シミュレーションを通じて日本全土にわたる経済循環の実態を把握し、将来の動向を予測するために設定されたものではあるが、研究の最終結果というよりは、むしろ、研究を行なうための作業仮設的な地域区分であり、したがって、一種の形式地域だといってよいだろう。

このような最近の動向は、しかし、一方における「実質地域」としての地域概念の重要性を否定するものではない。むしろ、このタイプの地域区分、すなわち、経済諸現象の空間的分布と相互関係の実態を可能なかぎり正確かつアップトゥデートなかたちで把握し、これにもとづいて、地域の設定を試みる、いわば、地域(区分)そのものを最終結果としてもつような研究の意義は、以前にもまして重要なものとなりつつあると思われる。

研究の成果として見出だされた実質地域が果してそのまま具体的な地域開発政策における計画地域として採用され得るか否か、又地域経済計量モデルの中に組み入れられるべきものか否か、これはあらためて議論すべき余地のあるところであるが、少なくとも、これら形式地域の設定にあたっては、実質地域の存在を十分考慮することが必要であり、望ましいことは否定し難いところであろう。こ

の意味で、最近のさまざまな形式地域の設定とともに、実質的な地域区分の役割は今後一層その重要性をたかめるものと思われる。

こうした地域区分に対する外的要請に対して、地域区分そのものの手法はどうかといえば、最近の統計的技法の発展とその導入によって、ますます技術的に高度化しているのが現状である。特に、注目すべきは、主成分分析や因子分析等のいわゆる多変量分析の導入によって、非常に多数の変数からいわゆる合成変数を抽出することが可能となったことであろう。これによって、本来多変量的性格をもつ基域特性が、その多変量的性格を失うことなく、いくつかの基本的成分ないし因子によって表現されるようになったことは注目に値するといわなければならない。又、マハラノビスの汎距離その他の距離概念の採用によって多次元空間への対象布置ならびに距離測定は理論的により厳密なものとなりつつあるし、かつ、これら多次元空間に布置された対象のグループ化もクラスタ分析その他の導入によって著しく効率化されつつある。更に、これら多変量分析と多次元空間への対象の布置およびグループ化の過程がコンピューターの導入によって、著しく高速化し、精度の高いものになりつつあることも附記しておくべきことであろう。

本稿の主要な目的は、このような最近の地域区分に関連するデータ処理過程のなかで最も中核的な位置を占めるとされる主成分分析を我国の府県別データに適用し、その結果抽出された合成変数——「都市化」をあらわすものと解釈される——をもとにして、実質的な地域区分を試みた結果について報告することにある。

2. 主成分分析と地域区分

主成分分析の基本的アイデアとそれが地域区分との関連でいかなる役割を果すかについては、既に Berry (1961, 1968), King (1969) が詳述しているところであり、又、我国では、高橋 (1967), 奥野 (1971) が触れているところであるが、両者の関係は要約して、次のようにいうことができるだろう。すなわち、いま、「調査対象となる全域が N の基域に区分され、各基域毎にその特性をあらわす n の変数に関する観察値 $x_{ij} (i=1 \dots N, j=1 \dots n)$ が得られたものとする。これら $N \times n$ の観察値 x_{ij} を基礎的データとして、 N の基域をそれより少数の K の可能なかぎり同質的な「連接域」に組みわけ作業をいわゆる同質地域区分の定義と考えるとすれば、この地域区分の作業は、基本的には、各変数を軸とする n 次元の距離空間を構成し、ここに各基域を点として布置し、これら N の点間の距離を測定し、これを規準にして、 N の点を K のグループに組みわけるといふより一般的な分類問題の特殊ケースと理解することができよう。ただこの場合、問題となるのは、

- (1) 各基域の特性をあらわす各変数がことなつた測定水準にある場合、これを適切な手段によ

て標準化しなければならない

- (2) 各変数間に相関関係がみられる場合、この相関の効果をとりぞくために、軸の回転をはからねばならない

ということであろう。更に、分類にあたって、

- (3) 各基域の特性をあらわす非常に多数の変数をそれよりはるかに少数の基本的な変数で代表することができる

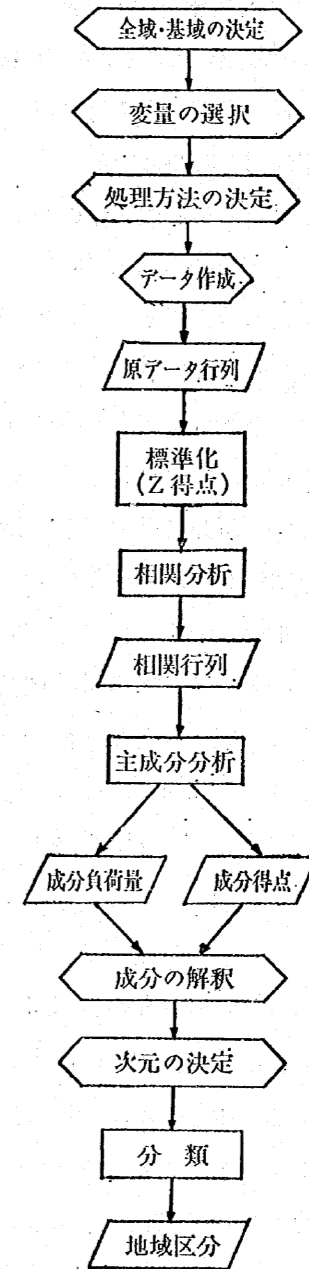
とすれば、分類の作業が著しく軽減するだろうことも容易に推察されることであろう。主成分分析は、この三つの問題をともに解決することによって、地域区分を行なう上で、基本となる少数の次元からなる特性距離空間を構成し、そこに、各基域を点として布置するに役立つのである。

高橋 (1967), King (1970) にみられるように、主成分分析の実際の計算過程はかなり複雑であるが、コンピューターの導入によって、基域の特性をあらわす $N \times n$ の原データ行列 $X = [x_{ij}]$ をいわゆる Z 得点を求めることによって標準化することからはじまり、各変数間の相関行列 R を求め、これを基礎に、固有方程式 $|R - \lambda I| = 0$ の解を得ることによって、いわゆる成分負荷量 component loadings を求め、更に、各基域の成分得点 component scores を求めるまでの全過程がプログラム化されるに至った現在、同分析はかつてそうであったより一般研究者にとって、はるかに近づき易い存在となったといえるであろう。換言すれば、これらデータ処理過程のプログラム化によって、研究者の作業は、事実上、計算過程へのインプットであるデータの作成とアウトプットである成分負荷量と成分得点の解釈に限定されることになったわけであり、この意味で、研究効率は著しくレベル・アップされることになった。

図1は、今回の地域区分において採用されたデータ処理過程をフロー・チャートの形で示したものであるが、研究者の意志決定のフェイズは、事実上、処理過程の前半では、原データ行列作成までの段階と後半では主成分分析の結果抽出された成分の解釈と次元の決定に限定されている。

研究者が長期にわたる煩雑な計算過程から解放され、本来の役割であるこうした意志決定にのみ集中し得ることになったのはまさに画期的なことだといわなければならない。

図1 基本的データ処理過程



しかしながら、他方、こうした計算過程の機械化・高速化が、研究者にとって従来経験しなかつたいくつかの問題をもたらしていることも事実であろう。この点、特に重視しなければならないのは、計算の全過程が機械化され、事実上、研究者の手をはなれてしまい、かつ、処理過程が高速化し、データをインプットすれば、数時間後にはアウトプットを入手することができるようになった結果、分析結果を解釈する上で、基本的に重要である研究対象としての全域と各基域に関する適切な「計量的イメージ」を研究者自身がつくことを著しく困難ならしめたことであろう。地域区分にかぎらず、あらゆる域データの計量的分析において、われわれが研究対象となる基域の特性およびその相互の関係について明確な「計量的イメージ」をもつことは、分析結果を適切に解釈するための不可欠の条件である。この場合、この計量的イメージが、単に、各基域の特性をあらわす諸変量——人口、所得、生産物出荷額等——に関する絶対値を記憶するというのではなく、むしろ、全域の平均値と各基域のそれからの偏差、又、各変量間の相関関係を明確に把握することによって得られるものであることはいままでもない。換言すれば、基域の特性をあらわすさまざまな変量間に見られる areal covariation の実態を明確に把握することが計量的イメージ形成の基礎となっているといえるのである。コンピューター導入以前の段階では、長時間にあたる計算過程を通じて、研究者はきわめて自然にこの計量的イメージを身につけ、したがって、分析結果を解釈する時点では、かなりまとまった研究対象に対するイメージを確立しており、結果を解釈する際十二分にこれを活用することができたわけであるが、前述のように、データ処理過程が機械化し、著しく高速化した現在では、研究者自身の中で計量的イメージが形成されるいとまもなく、分析結果がでてしまい、その解釈をしなければならないという事態があらわれつつあるといつてよいだろう。これは地域区分そのものに関する問題というよりは、すべてのデータ処理に関していえる一般的な問題であり、いわば、データ処理過程における man-machine system にみられるコミュニケーション・ギャップに関する問題ともいえるが、特に、主成分分析に代表される多変量分析を試みる際には、考慮されてしかるべき問題といえるであろう。

3. 予備的データ処理

あらゆる地域区分の出発点は、研究対象である「全域」の決定とこの全域をいくつかの地域すなわち同質的な接続域に区割るにあたって、基本となるデータの最小集計単位である「基域」の決定である。本稿では、全域を沖縄をのぞく北海道、本州、四国、九州とし、基域として都道府県を採用している。より厳格な意味での地域区分には、市町村ないしメッシュ・データを採用することが望ましいことはいままでもない。市町村に対して都道府県を基域として採用したのは、第一義的には利用可能なデータ上の制約（特に生産・分配所得等の地域会計上の諸変量の入手可能性）によるが、

市町村の数があまりに多く、かつ、その特性が多様であり、各変量の分散がきわめて大であり、データ処理上の困難が予想されたこともその一因となっている。都道府県を基域として採用したことは、分析結果である地域区分に対してあきらかな影響を及ぼしており、後述するように、最終的な二地域区分を行なう際に、本来ならば、関東から近畿に至る太平洋岸とその他地域とに区割されるべきものが、京都府と兵庫県の存在があるために、東北日本と西南日本とに区分されるという結果をもたらしている。

各基域の変量の選択は、地域区分の目的によって相違する。本稿における地域区分の目的はあくまでも経済地域の区分、すなわち、基域である都道府県の経済的諸特性にもとづく同質地域の設定にあるから、選択対象となる変量はそのかぎりでも自ずから限定されるが、しかしそうはいっても、さまざまな経済変量のうちいずれを選択し、原データ行列を作成するかはきわめて難しい問題である。主成分分析にもとづく地域区分そのものが帰納的性格をもつものである以上、収集される基礎データは可能なかぎり多数かつ多様であることが望ましいことはいままでもないが、データ処理能力に一定の限界があることからいって、これら多数のデータから県経済の実態を最もよくあらわす一定数の変量を選択することが必要である。と同時に、これら選ばれた変量が既存の地域経済理論にてらして十分理論的な意味をもち得ることも必要であろう。

又、変数の選択にあたって、無視できぬことの一つに、基礎的データの表現方法に関する決定がある。一般に域データの表現形式は、(1)規模（絶対量）、(2)構成比、(3)単位当り、(4)変化率等があり、それぞれ広範に採用されているが、各変量を示すに当ってこれらいずれの表現形式を採用するか、これまたきわめて大きな問題である。たとえば、人口特性をあらわすために、幼少年、生産年齢、老齢の各人口にわけるといことがしばしば行なわれるが、これら変量を、規模すなわち絶対数で表現するか、或いは、比率（構成比）で表現するか、それぞれのケースで相関行列は全くことになったものになり、分析結果もまたことなってくる。

本稿では、このような問題を考慮して、予備的なデータ処理として、

- (1) 現在利用可能な都道府県データのうちから、i 基域特性、ii 人口特性、iii 産業構造・職業構成、iv 所得——生産・消費支出、v 生活水準、vi 生産活動、vii 流通活動、viii 財政活動、ix 投資活動（公共投資）、x その他の 10 項目に該当する 60 変量を選び、

(2) これらを相関分析することによって、変量相互間の関連を把握するとともに、リダンダントな変量をあらかじめ除去するという操作を試みることによって、主成分分析を行なう上で適切と思われる変量を選定することにした。

表1は、これら 60 変量を、前述の各項目のもとにまとめたものであるが、民間投資に関する諸変量をのぞいて、県経済の特性を比較検討する上で、現在利用可能なデータはほぼカバーしてあ

主成分分析の地理学的应用

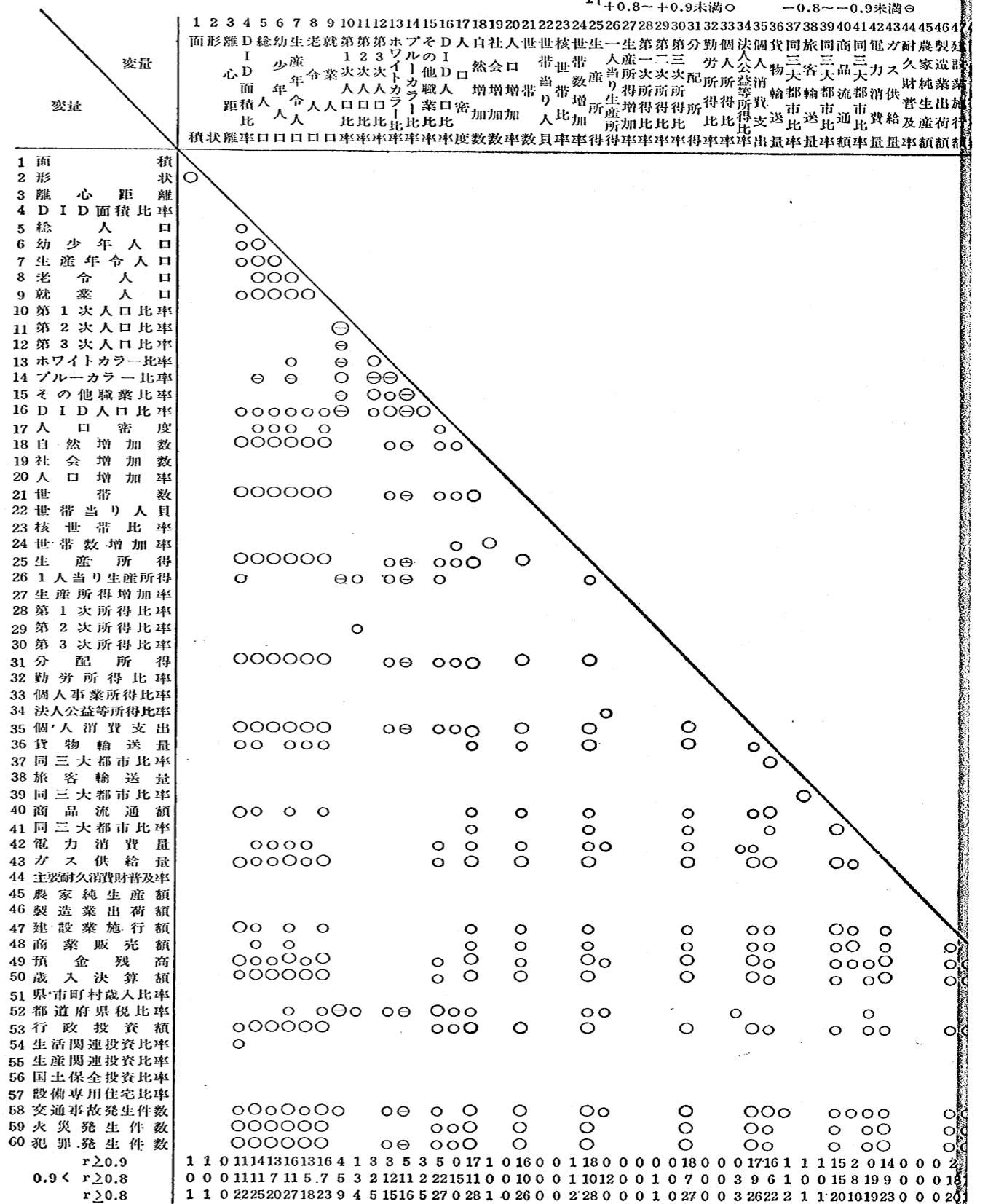
表 1 分析のための予備的データ表 (60 変量)

項目	変 量	資 料 (採用年次)	項目	変 量	資 料 (採用年次)	
i 基礎特性	面積	国調 40	v 生活水準	電力消費量	経済企画庁「地域経済要覧」40	
	形状	県域をかこむ最小円の半径		ガス供給量	通産省「ガス事業統計年報」40	
	離心距離	東京一県都間の直線距離		主要耐久財普及率	経済企画庁「消費者動向予調査報告書」43 (県庁所在都市)	
	D I D 面積比率	国調 40		設備専用住宅比率	総理府統計局「昭和38年住宅統計調査報告書」38	
ii 人口特性	総人口	国調 40	vi 生産活動	農家1戸当り農業純生産額	農林省「農業所得統計」40	
	幼年人口	国調 40		製造業一事業所当り出荷額	通産省「工業統計表」40	
	生産年齢人口	国調 40		一店舗当り卸・小売販売額	通産省「商業統計表」41	
	高齢人口	国調 40		金融機関一店舗当り預貯金残高	日本銀行「都道府県別経済統計」40	
	自然増加数	国調 40 (35~40)	建設業一事業所当り施行額	建設省「建設工事施行調査」42		
	社会増加数	国調 40 (35~40)	vii 流通活動	貨物輸送量	「地域経済要覧」40	
	世帯増加数	国調 40		同対3大都市比率	「地域経済要覧」40	
	世帯当り人員	国調 40		旅客輸送量	「地域経済要覧」40	
	核世帯比率	国調 40 (A-I-1, A-II-4, 6, 7)		同対3大都市比率	「地域経済要覧」40	
	世帯増加率	国調 40 (35~40)		商品流通額	「地域経済要覧」40	
人口増加率	国調 40 (35~40)	同対3大都市比率		「地域経済要覧」40		
iii 産業構造・職業構成	D I D 人口比率	国調 40	viii 財政活動	県歳入決算額	自治省「地方財政統計年報」40	
	人口密度	国調 40		県、市町村歳入比率	「地方財政統計年報」40	
	就業人口	国調 40		都道府県税比率	「地方財政統計年報」40	
	第1次就業比率	国調 40		ix 投資	行政投資額	自治省「都道府県別行政投資等実績調査報告」40
	第2次就業比率	国調 40			生活関連投資比率	「地域経済要覧」40
第3次就業比率	国調 40	産業関連投資比率	「地域経済要覧」40			
ホワイトカラー比率	国調 40 (専門的・技術的職業、管理的職業、事務従業者をふくむ)	国土保全投資比率	「地域経済要覧」40			
ブルーカラー比率	国調 40 (農林漁業、採鉱石、運輸通信技能工、生産工程従事者、単純労働者をふくむ)	x その他	交通事故発生件数		「地域経済要覧」40	
その他職業比率	国調 40 (その他従業者)		火災発生件数	「地域経済要覧」40		
iv 所得・消費支出	生産所得		経済企画庁「県民所得統計」40	犯罪発生件数	「地域経済要覧」40	
	1人当り生産所得		「県民所得統計」40			
	生産所得増加率		「県民所得統計」40 (35~40)			
	第1次生産所得比率	「県民所得統計」40				
	第2次生産所得比率	「県民所得統計」40				
	第3次生産所得比率	「県民所得統計」40				
	分配所得	「県民所得統計」40				
勤労所得比率	「県民所得統計」40					
個人所得比率	「県民所得統計」40 (個人業主、貸貸・利子所得をふくむ)					
法人公益等所得比率	「県民所得統計」40 (法人、公益事業剰余をふくむ)					
個人消費支出	「県民所得統計」40					

と考える。

表 2 は、これら 60 変量間の相関関係を示したものであるが、総人口をはじめとするデモグラフィックな特性をあらわす変量間に非常に高い相関関係がみられるとともに、生産・分配所得、消費支出額等、そして、交通事故、火災、犯罪発生件数もこれと密接に関連していることがうかがわれよう。特に、興味あることは、人口の自然増加数が D I D 面積比率、同人口比率と高い相関関係をもっていることであって、このことは昭和 30 年代にみられた人口の急激な都市集中による人口再生

表 2 予備的データ (60 変量) 相関表



主成分分析の地理学的应用

表 3 都道府県 主成分分析固有値および説明分散比率

成分	λ_i	$\sqrt{\lambda_i}$	分散比率	累積分散比率
第 1 成分	24.2250	4.9219	53.83 (%)	53.83 (%)
第 2 成分	4.0948	2.0236	9.10	62.93
第 3 成分	3.2082	1.7911	7.13	70.06
第 4 成分	2.4398	1.5620	5.42	75.48
第 5 成分	2.0794	1.4420	4.62	80.10
第 6 成分	1.3849	1.1768	3.08	83.18
第 7 成分	1.1419	1.0686	2.54	85.72
第 8 成分	1.0437	1.0216	2.32	88.04
第 9 成分	0.9388	0.9689	2.09	90.13
第 10 成分	0.7086	0.8418	1.57	91.70

のであるが、説明する分散比率は、第 1 成分が 53.83% ととび抜けて高く、第 2 成分以下を大きくひきはなしているのが注目される。第 10 成分までの累積比率は 91.70% におよんでいるが、第 5 成分までの同比率は 80.10%、結局、前述の 5 成分で全体の変動の約 8 割を説明していることになる。

主成分分析の結果、抽出された成分を解釈するに当って、未解決の問題の一つに何番目までの成分を解釈の対象とするか、があるが、5%以上の説明力をもつ成分までを限度とするというのが通常の rule of thumb であり (King 1969)、事実、第 6 成分以下の解釈はかなり難しい。

表 4 および表 5 は、これら 5 成分の成分負荷量と成分得点を示したものであるが、前者は自乗和が 1 になるように標準化してあり、(通常の成分負荷量を求めるには、各成分の固有値 λ_i の平方根を各負荷量に乗ずればよい) 又後者は平均が 0、分散が 1 となるようにこれまた標準化してある。これら両者にもとづく各成分の解釈は以下の通りである。

(1) 第 1 成分

原データの全分散の 53.83% を説明するこの成分には総人口、生産年齢人口、ホワイトカラー比率、D I D 人口比率、自然増加数、世帯数、生産所得、1 人当り生産所得、電力消費量、金融機関一店舗当り預金残、交通事故発生件数等、非常に多数の变量が強い正の関連をもっているが、このうち、最も高い負荷量を示すのは、D I D 人口比率 (0.9470)、次いで人口自然増加数 (0.9401)、生産所得 (0.9332)、交通事故発生件数 (0.9332) であり、他方、高い負の関連を示すものとして、第 1 次産業人口比率 (-0.9036) があげられる。正の関連をもつ变量は、そのほとんどが経済的集積特に都市的なそれをあらわすものであり、その意味で、この成分は都市化と解釈するのが妥当ではないかと思われる。

図 2 は、この成分に関して正の得点を得た 13 都道府県と負の得点を得た 33 県を示したものであるが、最も高い得点を得た東京都 (2.7653) から次位の大阪府 (2.2302) 以下、神奈川、兵庫、埼玉、京都、福岡、千葉、北海道、広島、静岡、山口まで、正の得点を得た基域が最後の山口県をのぞいていずれも

主成分分析の地理学的应用

表 4 45 変量成分負荷量

変量 \ 成分	第 1 成分	第 2 成分	第 3 成分	第 4 成分	第 5 成分
1	0.0111	0.1119	0.4090	-0.1768	-0.2819
2	0.0336	0.1093	0.3827	-0.1772	-0.3135
3	0.0681	0.2011	0.2095	0.3489	0.0567
4	-0.1798	0.1541	-0.0979	-0.0425	0.0636
5	-0.1855	0.1404	0.0731	-0.1017	-0.0520
6	-0.1873	0.1416	0.0525	-0.0948	-0.0499
7	0.1836	0.0901	-0.0161	-0.1336	0.0626
8	-0.1649	-0.1911	-0.0867	0.0726	-0.1638
9	-0.1563	0.0964	0.1690	0.1846	0.1153
10	-0.1838	0.0168	0.0870	0.1109	0.1142
11	-0.1924	0.0309	0.0577	0.0456	0.0211
12	-0.1753	-0.0106	0.1664	-0.1642	0.0975
13	-0.1910	0.1240	0.0330	-0.1064	-0.0171
14	-0.1549	-0.2124	-0.0640	-0.0844	0.1916
15	-0.1353	-0.2066	-0.0146	-0.1519	0.0684
16	-0.1855	0.1589	0.0540	-0.0722	-0.0411
17	0.1240	-0.0334	-0.0526	-0.3904	-0.0055
18	-0.1467	0.0510	0.2103	0.2029	0.1009
19	-0.1610	-0.1974	0.0632	-0.1678	0.1486
20	-0.1896	0.1479	-0.0051	-0.0740	-0.0376
21	-0.1837	-0.0488	-0.0404	0.0357	-0.1075
22	-0.0866	-0.1702	0.0425	-0.1389	0.3369
23	0.0683	0.1252	0.0893	0.1050	0.2622
24	-0.1320	-0.3105	0.0017	0.0749	-0.1952
25	-0.0173	0.3513	-0.0225	0.1493	0.2108
26	-0.1245	-0.1704	0.2401	0.2062	-0.0121
27	0.1577	0.1589	-0.1599	-0.1623	0.0941
28	-0.1656	-0.0913	-0.0127	0.0354	-0.2079
29	-0.1631	0.0250	-0.1320	-0.0031	-0.0502
30	-0.1199	-0.0832	-0.1424	-0.0565	-0.0282
31	-0.1632	0.1156	-0.1955	-0.0340	-0.1403
32	-0.1829	-0.0068	0.0428	-0.0743	-0.1281
33	-0.1730	0.2106	-0.1076	-0.0314	-0.0091
34	-0.0808	-0.1293	-0.2377	0.1166	-0.2734
35	0.0516	-0.0021	0.1639	-0.4739	0.0612
36	-0.1118	-0.2309	0.2038	-0.0015	0.0865
37	-0.1670	0.1683	-0.1323	-0.0663	-0.0281
38	-0.1687	0.1501	-0.1284	0.0313	-0.1647
39	-0.1895	0.0792	-0.1147	-0.0603	0.0254
40	-0.1110	-0.1693	0.1964	0.1397	0.0598
41	-0.1604	-0.0515	-0.0870	-0.0280	0.2733
42	0.0869	-0.0364	0.2251	0.0186	-0.1900
43	0.1215	0.0984	-0.1219	0.0859	-0.2339
44	0.1417	-0.2293	-0.1379	0.0735	-0.1048
45	-0.1916	0.0805	-0.0546	-0.0288	-0.0541

主成分分析の地理学的应用

表 5 都道府県成分得点

成分	第 1 成分	第 2 成分	第 3 成分	第 4 成分	第 5 成分
北海道 1	0.3073	0.4252	1.2740	-0.3079	-0.4992
青森 2	-0.4439	0.2646	0.0807	-0.2472	0.2144
岩手 3	-0.5126	0.1403	0.0285	-0.2464	0.0282
宮城 4	-0.2296	0.1832	0.2010	-0.2960	0.3519
秋田 5	-0.6190	0.0802	0.0413	-0.3092	-0.0124
山形 6	-0.6209	0.1458	-0.1141	-0.4312	0.0189
福島 7	-0.4444	0.1620	0.1408	-0.3786	-0.0415
茨城 8	-0.4022	-0.1023	-0.1371	-0.3832	0.0071
栃木 9	-0.2594	-0.1612	-0.2608	-0.2893	0.0228
群馬 10	-0.2008	-0.1329	-0.2129	-0.0881	-0.2225
埼玉 11	0.6817	-0.6733	-0.0418	-0.2421	0.3935
千葉 12	0.3929	-0.5079	0.1024	-0.3280	0.5034
東京 13	2.7653	1.0632	-0.3397	-0.1048	0.0041
神奈川 14	1.5675	-0.7424	-0.4927	-0.0743	0.3062
新潟 15	-0.2643	0.0287	0.0358	-0.2833	-0.3085
富山 16	-0.1519	-0.1459	-0.1698	-0.0004	-0.1416
石川 17	-0.1187	-0.1532	-0.1496	0.1205	-0.2119
福井 18	-0.3468	-0.0710	-0.3291	0.0247	-0.2276
山梨 19	-0.4378	0.0538	-0.2770	-0.0207	0.0285
長野 20	-0.3263	0.0137	-0.1009	-0.0907	-0.1641
岐阜 21	-0.0537	-0.1791	-0.1655	0.1546	-0.3030
静岡 22	0.2426	-0.3229	0.1346	-0.1449	-0.2663
愛知 23	1.2959	-0.2231	-0.2257	-0.0404	-0.3253
三重 24	-0.1005	-0.3125	0.0501	0.1549	-0.2408
滋賀 25	-0.3432	-0.2556	-0.0974	0.0646	-0.1559
京都 26	0.4909	-0.0285	0.0202	0.3542	0.0814
大阪 27	2.2302	0.2506	-0.4245	-0.0788	0.0003
兵庫 28	0.9949	-0.2469	0.2209	0.2340	-0.0317
奈良 29	-0.0966	-0.0990	-0.1453	0.0942	0.2643
和歌山 30	-0.1016	-0.1813	-0.0409	0.1359	-0.0322
鳥取 31	-0.5212	0.1257	-0.0745	0.0767	0.1649
島根 32	-0.6643	0.2613	-0.0642	0.1611	-0.0327
岡山 33	-0.1503	-0.1468	-0.1240	0.1444	-0.1421
広島 34	0.2729	-0.2726	0.1071	0.3821	-0.1131
山口 35	0.0566	-0.2400	0.1604	0.3968	-0.1659
徳島 36	-0.5382	0.1262	-0.2437	0.0631	0.0133
香川 37	-0.3022	0.0598	0.1235	0.1583	0.1029
愛媛 38	-0.2834	-0.0218	0.1017	0.1416	-0.0175
高知 39	-0.4722	0.1985	0.0288	0.2122	-0.0064
福岡 40	0.4797	0.1914	0.3546	-0.2849	0.0848
佐賀 41	-0.5551	0.1416	-0.1080	-0.0769	0.0879
長崎 42	-0.3168	0.3564	0.2841	0.4251	0.5092
熊本 43	-0.4640	0.2524	0.0037	-0.0285	0.1423
大分 44	-0.4420	0.1470	-0.0274	0.1565	0.1110
宮崎 45	-0.5066	0.2337	0.1466	0.2438	0.0603
鹿児島 46	-0.4879	0.3148	0.0591	0.3069	0.1605

主成分分析の地理学的应用

図 2 第 1 成分因子得点による区分

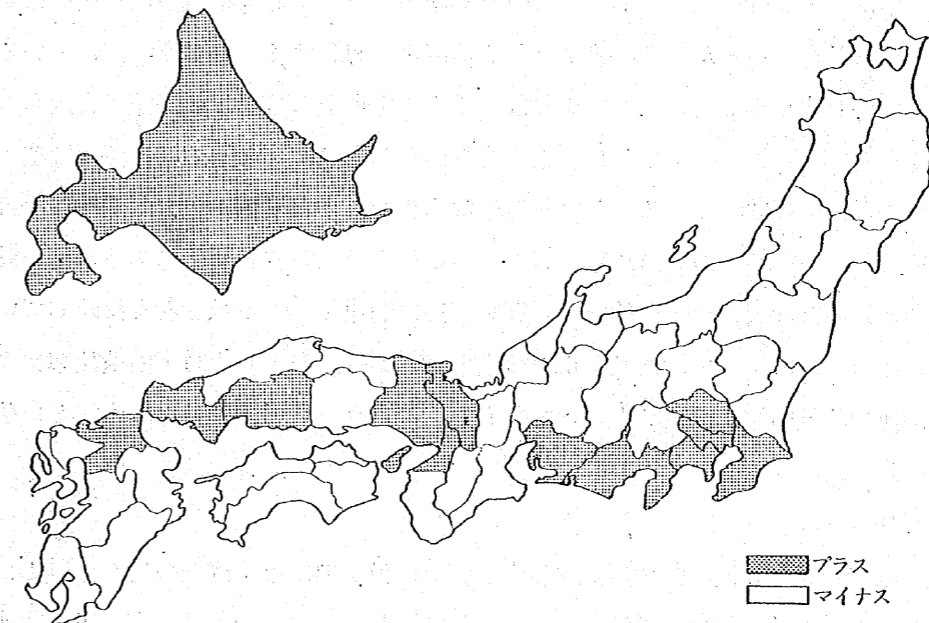
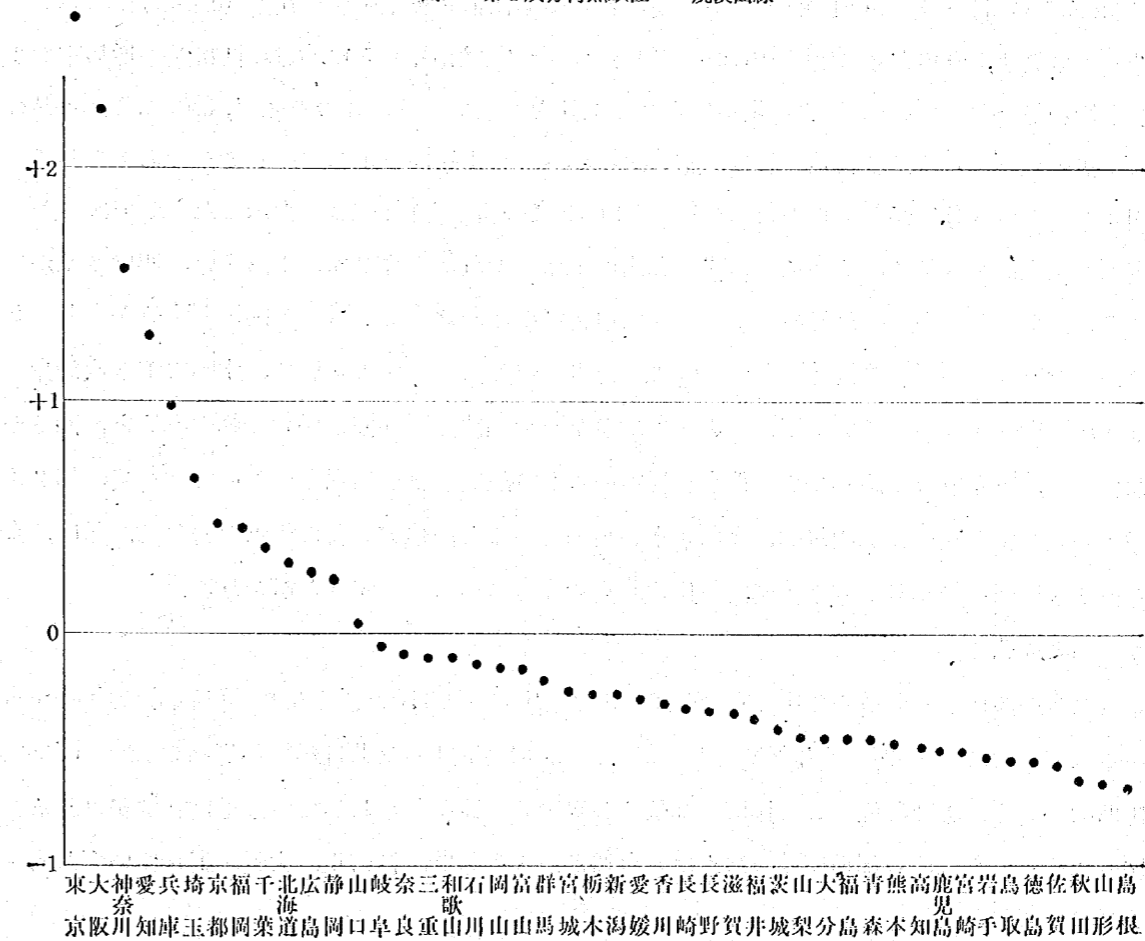


図 3 第 1 成分得点順位——規模曲線



都市化の進んだ府県であることは明瞭であろう。又、図3は、これらの各都道府県の成分得点をたて軸に、その順位をよこ軸にプロットしたものであるが、いわゆる都市の順位—規模曲線と非常によく近似した曲線を得ることができ、第1成分を都市化と解釈することをうらづけている。

ただ、ここで注意すべきことは、この成分はあくまでも多数の変量との関連で合成されたものであり、都市化を意味するとはいいながら、都市化をあらわす変量そのものだけでは説明され得ない、その他の変量の動きも反映していることである。都市化をあらわす第1成分が、一般に都市化の指標として認められているDID人口比率、DID面積比率と高い相関関係にあるのは当然であるが、成分はこれらの比率だけではなく、他の変量とも密接に関連し、その動きをも反映しているのである。山口県はその一例であって、同県が正の得点を得たのは、DID人口・面積比率によるよりは、むしろ同県が比較的低い第1次産業人口比率をもち、他方高い電力消費量をもつことに依存しているように思われる。

(2) 第2成分

第2成分は原データの分散の9.10%に因与する。その説明力は第1成分の約6分の1にすぎず、各変量の負荷量も第1成分に比べて低い値しか示していない。このうち、最も高い正の関連を示すのは、第3次生産所得比率(0.7110)、他方最も高い負の関連は第2次生産所得比率(0.6284)にみられ、製造業一事業所当り出荷額、設備専用住宅比率とマイナスの相関、さらに、人口、世帯数の増加率および社会増加等、デモグラフィックな増加率と負の関連をもっている。この意味で、この成分は各県の非農業活動にみられる産業格差、特に、第3次産業への特化傾向を示しているように考えられる。

図4は、この成分に関して正の得点を得た24県と負の得点を得た22県も示したものであるが、東京都と大阪府をはじめ、北海道、東北、九州の全県、それに中部3県、山陰2県、四国3県がプラスの得点を示している。周知のように、生産所得の産業別比率そのものを問題とするかぎり、我が国の都道府県は、わずかの例外(埼玉、神奈川、静岡、愛知、兵庫)をのぞいて、すべて第3次産業が第2次産業をりょう駕しているが、いま、平均値からの偏差という観点からこれをみると、図5のようになり、第2成分に関して正の得点を得ている24県は、大阪、新潟をのぞいてすべて、第2次産業に対して第3次産業が相対的に高い比率を占めているし、他方、負の得点を得た22県は、奈良をのぞき、すべて、逆に第2次産業の比重が高くなっていることがうかがわれる。

(3) 第3成分

第3成分の説明する分散の比率は7.3%、第2成分に比較して1段と低くなり、第1成分の約7分の1である。これまたいくつかの変量に関連しているが、面積と形状の負荷量がそれぞれ0.7325、0.6854と圧倒的に高く、集計単位として採用した県域の空間的特性をあらわしているものと解釈される。

図6にみるように、各都道府県の得点は北海道が最大であり、神奈川、大阪、東京はいずれもマイナスである。

図4 第2成分因子得点による区分

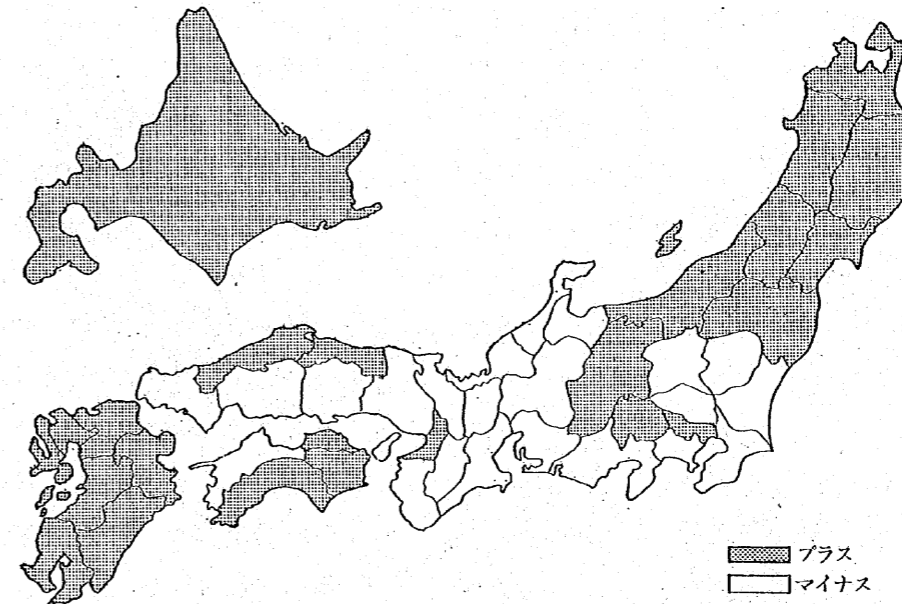


図5 第2次/第3次所得比率と成分得点(正・負)

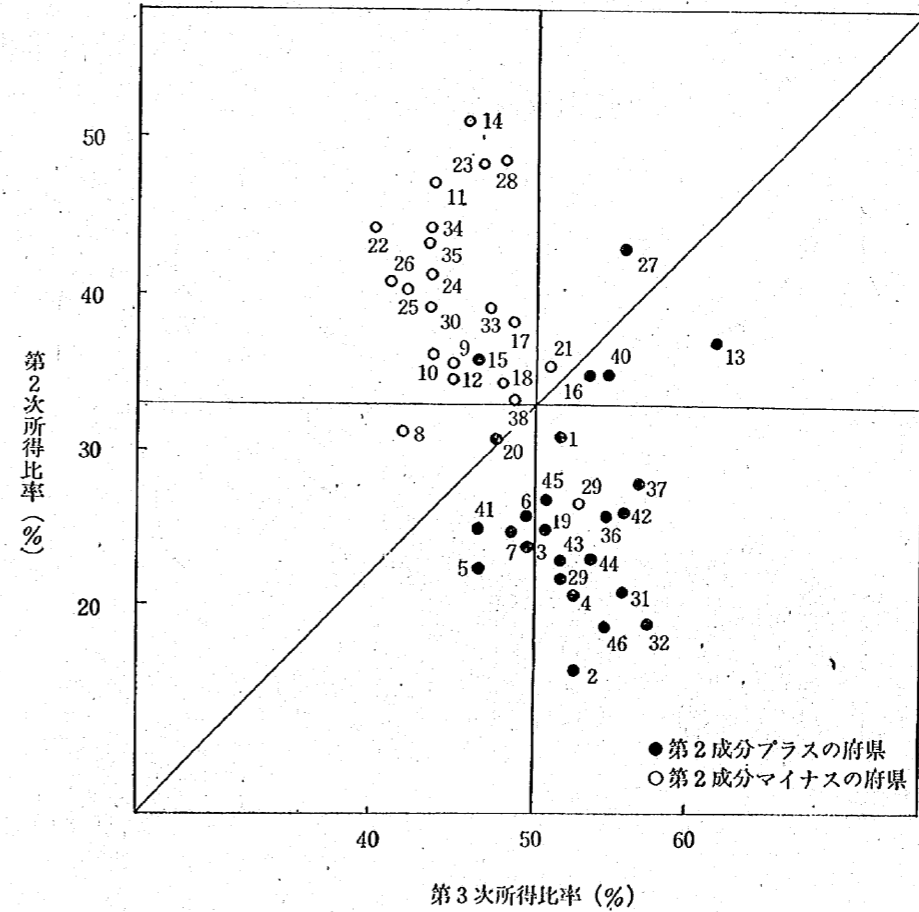


図6 第3成分因子得点による区分

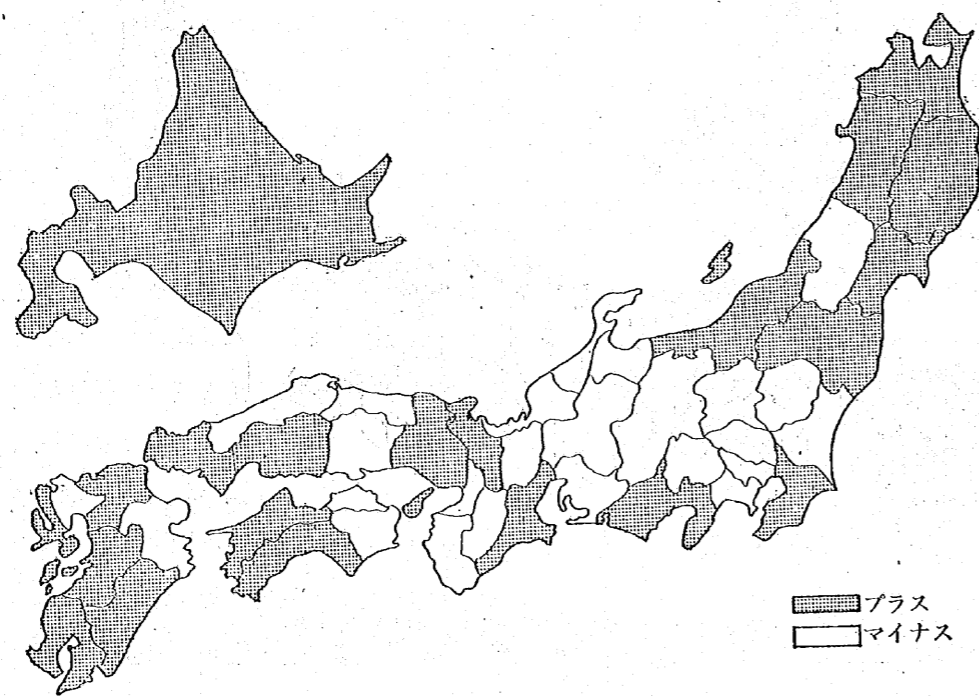
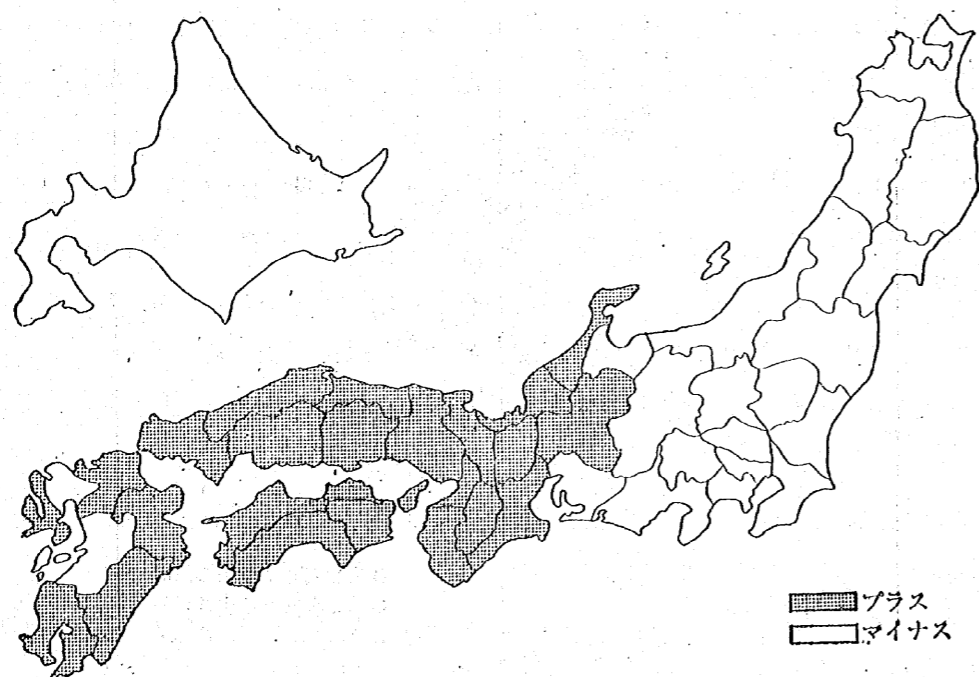


図7 第4成分因子得点による区分



(4) 第4成分

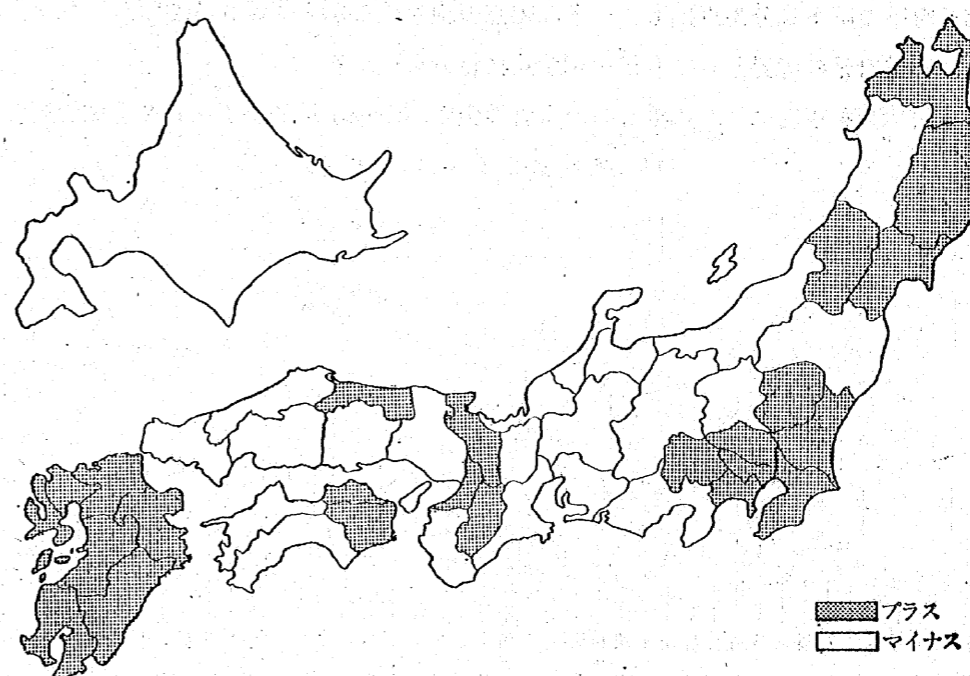
説明する分散比率はさらにさがり、5.42%、第1成分の約10分の1にすぎない。最も高い関連を示す変量は農家1戸当り農業純生産額(-0.7402)、次いで世帯当り人員(-0.6098)であり、農業活動における格差を示しているものと考えられる。この成分に関して、最も高い正の得点を得たのは、長崎県、次いで山口、広島県、逆に最も高い負の得点は山形県、次いで茨城、福島県の順である。

図7は、この成分得点の分布を示したものであるが、西日本と東日本とが割然とわかれており、農業活動における西南日本型と東北日本型の存在を暗示している。唯一の例外は九州における熊本と佐賀県であるが、前者は九州一の農業県、後者は「新佐賀段階」で知られる全国でも有数の米作経営の先進県であることを考えれば、これは当然といえよう。

(5) 第5成分

この成分の説明力はわずかに4.62%であり、各変量の負荷量も低く、著しく解釈が困難であるが、一応生産所得の増加率を示しているものと考えてよいだろう。図8からもあきらかなように、正の得点を得た県は、大都市周辺と東北、九州にかたまっている。東京、大阪はともに正の得点を得ているが、その値は低く、その周辺諸県、特に埼玉、神奈川、千葉が高い得点を示しているのが注目される。この成分には、行財政投資の構成比率も関連しており、その意味で解釈もし得ることをつけ加えておくべきであろう。

図8 第5成分因子得点による区分



5. 地域区分

以上、主成分分析の結果抽出された5成分に関する解釈を示したが、これら5成分によって示される諸特性は、都道府県を基域として地域区分する際に構成される特性距離空間の基本的次元として、考慮の対象となり得るものといえよう。その中でも、特に重視すべきが、都市的集積をあらわすものと解釈された第1成分であることはいうまでもない。もし、この第1成分に対する解釈が妥当であると仮定するならば、この成分を基本的次元として採用し、そこに各基域の得点値を布置し、各基域間の距離を測定することによって、可能なかぎり同質的な接続域を区劃するという通常のデータ処理を経ることによって、われわれは都市化にもとづく実質的な地域区分を得ることになる。この最終の過程において注意しなければならないことは、基域のグループ区分と地域区分とが全くことなつた概念であることだろ。前述のように、地域区分の基本は「Nヶの基域を可能なかぎり同質的なKヶの接続域に組みわけ」ることにあるのであって、単にNヶの基域をそれより少数のKヶの同質的なグループにわけることではない。同一地域を構成する基域は空間的に接続する——すなわち境界を接するないし共有している——ことが必要不可欠の条件となっているのである。

この意味での地域区分を行なうために、ここでは通常の基域間特性距離行列の他に、基域間の接続性をあらわす1-0行列を作成し、これら両行列を基礎にして、次のような分類過程を採用することにした。

1. 各都道府県の第1成分得点にもとづき、得点の絶対差を示す基域間特性距離行列を作成するとともに都道府県の境界共有の有無を示す接続行列をつくる。
2. これら両行列を基礎にし、接続ししかも特性距離(得点差)が最小である2基域を結合する。

図9 46都道府県の連鎖樹

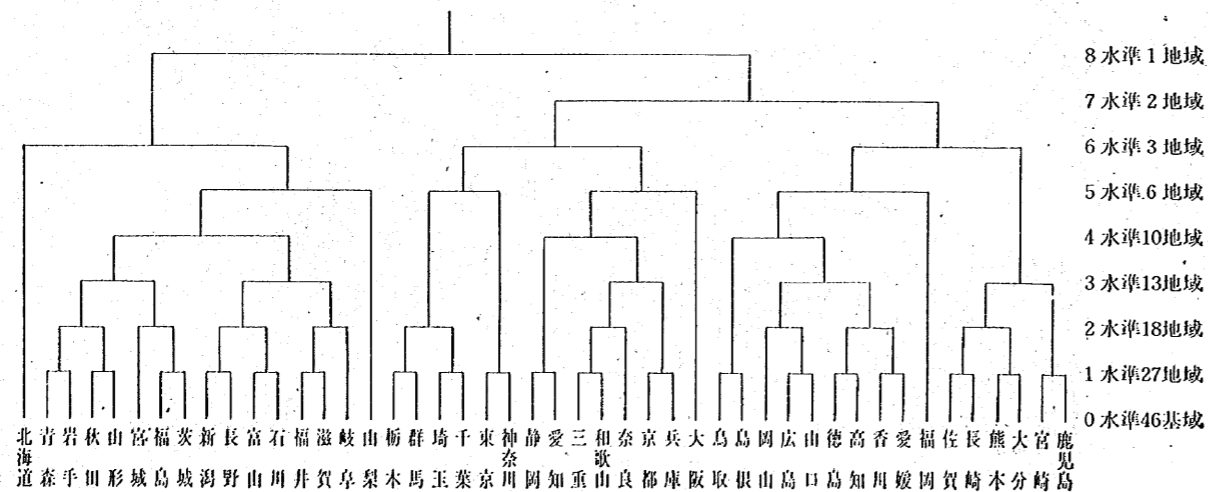


図 10-1 都市化にもとづく地域区分 (レベル1)

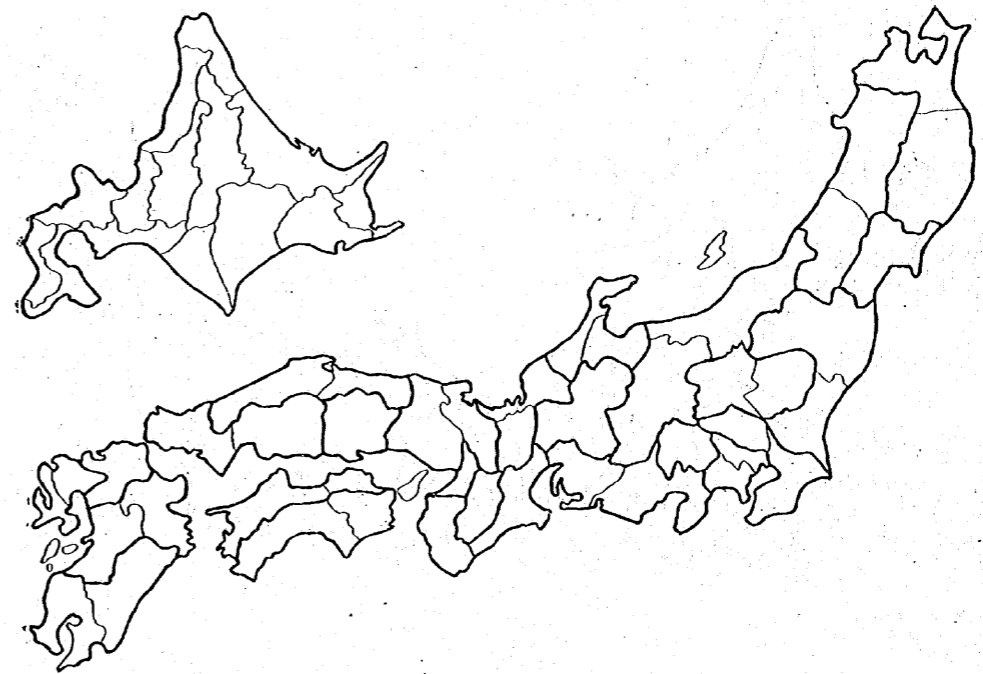


図 10-2 (レベル2)



図 10-3 (レベル3)

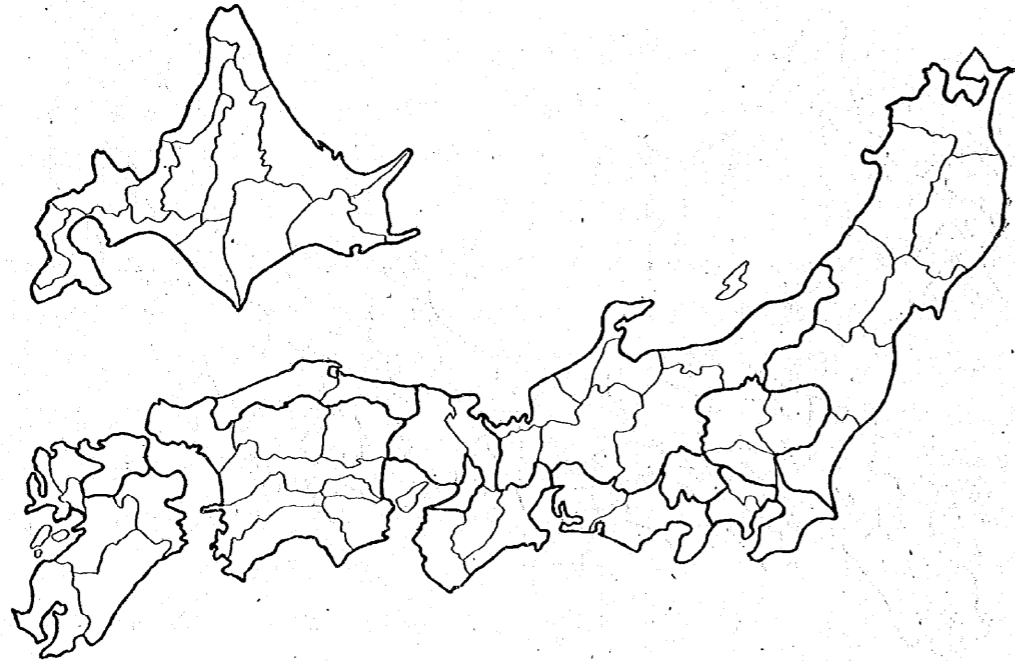


図 10-4 (レベル4)

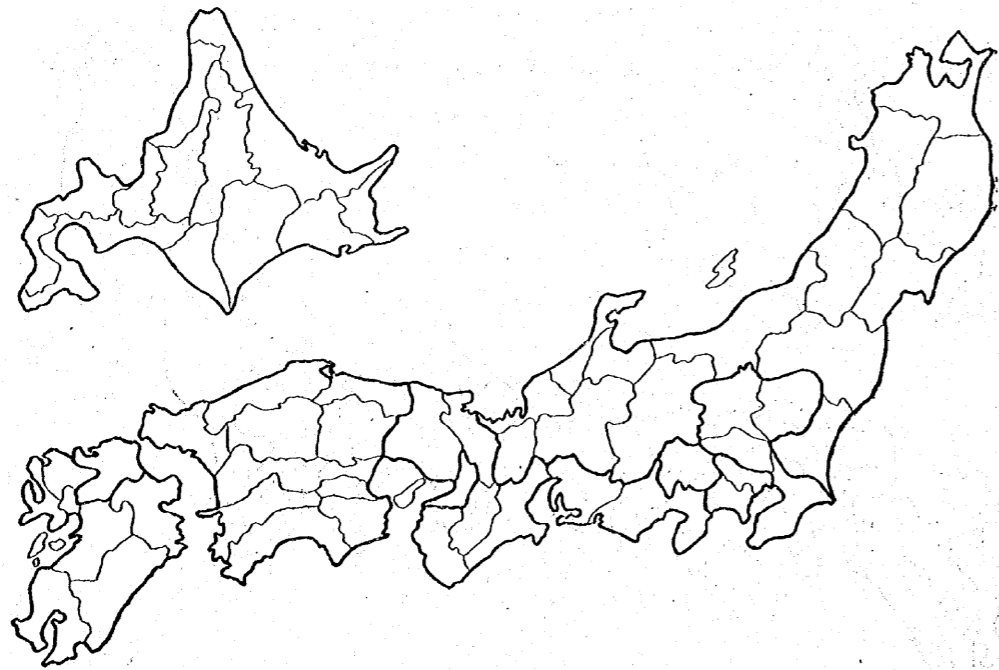


図 10-5 (レベル5)

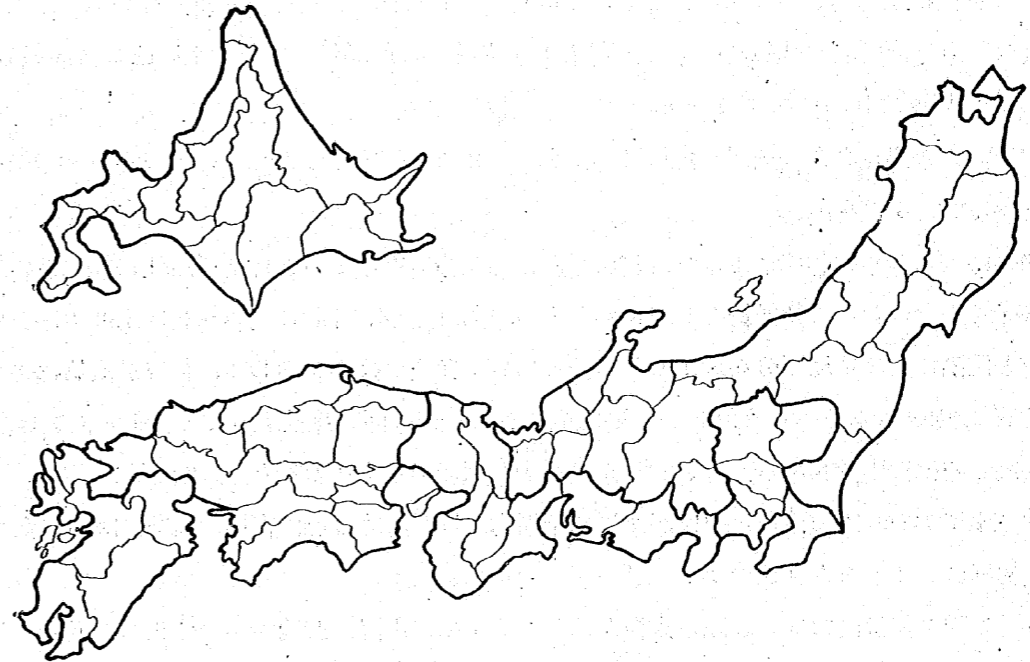


図 11 都市化にもとづく地域区分 (レベル6)



(この結果形成される地域を第1分類水準における地域と呼ぶ)

3. 各地域に結合された2基域間の得点を平均し、これら平均値を基礎にして、第1水準の地域間特性距離行列を作成するとともに、地域の境界共有の有無を示す接続行列をつくる。
4. これら両行列を基礎にして、接続ししかも特性距離が最小である2地域を結合する。(この結果形成される地域を第2分類水準における地域と呼ぶ)

以下、この操作を全基域が一地域にふくまれるまで継続することによって、各レベルでの地域区分を得ることができる。

図9は、これらの特性によって、基域である46都道府県がいかに結びつけられ、最終的に、一体化するかをいわゆる連鎖樹として示したものであり、又、図10-1~5は水準1から水準5までの各段階による地域区分を図示したものである。前述の操作を通じて、基域が次第により大きな接続域(地域)へと組みこまれてゆく過程がこれら一連の図を比較することによってあきらかとなる。この過程で興味ある点をいくつか列記すれば、およそ次のごとくとなる。

- (1) 水準1の段階で殆んどどの県が接続域を構成するが、北海道、宮城、山梨、岐阜、奈良、大阪、福岡はいずれとも結びつかない「ハナレザル」となる。
- (2) 水準2の段階では、北陸、東海をはじめ、山陰、山陽、四国等が一体化し、同時に東北が二分、九州が三分される等、地域の骨格がかなり明瞭化する。
- (3) 水準3の段階では、東北が一体化し、北陸が福井、岐阜、滋賀の三県をふくみ、九州が福岡をのぞいて一体化する。ここで興味のあることは、山陽が山陰とではなく、四国と結びつくことであろう。大阪をのぞいた近畿は既に一体化している。
- (4) 水準4の段階では東北と北陸が結びついた他、東海と近畿、中国と四国がそれぞれ一体化している。北陸が西にのびず、東北と結びつくのは、いうまでもなく、京都府と兵庫県があるため、もしこれらがなければ、北陸は山陰と結びつき、裏日本地域を形成した筈である。この段階におけるハナレザルは、北海道、福岡、大阪、山梨であり、東京-神奈川と関東四県も結びついていない。
- (5) 水準5の段階で、前記のハナレザルのうち後三者がそれぞれの地域に一体化した。ここで興味のあることは福岡が九州ではなく中国-四国へ、山梨が東北-北陸-中部へと合体したことである。大阪は東海-近畿に属し、関東もこの段階で一体化する。

このような段階を経て、最終的に図11に示す3地域区分を得ることができた。この3地域は、昭和40年現在の我国都市化の現況を最も端的に示すものであり、都市化にもとづく実質的な地域区分として十分一般性をもち得るものと考えられる。

6. おわりに

以上、我国都道府県を基域として、その経済的特性をあらわす45変量を対象にして主成分分析を行い、その結果抽出された5成分のうち、最も重要であると思われる第1成分に関する各府県の成分得点を基礎に、地域区分を試みた結果について報告した。この地域区分は「都市化」に焦点を合わせ、連鎖樹をつくることによって、多数の基域がいくつかの地域に分類されてゆく過程をフォローすることに主眼を置いたため、第1成分のみを基本軸とするいわば1次元空間における分類に止ったが、より一般的な地域区分においてはより多くの成分を基本軸として採用し、その結果構成される多次元空間に基域を布置し、基域間の距離を測定することが必要となることはいうまでもない。この多次元空間における分類に関して、現在最も有効であると思われる分析方法の一つとして、クラスター分析があげられる。今回試みた都道府県データに関する主成分分析の結果抽出された10成分を基本軸とするクラスター分析の結果、これとはまたちがった地域が導出されたが、これについては、機会をあらためて報告したい。

文 献

1. Berry B. J. L., "A Method for Deriving Multi-Factor Uniform Regions", *Przeegląd Geograficzny*, 33, (1961), 263-282.
2. — "A Synthesis of Formal and Functional Regions using a General Field Theory of Spatial Behavior", Berry, B.J.L. and Marble D.F. ed. *Spatial Analysis*, 1968, 419-428.
3. King, L. J., *Statistical Analysis in Geography*, 1968, 154-164.
4. 奥野隆史「人文地理学研究のための主成分、因子分析」*地理* 16 巻1号, 1971年 39-46.
5. 高橋潤二郎「地域区分のための主成分分析」*三田学会雑誌* 60 巻2号, 1967年 93-106.