

# Regional Clustering and Visualization of Industrial Structure based on Principal Component Analysis for Input-output Table Data

メタデータ	言語: jpn 出版者: 公開日: 2021-09-10 キーワード (Ja): キーワード (En): 作成者: YAMAMOTO, Keiko, Harada, Kaisei, Sagae, Masahiko メールアドレス: 所属:
URL	<a href="https://doi.org/10.24517/00064103">https://doi.org/10.24517/00064103</a>

This work is licensed under a Creative Commons Attribution-NonCommercial-ShareAlike 3.0 International License.



# 主成分分析に基づく地域クラスタリングと 産業構造の可視化

人間社会環境研究科 人間社会環境学専攻

山本 けい子

人間社会環境研究科 人間社会環境学専攻

原田 魁 成

人間社会研究域 経済学経営学系

寒河江 雅彦

## 要旨

地域の経済構造の把握や経済波及効果の分析に用いられる産業連関表（I/O表）の新たな活用法を提案する。I/O表は全国で一律かつ規則的に作成されるため、地域間の比較や産業構造の可視化、長期的な変化の抽出のための統計表として利用できると考えた。そこで、都道府県ごとに作成された産業連関表を多変量データとみなして、いくつかのデータ解析手法を適用した。本稿では、類似比較・可視化の観点から、複雑な高次元I/Oデータから主成分分析により低次元の特徴空間にデータを還元した後、クラスタリングとマッピングを試みた。クラスター分析では、縮小された次元を特徴づける産業グループで都道府県を分類している。マッピングでは、産業分野ごとに都道府県の相対的な位置関係を可視化している。さらに主成分分析では、都道府県の産業構造の概要と年次推移を把握することができる。最後に、これらの分析方法とその結果を通して多変量解析の枠組みを用いたI/O分析のプロセスをまとめる。

## キーワード

主成分分析, 産業連関表, 産業構造の可視化

## Regional Clustering and Visualization of Industrial Structure based on Principal Component Analysis for Input-output Table Data

Division of Human and Socio-Environmental Studies  
Graduate School of Human and Socio-Environmental Studies

YAMAMOTO Keiko

Division of Human and Socio-Environmental Studies  
Graduate School of Human and Socio-Environmental Studies

HARADA Kaisei

Faculty of Economics and Management, Institute of Human and Social Sciences

SAGAE Masahiko

## Abstract

We propose a new application of the input-output table (I/O table) to understand the economic structure and analyze the economic ripple effect in a region.

As the I/O table is created uniformly and regularly regardless of the study area, we thought that it could be used as a statistical table to obtain comparisons between regions, visualization of the industrial structure, and extraction of long-term change. We apply several data analysis methods by considering the I/O table created for each prefecture as multivariate data.

In this study, we attempt clustering and mapping, from the perspective of similar comparisons and visualizations after a principal component analysis is used to reduce data from complex high-dimensional I/O data to low-dimensional feature space. The cluster analysis shows the classification of prefectures with industry groups that characterize the reduced dimension. The mapping visualizes the relative positional relationship between prefectures by industry sector. In addition, the principal component analysis for a prefecture provides an overview of the annual trend of the industrial structure in the prefecture. Finally, through these analysis methods and results, we summarize the process of I/O analysis using a multivariate analysis framework.

## Keyword

Principal Component Analysis, Input-output table data, Visualization of Industrial Structure

## 1. はじめに

産業連関表<sup>1)</sup>は、国内経済において一定期間（通常1年間）に行われた財・サービスの産業間取引を1つの行列に示した統計表であり、5年ごとに作成・公開されている。一般的に、作成対象地域における経済構造の把握、経済波及効果の計算、各種経済指標の基礎資料として用いられているが、我々は、恣意的な選択なく、統一的に作成されるという特性から、従来のような地域内の分析だけでなく、地域間の比較や年次推移の可視化などにも用いることができると考えている。

本稿では、2005年(H17)と2011年(H23)の2時点における都道府県版産業連関表（取引基本表内の内生部門）データを用いて、産業取引の構造に関する都道府県比較や経年変化の抽出および可視化を試みる。我々の分析の関心は、個々の産業間の取引構造ではなく、産業全体の取引構造やそれに基づく都道府県の類似性の抽出であるため、(34×34で表される)産業部門間の複雑な取引を数値化した内生部門データに対して、主成分分析を用いて、多くの情報を反映させつつ解釈可能な扱い

やすい特徴空間へ変換を行い、各産業の特徴を保持した変換データを用いて産業構造の全体像を俯瞰的に見ることに徹した。

産業連関表を多変量データとして扱い、分析を試みた先行研究は、適用する解析手法とその目的および使用データに違いが見られる。解析手法においては、長沢(1988)<sup>2)</sup>や渡邊・下田・藤川(2009)<sup>3)</sup>、Feser&Bergman(2000)<sup>4)</sup>は解析手法に因子分析を採用しており、共通して、「産業クラスター」の特定を目的としている。特に渡邊他(2009)<sup>3)</sup>では、複雑な連関構造の背後に潜む相互に関連性の高い産業群を「産業クラスター」と定義し、共通因子の抽出を試みている。原田・寒河江(2019)<sup>5)</sup>は抽出された産業の特徴が正負で相殺されることがない非負値行列因子分解を適用し、特徴の解釈を容易にしながら、産業構造の比較を試みている。本研究と同じく産業連関表に主成分分析を適用した例は、Roepke,Adams&Wiseman(1974)<sup>6)</sup>、Czamanski(1974)<sup>7)</sup>、千葉(2019)<sup>8)</sup>などがあるが、いずれも特定の対象地域における産業連関表に適用し、地域内に限って議論されたものである。

分析対象とする産業連関表の構造上の違いとしては、我々が産業連関表の取引基本表データを対象としているのに対して、長沢(1988)<sup>2)</sup>は日本の製造業における投入産出構造の変化を分析するため、生産誘発依存度と生産誘発係数を対象データとしている。またFeser & Bergman(2000)<sup>4)</sup>や渡邊他(2009)<sup>3)</sup>は投入産出構造による産業クラスターの違いを議論することを目的に「投入係数」や「産出係数」を使用している。Vom & Bhatta(2007)<sup>9)</sup>は需要と供給の繋がりを分析するため、部門別の「投入係数」と「産出係数」における相関係数を利用した相関係数行列を使用している。入江(2018)<sup>10)</sup>は、生産額構成比に対してクラスター分析を行い、都道府県の類型化を試みている。このように産業連関表は、分析対象となるデータが複数存在し、どのデータを使用するかは解析者の分析目的に依存する。本研究では、産業連関表に主成分分析を適用することの有効性を検証できるように、どの地域においても必ず作成され、産業連関分析にも多用される取引基本表内データを使用した。加えて、生産額基準では見えない産業間取引構造を分析するため、各部門の列和を0、分散を1にする標準化処理(詳細は「4. データの概要」で後述している)を行い、それも合わせて使用データとして扱っている。

本稿では、全国版の産業連関表に主成分分析を適用し、その特徴空間上で各都道府県の位置づけを図るとともに、類似性や産業構造の抽出を行う。さらに、2期間分の産業連関表(H17, H23)を用いて、経年による産業構造の変化を視覚的に捉えるための分析例を示す。

## 2. 産業連関表

産業連関表は、総務省を中心に各府省庁において、西暦年の末尾が0又は5の年次(昭和30年(1955年)表以来)を対象として5年ごとに作成される。調査対象地域によって、全国版・地域版・都道府県版などがあり、基本的な構成要素は、取引基本表、投入係数表、逆行列係数表である。対

象となる産業部門は、分類の粒度によって小分類(190)・中分類(108)・大分類(37,平成17年は34分類)に分けられる(なお、平成23年は、13分類も存在する)。取引基本表は、各産業間で取引された財・サービスを金額で表示したもの、投入係数表は、取引基本表の内生部門において原材料等の投入額を当該産業の生産額で除した係数、逆行列係数表は、ある産業に対して、1単位の最終需要があった場合、各産業の生産が究極的にどれだけ必要となるかを示す係数となっている。

本稿で用いる産業連関表は、都道府県単位で作成された平成17年版(2005年)と23年版(「経済センサス-活動調査」の調査対象年次に合わせて例外的に2011年で作成)の取引基本表内の内生部門データ(大分類)を分析対象としている。

## 3. 分析方法

産業連関表(正確には、取引基本表の内生部門データであるが、投入係数表や逆行列係数表でも同様の分析が可能であることから、単に産業連関表と表記する)を、分類部門数×分類部門数の多変量データとみなした場合、もっとも分類数の少ない大分類であっても、34あるいは37次元の部門数の次元を有するデータとなる。

我々の分析目的が、個々の産業間取引ではなく、全体の取引構造に着目した類似性比較や可視化であることから、主成分分析を用いて、データを直感的理解が容易な低次元の特徴空間へ縮約する。次に、縮約されたデータにクラスター分析を適用することで、地域間の類似性を樹形図によって視覚的に確認することを目指す。

特徴空間上のデータの作成や分類、可視化のために、本稿で採用した主成分分析とクラスター分析の2つの分析手法について以下に述べる。

### 3.1 主成分分析

主成分分析は、多変量データの次元縮約手法として知られており、データの次元数が多い場合に、少ない情報の損失で少数の合成変数にデータを

変換するための分析手法である。多変量データ $X$  (供給側産業部門数 $n \times$  需要側産業部門数 $n$ ) に対する合成変数 $Z$  (供給側産業部門数 $n \times$  次元縮約後の次元数 $k$ ) は(1)式のように与えられる。

$$Z=XA \cdots (1)$$

ただし、行列 $A$  (需要側産業部門数 $n \times$  次元縮約後の次元数 $k$ ) はデータ $X$ の分散共分散行列の固有ベクトルに相当する。元のデータの情報を多く含む(固有値の大きさの)順に、第1主成分 $Z_1$ 、第2主成分 $Z_2$ 、第3主成分 $Z_3$ 、 $\cdots$ となる。本稿では、第1主成分と第2主成分を用いて2次元データに変換し、比較や可視化の分析を行う。

### 3.2 クラスタ分析

クラスタ分析は、データを類似度などの基準に従っていくつかの集団(クラスタ)に分類する手法である。クラスタ分析には、階層的クラスタリングと非階層的クラスタリングの2種類があるが、非階層的クラスタリングでは、あらかじめクラスタ数が必要となるため、本稿では、階層的クラスタリングの中でも、個々のデータを1つのクラスタとしてスタートし、クラスタ間の類似度に基づいて、2つのクラスタを逐次的に結合しながらデータを分類する凝集型階層的クラスタリングを使用し、樹形図を用いて階層構造を表現する。類似度の指標としては、ユークリッド距離、マンハッタン距離、マハラノビス距離、コサイン類似度など数多くあり、クラスタ間の結合方法としては、ウォード法、群平均法、最短距離法、最長距離法などがある。

なお、解析は、オープンソースの統計解析システムRを用いて実装した。

## 4. データの概要

分析対象として、以下のデータを収集した。

- ・対象地域：都道府県  
各都道府県ホームページからダウンロードした
- ・対象年：平成17年版と平成23年版  
分析時の最新データは23年版であった

・対象表：取引基本表内の内生部門内データ  
中間投入 $\times$ 中間需要の計を除く部分

・対象分類：大分類

平成17年版は34分類、平成23年版は37分類であったため、23年版を34分類になるよう次のように再構成した。

・再構成の方法：総務省「平成23年産業連関表作成基本要綱」の平成17年表との相違点[別表4]部門分類対応表(4)統合大分類を参考に、23年表にある「はん用機械」「生産用機械」は「一般機械」, 「水道」「廃棄物処理」は「水道・廃棄物処理」, 「業務用機械」は「一般機械」と「精密機械」, 「プラスチック・ゴム」は「その他の製造工業製品」とし、17年表の34分類に合わせた。

統一的に作成されている産業連関表であるが、都道府県によって、産業部門の設定に相違がみられたため、できる限り全国表の産業部門に合わせる形で人手によりデータを加工して用いた。具体的には、全国表では1つの産業部門であるが都道府県版では複数の部門に分割されているものは値を合算し、逆に、全国表では複数の産業部門となっているものが都道府県版では1つの部門として表されているものは、部門数に応じて値を等配分した。また、都道府県独自の産業部門(例えば、愛知県では「自動車」「航空機」「その他輸送機械」がある)を有する場合は、全国表にある最も近いと思われる部門(前述の例では、「輸送機械」)へ集約した。内生部門データは、34に分類された産業(行方向:供給部門)がどの産業(列方向:需要部門)にいくら販売したか、逆に言えば、34の産業(列方向:需要部門)は、どの産業(行方向:供給部門)からいくら購入したかを表している。全国表についても同様に収集し、全国版および47都道府県版のそれぞれについて、 $34 \times 34$ の行列データが得られた。図1は収集データのイメージである。

本稿では、これらの収集データを多変量データとみなし、分析を実施する際には、供給部門の各産業データが需要部門の産業の金額で与えられると解釈することとする。需要部門と供給部門を入

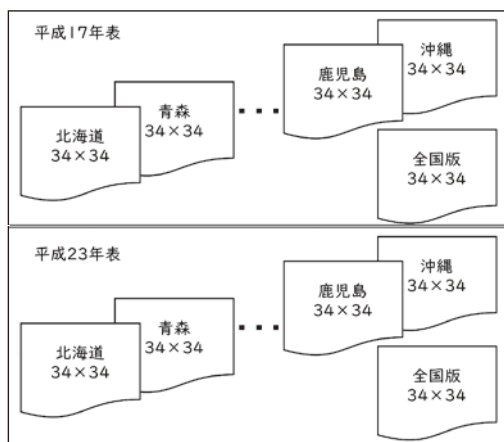


図1. 収集データのイメージ図

れ替えた（内生部門データ行列を転置した）場合も同様に分析が可能である。

取引基本表は、都道府県によって金額の単位が大きく異なることに加えて、産業部門によっても取引金額（生産額）に大きな差が見られるため、分析の際に取引基本表の値をそのまま使用することは、金額に応じた影響が反映されると考えてよい。一方で、産業間の取引構造が類似しているにもかかわらず、取引金額が小規模なためにそれらの情報が表面化しないという状況も考えられる。そこで、我々は、取引金額だけでなく、産業の取引構造にも着目した類似性の抽出を図るため、金額の大小（単位）の影響を排除するようデータを標準化して用いた分析も行う。分析の目的において、金額に着目する場合は、収集データを利用し、産業間の取引（需要/供給）構造に関心がある場合は、収集データを標準化（行列総和で除する/行和で除する/平均0分散1の正規化など）したものを利用すればよい。なお、本稿で用いる標準化データは、列ごとに平均0、分散1の処理を行ったデータを表す。

## 5. 分析結果

### 5.1 主成分分析による次元縮約

3. で述べたように、47都道府県ごとに存在する34×34の行列データに対して、主成分分析を

用いて34(供給部門)×2の2次元データへ縮約する。その際、変換基準が異なると、同一平面上での議論ができなくなるため、全国表の主成分分析で得られた第1主成分と第2主成分の各軸上に統一的に各都道府県データを変換した。

図2、図3<sup>注1</sup>は各都道府県データを変換するための全国表の主成分分析に関するプロットであり、平成17年、平成23年データを標準化の有無によって区別している。第1主成分と第2主成分を

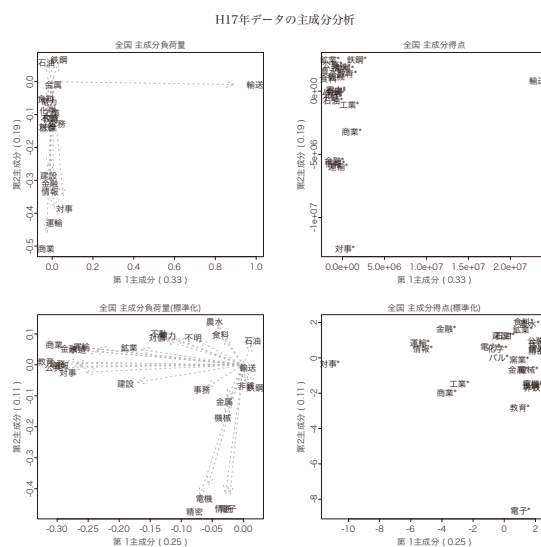


図2. H17年全国表の主成分分析結果のプロット

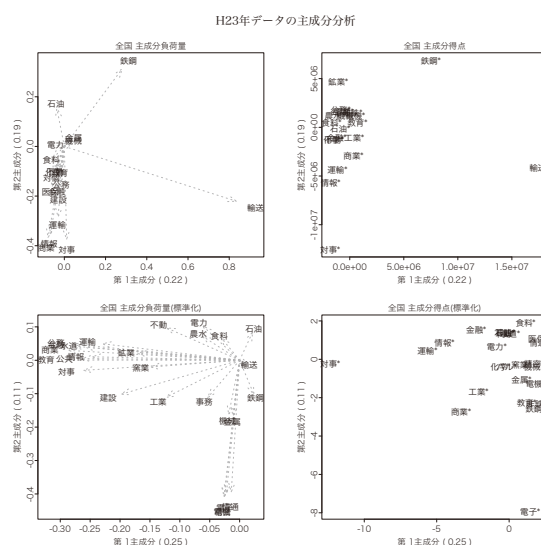


図3. H23年全国表の主成分分析結果のプロット

軸として、左上図が寄与する需要側産業部門（列方向：左側）の主成分負荷量（固有ベクトル）、右上図が供給側産業部門（行方向：右側）の主成分得点（合成変数Z）の散布図を示したものになっており、標準化したデータを用いた場合が、それぞれ左下図と右下図である。各軸の（）内の数値は寄与率である。例えば、図2に示した平成17年のプロットから、寄与率に関しては、第1主成分と第2主成分の合計が0.52であることから、縮約した第1、第2主成分得点によって、元の34次元データの約52%の情報を保持しているとみなすことができる。また、輸送機械が主成分負荷量（固有ベクトル）および主成分得点ともに第1主成分に大きな値として特出しており、第2主成分では、固有ベクトルとして商業、主成分得点として対事業所サービスが特出している。標準化によって、取引金額（生産額）の情報ではなく、取引構造に着目すると、第1、第2主成分については、いくつかの産業部門がまとまった方向にプロットされ、主成分得点としては対事業所サービスや電子部品が特徴的に現れていることがわかる。

## 5.2 都道府県クラスタリング

47都道府県ごとの34×34の産業連関表データを、5.1で示した全国表主成分分析で得られた第1、第2主成分軸へ統一的に変換し、各都道府県のデータを34×2の2次元データへ縮約する。これらを元に、クラスタ分析を用いて、都道府県間の類似度に基づくクラスタを形成し、樹形図によって可視化する<sup>注2</sup>。

前述のように、データの性質上、分析の視点が取引金額（生産額）の場合は収集データを、産業間の取引構造の場合は、標準化したデータ（本稿では、列ごとに平均0、分散1に正規化したもの）を用いる。例えば、 $n \times m$  行列で表されるA県のデータ $X^A$ とB県のデータ $X^B$ の類似度 $d(X^A, X^B)$ の指標は、(2)式で与えられるユークリッド距離で統一した。

$$d(X^A, X^B) = \sqrt{\sum_{i=1}^n \sum_{j=1}^m (X^A_{ij} - X^B_{ij})^2} \dots (2)$$

先に示した第1、第2主成分へ変換した場合は、 $n=2$ 、 $m=34$ となる。また、主成分ごとに類似度を測る場合は、 $n=1$ とすればよい。第1主成分と第2主成分を合わせた類似度((2)式で $n=2$ とした場合)を用いて、クラスタリングを行うことも可能であるが、本稿では全国表の各主成分軸に解釈可能な特徴(5.1で詳述)が見られたため、その特徴を反映するようなクラスタリングとしての位置付けとした。

図4、図5は平成17年版データ、図6、図7は平成23年版データで、さらに標準化の有無と第1主成分得点および第2主成分得点に分けて図示したものである。なお、クラスタの結合方法は、ウォード法を用いた。

図4、図6の上図（第1主成分得点による樹形図）をみると、愛知県が1つのクラスタを形成していることがわかる。これは、全国表の第1主成分（図2や図3左上図）に影響ある産業部門として「輸送機械」がみられることから、大手自動車産業を有する愛知県が特出していると考えられる。このことは、元の産業連関表の数値においても確認することができる。大手自動車産業を有する愛知県では、輸送機械部門の供給側総額が6.1兆円（供給側全国同比26.7%）、需要側総額が9.5兆円（需要側全国同比28.2%）であり、これは愛知県に次いで輸送機械が突出する神奈川県（供給側総額1.9兆円（8.5%）、需要側総額2.9兆円（8.5%））や静岡県（供給側総額1.8兆円（7.6%）、需要側総額2.7兆円（8.0%））などと比較しても極めて突出していることがわかる。また、下図の第2主成分得点では、東京都がクラスタとして出現しているが、こちらも図2や図3で確認すると、第2主成分は、「商業」「運輸」「対事業所サービス」などが特出しており、それらの特徴は東京都の産業として関連づけることができる。図5や図7の標準化データを用いた場合は、生産額の影響が現れないことから、各主成分を特徴づける産業群に対して、いくつかのクラスタの形成が確認できた。これらは、他の都道府県とは異なる構造を有する都道府県群がクラスタとして出現しているよう

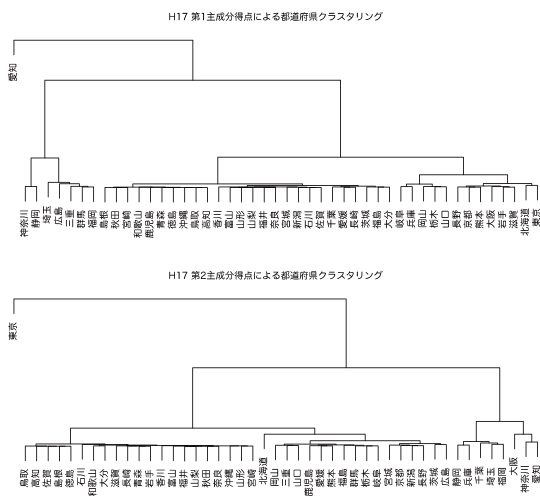


図4. H17年データの都道府県クラスタリング  
(上図：第1主成分得点，下図：第2主成分得点)

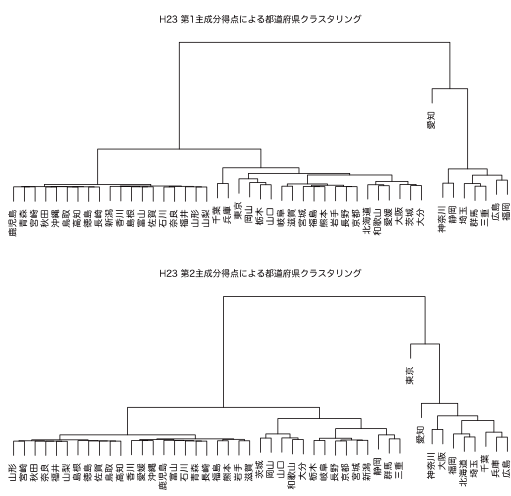


図6. H23年データの都道府県クラスタリング  
(上図：第1主成分得点，下図：第2主成分得点)

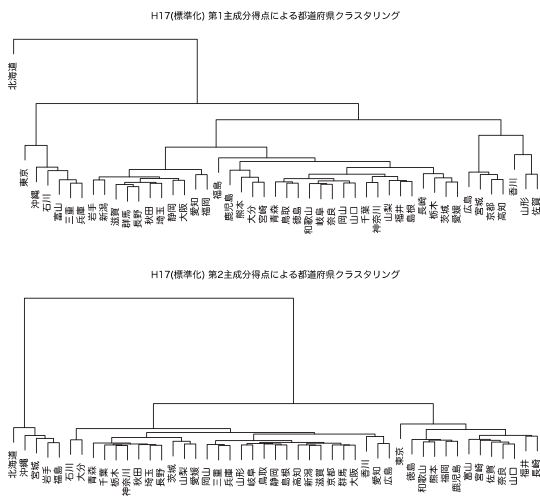


図5. H17年標準化データの都道府県クラスタリング  
(上図：第1主成分得点，下図：第2主成分得点)

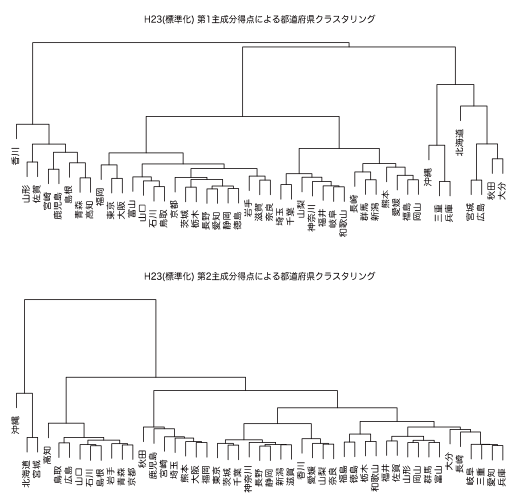


図7. H23年標準化データの都道府県クラスタリング  
(上図：第1主成分得点，下図：第2主成分得点)

にみえる。階層の低いクラスターに関しては、クラスターの結合方法などにも依存して変化するため、本稿では言及しない。

ここでは、(2)式のように全34産業部門を合算した類似度を用いてクラスタリングを行ったが、産業部門ごとの類似度を用いてクラスタリングすることも可能である。

### 5.3 主成分分析を用いた都道府県マッピング

低次元特徴空間（2次元の場合は平面）へ縮約した産業連関表データを用いて、平面上に各都道府県のデータをマッピングし、相対的な関係の可視化を試みる。これらのマッピングでは、都道府県ごとに、本来は34次元で表される供給部門のデータを同じ基準を用いて同一空間上に縮約する



ことで、供給部門における都道府県の関係性を同一マップ上に表現することができる。

2次元の同一特徴空間上で比較を行うために、5.1. に示した全国表における主成分軸上(図2, 図3の左)へ各都道府県のデータを変換する。例と

して、石油・石炭製品(石油炭)と電気機械(電機)、電力・ガス・熱供給(電力ガス)、情報通信部門ごとに、対象年と標準化の有無によってプロットしたものが、図8から図11である。横軸と縦軸は、それぞれ全国表に合わせた第1主成分得点, 第2

H17 都道府県マッピング

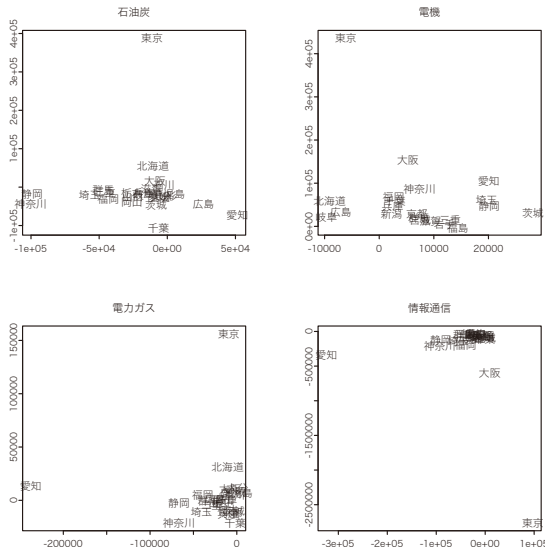


図8. 都道府県マッピング (H17年)

H23 都道府県マッピング

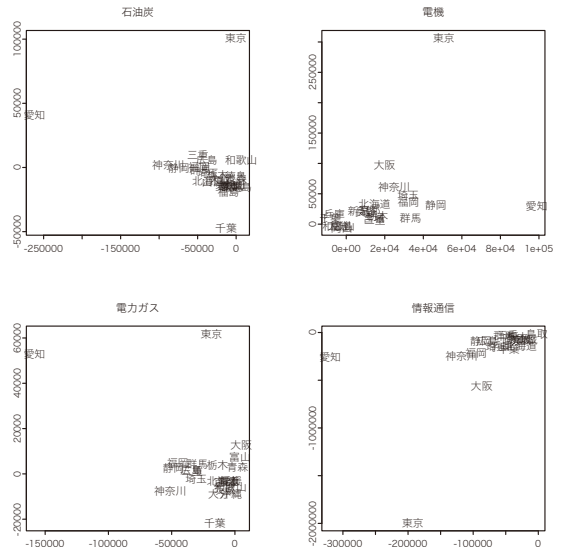


図10. 都道府県マッピング (H23年)

H17(標準化) 都道府県マッピング

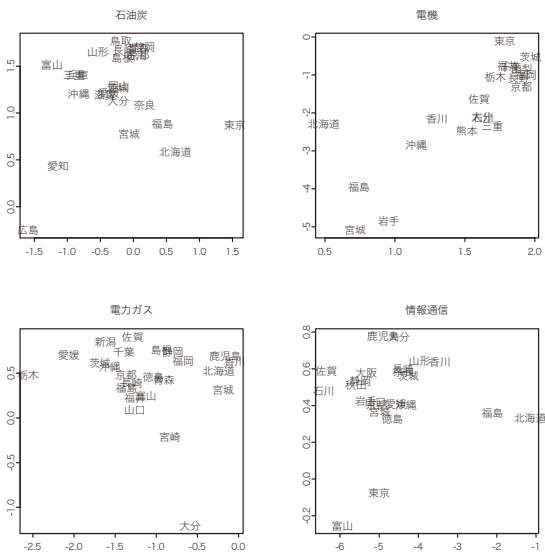


図9. 都道府県マッピング (H17年標準化)

H23(標準化) 都道府県マッピング

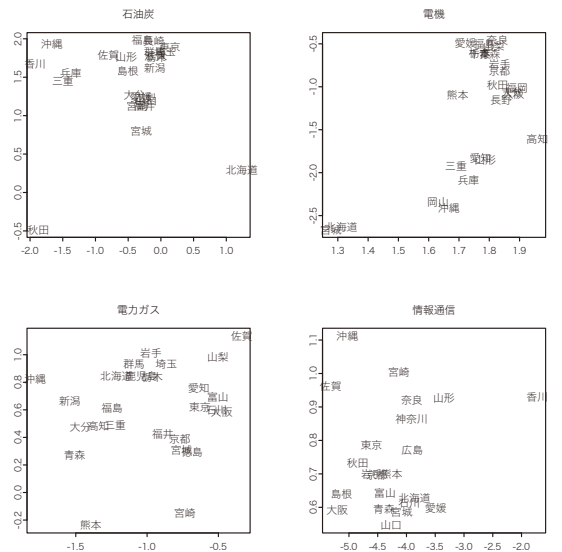


図11. 都道府県マッピング (H23年標準化)

主成分得点を表す。標準化しないデータを用いたプロット(図8, 図10)では, 各供給産業部門における取引金額の大きい都道府県が特出する傾向にある。また, 標準化データを用いたプロット(図9, 図11)では, 供給産業部門における取引構造の性質が似ている都道府県は近くに, そうでない場合は他から離れて配置されているように見える。

産業全体における都道府県の関係を示すには, 変換後の各都道府県 $34 \times 2$ の主成分得点データの総和を取ることで $47 \times 2$ データを生成し, 図示する方法も考えられるが, 都道府県間関係は, 産業部門ごとに特徴が異なることから, あえてまとめない方がよいだろう。

本稿では, 変換基準として全国表の主成分負荷量(各産業の固有ベクトル)を用いたが, 全国表に合わせることで, 全国表と似た産業構造の都道府県は過大に, そうでない都道府県は過少に特徴を抽出している可能性もあることから, 分析の目的に応じて基準を設定するとよい。

#### 5.4 都道府県ごとの産業構造の可視化

都道府県ごとの主成分分析を用いて, 当該都道府県における産業間取引の全体構造の把握と可視

化を行う。

図12から図17は, 例として福島県, 東京都, 石川県のH17, H23データに関する主成分分析の結果として, 第1主成分・第2主成分として寄与する需要側産業部門(列方向)の固有ベクトル(左図)と供給側産業部門(行方向)の主成分得点(右図)の散布図を示したものである。これまでと同様に, 横軸が第1主成分, 縦軸は第2主成分に対応している。

図中の主成分得点で表される産業部門(供給側)は相対的な位置関係を示しているため, 上下左右の反転(正負)の違いには意味がないことに注意が必要である。図16.石川県H17データの主成分分析の散布図

図12から図17において, 左上図の主成分プロットから, 各都道府県で大きな生産額(取引金額)を有する特徴的な産業が特出していることがわかる。また, 標準化したデータでは, 産業全体の取引構造を俯瞰的にみることができ, 右図の主成分得点プロットでは, 各都道府県において, 似たような取引構造を持つ産業がクラスターとして抽出されている。

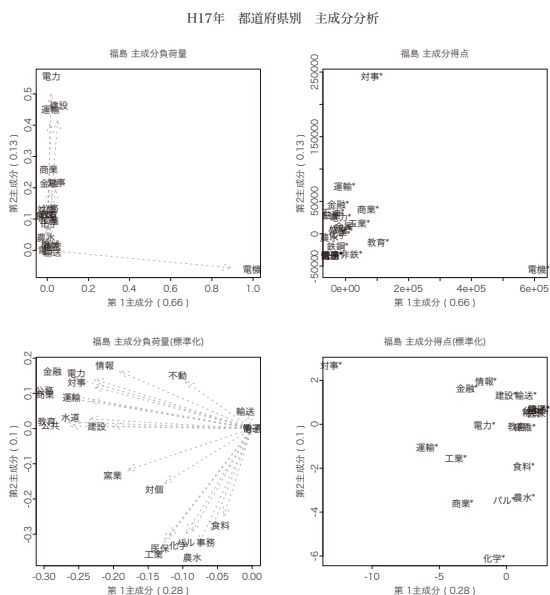


図12. 福島県 H17データの主成分分析の散布図

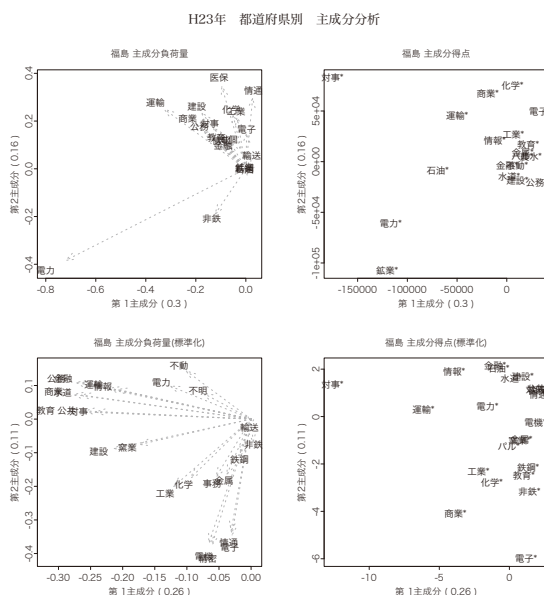


図13. 福島県 H23データの主成分分析の散布図

H17年 都道府県別 主成分分析

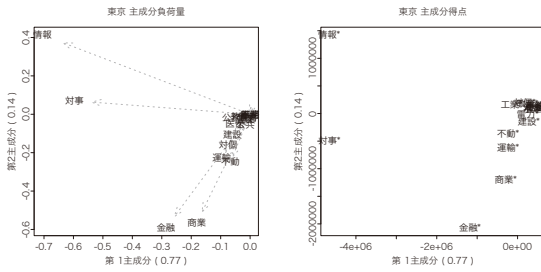


図14. 東京都 H17データの主成分分析の散布図

H17年 都道府県別 主成分分析

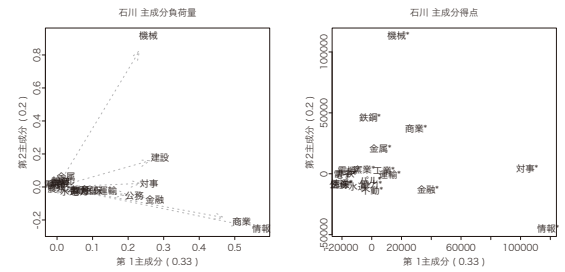


図16. 石川県 H17データの主成分分析の散布図

H23年 都道府県別 主成分分析

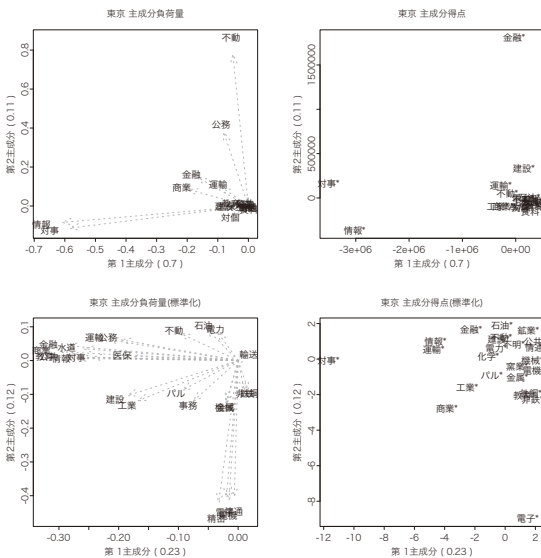


図15. 東京都 H23データの主成分分析の散布図

H23年 都道府県別 主成分分析

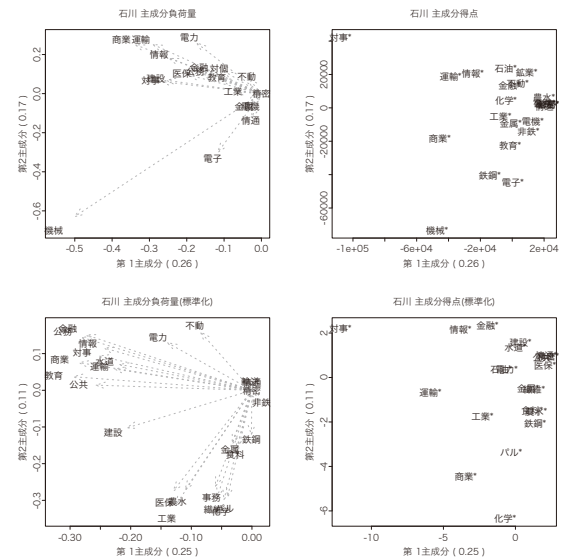


図17. 石川県 H23データの主成分分析の散布図

5.5 産業構造の推移の可視化

ここまで、平成17年と平成23年の2時点のデータをそれぞれで分析してきたが、同一地域における産業構造の経時的な変化を捉えるために、対象年ごとに得られる主成分負荷量（例えば図12，図

13の左上図）を比較する。これらの図は、ベクトルの方向や長さが、特徴空間に縮約された産業間の需給構造に対する需要側産業部門の影響の度合いを表現していると考えられることから、同一地域の2時点におけるこれらの構図を視覚的に比較

することで、産業構造の推移としてみなすことができると考えている。見やすさのために、取引金額における、第1主成分、第2主成分について絶対値の大きい順に6産業部門についてのみプロットしたものを図18に示す。各軸の( )内の数値は寄与率である。ベクトルの長さに着目して、福島県の場合は、平成17年に比べて平成23年では、「電力ガス」の影響、つまり、産業部門間の需給構造(取引構造)に対する需要側「電力ガス」部門の影響が大きいことがわかる。東京都の場合は、平成17年では「情報通信」と「金融」、「サービス」系に特徴的な傾向がみられたが、平成23年では、「サービス」系の増加に加えて、さらに「不動産」も一つの大きな特徴として現れていることがわかる。また、石川県においては、基盤産業の「機械」に

加えて、「電子部品」や「電力ガス」も長さが増している。

## 6. 考察

産業連関表における取引基本表内の内生部門データに対して、データ解析手法を適用した都道府県間の関係や産業の移り変わりについて可視化を軸に考察する。

5.2のクラスター分析では、都道府県をデータ間の距離に基づいて分類し、樹形図で表現した。産業を軸としたこの分類では、34の産業部門間の取引金額(収集データ)と取引構造(標準化データ)をもとに都道府県間の関係を可視化した。収集データを用いて分析する場合、生産額の大きな東京や愛知、大都市を有する都道府県がまとまったクラスターとして存在する形となった。標準化したデータで分析する取引構造の場合は、産業構造の特異な都道府県など、何らかの共通解釈ができるクラスターもあればそうでないクラスターもあった。今回、クラスター間の結合方式として、ワード法を用いたが、連結方法によってクラスターの違いが若干見られることから、分析の際にいくつか試行して、総合的に判断するとよいだろう。

5.3の主成分分析を用いた都道府県マッピングでは、取引金額でみた場合、産業部門ごとに額の大きな都道府県が特出するため、全体がわかりやすい。また、取引構造でみた場合は、各産業部門で同じ傾向のある都道府県は近くに、そうでない場合は単独で配置される結果となった。

5.4で示したように各都道府県のデータを主成分分析することによって、産業構造を概観することができ、様々な可視化や都道府県間での比較も可能となる。また、5.5のように年次で比較すれば、経年変化などに関する直感的な把握も容易となり、そのポイントから、産業構造の変化に関する詳細な要因について探っていくことも可能と考える。

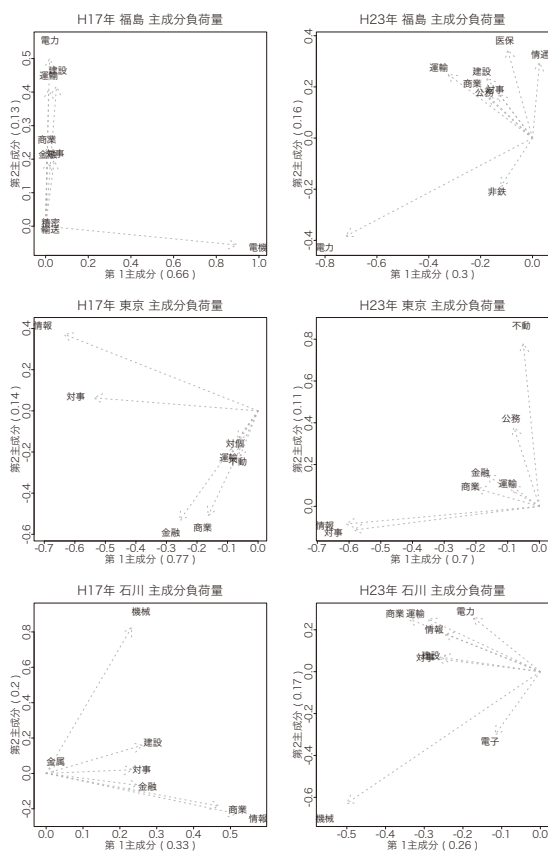


図18. 産業構造の推移の可視化  
(上から福島、東京、石川)

## 7. おわりに

本稿では産業連関表に主成分分析を適用し、特徴空間上で産業取引構造に関する都道府県間比較や経年変化の分析を行ってきた。

産業取引構造に関する都道府県間比較を行った分析について、本研究では47都道府県の産業構造の特徴を、全国版産業連関表を用いて同一空間上で比較することにより、ある主成分負荷量に対応した産業（製造業やサービス業など）を有する地域同士をクラスターとして抽出することができた。産業連関表を用いた既存の都道府県間比較において、産業連関表内のある項目、例えば製造業部門などの産業部門や粗付加価値、輸入・輸出などの項目のみ、あるいはそれらの組み合わせによる経済分析例は散見されるが、本研究の手法では、すべての産業部門の情報を加味したうえで、全都道府県の産業の特徴を統一した基準で客観的に比較できるという利点がある。分析目的に応じて、上述した粗付加価値項目や輸入・輸出等の影響も考慮した恣意性のない客観的な分析も可能である。

産業構造の経年変化を分析については、下田・藤川・渡邊(2005)<sup>11)</sup>の「比例成長からの乖離分析(DPG分析)」がある。これは最終需要等の変化に伴う各産業の生産量の変化を定量化した手法であり、使用データとして、最終需要に当たる消費・投資・輸出や、産業構造を見る投入係数などを用いていることから、本研究とは分析対象及び評価尺度が大きく異なるため一概に比較はできない。本研究における主成分分析を用いた例は、都道府県別産業部門ごとの主成分負荷量から、産業取引構造の経年変化をベクトルの変化量（長さ）で見ることにより特徴成分の大小関係が視覚的に理解しやすい点が挙げられる。

データ解析者の立場から、産業連関表の新たな活用法として、主成分分析やクラスター分析を適用し、それらの可視化に関して分析の有効性や可能性を検証した。分析結果における詳細な解釈は、経済学の立場から行うべきであり、現段階では、まだ十分できていない。しかし、本研究では、

産業連関表に関する個々のデータに興味があるのではなく、全体の経済構造や都道府県間の関係性に関心があるため、厳密な数値よりは、可視化というツールで全体の把握に徹してきた。特に、主成分分析を取り掛かりとして、都道府県ごとに34×34の行列で表される複雑な連関構造に対して、全国表における主成分軸へ変換することで、統一した空間上で類似性の議論を進めてきた。縮約した第1・第2主成分を用いた分析では、各主成分へ大きく寄与する産業群という新たな切り口で都道府県の主要なクラスターを確認することができた。取引金額の影響を除いた産業間の取引（需要/供給）構造に着目するために、標準化データを用いて分析したところ、各主成分でいくつかのクラスターの形成が確認できた。これらのクラスターに対する詳細な特徴や解釈には、専門的な知識が必要となるだろう。

### 【注】

注1 主成分分析を適用した作図（図2,3及び図8～18に該当する）では、各主成分の要素数が多いため、各主成分のメディアン値に対する差の絶対値が大きい順に選定し、大きな特徴を有する部門・地域を表すものとして作図している。特に図2,3及び図8～17では15部門及び都道府県、図18では6部門に関して選定している。

注2 次元縮約後のデータに対し、クラスター分析を行う2段階の分析方法に関して、Arabie & Hubert (1994)<sup>12)</sup>、岡太 (2015)<sup>13)</sup>、山本 (2015)<sup>14)</sup>、林 (2018)<sup>15)</sup>では、元のデータに対して適切なクラスターが得られない「Tandem clusteringの問題」が生じる可能性があるとして指摘している。本研究では、次元縮約後の各主成分得点のデータに階層的クラスタリングを適用し、都道府県間の比較を行っている点で、Tandem clusteringには該当しない。

### 【参考文献】

- 1) 産業連関表 総務省  
<https://www.soumu.go.jp/> 2021年3月25日閲覧
- 2) 渡邊隆俊・下田充・藤川清史. 「投入構造と産出構造からみた産業クラスターの地域別特性-2000

- 年の関東・中部・近畿を例にとって-」. 経営経済, 44, pp.39~64. 2009
- 3) 長沢克重「因子分析による投入産出構造変動の分析—昭和45-50-55年接続産業連関表による—」. 経済統計学会, 55, pp.52~63. 1988.
  - 4) Feser, E. J. and E. M. Bergman (2000). National Industry Templates : A Framework for Applied Regional Cluster Analysis. *Regional Studies*, 34.1, 1-19.
  - 5) 原田魁成・寒河江雅彦. 非負値行列因子分解法の産業連関構造分析への応用～地域と産業特性の可視化～. 2019年度統計関連学会連合大会講演報告集, pp.320. 2019
  - 6) Roepke H. D., Adams, D. and Wiseman, R. "A New Approach to the Identification of Industrial Complexes Using Input-Output Data". *Journal of Regional Science*, 14.1, 1529. 1974
  - 7) Czamanski, S. "Study of Clustering of Industries". Halifax, Nova Scotia, Canada: Institute of Public Affairs, Dalhousie University. 1974
  - 8) 千葉雄二 (2019). 市町村間産業連関表の作成と町村の存続. 2019年度日本地域学会年次大会.
  - 9) Vom Hofe, R. and Chen, K. "Method for identifying local and domestic industrial clusters using interregional commodity trade data". *The Industrial Geographer*, 4, pp.2-46. 2007
  - 10) 入江啓彰 (2018). 2011年産業連関表から見た都道府県の産業構造. 産研論集, 45, 35-43.
  - 11) 下田充・藤川清史・渡邊隆俊 (2005) 規模別産業連関表から見た日本の産業構造, 産業連関 vol.13, No.3 52-65.
  - 12) Arabie, P., and Hubert, L. "Cluster analysis in marketing research. In Bagozzi, R.P. (Ed.) *Advanced method of marketing research.*" Oxford: Blackwell. pp.160-189. 1994
  - 13) 岡太彬訓 (2015). タンデム分析の背景と問題. 日本行動計量学会第43回大会抄録集, 134-135.
  - 14) 山本倫生 (2015). 次元縮小とクラスタリングの同時分析法とその問題点. 日本行動計量学会第43回大会抄録集, 136-139.
  - 15) 林拓也 (2018). Reduced K-means法によるクラスター分析の社会調査データへの応用. データ分析の理論と応用, Vol.7, No.1, 9-19.