Title	地域区分のための主成分分析
Sub Title	Principal component analysis for regionalization problem
Author	高橋, 潤二郎
Publisher	慶應義塾経済学会
Publication year	1967
Jtitle	三田学会雑誌 (Keio journal of economics). Vol.60, No.2 (1967. 2) ,p.221(93)- 234(106)
JaLC DOI	10.14991/001.19670201-0093
Abstract	
Notes	研究ノート
Genre	Journal Article
URL	https://koara.lib.keio.ac.jp/xoonips/modules/xoonips/detail.php?koara_id=AN00234610-19670201-0093

しかし、筆をすすめるうちに、そのためには組織のみならず、それをめぐる問題にもひろくふれねばならなくなり、必ずしも組織問題のみに十分な紙面をさくことができなくなってしまった。従って、組織形態に限っても十分整理をしつくしたとはいえない。とはいえ、本稿においても、造機船工労組合中心に、その生成、発展、消滅等について新たに解明された点も少くはない。同時に、これから掘りさげるべき点も決して少くはないであろう。そして、本稿でとりあげた問題をさらに整理し一般化するのもこれからの作業であると考えている。

ただ、心のこりなことは、四章に「戦前日本の労働組合——その機能と性格——」という一節を用意していたが、頁数の都合で掲載できなかったことである。戦前の組合については本論中にもかなりふれているし、この節は独立的な部分でもあるので、今回は省き、いずれ機会をみて発表することにした。

単に一企業の歴史と思っても、いざとりかかってみると、資料的に利用しうるものが少く、正確に事実を追うだけでも難事業であることに気づく。それ故に、このような小稿でも資料の点で各々専門の方々の御教示を仰がねばならなかった。斎藤忠利氏をはじめ、松尾洋氏、森長英三郎氏、それに法政大学大原社会問題研究所所蔵の資料の利用にあたって便をはかって下さった二村一夫氏等にはその点謝意を表したい。

「序」でも記したとおり、本稿は一連の作業の最初の仕事であり、これからさらに研究を深めてゆく過程で、本稿も当然補正し、よりよいものに仕上げてゆくつもりである。大方の御教示をお願いしたい次第である。

研究ノート

地域区分のための主成分分析

高橋潤二郎

本稿の主要な目的は、最近、地域区分をするための効果的な統計的技術として注目を集めている、いわゆる主成分分析について、その基本的アイディアと計算過程について報告することにある。

主成分分析の手法を地域区分に適用した事例は、筆者の知る限りでは、M. J. Hagood (1943) が最初であろう。Miller Kahn や P. Greig-Smith によれば、主成分分析ないし因子分析の地域区分への適用は、植物生態学、地質学等においては既に一九五〇年代にあらわれているが、地理学の分野で、これを最も明瞭な形で紹介したのは、いうまでもなく、Berry (1961) であった。PRZEGLAD GEOGRAFICZNY に発表されたその論文で、彼は、地域区分の問題 regionalization problem を「より一般的な分類問題の特殊ケース」であると規定し、この意味での地域区分をする過程で direct factor analysis（ここでいう主成分分析）が極めて有効であることを強調し、その適用のデモンストレーションを行っているが、地域区分という作業において、主成分分析の果す役割をこれ程明瞭に説明したものは他にないといってよかろう。

ところで、地域区分という概念であるが、本稿では、これを次のように規定しておくことにする。即ち、

調査対象となる「全域」がN箇の「基域」に区分され、各基域毎にn箇の変量に関する観察値が得られたものとする。これら$N \times n$箇の観察値X_{ij}を基礎的データーとして、N箇の基域をそれより少数のK箇の同質的な「連接域」に組みわける作業を一般に地域区分という。

ただしここで「全域」、「基域」は、夫々 Duncan, Cuzzort, Duncan (1961) の "universe of territory", "basic set of areal units" を意味するものとする。いうまでもなく、「可能なかぎり同質的な」という表現からも明らかなように、これはいわゆる同質地域に関する規定であり、又、「連接域」という表現からも明らかなように各地域に含まれる基域は空間的に連接 contagious していなければならないという意味で、地理学において伝統的地域概念を継承するものである。

ところで、この意味での地域区分をするに当って基本となるのが

同質性という概念であることはいうまでもない。ここで注意すべきことは、一般的な分類問題における同質性（より適切には、異質性）がしばしば「距離」との関係において規定されることであろう。即ち、われわれの基礎的データーは、夫々n箇の要素をもつN箇のベクターより成ると考えることができるが、このことは、各基域が各変量を軸とするn次元の空間の一点としてあらわせることを意味している。この場合、基域の属する空間が距離空間である限り、個々の基域間の同質性（ないし異質性）をその空間に定義された距離との関係において規定することが可能であろう。例えば、いま、分類の場である空間を集合R^nと距離函数、

$$d^n(x, y)=\left[\sum_{i=1}^{n}(x_i-y_i)^2\right]^{\frac{1}{2}}$$

とで規定されるものとする。即ちn次元のユークリッド空間とするならば、われわれは、ことなった基域xとy間の同質性を$d^n(x, y)$との関係において、次のように定義することができる。

任意の基域xとyについて、$d^n(x, y)=0$のとき、そのときにのみxとyは同質的である。

任意の基域x、y、zについて、もし$d^n(x, y)<d^n(y, z)$ならば、xとyは、yとzがそうであるよりもより同質的である。

この結果、地域区分の問題は、基本的にはN×n箇の観察値があたえられたとき、これを基礎にして、㈠いかに距離空間としてのn次元空間を構成し、そこにN箇の点を布置するか、㈡いかにしてこれらN箇の点を距離を規準として、K箇のグループに組みわけるか、という純粋に技術的な問題に還元されることになる。

主成分分析の地域区分への適用、少くとも従来行われてきたそれは、これら二つの問題のうち、前者に関するものだといってよかろう。即ち、それは基域の特性を示すn箇の変量を、それより少数のm箇の因子によって説明することによって、m次元の空間を構成し、これに各基域を点として位置せしめるに役立つのである。ここで重要なのが、n箇の変量をそれより少数のm箇の因子によっておきかえるという操作であることはいうまでもない。n箇の変量をもって、n次元の距離空間を構成することは、たとえ、n箇の変量に関する尺度が相違していてもさしたる困難はない。個々の観察値X_{ij}をいわゆる標準化されたZ_{ij}に変換することによって、或る程度、説明可能なn次元空間を構成することができる。しかし、このような単純な空間構成は必ずしも好ましいものではない。もしnが非常に大であればこの空間を分類の場として利用することは、点のグルーピングということで面倒を倍加するであろう。又、個々の基域が或る程度のまとまりをもつ entity として考えられるとき、n箇の変量がすべてことなった動きをするとは限らない。むしろそのうちのいくつかはある程度似通った動きをすると考えるのが妥当であろう。換言すれば、n箇の変量のうちいくつかは、基域の同質性を記述する上でリダンダントなものと考えられる場合が往々にしてある。このような場合、n箇の変量をそれよりはるかに少数のm箇の因子で表現し得るということは、地域区分の作業を著しく容易ならしめるであろう。主成分分析の手法が地域区分に役立つというのはまさにこの点に他ならない。

このような意味で、分析を解釈する限り、主成分分析に限らず、因子分析一般を適用することが可能であろう。主成分分析がいわゆる因子分析の一つとして理解されるべきか否かについて議論が多いが、ここでは Lawley-Maxwell (1963) にしたがって、一応両者を全くことなった目的をもつものとして扱うことにした。因子分析におけるターミノロジイはかなり混乱しているが、一応、ここでは、形式的には基本方程式として $Z_{ji}=\sum_{p=1}^{m}a_{jp}F_{pi}+a_jV_{ji}$ のかたちをとり、変量の変動を共通因子と独自因子とで説明するものを因子分析といい、後述するように、独自因子を含まぬものを主成分分析として区別している。主成分分析は principal component analysis の訳であり、この他にも主因子分析、主成因分析と呼ばれることがあるが、すべて同じ手法を指していると理解してさしつかえない。その開発系譜は既によく知られているように、Karl Pearson (1901) に始まり、H. H. Hotelling (1933) に至って一応の形をととのえたものである。

基本的アイディア

さて、主成分分析は、前述のように、地域区分のための分類の場をあたえるn次元空間を基本因子の抽出という操作によってはるかに少数のm次元の空間におきかえることがその主要な役割であるが、一般に、地域区分をするに当って、基礎的観察値はN×nの行列 $X=[X_{ij}]$, $i=1\cdots\cdots N$, $j=1\cdots\cdots n$ として考えられるであろう。

この原資料に若干の統計的操作を加えることによって、各基域に関する観察値X_{ij}を標準化された観察値Z_{ij}に変換するものとしよう。ここで標準化ないし基準化は通常の方法でなされている。即ち、

$$\bar{Z}_j=\sum_{i=1}^{N}Z_{ij}/N=0$$

$$\sigma_j^2=\sum_{i=1}^{N}Z_{ij}^2/N=1$$

主成分分析の基本的アイディアはこの基域iの変量jに関する観察値Z_{ij}を次のような一次式として表現することにある。

$$Z_{ji}=\sum_{p=1}^{m}a_{jp}F_{pi} \qquad (1)$$

或いは、いま変量jのみを問題にすれば、

$$Z_j=\sum_{p=1}^{m}a_{jp}F_p \qquad (1')$$

となる。ここで、F_pは二つ以上の変量の動きを説明する成分、いわゆる共通因子であり、観察値と同じく、次のように標準化されているものとする。即ち、

$$\bar{F}_p=\sum_{i=1}^{N}F_{pi}/N=0$$

$$V(F_p)=\sum_{i=1}^{N}F_{pi}^2/N=1$$

である。

各因子の係数a_{jp}は、いわゆる各変量Z_jの因子負荷量であり、F_pのZ_jの動きに対する説明力を示しているものである。

即ち、(1')式を Z_j に想定することによって、Z_j の分散は次のように表現される。

$$\sigma_j^2=\sum Z_{ij}^2/N=a_{j1}^2\sum F_{1i}^2/N+a_{j2}^2\sum F_{2i}^2/N$$
$$+\cdots\cdots+a_{jm}^2\sum F_{mi}^2/N+2(a_{j1}a_{j2}\sum F_{1i}F_{2i}/N$$
$$+\cdots\cdots+a_{jm-1}a_{jm}\sum F_{m-1i}F_{mi}/N) \qquad (2)$$

前述のように、Z_j の分散は1、F^p の分散もまた仮定によって1であるから、(2)は次のように表現される。

$$1=\sigma_j^2=a_{j1}^2+a_{j2}^2+\cdots\cdots+a_{jm}^2$$
$$+2(a_{j1}a_{j2}r_{F_1F_2}+\cdots\cdots+a_{jm-1}a_{jm}r_{F_{m-1}F_m}) \qquad (3)$$

更に、各因子が相互に独立であるならば、

$$1=\sigma^2=\sum_{p=1}^{m}a_{jp}^2 \quad (p=1\cdots\cdots m) \qquad (4)$$

となる。即ち、a_{jp}^2 は Z_j の分散 $(\sigma_j^2=1)$ に対する因子 F_p の寄与を示している。因子分析における一般的定義にしたがって、この寄与の総和を共通性 h_j^2 と呼ぶことにしよう。主成分分析の基本的特徴はこの共通因子間の独立性にあるわけであって、この結果、Z_j の分散はすべて m 箇の共通因子のみによって説明されることになる。

これらのことから、上掲の基本方程式の意味は明瞭であろう。即ち、主成分分析の目的は、観察値として得られる変量 Z_j の動きを相互に独立な m 箇の因子 F_p によって(1')式の形で説明することにある。換言すれば、変量 Z_j における基域差 Z_{ij} は、共通因子における基域差 F_{pi} に「還元」され、かつ、任意の基域 i の変量 Z_j、Z_k における特徴をあらわす Z_{ij} と Z_{ik} は、共通因子の負荷量である a_{jp} と a_{kp} との相違なる条件を満たしていなければならない。ここで $r_{jk}=r_{kj}$, r_{jj} は Z_j の共通性 h_j^2 と等しいことはいうまでもない。

さて、通常のラグランジュ乗数の適用によって(6)の制約式のもとに(5)の最大化をはかり、$\lambda_1=\sum_{j=1}^{n}a_{j1}^2$ とおくことによって、(Harman, 1960)

$$\sum_{k=1}^{n}r_{jk}a_{k1}-\lambda_1 a_{j1}=0 \qquad (7)$$

が導かれるが、$j=1\cdots\cdots n$ であるから、これは展開して、

$$\begin{array}{l}(r_{11}-\lambda)a_{11}+r_{12}a_{21}+r_{13}a_{31}+\cdots\cdots+r_{1n}a_{n1}=0\\ r_{21}a_{11}+(r_{22}-\lambda)a_{21}+r_{23}a_{31}+\cdots\cdots+r_{2n}a_{n1}=0\\ r_{31}a_{11}+r_{32}a_{21}+(r_{33}-\lambda)a_{31}+\cdots\cdots+r_{3n}a_{n1}=0\\ \cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots\\ r_{n1}a_{11}+r_{n2}a_{21}+r_{n3}a_{31}+\cdots\cdots+(r_{nn}-\lambda)a_{n1}=0\end{array}$$

ともかける。ただし、$r_{jj}=h_j^2$ である。いうまでもなく、これは n 箇の未知数 a_{j1} をもつ n 箇の同次方程式に他ならない。更に行列式で表現すれば、$(R-\lambda I)a=0$ となるであろう。この連立方程式がノントリヴィアルな解をもつ必要十分条件は a_{j1} の係数の行列式が消滅すること、即ち $|R-\lambda I|=0$ である。ところでこのかたちは、通常の線型代数で固有方程式と呼ばれているものに他ならない。即ち、正方行列Rの対角要素からある値 λ を減ずることによって、得られる行列の行列式が0となることを意味する。このことから、ここでの問題がより一般的にはいわゆる固有値問題として知られる線型問題の一変型であることが明らかとなる。即ち、問題は、$|R-\lambda I|=0$ に帰せられることになる。

基本方程式(1')は、その形式からいって、Z_j を従属変数とし、F_p を独立変数とする誤差項を含まない線型回帰方程式を連想させるが、ここで重要なことは、回帰方程式であれば、従属、独立変数がともにあたえられるのに対し、主成分分析では、観察値は Z_{ij} のみで、これを説明する各因子は決して観察されない。その内容すらわかっていないことであろう。F_p とその係数である a_{jp} は、われわれの基礎的データー行列であるZから直接導出される変量間の相関行列、

$$R=\begin{bmatrix} r_{11} & r_{12} & \cdots\cdots & r_{1n}\\ \vdots & & & \vdots\\ r_{n1} & r_{n2} & \cdots\cdots & r_{nn}\end{bmatrix}. \qquad r_{jk}=\sum_{i=1}^{N}Z_{ij}Z_{ik}/N$$

を基礎にして決定されねばならないのである。

さて、前述のように、主成分分析の基本的問題は相関行列Rを基礎に、n 箇の変量 Z_j を説明する m 箇の F_p の係数 a_{jp} $(j=1\cdots n,\ p=1\cdots m)$ を求めることにあるが、この場合、主成分分析の解法の特徴は a_{jp} の選択にあたって、各因子 F_p の共通性に対する寄与の総和が全体として最大となるような a_{jp} を常に選ぶ。そして、この寄与の総和が F_1 で最も大きく、次いで F_2、そして F_3、$F_4\cdots\cdots F_m$ まで漸減するようになっていることである。即ち、いま、F_1 の h_j^3 に対する総寄与を、

$$V_1=\sum_{j=1}^{n}a_{j1}^2 \qquad (5)$$

とするならば、V_1 を最大化するような a_{j1} が選ばれるわけであるが、ただし、この a_{j1} は

$$r_{jk}=\sum_{p=1}^{m}a_{jp}a_{kp} \quad (j,\ k=1\cdots\cdots n) \qquad (6)$$

の形であたえられる固有方程式の根、即ちRの固有値 λ_p を求め、更に、各 λ_p に対応するかたちで、$(R-\lambda I)a=0$ より一般的には、$Ra=\lambda a$ の形であたえられた n 次同次方程式を成立させる a_p、$(a\neq 0)$ 即ちRの固有ベクターを求めることに他ならない。

ところで、相関行列は対称であるから、Rの要素が実数である限り、Rの固有値は実数、又、固有方程式の λ に代入される q 次の根は行列式の階数を $(n-q)$ に減ずるであろう。したがって、固有方程式の一つの根が(7)に代入されるならば、階数 $(n-1)$ の一次同次方程式体系を得ることになるが、ここで基本となるのはいうまでもなく、$\lambda_1=\sum_{j=1}^{n}a_{j1}^2$、そして、これは最大化さるべき V_1 に等しい。換言すれば、V_1 は固有方程式の根のうち最大のものに等しいことになる。

以上の推論によって、第一因子 F_1 の負荷量 a_{j1} を求めるという基本的問題は次のようにして解かれることになる。即ち $|R-\lambda I|=0$ の最大根 λ_1 を $(R-\lambda I)a=0$ に代入し、解 $a_{11}, a_{21}\cdots\cdots a_{n1}$ を得、そして(5)の条件を満たすために、各値をこれら値の自乗和の平方根で除し、最大根の平方根を乗ずる。即ち、

$$a_{j1}=a_{j1}\left(\frac{\lambda_1}{\sum_{j=1}^{n}a_{j1}^2}\right)^{\frac{1}{2}} \qquad j=1\cdots\cdots n \qquad (8)$$

となる。

このようにして、第一因子の負荷量を決定したら、次の作業は残余の共通性を最大に説明する第二の因子を見いだし、その負荷量を

決定することである。このために、もとの相関行列から第一因子による影響（F_1の共通性に対する寄与）を除去し、第一因子残差相関を得、これをもとにして、前と同じ操作によってF_2とa_{j2}を求めればよい。そして、あとは同一の作業を繰返すことによって、m番目の因子まで至るのである。いまS箇の因子の影響が除去されたr_{jk}の残差相関を${}_{s}r_{jk}$であらわすならば、第一因子残差相関は、

$$ {}_{1}r_{jk}=r_{jk}-a_{j1}a_{k1}=\sum_{p=2}^{m}a_{jp}a_{kp} \qquad (9) $$

となり、これはより一般的には、

$$ R_1=R-Q_1 \qquad (9') $$

ただし $Q_1=a_1 a'_1$ となる。F_2の負荷量a_{j2}の決定は、再び、

$$ V_2=\sum_{j=1}^{n}a_{j2}{}^{2} \qquad (10) $$

を最大化することによってなされるが、ただしこの場合の制約条件は（8）となる。いうまでもなく、ここでも基本となるのはR_1の固有値と固有ベクターであるが、ここで便利なことは、R_1の最大固有値は原の相関行列Rの二番目に大きい根、即ちλ_2である。したがって、F_2の負荷量を決定するには、Rの二番目に大なる固有値を求め、それに対する固有ベクターを求めればよいということになる。このことはF_3、F_4……以下F_mに至るまで同様である。こうして、すべての共通性はm箇の因子によって説明されることになるが、このとき、相関行列Rの対角要素がすべて1であれば、一般に $m=n$、即ち、因子の数は変量の数に一致するであろう。又、いくつかの対角要素が1より小さくなれば、$m<n$ となろう。

さて、このような過程によって、われわれは変量Z_jを説明するm箇の共通因子とその負荷量を決定することができたが、いうまでもなく、これは、Z_jを（1'）の形で表現できることを意味している。もし、Z_{ij}を陽表的にあらわすならば、$Z'=AF$となるであろう。前述のように、個々の基域差Z_{ij}は共通因子における基域差F_{pi}に「還元」されるわけであるが、もしこの共通因子が基礎的DMを構成する変量の数に比して、はるかに少く、かつ、それにも拘らず、Z_{ij}の動きを殆んど説明するならば、主成分分析を用いて基域の特性を説明する基本的因子を抽出する、ないしこれら特性を説明する次元を減ずるという、われわれの目的は一応達成されたといってよかろう。

しかしながら、われわれの基本的目的は地域区分にある。即ち、N箇の基域を以上の分析によって抽出された基本的因子を考慮して、可能なかぎり同質的なK箇の地域へと組みわけることにある。このためには、前述の形であたえられる情報は、必ずしも十分なものとはいえない。各基域が抽出された基本因子との関係において、その同質性を判定されるものならば、われわれの判定規準となるのは、現実の基域差が共通因子において還元された値でなければならない。即ち、われわれにとって必要なのは一般に因子得点と呼ばれているF_{pi}に他ならない。主成分分析において、特に相関行列Rが対角要素として1をもつとき、F_{pi}は簡単に求められる。即ち、前述の$Z'=AF$を前提として、F_{pi}は

$$ F=A^{-1}Z' $$

であたえられるであろう。

主成分分析の計算過程

前節において、主成分分析の基本的アイディアを略述したが、ここでは、実際の計算手続きについて述べることにしよう。主成分分析の基本方程式の形式、又、因子負荷量a_{jp}の性格からいって、この計算過程が、一般的な数値計算の問題としては、いわゆる線型計算、そのうちでも、固有値と固有ベクターを求めるそれであることはいうまでもない。一般に行列Aの固有値計算法には、直接法、反復法、行列変換法等いくつかの方法があるが、われわれの扱う相関行列Rが実対称行列であるという基本的特性からいって、反復法と行列変換法について説明するのが最適であろう。前者は試値としてのベクターを次第に改良して固有ベクトルに近づけてゆくものであり、理論的には既に一八八五年 H. A. Schwarz によって見出されていたが、主成分分析の解法として、最初に適用したのは、H. Hotelling (1933, 1936) であった。ここでは、H. H. Harman にしたがって、その計算過程を説明することにする。後者、即ち行列変換法は、いわゆる Jacobi 方法であり、直交変換をくりかえすことによって行列Aを対角型に変換してゆくものである。最初の適用は、T.L. Kelley (1935) によるが、ここでは、やはり Harman にしたがってその方法を解説することにしよう。Harman はこれら二方法のうち、前者を卓上計算機、後者を電子計算機用に夫々リコメンドしている。この意味で、後者の計算手続きについては、本来、プログラミングの形で提出すべきであるが、それは現在の筆者の能力をこえているわけであり、ここでは全く触れなかった。

（一） 反復法

前述したように、ここで述べる方法は基本的に、Hotelling—Harman のそれであるが、内容を完全に self-sufficient かつ operational なものにし、更に記述の一般化と明確化を図るために、計算過程を次の五項目、

Ⅰ　基礎的データーの処理（相関行列Rの算出）

Ⅱ　第一因子負荷量の推定

Ⅲ　第一因子残差の推定

Ⅳ　第二因子負荷量の推定

Ⅴ　因子得点の算出

にまとめ、各項目を更にいくつかの手順に細分し、更にそれぞれの段階にあらわれる基本的数値をまとめた表を附記することにした。

Ⅰ　基礎的データーの処理

地域区分に用いられる基礎的資料であるn箇の変量に関するN箇の基域の観察値行列 $X=[X_{ij}]$ を標準化し、主成分分析にとって必要となる相関行列 $R=[r_{jk}]$ を導く。

Ⅰ—1　基礎的観察値行列（変量n、基域Nより成る総計$n\times N$箇の観算値X_{ij}を要素とする行列）$X=[X_{ij}]$ $(i=1……N,\ j=1……n)$ をつくる。

① 各列毎に列和を集計する。$\sum X_{ij}=\sum_{i=1}^{N}X_{ij}$

② 各列の平均値を算出する $\bar{X}_j=\sum_{i=1}^{N}X_{ij}/N$

I－2 平均偏差行列 $x_{ij}=X_{ij}-\bar{X}_j$ によってつくる。

⑴ x_{ij}を自乗し、列和を集計する。 $\sum x_{ij}^2=\sum_{i=1}^{N}x_{ij}^2$

⑵ 各変量の分散を算出する。$\sigma^2=\sum x_{ij}^2/N$

変量＼基域	1	2	…………	n
1	X_{11}	X_{12}	…………	X_{1n}
2	X_{21}	X_{22}	…………	X_{2n}
⋮	⋮	⋮		⋮
N	X_{N_1}	X_{N_2}	…………	X_{N_n}
計	$\sum X_{1i}$	$\sum X_{i2}$	…………	$\sum X_{in}$
平均	$\bar{X}_1$	$\bar{X}_2$	…………	$\bar{X}_3$

変量＼基域	1	2	…………	n
			……	X_{1n}
			……	X_{2n}
				⋮
				X_{N_n}
			……	$\sum X_{in}$
				$\bar{X}_3$

(1) X_{ij} 表

変量＼基域	1	2	…………	n
1	x_{11}	x_{12}	…………	
2	x_{21}	x_{22}	…………	
⋮	⋮	⋮		
N	x_{N_1}	x_{N_2}	…………	
自乗和	$\sum x_{i1}^2$	$\sum x_{i2}^2$	…………	
分散	σ_1^2	σ_2^2	…………	

(2) x_{ij} 表

I－3 標準値行列 $[z_{ij}]$; $z_{ij}=x_{ij}/\sigma_j$ をつくる。各列が分析の対象となるZ_jを構成する。

変量＼基域	1	2	…………	n
1	z_{11}	z_{12}	…………	z_{1n}
2	z_{21}	z_{22}	…………	z_{2n}
⋮	⋮	⋮		⋮
N	z_{N_1}	z_{N_2}	…………	z_{N_n}

(3) z_{ij} 表

I－4 相関行列 $R=[r_{jk}]$, $r_{jk}=\sum_{i=1}^{N}z_{ij}z_{ik}/N$ $(j,k=1\cdots\cdots n)$ をつくる。

① 共通性の推定は $r_{jj}=1$とする。

変量＼変量	1	2	…………	n
1	r_{11}	r_{12}	…………	r_{1n}
2	r_{21}	r_{22}	…………	r_{2n}
⋮	⋮	⋮		⋮
n	r_{n1}	r_{n2}	…………	r_{nn}

(4) r_{jk} 表

II 第一因子負荷量の推定

相関行列Rを基礎とし、第一因子負荷量推定のための試値を推定し、これにもとづいて負荷量を決定する。

II－1 Rの累乗により試値を求め、その安定性を規準に最終試値を求める。

(イ) 第一試値$\alpha_{j1}^{(1)}$の算出

① Rの行和を集計する $s_j=\sum_{k=1}^{n}r_{jk}$

② 最大の行和で各行和を除す $\alpha_{j1}^{(1)}=s_j/\max(s_j)$

(ロ) 第二試値の算出

① Rの各要素を自乗する。$R^2=[r_{jk}^2]$

② R^2の行和を求める $s_j^{(2)}=\sum_{k=1}^{n}r_{jk}^2$

③ 最大の行和で各行和を除す $\alpha_{j1}^{(2)}=s_j^{(2)}/\max(s_j^{(2)})$

④ 検算のため$T_j^{(2)}$を計算する $T_j^{(2)}=\sum_{k=1}^{n}r_{kj}s_k$

⑤ 検算は $\alpha_{j1}^{(2)}=s_j^{(2)}/\max(s_j^{(2)})=T_j^{(2)}/\max(T_j^{(2)})$ でなされる。

(ハ) 試値の比較

① $\alpha_{j1}^{(1)}$と$\alpha_{j1}^{(2)}$を比較して、もし $\alpha_{j1}^{(1)}\approx\alpha_{j1}^{(2)}$ ならば$\alpha_{j1}^{(2)}$を（II－2）の試値$\alpha_{j1}^{(l)}$として採用する。* もし $\alpha_{j1}^{(1)}\neq\alpha_{j1}^{(2)}$ならば、(ロ)の手続きによって$R^3$、$R^4$……を求め、$\alpha_{j1}^{(3)}$、$\alpha_{j1}^{(4)}$……を求める。

* 誤差の水準はr_{jk}が小数点以下3ケタならば 0.005 以下。

II－2 $\alpha_{j1}^{(l)}$より最終試値α_{j1}を求め、その安定性にもとづき第一固有値λ_1を求め、α_{j1}、λ_1で第一因子負荷量a_{j1}を決定する。

	1	2	…………	n	s_j	$\alpha_{j1}^{(1)}$
1	r_{11}	r_{12}	…………	r_{1n}	s_1	$\alpha_{11}^{(1)}$
2	r_{21}	r_{22}	…………	r_{2n}	s_2	$\alpha_{21}^{(1)}$
⋮	⋮	⋮		⋮	⋮	⋮
n	r_{n1}	r_{n2}	…………	r_{nn}	s_n	$\alpha_{n1}^{(1)}$

(5) R行列と第一試値

	1	2	…………	n	$s_j^{(2)}$	$T_j^{(2)}$	$\alpha_{j1}^{(2)}$
1	r_{11}^2	r_{12}^2	…………	r_{1n}^2	$s_1^{(2)}$	$T_1^{(2)}$	$\alpha_{11}^{(2)}$
2	r_{21}^2	r_{22}^2	…………	r_{2n}^2	$s_2^{(2)}$	$T_2^{(2)}$	$\alpha_{21}^{(2)}$
⋮	⋮	⋮		⋮	⋮	⋮	⋮
n	r_{n1}^2	r_{n2}^2	…………	r_{nn}^2	$s_n^{(2)}$	$T_n^{(2)}$	$\alpha_{n1}^{(2)}$

(6) R^2 行列と第二試値

① $R\longrightarrow R^l$によりα_{j1}^lを求める。

② α_{j1}^lを相関行列Rに乗じ、γ_{j1}を求める。$\gamma=R\alpha$

③ γ_{j1}を最大のγ_{j1}で除し、最終試値α_{j1}を得る。$\alpha_{j1}=\gamma_{j1}/\max(\gamma_{j1})$

④ α_{j1}^lとα_{j1}を比較して、最終試値α_{j1}の安定性を検討する。$\alpha_{j1}^l\approx\alpha_{j1}$

⑤ $\alpha_{j1}=1.000$ に対応するγ_{j1}を、第一固有値λ_1とする。 $\alpha_{j1}=1.000\longrightarrow\gamma_{j1}=\lambda_1$

⑥ a_{j1}をα_{j1}、λ_1を用いて計算する。

	α_{j1}^l	r_{j1}	α_{j1}	a_{j1}
1	α_{11}^l	r_{11}	α_{11}	a_{11}
2	α_{21}^l	r_{21}	α_{21}	a_{21}
⋮	⋮	⋮	⋮	⋮
n	α_{n1}^l	r_{n1}	α_{n1}	a_{n1}

(7) r_{j1}, α_{j1}, a_{j1} 表

$$a_{j1}=\alpha_{j1}\sqrt{\lambda_1}\Big/\sqrt{\sum_{j=1}^{n}\alpha_{j1}^2}$$

II－3 第一因子負荷量の検算は、$\sum_{j=1}^{n}a_{j1}^2=\lambda_1$ でなされる。

III 第一因子残差の推定

第一因子の共通性に対する寄与を示す第一因子負荷量の積行列Q_1を算出し、これを原相関行列Rから減ずることによって、第一因子残差を求める。

III－1 積行列 $Q_1=[{}_1q_{jk}]$ をつくる。$Q_1=a_1a_1'$

① Q_1の各行和E_{j1}を求める。$E_{j1}=\sum_{k=1}^{n}{}_1q_{jk}$

② 検算のためD_1を求める。$D_1=\sum_{k=1}^{n}a_{k1}$

③ 検算は、$a_{j1}D_1=E_{j1}$ でなされる。

	1	2	…………	n	E_{j1}	a_{j1}	D_1
1	${}_1q_{11}$	${}_1q_{12}$	…………	${}_1q_{1n}$	E_{11}	a_{11}	D_1
2	${}_1q_{21}$	${}_1q_{22}$	…………	${}_1q_{2n}$	E_{21}	a_{21}	D_1
⋮	⋮	⋮		⋮	⋮	⋮	⋮
n	${}_1q_{n1}$	${}_1q_{n2}$	…………	${}_1q_{nn}$	E_{n1}	a_{n1}	D_1

(8) Q_1 表

III－2 原相関行列Rから積行列Q_1を差引き、第一因子残差行列 $R_1=[{}_1r_{jk}]$ を得る。$R_1=R-Q_1$

	1	2	…………	n	s_{j1}	$\alpha_{j2}^{(1)}$
1	${}_1r_{11}$	${}_1r_{12}$	…………	${}_1r_{1n}$	s_{11}	$\alpha_{12}^{(1)}$
2	${}_1r_{21}$	${}_1r_{22}$	…………	${}_1r_{2n}$	s_{21}	$\alpha_{22}^{(1)}$
⋮	⋮	⋮		⋮	⋮	⋮
n	${}_1r_{n1}$	${}_1r_{n2}$	…………	${}_1r_{nn}$	s_{n1}	$\alpha_{n2}^{(1)}$

(9) R_1 表

① R_1の各行和を求める。 $s_{j1}=\sum_{k=1}^{n}r_{jk}$

② 最大の行和で各行和を除し、第二因子負荷量推定量のための試値$\alpha_{j2}^{(1)}$を得る。 $\alpha_{j2}^{(1)}=s_{j1}/max(s_{j1})$

③ 検算は $s_{j1}=s_j-E_{j1}$ でなされる。

Ⅳ 第二因子負荷量の推定

第二因子負荷量 α_{j2} の推定は、 α_{j1} 推定と同じ過程、即ち相関行列の累乗を通じて、試値 $\alpha_{j2}^{(1)}$ $\alpha_{j2}^{(2)}$ …… $\alpha_{j2}^{(l)}$ を求め、その安定性を検定し、 $\alpha_{j2}^{(l)}$ に相関行列を乗ずることによって γ_{j2} を得て、第二固有値 λ_2 と最終試値 α_{j2} を用いて a_{j2} を得るという過程を経てなされる。試値の決定には、いうまでもなく、 R_1 の累乗が使われるが、実際に R_1 の累乗を計算する代りに、 $R_1^{\,l}=R^l-\lambda_1^{\,l-1}Q_1$ を用いて $R_1^{\,2}$ $R_1^{\,3}$ …… が R^l と Q_1 から計算される。又、試値の決定には、 $R_1^{\,2}$ $R_1^{\,3}$ …… の行和さえ知ればよく、その各要素を知る必要はない。 $s_j^{(2)}$ と E_{j1} を夫々 R^2 と Q_1 の要素から成ると考えれば、前掲式にしたがって、 $s_{j1}^{(2)}=s_j^{(2)}-\lambda_1E_{j1}$ となり、 $R_1^{(2)}$ の各行和 $s_{j1}^{(2)}$ は $R_1^{(2)}$ の各要素を算出することなく得られる。

Ⅳ－1 試値の推定　前述の議論にしたがって、 $\alpha_{j1}^{(1)}$ $\alpha_{j2}^{(2)}$ …… $\alpha_{j2}^{(l)}$ を推定する。

① s_j、 E_{j1} は既出

② s_j から E_{j1} を差引くことによって s_{j1} を得る。 $s_{j1}=s_j-E_{j1}$

③ 最大の s_{j1} で各 s_{j1} を除して第一試値を得る。 $\alpha_{j2}^{(1)}=s_{j1}/max(s_{j1})$

④ $s_j^{(2)}$ は既出。 λ_1、 E_{j1} (ともに既出)で $\lambda_1 E_{j1}$ を求める。

⑤ $s_j^{(2)}$ から λ_1E_{j1} を差引いて $s_{j1}^{(2)}$ を求める。 $s_{j1}^{(2)}=s_j^{(2)}-\lambda_1E_{j1}$

⑥ 最大の $s_{j1}^{(2)}$ で各 s_{j1} を除して、第二試値を得る。 $\alpha_{j2}^{(2)}=s_{j1}^{(2)}/max(s_{j1}^{(2)})$

⑦ $\alpha_{j2}^{(1)}$ と $\alpha_{j2}^{(2)}$ を比較して $\alpha_{j2}^{(1)}\fallingdotseq\alpha_{j2}^{(2)}$ ならば $\alpha_{j2}^{(2)}$ を $\alpha_{j2}^{(l)}$ としてⅣ－2へ、もし $\alpha_{j1}^{(1)}\neq\alpha_{j1}^{(2)}$ ならば、④⑤⑥の手続きで $\alpha_{j2}^{(3)}$ $\alpha_{j2}^{(4)}$ …… を算出する。

Ⅳ－2

① Ⅳ－1により $\alpha_{j2}^{(l)}$ より α_{j2} を求め、 a_{j2} を求める。

② $\alpha_{j2}^{(l)}$ に R_1 を乗じ γ_{j2} を得る。 $\gamma=R_1\alpha$

③ 最大の γ_{j2} で γ_{j2} を除し、 α_{j2} を得る。 $\alpha_{j2}=\gamma_{j2}/\max(\gamma_{j2})$

④ $\alpha_{j2}^{(l)}$ と α_{j2} を比較して安定性を認めれば、 α_{j2} を最終試値とする。 $\alpha_{j2}^{(l)}\fallingdotseq\alpha_{j2}$

	s_j	E_{j1}	s_{j1}	$\alpha_{j2}^{(1)}$	$s_j^{(2)}$	λ_1E_{j1}	$s_{j1}^{(2)}$	$\alpha_{j2}^{(2)}$
1	s_1	E_{11}	s_{11}	$\alpha_{12}^{(1)}$	$s_1^{(2)}$	λ_1E_{11}	$s_{11}^{(2)}$	$\alpha_{12}^{(2)}$
2	s_2	E_{21}	s_{21}	$\alpha_{22}^{(1)}$	$s_2^{(2)}$	λ_1E_{21}	$s_{21}^{(2)}$	$\alpha_{22}^{(2)}$
⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮	⋮
n	s_n	E_{n1}	s_{n1}	$\alpha_{n2}^{(1)}$	$s_n^{(2)}$	λ_1E_{n1}	$s_{n1}^{(2)}$	$\alpha_{n2}^{(2)}$

(10) $\alpha_{j2}^{(1)}, \alpha_{j2}^{(2)}$ 表

	$\alpha_{j2}^{(l)}$	γ_{j2}	α_{j2}	a_{j2}
1	$\alpha_{12}^{(l)}$	γ_{12}	α_{12}	a_{12}
2	$\alpha_{22}^{(l)}$	γ_{22}	α_{22}	a_{22}
⋮	⋮	⋮	⋮	⋮
n	$\alpha_{n2}^{(l)}$	γ_{n2}	α_{n2}	a_{n2}

(11) γ_{j2}、 α_{j2}、 a_{j2} 表

⑤ $\alpha_{j2}=1.000$ に対応する γ_{j2} の値を λ_1 の値とする。

⑥ α_{j2}、 λ_2 で a_{j2} を計算する。 $a_{j2}=\alpha_{j2}\sqrt{\lambda_2}\Big/\sqrt{\sum_{j=1}^{n}\alpha_{j2}^{\,2}}$

⑦ a_{j2} の検算は $\sum_{j=1}^{n}a_{j2}^{\,2}=\lambda_2$ でなされる。

以上の計算過程を繰返すことによって、第三、第四……の因子負荷量 a_{j3}、 a_{j4} …… a_{jm} が決定される。これら計算のアウトプットは最終的に因子負荷量表にまとめられる。 V_p はいうまでもなく、 p 番目の因子の総寄与、 P_p は全共通性の p 番目の因子によって説明された部分の割合を示している。

Ⅴ 因子得点の測定

各因子の F_{pt} 得点は、負荷量行列の逆行列を求め、これに Z' 行列を乗ずることによって得られる。 $F=A^{-1}Z'$ この結果として得られる数値は因子得点行列の各要素を構成し、 Z_{ij} を m 次元空間の一点としてあらわすことを可能にする。

因子／変量	Ⅰ	Ⅱ	……	M
z_1	a_{11}	a_{12}	……	a_{1m}
z_2	a_{21}	a_{22}	……	a_{2m}
⋮	⋮	⋮		⋮
z_n	a_{n1}	a_{n2}	……	a_{nm}
V_p	V_1	V_2	……	V_m
P_p	P_1	P_2	……	P_m

(12) 因子負荷量表

因子／基域	Ⅰ	Ⅱ	……	M
1	f_{11}	f_{12}	……	f_{1M}
2	f_{21}	f_{22}	……	f_{2M}
⋮	⋮	⋮		⋮
N	f_{N1}	f_{N2}	……	f_{NM}

(13) 因子得点表

(二) Jacobi 法

行列変換、即ち変形 Jacobi 法の基本的アイディアは、一連の直交変換をほどこすことによって、対角要素以外の要素を0として、相関行列Rを対角型に変換することにある。一般に、実 n 次元空間の R^n の直交変換、即ち回転は $(n-2)$ 本の座標軸を固定して残りの二本の座標軸 x_j、 x_k の平面 π_{jk} 内の回転によって生成される。即ち、いま、平面 π_{jk} 内の角 θ の回転を

$$B_{jk}=\begin{pmatrix}1&&&&&&\\&\ddots&&&&&\\&&\cos\theta_{jk}&&-\sin\theta_{jk}&&\\&&&\ddots&&&\\&&\sin\theta_{jk}&&\cos\theta_{jk}&&\\&&&&&\ddots&\\&&&&&&1\end{pmatrix}$$

(空白はすべて0)

と表わせば、各直交変換は $B_{jk}RB_{jk}'$ という形であらわされる。ここで、 B_{jk} の $\cos\theta_{jk}$、 $-\sin\theta_{jk}$、 $\sin\theta_{jk}$、 $\cos\theta_{jk}$ は夫々 j と k の行と列の交錯要素、その他の対角要素はすべて1、その他の要素はすべて0である。いうまでもなく、Rを対角行列に変換する回転を正確に B_{jk} の結合であらわすのは難しいが、これは逐次近似になされる。即ち、Rの対角要素以外の要素で絶対値の最大な要素を r_{jk} とするならば、それを j、 k に選び、 θ を、

$\tan 2\theta=2r_{jk}/r_{jj}-r_{kk}$

ただし $r_{jj}=h_j^2$, $r_{kk}=h_k^2$ として、

$$\begin{pmatrix} r_{jj} & r_{jk} \\ r_{kj} & r_{kk} \end{pmatrix}$$

が対角行列になるようにする。次に、再び導かれた行列の対角要素以外の要素で絶対値が最大な要素を選び、同様な手続きをくりかえす。この直交変換のくりかえしの段階にあって、既に変換された r_{jk} が0に止まるという保証はない。しかしながら、変換の度に対角要素以外の要素の自乗和が $2r_{jk}^2$ だけ減ずるということ、したがってRが対角行列に収束することは証明されており、対角要素以外の要素が要求される精度で $r_{jk}\approx 0$ となれば、計算を終了すればよい。対角要素が求める解即ち固有値をあたえることになる。

いま対角行列をD、Rに最初の直交変換をほどこして得られる行列をDに v をつけて表記すれば、各変換は、

$${}_vD=B_{jk\,(v-1)}DB_{jk}'$$

と表記される。ただし、変換順序は、

$$v=n(j-1)-j(j+1)/2+k;\ (j=1,\ 2,\cdots n-1,\ k=j+1,\ j+2,\cdots n)$$

で得られるものとする。

各段階における v 箇の変換の積を

$${}_vB=B_{jk}\cdots\cdots B_{13}B_{12}$$

とあらわし、又、$n(n-1)/2$ 箇の可能な組合せ $(j,\ k)$ すべてについて変換をほどこしたとき、i 番目のくりかえしにおける変換の積を、

$$B_i=\Pi(B_{jk};\ j<k=1,\ 2,\cdots\cdots n),\ i=1,\ 2,\cdots\cdots\mu$$

とあらわせば、i 番目のくりかえしによって導出される対角行列は、

$$D_i=B_iD_{i-1}B_i',\quad i=1,2,\cdots\cdots\mu$$

であたえられる。ここで i 番目のくりかえしまでの変換行列の積を、

$$P_i=B_iB_{i-1}\cdots\cdots B_2B_1$$

とすれば、前掲式は、

$$D_i=P_iRP_i'$$

ともあらわせる。即ち D_i は相関行列Rとの関係で明記できる。

μ 回のくりかえしによって、解は次の形であたえられる。

$$D_\mu=B_\mu D_{\mu-1}B'=BRB'=D$$

ただし $B=P_\mu=B_\mu B_{\mu-1}\cdots\cdots B_2B_1$, Dの対角要素 λ_p はRの固有値であり、最終の変換行列Bの行はそれに対応するRの固有ベクタ $\alpha_p'=(\alpha_{1p}\ \alpha_{2p}\cdots\cdots\alpha_{np})$ を構成することになる。各因子負荷量は

$$a_p=\alpha_p\sqrt{\lambda_p}$$

によって得られる。

若干の問題点

以上、主成分分析を地域区分に適用するという観点から説明してきたが、最後に、本稿では意識的に論じなかった、ないし論及し得なかったいくつかの問題（それはすべて重要なものである）があることを指摘しておくことが必要であろう。そのうち、特に重要なものは、(一)因子の回転、(二)共通性の推定とであろう。

(一) 因子回転の目的はいうまでもなく、抽出された因子の適切な解釈にある。既にくりかえし述べてきたように、分析の目的は、n 箇の変量を m 箇の因子で説明することにあるが、この場合、各因子がどのような意味内容をもつものであるかが解釈されなければならない。基域の特性を可能なかぎり多数の変量との関係において観察するというのは、いうまでもなく、われわれが基域を本来多面的性格をもっているとし、したがって、その結果形成される地域を一つのまとまりある entity として考えて、それがどのような側面からみても可能なかぎり同質的であることを望むからであるが、主成分分析は、この意味からいえば、基域の性格の基本的側面を因子というかたちで浮彫することに他ならない。したがって、抽出された因子がどのような内容をもち、たとえば、基域の経済的側面或いは政治的側面を示すものであるかを解釈し、これら各側面の関係を知ることは、地域分析という点からいって基本的に重要であることはいうまでもない。地域区分を単に統計的処理に終らせないためにも、因子回転による各因子の解釈は是非とも必要であろう。

(二) 問題は共通性の推定、より明確にいえば、Rの対角要素として $h^2=1$ を用いるか $h^2<0$ を用いるかについても基本的には同じであろう。実際に因子分析においてこのいずれを用いるかは、Thusstone (1947) が述べているように、調査の問題意識に依存するわけであって、分析の目的が単に m 次元空間における基域の点としての位置づけである場合には、対角要素には1が使用さるべきであるが、他方、目的が変量間の相関関係の適格な再生にあるならば、推定された共通性が用いられるべきであろう。又、こうした基本的認識とは別に、いずれを利用した方がより計算過程上有効かという技術上の問題もある。Greig-Smith (1964) はこの点について、Goodall の Victorian Mallee の分析を基礎にして、この問題について、「分析のこまいかい方法上の差異は、生態学的結論に重大な影響をあたえない」と述べているし、又、Cooley, Lohnes (1962) は、少くとも主成分分析に関する限り、対角要素に1をおくことはさしつかえない。いやむしろ、それが望ましいという H. F. Kaiser の報告をあげ、それにしたがっている。

以上、本稿では扱わなかったが、特に重要な問題点を二つあげたが、要するに、これは、地域区分という作業を限定的に考えるか、或いはより広範囲に地域の理解という立場から考えるかによって、これらをどう扱うかということがきまってくるといってよかろう。これとは別に、主成分分析の包括的理解のためにこれらを吟味検討することが必要なことはいうまでもない。

文献

1) Berry, Brian J. L. A method for Deriving Multi-Factor Uniform Regions, PRZEGLAD GEOGRAFICZNY t. xxxiii, z. 2, 1961.
2) Duncan, Cuzzort, Duncan, Statistical Geography, 1961.
3) Cooley, Lohnes, Multivariate Procedures for the Behavioral Sciences, 1962.

4) Greig-Smith P. Quantitative Plant Ecology, 1964, 2nd ed.
5) Hagget, Peter, Locational Analysis in Human Geography, 1965.
6) Hagood M.J., Statistical Method for Delineation of Regions, Applied to Data on Agriculture and Population, Social Forces, 21 (March 1943), pp. 288-297.
7) —— and Price, Statistics for Sociologists, 1952.
8) Hadley G., Linear Algebra, 1961.
9) Harman H.H., Modern Factor Analysis, 1960.
10) 一松信 数値計算 昭和三八年
11) Lawley, Maxwell, Factor Analysis as a Statistical Method, 1963.
12) Miller, Kahn, Statistical Analysis in the Geological Sciences, 1962.

次号目次

書評

大前朔郎・池田信共著
『日本労働運動史論――大正一〇年の川崎・三菱神戸両造船所争議の研究』

飯田鼎

一

わが国の労働運動史にかんする本格的研究は、はじめられたばかりである。史料の蒐集および編纂とならんで、これらの上に立った科学的研究の成果が次第にあらわれつつあるとはいえ、それらは明治期の黎明期に属するものか、あるいは戦後の運動史が多く、大正期、まさしく、日本における独占資本主義の成熟とこれを背景としての労働組合運動の全国的な発展という時点での研究は、実に寥々たる有様であった。その大きな理由は、何といっても、当時の労働者階級の運動の生々しい姿を正確に伝える史料が絶対的に不足していることがあげられなければならない。官憲の苛酷な弾圧による史料保存の困難と、これに加えて第二次世界大戦による焼失の被害が、研究者のこの時期の労働運動史への接近をいちじるしく制約していたことは疑いえない。このような悪条件のなかで、これらを克服し、第一次的史料の充分な評価の上に立って、この時期の労働運動史の一断面を明らかにすることは容易な業ではない。ここにとりあげた大前・池田両氏の共同労作は、この意味でまことに意義深い研究成果であるといわなければならない。わたくしはいま、わが国の労働組合運動の研究をはじめたばかりであり、本書から実に多くのことを学ぶことができた。そこで本書について、その内容の紹介とともに、所見の一端を披瀝してみたいと考える。

二

著者は、その「はしがき」の冒頭でつぎのようにのべている。「本書の中心テーマは、神戸の二大造船所、川崎造船所と三菱神戸造船所における大正一〇年（一九二一）の大争議である。それは、戦前の日本労働運動史上における画期的な争議であった。参加人員三万人、直接の争議日数五〇日におよぶ大規模なものであった。また、おなじころおこなわれた大阪地方の争議とともに、わが国においてはじめて、明確に団体交渉権を要求して闘われた争議であった。そしてついには、争議団が工場管理宣言を発表し、これを弾圧するために軍隊が出動した深刻な争議でもあった。この争議は第一次大戦ブーム期から戦後恐慌期へ移行するにあたって、経営者がとった労働者支配の強化――いわゆる飴と鞭との二面をもった労務管理――にたいする労働者の強力で創意ある闘争であった。」この叙述からも明らかなように、本書は、その規模の大きさも画期的であったが、その要求は団体交渉権、闘争方式は工場管理という点でまこ