

# 第 1 部

## Part 1

### 第 1 章 商品分類の改定に伴う貿易統計の変換

－ 日本および韓国を例として －

Chapter 1 Conversion of Trade Statistics with Reversion to Commodity Classification:  
Case Study of Japan and Korea

### 第 2 章 商品分類 SITC3 桁分類コード変換のための配分ウエイト推計

－ ニューラル・ネットワークを用いて －

Chapter 2 Estimation of Distributed Weight to Conversion for 3 Digits Level of SITC

### 第 3 章 商品分類の改定に伴う貿易統計の整合性評価

Chapter 3 Evaluation to Consistency of Trade Statistics with Reversion to  
Commodity Classification

### 第 4 章 貿易統計データベースにおける国コードの利用

Chapter 4 Application of Country code in Database of Trade Statistics

# 第1章

## 商品分類の改定に伴う貿易統計の変換

### 一 日本および韓国を例として一

野田容助

#### はじめに

国連の標準国際貿易商品分類 (SITC) 体系は経済分析に適した体系を保ちつつ関税協力理事会 (CCC) の品目表の改定に合わせてそれとの対応関係を維持することを目的として改訂第1版 (SITC-R1)、改訂第2版 (SITC-R2)、改訂第3版 (SITC-R3) まで順次改訂されてきている。そのため、過去に溯って時系列データとして貿易統計を利用するときには分類の改訂のおこなわれた前後では商品に対する定義およびそのカバレッジが必ずしも同じではないため改訂年を含む連続年では取引金額や数量を直接利用できなくなることがある。しかし、商品分類の改訂により貿易統計の利用ができなくならないように国連ではSITC-R1とSITC-R2、SITC-R2とSITC-R3のように改訂ごとに対応関係の一覧表を作成してきている。

SITC-R1とSITC-R2の対応関係のように、異なる分類体型をもつ分類Aと分類Bの間に相互にかかわりをもつ個別分類コードの対応関係があるとき、この分類AとBのコード間の閉じた対応関係の集まりをグループという。商品分類の場合には商品グループとなる<sup>(注1)</sup>。グループ化された対応関係を使用するには2つの分類どうしがどのような対応関係にあるかを知ることが重要な問題になる。グルー

プ化された対応関係の結合の組み合わせは基本的にはタイプ1、タイプ2、タイプ3およびタイプ4の4つのタイプに分割することができる。分類Aから分類Bへの方向をもった対応関係を考えるとき、タイプ1は分類Aと分類Bにおける個別分類コードの結合の関係が1対1の対応関係である。タイプ2はそれが1対多の対応関係である状態をいい、複数個ある分類Bの個別分類コードが1本ずつの結合の手を持つのに対して、分類Aは1個の個別コードが分類Bの個別コードに対応する本数の結合の手をもつ対応関係の集まりである。タイプ3は多対1の対応関係であり、タイプ2とは逆に複数個ある分類Aの個別コードが1本ずつの結合の手を持つのに対して、分類Bの1個の個別コードは分類Aの個別コードに対応する個数の結合の手を持つ。タイプ4は多対多の対応関係を表し、タイプ1からタイプ3以外のタイプの集まりである。タイプ4はさらにタイプ4aとタイプ4bの2つに識別することができる<sup>(注2)</sup>。

このようにしてグループ化された対応関係の中で分類Aで表わされた統計値が対応する分類Bではどのように変換されて推計されるか、また逆に分類Bで表わされた統計値が対応する分類Aではどのように変換されて推計されるかを貿易統計の商品分類の改定に伴う分類の違いに対して適用を試みたのが

本章の内容である。対応関係における4つのタイプの中でタイプ1とタイプ3はAの統計値を分類Bに対して直接対応させるかあるいは統合するだけで変換が可能である。ところが、タイプ2およびタイプ4ではある条件の下で分類Aの統計値を分類Bに対してウェイトを考慮して配分することが必要になる。このウェイトを配分ウェイトといい、グループ化された対応関係に存在する分類Aの個別分類コードの統計値をBへのそれぞれの個別分類コードへ適当なウェイトを考慮して配分するための係数である。図1に配分ウェイトの関係が示されている。

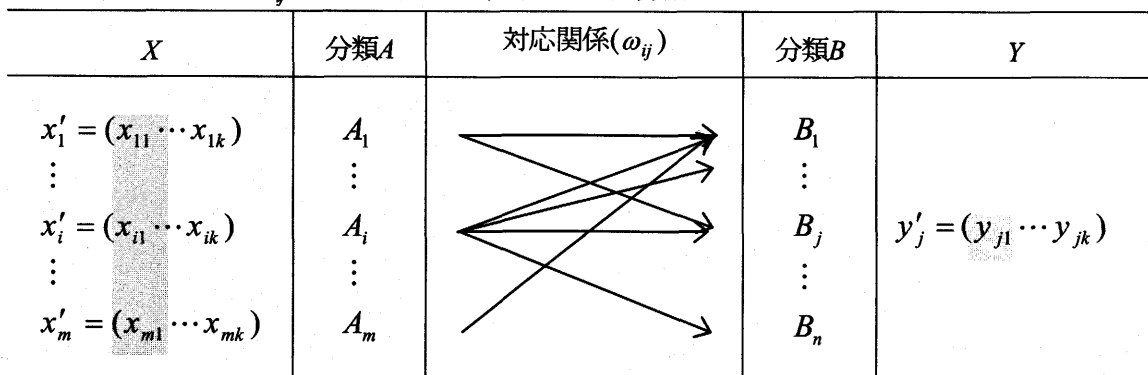
配分ウェイトの推計にあたってはデータ数との関係から本章で紹介した方法を直接用いるのではなく第2章の「SITC3桁分類コード変換のための配分ウェイト推計—ニューラル・ネットワークを用いて—」で説明されるニューラル・ネットワークによるデータマイニングの方法により商品グループごとに計算する。データの安定性を考慮して商品分類をSITCそれぞれの3桁レベルとする。また、データ数を考慮して輸出入の区別をせずに同一のものと仮定する。この貿易統計の変換

は分類Aから分類Bへ方向を持った配分ウェイトの推計である。したがって、この逆の方向の配分ウェイトを得るためにはAとBを入れ替えた上で同じ推計をおこなう必要がある。

本章ではこの得られた配分ウェイトにもとづいてSITC-R1、SITC-R2、SITC-R3における輸出入の各系列の時系列データを作成し、貿易統計を推計している。貿易統計はOECD編集による貿易統計の日本、国連編集による貿易統計の韓国について1962年から99年までを対象とする。商品分類は日本では1962年から77までがSITC-R1、78年から87年までがSITC-R2、87年から99年までがSITC-R3であり、韓国は62年から84年までがSITC-R1、85年から87年までがSITC-R2、88年から99年までがSITC-R3により編集されている。

ここで作成されたSITCの各系列で表された貿易統計は商品分類の改訂年の前後において必ずしも整合性のある状態で接続されているとは限らない。貿易統計のこのような商品分類の改訂年前後の構造変化を検討しているのが第3章の「商品分類の改訂に伴う貿易統計の整合性評価」である。また、貿易統

図1 配分ウェイト  $\omega_{ij}$  による分類Aから分類Bへの配分構造



(出所) 筆者作成

(注)  $\omega_{ij}$ は分類コード $A_i$ から $B_j$ への方向に対する配分ウェイトである。統計値 $X$ と $Y$ は標本数をともに $k$ 個とする。

表1 SITC改訂版ごとの商品分類の項目数

分類のレベル 改訂版の種類	大分類 (1桁)	中分類 (2桁)	小分類 (3桁)	細分類 (4桁)	基本項目 (4,5桁)
SITC-R1	10	56	177	625	1,312 (368)
SITC-R2	10	63	233	786	1,832 (365)
SITC-R3	10	67	261	1,033	3,121 (299)

(出所) 野田容助「商品分類の改定に伴う対応関係の連結」(古河俊一・野田容助共著『標準国際貿易統計と産業分類の対応関係』アジア経済研究所 1998) の表1を引用

(注) 基本項目の後ろの( )の中の数字は4桁コードを表わす。

計の推計は38年間に渡っておこなわれるため、相手国の変遷が問題となるが、それについては第3章の「貿易統計データベースにおける国コードの利用」において説明されている。

### 1. 3桁レベルによる商品分類のグループ化

商品分類SITC-R1とSITC-R2の詳細な分類コードである基本項目による対応関係コード表はUN統計局発行の*Standard International Trade Classification, Revision 2* から得ることができる。この対応関係コード表は1,309個のSITC-R1の個別分類コードと1,832個のSITC-R2の個別分類コードの間のいくつかの組み合わせにより2,002個の対応関係から構成されている。山本の「貿易統計における商品の分類」によれば、SITC-R1を改訂したSITC-R2は1960年以降の対外貿易の構造的変化を考慮して分類の有用性を増大するための修正あるいは技術の進歩によってSITC-R1を拡大するために必要になったための修正を除いては基本的にはSITC-R1の構造を保っている。しかし基本項目数ではかなりの改訂がおこなわれている。

表1によれば、SITC-R1の基本項目数1,312

に対してSITC-R2のそれは1,832と大幅に増加している。大項目を除く各項目においてもその傾向は同じである。この対応関係コード表は必ずしも整合性のあるものばかりではなく、問題となるところが随所に見られる。アジア経済研究所ではこれらの問題点について調整した対応関係コード表を作成し、貿易統計の利用に供してきている。対応関係のモデル化にさいしてアジア経済研究所が調整した対応関係コード表を対応関係の基本モデルとしている。

商品分類SITC-R2とSITC-R3の基本項目における対応関係コード表はUN統計局発行の*Standard International Trade Classification Revision 3* (Statistical Papers Series M no.34/Rev3, United Nations 1986)から得られた対応関係を基本とする。この対応関係コード表は2,654個のSITC-R3の個別分類コードと1,831個のSITC-R2の個別分類コードの間のいくつかの組み合わせにより5,245個の対応関係から構成されている。

SITCの基本項目をもとにして作成された対応関係からSITCの上位の1,2,3桁レベルを対象として改めて作成したのが各桁レベルにおける対応関係コード表である。第2部の表1にSITCの1,2,3桁レベルにおける商品分類コードとその名称が示されている。同一桁レ

表2 SITC-R1とSITC-R2における3桁レベルの対応関係の例

$G_i(j):type$	$X_A$	$X_B$	$X_{A-f}$	$X_{B-f}$	$X_{A-Q}$	$X_{B-Q}$	日本		韓国	
							$X_A \rightarrow X_B$	$X_A \leftarrow X_B$	$X_A \rightarrow X_B$	$X_A \leftarrow X_B$
39 1 1	263	263	1	1	49	52	1.000000	1.000000	1.000000	1.000000
40 1 4a	264	264	1	2	50	53	1.000000	0.974560	1.000000	0.999990
40 1 4a	265	264	2	2	51	53	0.049850	0.025440	0.061720	0.000010
40 1 4a	265	265	2	1	51	54	0.950150	1.000000	0.938280	1.000000
:										
47 1 4a	275	277	2	1	57	62	0.479860	1.000000	0.344860	1.000000
47 1 4a	275	278	2	3	57	63	0.520140	0.096701	0.655140	0.002540
47 1 4a	276	278	1	3	58	63	1.000000	0.903289	1.000000	0.997450
47 1 4a	283	287	1	2	61	67	1.000000	0.472140	1.000000	0.402640
47 1 4a	321	322	2	1	67	72	0.257220	1.000000	0.946820	1.000000
47 1 4a	321	323	2	2	67	73	0.742780	0.223910	0.053180	0.551490
47 1 4a	513	278	4	3	77	63	0.116299	0.000010	0.000010	0.000010
47 1 4a	513	287	4	2	77	67	0.011950	0.527860	0.138930	0.597360
47 1 4a	513	323	4	2	77	73	0.313257	0.776090	0.032280	0.448510
47 1 4a	513	522	4	1	77	89	0.558494	1.000000	0.828780	1.000000

(出所) 第2部の表3にもとづき筆者作成。

(注)  $X_A$  はSITC-R1、 $X_B$  はSITC-R2をそれぞれ表わす。

ベルの分類コードであっても改訂版によっては内容が異なっているので、その違いが分かるように一覧表にして表わされている。同一桁レベルの分類コードであっても改訂版によっては内容が異なっているので、その違いが分かるように一覧表にして表わされている。

SITCの1桁レベル、2桁レベルにおけるSITC-R1とSITC-R2、SITC-R2とSITC-R3の対応関係がそれぞれ第2部の表2の「SITC-R1:SITC-R2」と「SITC-R2:SITC-R3」に示されている。また、SITCの3桁レベルの対応関係は第2部の表3に示されているが、この表には個別分類コードに割振られる日本および韓国の配分ウエイトも同時に示されている。第2部の表2および第2部の表3では対応関係のモデルを表現するのに必要とされる関連情報が示されている。関連情報の各項目の記号とそれが示す内容は次のように表される。 $G_i(j)$ : グループ

およびサブグループを表し、 $i$ はグループの一連番号、 $j$ はそのサブグループの一連番号である<sup>(注3)</sup>。基本モデルの対応関係ではサブグループは存在しないので、グループ化された $j$ はすべて1となっている。 $Type$ : サブグループの対応関係のタイプを表す。商品グループは、商品グループ $i j: type$ 、 $G_i(j):type$ 、 $G_i-j: type$ と表わされる。 $Type$ は省略することが多い。 $X_A$ : 分類Aの分類コード、 $X_B$ : 分類Bの分類コード、 $X_{A-f}$ : 分類Aの分類コードの頻度、 $X_{B-f}$ : 分類Bの分類コードの頻度を表す。 $X_{A-Q}$ : 分類A内で分類コードを昇順に並べたときの一連番号、 $X_{B-Q}$ : 分類B内で分類コードを昇順に並べたときの一連番号を表す。

SITC-R1とSITC-R2の対応関係の例として表2に商品グループ $G_{39}(1):1$ から $G_{47}(1):4a$ が示されている。これらの商品グループの対応関係を図示すると図2のように表わされる。

図2 SITC-R1 と SITC-R2 の対応関係

SITC-R1		SITC-R2	
Cotton	263	263	Cotton
Jute incl cuttings and waste	264	264	Jute, oth textile bast fibres
Other vegetable fibres	265	265	Vegetable fibre excl cotn jute
:	:	:	:
Natural abrasives	275	277	atural abrasives incl diamond
Other crude minerals	276	278	Other crude minrrals
Inorganic chemicals	513	522	Inorganic elements
Ores, conc of non-fer metals	283	287	Base metal ores, concentr nes
		323	Briquettes, coke and semi-coke
Coal, coke and briquettes	321	322	Coal, lignite and peat

(出所) 第2部の表1と第2部の表3にもとづき筆者作成

表3 対応関係モデルの結果表 (対応関係の個数)

対応関係の タイプ	SITC-R1 と SITC-R2 の対応関係			SITC-R2 と SITC-R3 の対応関係		
	1桁レベル	2桁レベル	3桁レベル	1桁レベル	2桁レベル	3桁レベル
1	2 ( 8.7)	29 ( 26.4)	103 ( 33.2)	0 ( 0.0)	12 ( 5.4)	69 ( 11.6)
2	0 ( 0.0)	0 ( 0.0)	19 ( 6.1)	0 ( 0.0)	0 ( 0.0)	8 ( 1.3)
3	0 ( 0.0)	0 ( 0.0)	2 ( 0.6)	0 ( 0.0)	0 ( 0.0)	8 ( 1.3)
4a	0 ( 0.0)	6 ( 5.5)	49 ( 15.8)	0 ( 0.0)	0 ( 0.0)	8 ( 1.3)
4b	21 ( 91.3)	75 ( 68.2)	137 ( 44.2)	36 (100.0)	211 ( 94.6)	497 ( 83.5)
total	23 (100.0)	110 (100.0)	310 (100.0)	36 (100.0)	213 (100.0)	595 (100.0)

(出所) 筆者作成

(注) 括弧の中の数字はパーセントを表わす。

$G_{39}(1):1$ はタイプ1であり、SITC-R1の個別分類コード263がSITC-R2の263に対応している。 $G_{40}(1):4a$ はタイプ4aであり、SITC-R1には264と265の2つの個別分類コードが所属し、この2つのコードが共にSITC-R2の264に対応している。SITC-R2は2つの個別分類コードが所属しており、SITC-R1の265はもう一方でSITC-R2の265とも対応している。分類AおよびBとそれらの個別分類コードの頻度 $X_{Af}$ と $X_{Bf}$ を参照することにより対応関係を図示することが可能である。個別分類コードの数が多くなければ表から作図するのはそれほど困難ではない。SITCの1,2,3桁レベルをもとにし

た対応関係の基本モデルの概要が表3に示されている。この表によれば、SITC-R1とSITC-R2の対応関係において3桁レベルの対応関係の総数は310であり、そのうちの33.2%にあたる103の分類コードがタイプ1の関係にある。タイプ4は4aと4bを合せて60.0%を占める186である。

SITC-R2とSITC-R3の対応関係において3桁レベルの対応関係の総数は595あり、そのうちの11.62%にあたる69の分類コードがタイプ1の関係にある。タイプ4は4aと4bを合せて87.0%を占める497である。対応関係が複雑になればタイプ4の割合が増加するが、SITC-R1と

SITC-R2の対応関係に比べてSITC-R2とSITC-R3の対応関係が複雑になっているのをこの関係からも確かめることができる。

## 2. 配分ウエイトとその推計方法

グループ内における分類Aの個別分類コードの個数を $m$ 、分類Bの個別分類コードの個数を $n$ とする。本章では分類Aから分類Bへの方  
向を持つ変換を考える。分類Aと分類Bの個別  
分類コードがすべての結合の手でつながっ  
ている対応関係を考えるとこの対応関係は  
多対多の特殊な状態であり、タイプ4bになる。  
この対応関係と配分ウエイトの配分構造は  
図1にまとめられており、 $A_i$ は分類Aの個別分  
類コード、 $x_i$ は個別分類コード $A_i$ に関連する統  
計値を表すものとする。同じように $B_j$ は分類B  
の個別分類コード、 $y_j$ は個別分類コード $B_j$ に  
関連する統計値を表すものとして、観測値とし  
て $A_i$ と $B_j$ からそれぞれ $k$ 個のデータが得られ  
るとすれば、分類Aの統計値 $X$ と分類Bの統計  
値 $Y$ は、

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1k} \\ \vdots & & \vdots \\ x_{m1} & \cdots & x_{mk} \end{bmatrix} \quad Y = \begin{bmatrix} y_{11} & \cdots & y_{1k} \\ \vdots & & \vdots \\ y_{n1} & \cdots & y_{nk} \end{bmatrix}$$

と表わされる。 $A_i$ から $B_j$ への方  
向に対する配分のためのウエイトを、

$$\omega_{ij} \quad i=1 \cdots m, j=1 \cdots n$$

とする。配分の方法は個別分類コード $A_i$ に  
対応する統計値 $x_i$ が $n$ 個の $y_j$ に分割されるよう  
にしておこなわれる。

これを $y_j$ についてまとめたのが次の式であ  
る。すなわち、分類Aの $m$ 個ある要素 $A_1 \cdots A_m$   
におけるそれぞれの最初の観測値 $x_{11} \cdots x_{m1}$   
に分類Bの要素 $B_j$ の最初の観測値 $y_{j1}$ が対応  
しているとするとき、図1に影を付けて示さ  
れているように、

$$y_{j1} = x_{11}\omega_{1j} + \cdots + x_{m1}\omega_{mj} + u_{j1}$$

と表すことができる。分類Bは $n$ 個の要素から  
成っており、 $j=1 \cdots n$ なので、これを行列表  
示でまとめると、

$$\begin{bmatrix} y_{11} \\ \vdots \\ y_{n1} \end{bmatrix} = \begin{bmatrix} \omega_{11} & \cdots & \omega_{m1} \\ \vdots & & \vdots \\ \omega_{1n} & \cdots & \omega_{mn} \end{bmatrix} \begin{bmatrix} x_{11} \\ \vdots \\ x_{m1} \end{bmatrix} + \begin{bmatrix} u_{11} \\ \vdots \\ u_{n1} \end{bmatrix}$$

となる。配分ウエイト行列を $W$ として、さら  
に $\omega_i' = (\omega_{i1} \cdots \omega_{mi})$ とすれば、

$$W = \begin{bmatrix} \omega_{11} & \cdots & \omega_{m1} \\ \vdots & & \vdots \\ \omega_{1n} & \cdots & \omega_{mn} \end{bmatrix} = \begin{bmatrix} \omega_1' \\ \vdots \\ \omega_n' \end{bmatrix}$$

である。 $k$ 個の観測値を考慮し、同じように行  
列として $U$ も定義すれば、分類A、分類Bおよ  
び配分ウエイト $W$ の関係は、

$$(1) \quad Y = WX + U$$

と表わすことができる。また、すべての要素  
が1からなる $m$ 次元のベクトルを $l_m$ とすると、

(1) 式において配分ウエイトの条件から、

$$W'l_n = l_m$$

が満たされる。(1) 式はニューラル・ネット  
ワークをモデル化するとき用いられる表  
現形式である。

通常の線形回帰モデルの形式は(1) 式の  
ように表わさないことが多い。通常の線形回  
帰モデルの形式に表現しなおすために、 $j=1$   
 $\cdots n$ に対して $y_i' = (y_{i1} \cdots y_{ik})$ とする。行列 $Y$   
は、

$$Y = \begin{bmatrix} y_{11} & \cdots & y_{1k} \\ \vdots & & \vdots \\ y_{n1} & \cdots & y_{nk} \end{bmatrix} = \begin{bmatrix} y_1' \\ \vdots \\ y_n' \end{bmatrix}$$

となる。同じように行列 $U$ も $u_1 \cdots u_n$ で表わす  
ことにする。(1) 式を転置すれば、

$$Y' = X'W' + U'$$

となり、この式は、

$$[y_1 \cdots y_n] = X'[\omega_1 \cdots \omega_n] + [u_1 \cdots u_n]$$

なので、各要素をまとめると、

$$(2) \quad y_i = X'\omega_i + u_i \quad i=1 \cdots n$$

と表わされる。(2) 式をまとめて表わせば、

$$\begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix} = \begin{bmatrix} X' & & \\ & \ddots & \\ & & X' \end{bmatrix} \begin{bmatrix} \omega_1 \\ \vdots \\ \omega_m \end{bmatrix} + \begin{bmatrix} u_1 \\ \vdots \\ u_n \end{bmatrix}$$

となる。さらに、ベクトル $y$ を $y'=(y_1' \cdots y_n')$ 、行列 $X^*$ を、

$$X^* = \begin{bmatrix} X' & & \\ & \ddots & \\ & & X' \end{bmatrix}$$

ベクトル $\omega$ を $\omega'=(\omega_1' \cdots \omega_m')$ 、ベクトル $u$ を $u'=(u_1' \cdots u_n')$ とする。これをまとめると、

$$(3) \quad y = X^* \omega + u$$

となる。

$m$ 次元の単位行列を $I_m$ として、 $I_m$ を $n$ 個横に並べた行列を $C=(I_m \cdots I_m)$ とする。ベクトル $\omega$ はウェイトを表わしているのので、

$$(4) \quad C\omega = l_m$$

となる。すなわち、 $\omega_1 \cdots + \omega_n = l_m$ であり、

$$\begin{bmatrix} \omega_{11} \\ \vdots \\ \omega_{m1} \end{bmatrix} + \cdots + \begin{bmatrix} \omega_{1n} \\ \vdots \\ \omega_{mn} \end{bmatrix} = \begin{bmatrix} 1 \\ \vdots \\ 1 \end{bmatrix}$$

と表現される。

商品分類の対応関係では商品グループ内に対応関係がないウェイトが存在するのが一般的である。すなわち、 $\omega_i$ の要素の中に0となるものが含まれるということである。

(3)式および(4)式は $\omega_i$ の要素の中に0含まれていないという前提で作られているが、この部分に0も含むように設定を変更する必要がある。 $\omega_i$ においてその要素が0であるものを取り除き、0でない要素のみからなるベクトルをつくる。 $\omega_i$ の0である要素を $\omega_{ij}$ とすると、 $m$ 次元の単位行列から $j$ 行を取り除いた行列を $D_i$ とする。取り除く $j$ は複数個あってもかまわない。0を含まないベクトルは1次変換により、

$$\omega_i^D = D_i \omega \quad i=1 \cdots m$$

として得ることができる。すべてが0でない

要素からなる $\omega_i^D$ では $D_i = I_m$ であるので、

$$\omega^D = (\omega_1^{D_1} \cdots \omega_m^{D_m})$$

と、

$$D = \begin{bmatrix} D_1 & & \\ & \ddots & \\ & & D_m \end{bmatrix}$$

とすれば、

$$\omega^D = D\omega$$

となる。ベクトル $\omega$ に対応する観測値行列 $X^*$ および行列 $C$ に対しても同じような操作が必要となる。すなわち、 $\omega$ から取り除く要素に対応する $X^*$ のすべての列を取り除かなければならない。それは、 $X^{*D} = X^* D'$ と $C^D = C D'$ により処理できる。なお、 $C^D$ に関してはすべてが0である要素からなる行は削除する。

ここで改めて、 $\omega^D, X^{*D}, C^D$ をそれぞれ $\omega, X, C$ とする。(4)式を満足するという制約条件のもとで(2)式の残差平方和 $u'u$ を最小にする $\omega$ の値 $\hat{\omega}$ を求めるためにラグランジュ乗数法を利用する。ベクトルを $\hat{e}$ とおいて、スカラー $s$ を次のようにする。

$$(5) \quad \begin{aligned} s &= u'u + \lambda'(C\omega - l_m) \\ &= (y - X\omega)'(y - X\omega) + \lambda'(C\omega - l_m) \end{aligned}$$

この式を $\omega$ に関して偏微分して0とおく。

$$\partial s / \partial \omega = -2X'(y - X\omega) + C'\lambda$$

この式より、 $X'X$ 行列が正則行列であれば、 $\partial s / \partial \omega = 0$ を満足する $\hat{\omega}$ が得られる。

$$(6) \quad \hat{\omega} = \hat{\omega} - (X'X)^{-1} C'\lambda$$

ここで、 $\hat{\omega}$ は(5)式において(4)式の制約条件がないときに得られる $\omega$ の最小2乗法推定量である。

$$(7) \quad \hat{\omega} = (X'X)^{-1} X'y$$

さらに、 $\partial s / \partial \hat{e} = 0$ を満足する $\tilde{\omega}$ は、

$$(8) \quad C\tilde{\omega} - l_m = 0$$

も同時に満足する必要がある。これは(4)式の制約条件である。(6)式の両辺に右から $C$ をかけ、(8)式を利用すれば、



表4 数量単位のコードとその内容

B0 square metres	M3 1,000 metres	U0 KWH (killo watt hour)
B3 1,000 squre metres	M6 million metres	U3 1,000 KWH
B6 million square metres	N0 number	U6 million KWH
C0 carats	N3 1,000 number	V0 metre cube
G0 gramme	P0 pairs	V3 1,000 metre cube
K0 killogramme	P3 1,000 pairs	W0 metric tons
L0 litres	P6 million pairs	W3 1,000 metric tons
MO metres	S0 sets	X3 1,000 boxes

(出所) 『AIDXT (UN・OECD貿易統計検索システム) II コード表』 (経済協力基礎指標 No.-029(57-4) アジア経済研究所 1983) にもとづき筆者作成

(注) 同コード表には上記の数量単位以外にもコードが存在するが、本章で必要な数量単位のみを示す。

$$C\tilde{\omega} = C\hat{\omega} - C(X'X)^{-1}C'\lambda \\ = l_m$$

となる。この式から  $\tilde{\omega}$  を求めて、

$$\lambda = [C(X'X)^{-1}C']^{-1}(C\hat{\omega} - l_m)$$

を得る。この式を (6) 式に代入すれば、制約条件付きの最小2乗推定量、

(9)

$$\tilde{\omega} = \hat{\omega} - (X'X)^{-1}C'[C(X'X)^{-1}C']^{-1} \cdot \\ (C\hat{\omega} - l_m) \\ = \{I - (X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}C\}\hat{\omega} \\ + (X'X)^{-1}C'[C(X'X)^{-1}C']^{-1}l_m$$

が得られる。ここで、 $\hat{\omega}$  は (7) 式より計算される<sup>(注4)</sup>。

本書では対象とする日本および韓国の貿易統計の推計にあたっては、配分ウエイトの構造は基本的には (3) 式および (4) 式に従っていると仮定するが、結果の安定性を考慮して推計には (9) 式を適用せずに第2章で説明されるニューラル・ネットワークによるデータマイニングの方法を用いて (1) 式の  $W$  を直接推計する。

### 3. SITC各系列の貿易統計の推計

一般に貿易統計データは6個の分類カテゴリーである報告国、輸出入区分、商品分類、相手国、年、数量単位と2個の統計値の金額および数量から構成される。本章で使用される貿易統計データはアジア経済研究所が所有している世界貿易統計データベースシステム: AID-XT (Ajiken Indicators on Developing Economies: Extended for Trade Statistics) を基礎データとしており、日本はOECD作成の貿易統計から、韓国はUN作成の貿易統計からそれぞれ変換して作成されたものである<sup>(注5)</sup>。この基礎データは報告国および相手国には6桁で表される「アジア統一国コード」を使用しており、第2部の表6から必要な国および関税地域のコードが見つけれられる。本章で対象としている報告国の日本のアジア統一国コードは117310、韓国は118140である。同じようにして相手国も第2部の表6から得られる。本章ではこの基礎データの1962年から99年までの期間を対象としており、この間には相手国の分離、統合、新生等によりさまざまな変遷

表5 貿易統計の変換プロセス

SITCの種類 (系列) \ 年	62 ... 77	78 ... 87	88 ... 99
日本			
SITC-R1 (import & export)	$\boxed{R1}$	( R2 → R1 )	( R2* → R3 )
SITC-R2 (import & export)	( R1 → R2 )	$\boxed{R2}$	( R3 → R2 )
SITC-R3 (import & export)	( R2* → R3 )	( R2 → R3 )	$\boxed{R3}$
韓国			
SITC-R1 (import & export)	$\boxed{R1}$	( R2 → R1 )	-
SITC-R2 (import & export)	( R1 → R2 )	$\boxed{R2}$	-
SITC-R3 (import & export)	-	-	$\boxed{R3}$

(出所) 著者作成

を辿る国がある。国の変遷を長期時系列で捉えようとすればかなり厄介な処理が必要になるが、国コードの利用については第4章「貿易データベースにおける国コードの利用」で説明される。がある。国の変遷を長期時系列で捉えようとすればかなり厄介な処理が必要になるが、国コードの利用については第4章「貿易データベースにおける国コードの利用」で説明される。

輸出入区分については輸入を1、輸出を2で表す。年は西暦年の下2桁の数字で表される。数量単位は「アジア統一数量単位」を使用しており、数量単位コードとその名称が表4に示されている。金額は1,000US\$を単位として、数量は数量単位にもとづいて表示される<sup>(注6)</sup>。本節では対象とする分類カテゴリーは商品分類であり、SITCのそれぞれの改訂版をもとにした3桁レベルにおける商品分類に従う長期時系列データを作成することを目的とする。すなわち、本書の第2章「SITC3桁分類コード変換のための配分ウェイト推計—ニューラ

ル・ネットワークを用いて—」のニューラル・ネットワークを使用して得られた配分ウェイトにもとづいて日本および韓国の貿易統計をSITCの系列ごとにその3桁レベルで推計することである。配分ウェイトの推計値は第2章の表3に示されている。この配分ウェイトにもとづいて作成されたSITCの系列を*NN*で表す。一般に配分のためのウェイト情報が事前情報として得られないとき等配分することで暫定的なウェイトを設定することが多い。等配分のウェイトは事前情報なしのエントロピーの最大化に根拠を置いており、ウェイトに一様分布が適用されるからである<sup>(注7)</sup>。等配分のウェイトにより作成された同系列を*ED*で表すことにする。

まず最初、日本に関するSITC-R1の輸出入について*NN*および*ED*のそれぞれの系列を推計する。表5に示されているように日本ではSITC-R1の系列は1962年から77年までの範囲は原系列が存在するのでこれを使用する。原系列は囲って示されている。1978年から87年まで

表6 日本における SITC3 桁レベル以上で不整合のある系列

(1) SITC-R1 と SITC-R2 の対応関係に不整合のある系列

---

SITC-R1 (import) :
51, 025, 053, 266, 275, 697
SITC-R1 (export) :
8, 27, 29, 43, 55, 59, 89, 273, 276, 331, 431, 551, 599, 692, 697, 714, 717, 729, 733, 894
SITC-R2 (import) :
26, 63, 77, 88, 034, 035, 036, 056, 058, 232, 266, 268, 277, 322, 323, 335, 522, 551, 591, 592, 628, 634, 654, 659, 712, 713, 714, 716, 721, 726, 751, 761, 771, 772, 783, 785, 884, 898
SITC-R2 (export) :
05, 55, 59, 74, 75, 77, 035, 071, 246, 273, 335, 513, 522, 551, 582, 583, 591, 655, 693, 712, 724, 726, 727, 737, 744, 749, 751, 752, 759, 762, 764, 771, 772, 776, 783, 785, 842, 847, 874, 884
SITC-R3 (import) :
23, 41, 63, 77, 017, 035, 036, 059, 062, 231, 264, 266, 277, 325, 335, 411, 515, 522, 551, 572, 597, 598, 635, 655, 657, 686, 712, 721, 726, 747, 748, 751, 752, 761, 764, 772, 775, 776, 781, 783, 791, 844, 872, 873, 891, 898
SITC-R3 (export) :
03, 07, 41, 43, 55, 57, 59, 74, 77, 78, 87, 016, 035, 037, 071, 246, 273, 291, 411, 431, 522, 533, 551, 554, 573, 575, 579, 581, 593, 597, 598, 625, 659, 693, 713, 724, 737, 744, 747, 751, 759, 762, 764, 772, 773, 776, 781, 783, 784, 791, 841, 844, 872, 884, 891

---

(2) SITC-R2 と SITC-R3 の対応関係に不整合のある系列

---

SITC-R1 (import) :
1, 6, 7, 08, 09, 11, 25, 52, 56, 58, 59, 62, 65, 66, 67, 68, 69, 71, 73, 86, 94, 013, 022, 023, 024, 048, 061, 062, 073, 074, 081, 099, 211, 212, 251, 261, 271, 283, 292, 514, 521, 561, 581, 599, 621, 629, 632, 633, 652, 653, 654, 655, 661, 662, 663, 664, 666, 667, 671, 674, 675, 676, 678, 679, 682, 683, 684, 688, 691, 693, 696, 698, 718, 719, 725, 892, 941
SITC-R1 (export) :
9, 01, 07, 08, 11, 22, 24, 53, 63, 68, 69, 81, 83, 93, 95, 011, 013, 047, 048, 051, 054, 073, 074, 081, 111, 221, 243, 244, 292, 515, 533, 553, 554, 611, 613, 631, 632, 633, 654, 655, 664, 672, 673, 674, 675, 679, 683, 685, 691, 696, 698, 712, 723, 734, 812, 831, 863, 895, 899, 931, 951
SITC-R2 (import) :
1, 6, 7, 01, 08, 09, 11, 25, 28, 29, 41, 43, 56, 58, 62, 65, 66, 67, 68, 69, 72, 74, 87, 94, 95, 014, 022, 023, 024, 048, 061, 062, 073, 074, 081, 098, 211, 212, 251, 261, 271, 287, 292, 411, 424, 431, 516, 523, 562, 582, 583, 584, 621, 625, 633, 635, 652, 656, 658, 661, 662, 663, 664, 666, 667, 671, 674, 675, 676, 678, 679, 682, 683, 684, 688, 691, 692, 693, 696, 699, 723, 728, 741, 743, 744, 749, 759, 778, 786, 842, 881, 892, 899, 941, 951
SITC-R2 (export) :
8, 9, 07, 08, 09, 11, 22, 24, 52, 53, 63, 68, 69, 83, 93, 011, 034, 037, 047, 048, 054, 057, 073, 074, 081, 098, 111, 222, 233, 244, 248, 292, 524, 533, 553, 554, 584, 611, 613, 633, 635, 656, 657, 664, 672, 673, 674, 675, 679, 683, 685, 691, 696, 699, 714, 742, 743, 773, 781, 791, 792, 831, 843, 883, 895, 931
SITC-R3 (import) :
1, 4, 6, 7, 02, 05, 08, 11, 25, 28, 42, 56, 57, 66, 67, 68, 69, 72, 74, 75, 85, 87, 023, 024, 048, 054, 058, 061, 073, 081, 091, 211, 212, 251, 261, 268, 283, 285, 292, 342, 343, 344, 422, 513, 524, 531, 532, 542, 553, 562, 575, 579, 592, 612, 621, 625, 633, 652, 653, 654, 659, 661, 662, 663, 664, 665, 666, 667, 671, 675, 676, 677, 678, 679, 682, 683, 684, 691, 693, 694, 696, 697, 699, 714, 718, 723, 724, 727, 728, 735, 743, 744, 749, 784, 786, 841, 842, 848, 851, 881, 895, 899
SITC-R3 (export) :
2, 9, 02, 05, 09, 11, 22, 23, 58, 61, 62, 79, 81, 82, 83, 85, 89, 93, 022, 036, 047, 054, 056, 057, 074, 098, 111, 222, 223, 232, 244, 285, 287, 422, 523, 525, 531, 532, 583, 611, 612, 613, 633, 634, 642, 652, 654, 661, 663, 664, 666, 671, 672, 673, 674, 675, 676, 678, 683, 685, 692, 694, 696, 697, 699, 714, 726, 735, 742, 745, 763, 774, 775, 785, 792, 813, 821, 831, 845, 846, 851, 883, 892, 893, 894, 895, 896, 899, 931

---

(3) 両方の対応関係に不整合のある系列

SITC-R1 (import) :

29, 43, 95, 284, 291, 431, 513, 861, 951

SITC-R1 (export) :

86, 052, 055, 262, 291, 621, 656, 842, 861, 893

SITC-R2 (import) :

75, 223, 233, 246, 288, 291, 585, 653, 655, 657, 718, 737, 745, 782, 872

SITC-R2 (export) :

01, 81, 88, 95, 014, 036, 056, 232, 268, 585, 621, 628, 654, 658, 692, 711, 718, 721, 722, 723, 778, 782, 786, 812, 844, 846, 848, 871, 873, 881, 893, 899, 951

SITC-R3 (import) :

26, 29, 59, 223, 246, 288, 291, 322, 516, 554, 571, 573, 574, 629, 673, 711, 737, 741, 745, 746, 759, 763, 778, 782, 871, 874

SITC-R3 (export) :

01, 29, 63, 66, 69, 88, 012, 017, 025, 034, 058, 059, 075, 231, 268, 292, 513, 514, 516, 553, 571, 572, 582, 592, 629, 635, 653, 655, 656, 658, 721, 722, 723, 727, 731, 746, 748, 749, 778, 782, 786, 811, 812, 848, 871, 873, 881, 897

(出所) 第2部の表5にもとづき著者作成

はSITC-R1の系列は存在せず、原系列としてSITC-R2が存在するのでSITC-R2からSITC-R1へ  
の方向を持つ配分ウェイトを利用してSITC-R1の系列を推計する。矢印は変換する方向を示す。1988年から99年まではSITC-R1の系列は存在せず、原系列としてSITC-R3が存在するので、最初はSITC-R3からSITC-R2を推計する。続いて、この得られたSITC-R2の系列をSITC-R1へと変換する。この変換された系列を集めて1962年から99年までの期間におけるSITC-R1の系列の時系列データが得られる。

つぎに、SITC-R2の輸出入についてNNおよびEDのそれぞれの系列を推計する。表4に示されているように日本ではSITC-R2の系列は1962年から77年までの範囲は原系列が存在しないので原系列であるSITC-R1をSITC-R1からSITC-R2へ  
の方向を持つ配分ウェイトを利用してSITC-R2の系列を推計する。また、1988年から99年までは原系列としてSITC-R3が存在するので、SITC-R3からSITC-R2を推計する。

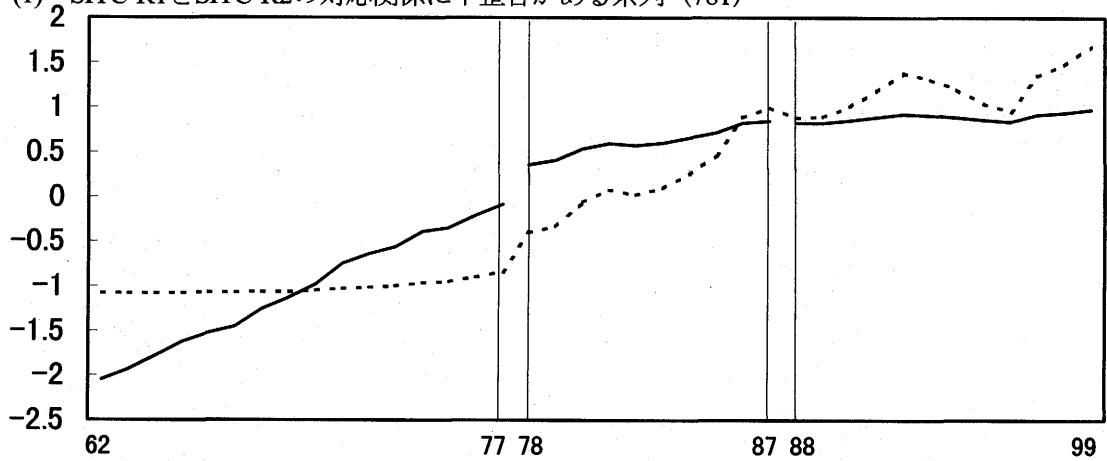
この変換された系列を集めて1962年から99年まで期間におけるSITC-R2の系列の時系列データが得られる。SITC-R3の系列も同じようにしてSITC-R1およびSITC-R2から変換して求めることができる。このようにして、SITC-R1、SITC-R2、SITC-R3の輸出入の各系列が求められる。

韓国についてはSITC-R2とSITC-R3の対応関係における最大規模となる商品グループ61:  $G_{61}(1):4b$ の配分ウェイトが推計できなかったため、1962年から99年までの期間を対象とした時系列データは作成されない。代わりに、1962年から87年までのSITC-R1における輸出入のNNおよびEDの系列と同じくSITC-R2の各系列が作成されるが、SITC-R3は原系列のまま維持されている。表5には韓国の系列も示されている。

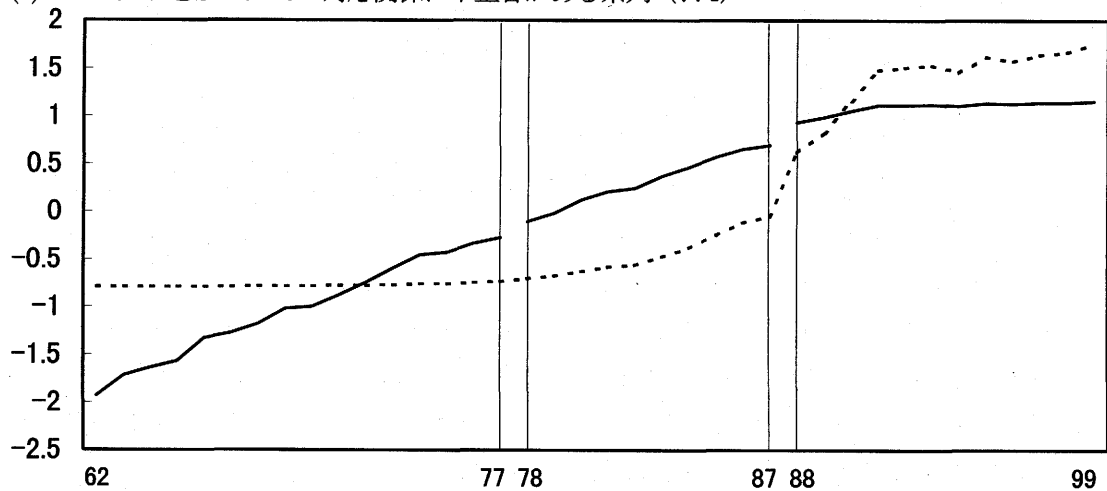
SITC-R2の系列は1回の変換で求めるものが得られるが、SITC-R1とSITC-R3のそれぞれの系列については2回の変換が必要なところが

図3 対応関係に不整合のある系列 (SITC-R3日本の輸出系列の例)

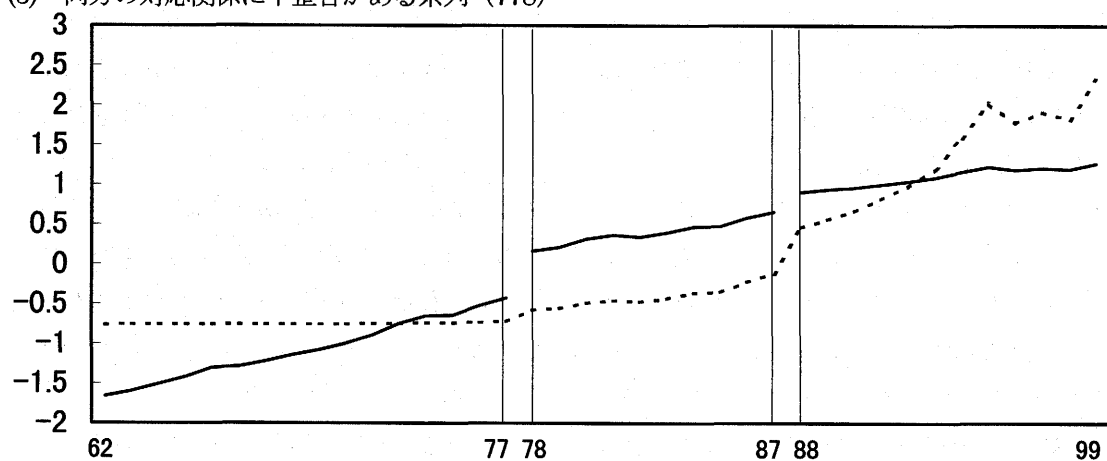
(1) SITC-R1とSITC-R2の対応関係に不整合がある系列 (781)



(2) SITC-R2とSITC-R3の対応関係に不整合がある系列 (774)



(3) 両方の対応関係に不整合がある系列 (778)



(出所) 筆者作成

(注) 破線は原系列、実線は原系列を対数変換した系列である。同一のグラフに表示するため、両系列は標準化されている。

出てくる。しかも、配分ウェイトの推計は対応関係のSITC-R1とSITC-R2においては改訂年の1978年の近傍、対応関係のSITC-R2とSITC-R3では改訂年の88年近傍のデータを使用している。すなわち、1988年から99年を範囲とするSITC-R1の系列の推計には78年前後のデータをもとにして推計された配分ウェイトを基礎として計算されているということである。SITC-R3の系列についても同じことがいえる。そのため、推計された貿易統計の利用にあたっては商品分類の改訂時期に生じた不整合が問題になる。この不整合性に関しては相手国を世界計に限定はしているものの、本書の第3章「商品分類の改訂に伴う貿易統計の整合性評価」において説明される。

#### おわりに

貿易統計を時系列データとして利用するときには分類の改訂のおこなわれた前後は商品に対する定義およびそのカバレッジが必ずしも同じではなく、改訂年を含む連続年では取引金額や数量は必ずしも整合性のある状態とはいえない。そのため、商品分類の改訂年の前後における対応関係にもとづいて配分ウェイトを推計し、このウェイトでそれぞれの分類コードに対応する統計値を再配分することが必要になる。

配分ウェイトの推計にあたっては、第2章で説明されるニューラル・ネットワークによるデータマイニングの方法によりSITC3桁レベルで作成された商品グループごとに計算している。本章では日本および韓国を対象として、この得られた配分ウェイトにもとづいてSITC-R1、SITC-R2、SITC-R3における1962年から99年までの期間の輸出入のNNおよびED系列の時系列データを作成し、貿易統計を推計した。作成されたSITCの各系列で表された

貿易統計は商品分類の改訂年の前後において必ずしも整合性のある状態で接続されていないため、第3章「商品分類の改訂に伴う貿易統計の変換評価」ではその整合性を検討している。その結果によれば作成された時系列は整合あるいは不整合のタイプに分けられ、(1) 整合性のある系列、(2) SITC-R1とSITC-R2の対応関係に不整合がある系列、(3) SITC-R2とSITC-R3の対応関係に不整合がある系列、(4) その両者に不整合がある系列、となる。これらの3つの不整合性のタイプは表6にまとめて示されている。また、それぞれのタイプの中の典型的な系列を図3により表示している。

配分ウェイトの本来の課題は分類コードが改訂されたときに対応している統計量をできるだけなめらかにその改訂時点でつなげることである。今回の配分ウェイトの推計にあたってはデータの制約上かなり厳しい仮説を設定しており、それが改訂年に構造変化を生じさせた大きな原因ではないかと考えられる。したがって、今度の課題はできるだけ仮説を緩め、実態に合ったモデルを設定することがあげられる。さらに、改訂年に構造変化が生じている分類コードについては個別にその対応関係を検討することが必要である。

---

(注1) 一般に分類はカテゴリーと呼ばれる抽象的な個別主体を要素とする集合で表すことができ、2つの分類間のカテゴリーの対応関係はこれらの分類間のカテゴリーを結びつける関数によって表された分類規則で表すことができる。このような分類規則によって定まる分類間の対応関係についての考え方は基本的には佐藤の*Fundamental Concept of Social/Regional Summary Data and Inference in Their Database* およびその要約である「要約

データの基礎概念とデータベース内での推論—世界貿易統計データベースを例として—の中にあるデータの記述対象の分類方法に関するデータの導出およびデータ間の比較にもとづいて考えることができる。野田、山本の「体型の異なる分類の対応関係と変換—グループ化および切断による商品分類の変換の試み—」において商品分類の対応関係にこの考えを取り入れて商品グループという考えを示している。また、野田の「商品分類の改訂に伴う対応関係の連結」はこの考えを貿易統計における商品分類体系の改訂に応用してより具体的な形で例示したものである。

(注2) 野田、山本の「体型の異なる分類の対応関係と変換—グループ化および切断による商品分類の変換の試み—」では注釈で「…説明は省略するが、タイプ4には配分ウエイトが代数方程式により一意的な解を持つタイプ4aと一意的に解を持たないタイプ4bに分けることができる。」(p79)と説明している。

(注3) ここでいうサブグループとは厳密な意味におけるグループ内の分割ではない。サブグループ間に互いに少数の共通の分類コードが含まれていてもこの分類コードを取り除けばグループ化されるという状態であり、比較的類似のものがまとまっているという意味でのゆるやかな対応関係の集まりをさす。したがって、サブグループは一意的には決まらない。商品分類SITC-R2からSITC-R3への変更における分類基準の変更および新設による複雑な絡み合いにより、基本モデルは対応関係がタイプ4に属する関係を多く生み出している。商品グループのタイプがグループ4の中で対応関係の要素数が大きなものについては切断によりサブグループ化をおこないサブグループの特性を明確にする必要がある。基本モデルからこのような切断によって切断の要素と商品サブグループを持つ切断モデルGRT<sub>23</sub>[IDE]がアジア経済研究所作成の切断モデルである。基本モデルには対応関係が2,766個の商品グループG<sub>888</sub>があり、このグループはSITC-R2の1桁は0から9までのすべてのセクションとSITC-R3の1桁は3と9を除くすべてのセクションの対応関係から構成され

ている。しかも、いろいろなものが混在しているため、特性を出すためのサブグループ化にはかなり大胆な切断が必要とされる。切断モデルでは結果的には繊維系や織物類はある程度原料別にサブグループ化されており、衣類、プラスチック関係品、ゴム製品などは1つにサブグループ化されている。商品グループの対応関係の要素数が多くてもそのグループの特性が明確なグループの衣類や鉄鋼製品はあえて切断はおこなっていない。

(注4) 一般に商品グループ内では対応関係がないウエイトが存在する。0となるウエイトの状態により対応関係のタイプが決まる。簡単のために $m=n=2$ として配分ウエイトの計算例を示す。対応関係のタイプ4bにおいて対応関係のウエイトベクトル $\omega$ は、

$$\omega = (\omega_1, \omega_2)' \\ = \begin{bmatrix} \omega_{11} \\ \omega_{21} \\ \omega_{12} \\ \omega_{22} \end{bmatrix}$$

である。 $\omega_i$ において0となる要素がないので $D_i = I_2$ となる。 $X^* = \begin{bmatrix} x_1 & x_2 \\ 0 & 0 \end{bmatrix}$ 、 $O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ とするとき行列 $X$ は、

$$X = \begin{bmatrix} X^* & O \\ O & X^* \end{bmatrix} = \begin{bmatrix} x_1 & x_2 & 0 & 0 \\ 0 & 0 & x_1 & x_2 \end{bmatrix}$$

と表わされ、行列 $C$ は、

$$C = (I_2 \quad I_2) = \begin{bmatrix} 1 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \end{bmatrix}$$

となる。(9)式にもとづいて計算すれば求める解が得られる。 $\hat{\omega}$ は(7)式で計算される。対応関係のタイプ4aはタイプ4bにおいて $\omega_1$ における2番目の要素の $\omega_{21}$ が0となった関係である。したがって、2次元の単位行列から2番目の行を取り除いた行列は $D_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$ となる。 $\omega_2$ は0の要素がないので、 $D_2 = I$ である。ベクトル $\omega$ は、

$$\omega^D = \begin{bmatrix} D_1 & \\ & D_2 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \begin{bmatrix} D_1 \omega_1 \\ \omega_2 \end{bmatrix} \\ = \begin{bmatrix} \omega_{11} \\ \omega_{12} \\ \omega_{22} \end{bmatrix}$$

と変更される。行列 $X$ についても同じように変更が必要になる。

$$\begin{aligned} X^D &= XD' \\ &= \begin{bmatrix} X^* & \\ & X^* \end{bmatrix} \begin{bmatrix} D_1' & \\ & I_2 \end{bmatrix} = \begin{bmatrix} X^* D_1' & \\ & X^* \end{bmatrix} \\ &= \begin{bmatrix} x_1 & \\ & X^* \end{bmatrix} = \begin{bmatrix} x_1 & 0 & 0 \\ 0 & x_1 & x_2 \end{bmatrix} \end{aligned}$$

となる。同じようにして、

$$C^D = CD' = \begin{bmatrix} 1 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

となる。最後に $\omega^D, X^D, C^D$ を改めて $\omega, X, C$ と置き直す必要がある。(9)式にもとづいて計算すれば、

$$\begin{bmatrix} \tilde{\omega}_{11} \\ \tilde{\omega}_{21} \\ \tilde{\omega}_{22} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 & -a \\ -1 & 1 & a \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} \hat{\omega}_{11} \\ \hat{\omega}_{21} \\ \hat{\omega}_{22} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1+a \\ 1-a \\ 2 \end{bmatrix}$$

が得られる。ここで、 $a = (x_1' x_1)^{-1} x_2' x_2$ であり、 $\hat{\omega}$ は(7)式で計算される。対応関係のタイプ3はタイプ4bにおいて $\hat{u}_2$ の2つの要素が0となった関係である。2次元の単位行列から1、2番目の行を取り除くと行列 $D_2$ の要素はなくなり存在しなくなる。したがって、 $D = (D_1 \ 0)$ となる。ここで、

$$O = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$$

とする。 $\hat{u}_1$ は0の要素がないので、 $D_1 = I$ である。ベクトル $\omega$ は、

$$\omega^D = D\omega = \begin{bmatrix} I_2 & O \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \omega_1 = \begin{bmatrix} \omega_{11} \\ \omega_{21} \end{bmatrix}$$

と変更される。行列 $X$ は、

$$X^D = XD' = \begin{bmatrix} X^* & \\ & X^* \end{bmatrix} \begin{bmatrix} I_2 \\ O \end{bmatrix} = X^* = (x_1 \ x_2)$$

となり、行列 $C$ は、

$$C^D = CD' = \begin{bmatrix} I_2 & I_2 \\ & O \end{bmatrix} = I_2$$

となる。 $\omega^D, X^D, C^D$ を改めて $\omega, X, C$ と置き直した後、(9)式にもとづいて計算すれば、

$$\begin{bmatrix} \tilde{\omega}_{11} \\ \tilde{\omega}_{21} \end{bmatrix} = \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

が得られる。対応関係のタイプ2はタイプ4bにおい

て $\omega_1$ の2番目の要素の $\omega_{21}, \hat{u}_2$ の2番目の要素の $\hat{u}_{22}$ が共に0となった関係である。 $i=1,2$ に対して2次元の単位行列から2番目の行を取り除きいた行列を $D_i = (1 \ 0)$ とする。ベクトル $W$ は、

$$W^D = DW = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} \omega_1 \\ \omega_2 \end{bmatrix} = \begin{bmatrix} \omega_{11} \\ \omega_{12} \end{bmatrix}$$

行列 $X$ は、

$$\begin{aligned} X^D &= XD' \\ &= \begin{bmatrix} X^* & \\ & X^* \end{bmatrix} \begin{bmatrix} 1 & 0 \\ 0 & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} X^* \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} \\ X^* \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \end{bmatrix} \\ &= \begin{bmatrix} x_1 & 0 \\ 0 & x_1 \end{bmatrix} \end{aligned}$$

行列 $C$ は、

$$\begin{aligned} C^D &= CD' = \begin{bmatrix} I_2 & I_2 \\ & O \end{bmatrix} D' \\ &= \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ 0 & 0 \end{bmatrix} \end{aligned}$$

である。 $C$ の2行目はすべて0の要素なので2行目を削除して、 $C^D = [1 \ 1]$ とおく。最後に $\omega^D, X^D, C^D$ を改めて $\omega, X, C$ と置き直した後、(9)式にもとづいて計算する。標本数を $k$ とすれば、タイプ2のときは、 $x_1' = (x_{11} \cdots x_{1k}) = (1 \cdots 1)$ となり、

$$\begin{bmatrix} \tilde{\omega}_{11} \\ \tilde{\omega}_{12} \end{bmatrix} = \frac{1}{2} \begin{bmatrix} 1 & -1 \\ -1 & 1 \end{bmatrix} \begin{bmatrix} \hat{\omega}_{11} \\ \hat{\omega}_{12} \end{bmatrix} + \frac{1}{2} \begin{bmatrix} 1 \\ 1 \end{bmatrix}$$

となる。 $\hat{\omega}$ は(7)式で計算され、 $\hat{\omega}' = (\bar{y}_1, \bar{y}_2)$ となる。ここで、 $\bar{y}_i = (y_{i1} + \cdots + y_{ik})/k$ ,  $i=1,2$ である。これを $\tilde{\omega}$ へ代入すれば、 $\tilde{\omega}' = (\bar{y}_1, \bar{y}_2)$ となる。すなわち、タイプ2のときの配分ウエイトは分類 $B$ の統計値 $y$ の平均で得られる。

(注5) アジア経済研究所作成による世界貿易統計データベースシステム(AID-XT)はOECD加盟国についてはOECD作成の貿易統計、OECD加盟国を除いたUN加盟国はUN作成の貿易統計、台湾は台湾作成の貿易統計から構成され、1962年から99年までの38年間のデータがあり、年平均80カ国の貿易データが利用可能である(2001年2月現在)。UN貿易統計では71年以降報告されなくなった台湾については台湾から入手した統計書あるいは磁気テープを当研究所でUN貿易統計に準拠した形



式・内容に変換して国際比較可能にしている。貿易統計の主要な分類カテゴリーとして報告国、輸出入区分、商品分類、相手国、年、数量単位があり、統計値は金額と数量がある。商品分類はUN貿易統計およびOECD貿易統計ともに共通にUN作成のSITCあるいは最近では関税協力理事会作成のHS (Harmonized Commodity Description and Coding System) であるが、当該国の商品分類がSITCあるいはHSのどの改訂版を採用しているかに依存するため、利用にあたっては個別の対応が必要である。国と数量単位についてはUNとOECD貿易統計の定義やカバレッジが必ずしも整合性のある関係ではない。そのため、両者の分類カテゴリーを統一するための工夫がなされ、国については「アジア統一国コード」、数量単位は「アジア数量単位」が使用されている。

(注6) 基礎データは48桁からなる固定長のデータで構成されおり、そのフォーマットは次のとおりである。報告国 (1-6)、輸出入区分 (7)、商品分類 (8-12)、相手国 (13-18)、年 (19-20)、金額 (21-33)、数量単位 (34-35)、数量 (36-48)。ここで、( ) の中の数字はそれぞれの項目のカラムを示す。

(注7) パラメータ空間を $\Theta$ としてその要素を $\theta_1 \dots \theta_m$ とする。事前情報がまったくない場合には $\Theta$ 上の分布、 $\{f(\theta_1), \dots, f(\theta_m)\}$ のエントロピーを最大にする分布が利用される。すなわち、そのような分布は最大の不確実性をもつ分布である。エントロピーは、

$$e\{f(\theta)\} = -\sum_{j=1}^m f(\theta_j) \log f(\theta_j)$$

で与えられる。分布の性質からその和は1となる。

$$\sum_{j=1}^m f(\theta_j) = 1$$

エントロピーの最大化は $\partial e\{f(\theta)\} / \partial f(\theta) = 0$ より得られる。これを計算して、

$$f(\theta_j) = \frac{1}{m} \quad j=1, \dots, m$$

となり、エントロピーを最大化する分布として一様分布が得られる。

## 【参考文献】

- [1] 黒子正人「貿易統計データベースにおける国コードの利用」(野田容助編『商品分類の改訂に伴う貿易統計の変換』統計資料シリーズ No.83 アジア経済研究所 2001)
- [2] 佐藤英人「要約データの基礎概念とデータベース内での推論—世界貿易統計データベースを中心として—」(木下宗七、野田容助編『世界貿易データシステムの整備と利用』統計資料シリーズ No.67 アジア経済研究所 1995)
- [3] 城坂晃正「SITC3桁分類コード変換のための配分ウエイト推計—ニューラル・ネットワークを用いて—」(野田容助編『商品分類の改訂に伴う貿易統計の変換』統計資料シリーズ No.83 アジア経済研究所 2001)
- [4] 野田容助、山本泰子「体型の異なる分類の対応関係と変換—グループ化および切断による商品分類の変換の試み—」(木下宗七、野田容助編『世界貿易統計データシステムの整備と利用』統計資料シリーズ No.67 アジア経済研究所 1995)
- [5] 野田容助「対応関係におけるタイプの識別」(河村鎰男、野田容助共編著『国際産業データシステムの利用と応用』統計資料シリーズ No.69 アジア経済研究所 1996)
- [6] 野田容助「商品分類の改定にともなう対応関係の連結」(古河俊一、野田容助共著『標準国際商品分類と産業分類の対応関係』統計資料シリーズ No.80 アジア経済研究所 1998)
- [7] 野田容助「商品分類の改訂に伴う貿易統計の整合性評価」(野田容助編『商品分類の改訂に伴う貿易統計の変換』統計資料シリーズ No.83 アジア経済研究所 2001)
- [8] 山本泰子「貿易統計における商品の分類」(木下宗七、野田容助編『世界貿易統計データシステムの整備と利用』、統計資料シリーズno.67、アジア経済研究所、1995)
- [9] Hideto Sato, *Fundamental Concepts of Social/Regional Summary Data and Inference in Their Database*, Doctoral thesis, Tokyo University, 1983