

第2章

長期時系列貿易データ作成のための商品分類統一化の方法

野田容助

要約

貿易データを共通の商品分類概念で長期時系列に利用するためには、商品分類の改訂された前後のどちらかの商品分類に統一することが必要である。本章は商品分類の改訂に伴って生ずる商品分類の違いを接続するため、新旧商品分類コードから構成される対応関係コード表の構造と両分類の取引額を考慮して、両者の対応関係の配分ウェイトを推計し、それに基づいて貿易データを変換する方法を示している。本章で試みている変換方法は新商品分類コードが旧商品分類コードの複数個に対応しているときにはその配分構造に従って推計していることに特徴がある。この方法により貿易データの変換にはすべての報告国に対して一律に同一の対応表が適用されるのではなく、報告国ごとに推計された配分ウェイト付きの対応関係コード表の適用が可能となる。

キーワード

商品分類の改訂、対応関係コード表、配分ウェイト、貿易データの変換

はじめに

貿易データを共通の商品分類概念で長期時系列に利用するためには、商品分類の改訂された前後のどちらかの商品分類に統一することが必要である。国連（UN）統計局国際貿易統計部（International Merchandise Trade Statistics Section, United Nations Statistics Division; 以下、UN 統計局と略記）作成による、同 Web サイトより得られる United Nations Commodity

Trade Statistics Database の貿易データ（以下、Comtrade データと略記）では商品分類改訂に合わせて改訂前後で接続できるような新旧商品分類コードにもとづく対応関係コード表を作成している。これらの対応関係コードを参照することにより、同一商品分類による貿易データの利用が可能となる。Comtrade データは商品分類の改訂後の新商品分類コードから旧商品分類コードへの変換がおこなわれており、旧商品分類による貿易データの長期時系列的な利用を可能にしている。

Comtrade データが採用している新商品分類コードから旧商品分類コードへの変換方法は、新旧商品分類コードに基づく対応関係コード表がそのまま利用されているのではなく、この対応表とは別な変換のための対応表が利用されている。この方式は新商品分類コードが旧商品分類コードの複数個に対応しているときには、その中から最適と思われるものを1つだけ選択するという方法を採用している。しかも、この方法は報告国ごとに固有の選択がおこなわれるのではなく、すべての報告国に対して一律に同一の選択がおこなわれている。

本章は商品分類の改訂に伴って生ずる商品分類の違いを接続するため、新旧商品分類コードから構成される対応関係コード表の構造と両分類の取引額を考慮して、両者の対応関係の配分ウェイトを推計することを目的としている。この配分ウェイトをもとに商品分類の SITC-R1 系列の長期時系列貿易データを作成している。すなわち本章で試みている変換方法は Comtrade データの方法とは違って、新商品分類コードが旧商品分類コードの複数個に対応しているときにはその配分構造に従って推計していることに特徴がある。この方法により貿易データの変換にはすべての報告国に対して一律に同一の対応表が適用されるのではなく、報告国ごとに推計された配分ウェイト付きの対応表が適用される。

本章ではまた Comtrade データから得られた報告国日本の輸出貿易データを利用して、SITC-R2 で編集されている 1976 年から 1987 年までの貿易データを同期間の SITC-R1 への変換を試み、推計方法の違いによる取引額の比較および検討もおこなっている。本章は第 1 節として商品分類を変換するために必要な対応関係コード表、第 2 節は商品分類を変換するための配分ウェイト行列の推計、第 3 節は貿易データにおける SITC-R1 への変換、から構成されている。

1. 商品分類を変換するために必要な対応関係コード表

商品分類体系の異なる分類どうしを結び付けるには両者の対応関係を明らかにした対応関係コード表が必要であり、対応関係コード表を使用する場合には 2 つの分類がどのような対応関係にあるかを検討することが重要な課題となる。例えば異なる商品分類の A と B において、貿易データを A から B へと分類変更をおこなう場合には、貿易データを統合

図1 分類AとBにおける詳細分類コードの対応関係から得られた商品グループ

分類A \ 分類B	$a_1 \cdots a_{n_1}$	$a_{n_1+1} \cdots a_{n_1+n_2}$	\cdots	$a_{n_1+\cdots+n_{p-1}+1} \cdots a_n$
b_1 : b_{m_1}	商品グループ 1			
b_{m_1+1} : $b_{m_1+m_2}$		商品グループ 2		
:			⋮	
$b_{m_1+\cdots+m_{p-1}+1}$: b_m				商品グループ p

(出所) 野田容助編『東アジア諸国・地域の貿易指数—作成から応用までの基礎的課題—』(SDS No.88 アジア経済研究所 2005) の第2章の表1を引用

(注) 影の部分は商品グループを表わす。

するかまたは配分するのかわかるような操作が必要となる。この操作は両分類におけるいくつかの分類コードの集まりを対象におこなわれる。すなわち、対応関係コード表の中で分類の核になる閉じた対応関係にある分類コードの集まりを「商品グループ」あるいは簡単に「グループ」ということにする。閉じた対応関係あるいはグループは佐藤 [1985] およびその要約である佐藤 [1995] によれば、2つの分類から「共通に導出可能な最も詳細な分類 (Finest Common Derivative: FCD) に対応する分類コードの集まりである。すなわち、対応関係を構成する分類コードに少なくとも1つの共通した結合があればそれをつないでいき、結合が無くなったところまでの構成要素でグループを決めるという方法で得られた対応関係の集まりである。貿易データの変換に対する操作はグループごとにそこに属する分類コード間でおこなわれる。

分類 G が A と B から得られる FCD に対する分類となる時、 $fcd(A, B) \rightarrow G$ と表わすことにする。本章ではこのようにして得られた分類 G を対応関係のグループと呼ぶ。 A と B の対応関係をグループに分割することを対応関係のグループ化といい、 G はグループ化された分類ともいう。図1は分類 A と B における詳細分類コードの対応関係から得られた商品グループであり、グループが p 個に分割された状態を示している。この図は G を昇順に並べ、さらに A と B を昇順に並べることにより作成される。本章ではそれぞれの商品グループの中で配分ウェイト行列の推計と貿易データへの変換がおこなわれる。例えば商品グループ1では、分類 A における分類コード $a_1 \cdots a_{n_1}$ の取引額が B の $b_1 \cdots b_{m_1}$ へとグループ内

で変換される。

佐藤の FCD による対応関係をグループ化するという考え方を基礎として、野田[2010]は2つの対応関係コード表の連結の方法を拡張して、複数存在する対応関係コード表に対しても共通に存在する分類コードを基礎に対応関係をグループ化し、さらに連結する考えとその方法を紹介している。この改訂版は野田 [2012] である¹。グループ化された対応関係コード表に含まれる任意の分類間の対応関係コード表の作成も可能となっている。

個別主体の集合 X の分類として A_1, A_2, \dots, A_n が複数個存在しているとする。それぞれの分類は X から対応する関数としての分類規則により関連付けられて分類体系を構成しているため、 $i \neq j$ とするとき、 A_i と A_j の関係は直接的には無関係であっても間接的に C あるいは X を通して関連付けられていると考えられる。分類体系が定期的あるいは不定期であって商品分類のように順次改訂されていくような場合には一般的な分類基準の X に相当するものが存在しているのは稀である。そのため実際に存在ししかも隣り合う分類の A_i と A_{i+1} により次々と関連付けていかなければならない。野田 [2010] は A_1, A_2, A_3 の分類の対応関係の連結を拡張してグループの連結における一般形を漸化式として導くことを示している。

共通に存在する分類 A_2 に対する A_1, A_2, A_3 の対応関係の連結は、分類基準の X あるいは C が存在するときは A_1 と A_2 、同じく A_2 と A_3 がそれぞれグループ化された対応関係コード表であることを前提とする。分類基準が存在しないときはそれらの代わりに ϕ とし、グループ化された対応関係コード表があることを前提とする。以下、 A_1, A_2, A_3 の対応関係における連結のための処理過程である。 A_2 に基づいて得られる A_1 と A_3 の FCD を求め、グループの分類を CG_3 とする。この FCD は $fcd(A_1, A_3 : A_2) \rightarrow CG_3$ 、で表わされ、 CG_3 は A_1, A_2, A_3 における連結されたグループの分類である。

[1] 分類 A_1 と A_2 のグループ化は ϕ を分類基準として、 A_1 と A_2 から得られた FCD を G_2 とし求める。この FCD は、 $fcd(A_1, A_2 : \phi) \rightarrow G_2$ 、で表わされる。

[2] 分類 A_2 と A_3 のグループ化は ϕ を分類基準として、 A_2 と A_3 から得られた FCD を G_3 とし求める。この FCD は、 $fcd(A_2, A_3 : \phi) \rightarrow G_3$ 、で表わされる。

[3] G_2 を CG_2 とする。

[4] A_2 に基づいて CG_2 と G_3 の FCD を求め、グループの分類を CG_3 とする。この FCD は、(1-1) $fcd(CG_2, G_3 : \phi) \rightarrow CG_3^*$

で表わされる。ここで、 CG_3 と CG_3^* は一致する。

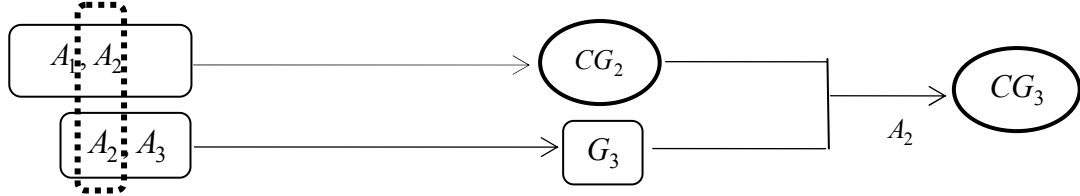
[5] A_1, A_2, A_3 のグループ化された対応関係コード表は CG_3 あるいは CG_3^* により求めることができる。

この処理過程で重要なことは CG_3 と CG_3^* が一致することである。商品分類を始め多くの分類間の対応関係コード表において分類基準が存在していることはほとんど考えられない。したがって、分類基準の存在しない A_1 と A_2 の対応関係コード表と A_2 と A_3 の対応関係コー

図2 CG_{k-1} と G_k から作成される連結されたグループの CG_k の作成過程

(1) k が 2 のときの CG_2 作成過程 :

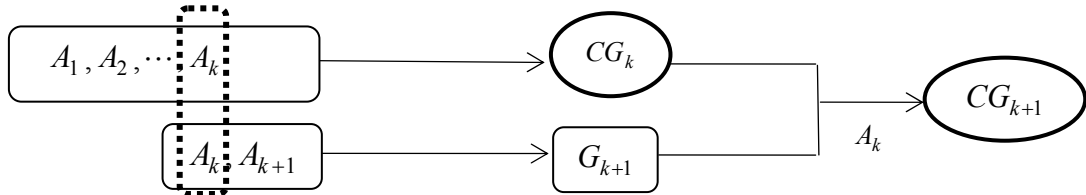
$$fcd(A_1, A_2 : \phi) \rightarrow G_2 (= CG_2), \quad fcd(A_2, A_3 : \phi) \rightarrow G_3, \quad fcd(CG_2, G_3 : A_2) \rightarrow CG_3$$



(2) $CG_3 \rightarrow \dots \rightarrow CG_k$ の作成

(3) 一般的な k に対する CG_{k+1} 作成過程 :

$$fcd(A_k, A_{k+1} : \phi) \rightarrow G_{k+1}, \quad fcd(CG_k, G_{k+1} : A_k) \rightarrow CG_{k+1}$$



(出所) 著者作成

(注) 波線の枠で囲まれている A_k は CG_k と G_k に対する共通の分類基準であることを示している。基準となる A_k に基づいて、 $fcd(CG_{k-1}, G_k : A_k) \rightarrow CG_k$ が作成される。(1) において分類 A_2 を基準とした A_1, A_2, A_3 の対応関係は、 $\xi_i^{(3)} \in CG_3$ に対して、 $R_i(A_1 \cdots A_3 : A_2)$ である。(2) において分類 A_k を基準とした $A_1 \cdots A_{k+1}$ の対応関係は、 $\xi_i^{(k)} \in CG_k$ に対して、 $R_i(A_1 \cdots A_{k+1} : A_k)$ である。

ド表が準備されているとき、共通に存在する A_2 を基にして A_1, A_2, A_3 を連結することは可能である。さらに両対応関係コード表のグループ化された分類 G_2 と G_3 に対して A_2 を基にしてグループ化することも可能である。しかもグループ化された両者は一致することである。

グループの分類の個数を n_3 として、 $i=1 \cdots n_3$ に対して、 $\xi_i^{(3)} \in CG_3$ とする。野田 [2012] によれば、処理過程の [4] において A_2 を基準とする CG_2 と G_3 の対応関係は、 $c \in A_2$ に対して、 $R_i(CG_2, G_3 : A_2) = R_i$ とすれば、

$$(1-2) \quad R_i = \{(p_1 h_2^{-1}(c) \times p_3 h_3(c)) \mid p_3^* h_3(c) = \xi_i^{(3)}\}$$

として求められる。グループ化された対応関係コード表は対応関係のすべての集まりとして求められる。処理過程の [5] において A_1, A_2, A_3 のグループ化された対応関係コード表は、 $R_i(A_1, A_2, A_3 : A_2) = R_i$ とすれば、

$$(1-3) \quad R_i = \{(h_2^{-1}(c) \times c \times h_3(c)) \mid p_3^* h_3(c) = \xi_i^{(3)}\}$$

と表わされる。対応関係コード表は対応関係のすべての集まりとして求められる。さらに、連結されグループ化された A_1, A_2, A_3 の対応関係コード表の中から取り出した A_1 と A_3 の対応関係は、 $R_i(A_1, A_3 : A_2) = R_i$ とすれば、

$$(1-4) \quad R_i = \{(h_2^{-1}(c) \times h_3(c)) \mid p_3^* h_3(c) = \xi_i^{(3)}\}$$

として得られる。

対応関係のグループ化における連結方法として野田 [2010] は分類 A_2 を基準とした A_1, A_2, A_3 の対応関係の連結方法を拡張してグループの連結における一般化を漸化式として導いている。これによれば、分類 A_1, A_2, \dots, A_{n+1} が $n+1$ 種類あり、 $k=1 \dots n$ に対して分類基準の存在しない A_k と A_{k+1} の対応関係コード表が存在する。 A_k と A_{k+1} の対応関係コード表が存在するので、分類基準の存在しない両者の FCD は、

$$(1-5) \quad fcd(A_k, A_{k+1} : \phi) \rightarrow G_{k+1}$$

となる。この $n+1$ 種類の対応関係の連結を CG_n としたとき、この関係を漸化式としてまとめることができる。以下がその漸化式における処理過程である。

[1] 初期値として CG_2 を設定するため、分類基準の存在しない FCD の求め方に従って、

$$(1-1) \text{ 式において } k \text{ を } 1 \text{ とおき、 } fcd(A_1, A_2 : \phi) \rightarrow G_2 \text{ とする。}$$

[2] $CG_2 = G_2$ とする。

[3] $k=2 \dots n$ に対して、 A_1, A_2, \dots, A_k を連結した CG_k が得られているとする。分類基準の存在しない A_k と A_{k+1} の対応関係は存在するので、(1-5) 式から FCD の G_{k+1} を求めることができる。これでグループ化に必要となる CG_k 、 G_{k+1} と基準となる A_k が揃ったことになる。

[4] 連結の基準となる分類 A_k に基づいて CG_{k-1} と G_k のグループ化をおこなう。この FCD は (1-1) 式の一般化であり、

$$(1-6) \quad fcd(CG_k, G_{k+1} : A_k) \rightarrow CG_{k+1}$$

となる。

[5] $k=n$ となるまで [3] から [5] までの処理過程を繰り返す。

[6] $k=n$ となったときに得られた (1-6) 式の CG_{n+1} が求める A_1, A_2, \dots, A_{n+1} の連結されたグループの分類である。

これらの処理過程は図 2 にまとめられており A_1 と A_2 のグループ化された対応関係コード表とその一般形である A_k と A_{k+1} のグループ化された対応関係コード表の作成のための処理過程が示されている。処理過程の [3] における分類 A_k に基づいた CG_k と G_{k+1} の対応関係は (1-6) 式を求めるためには重要である。

グループ化された分類 A_1 と A_2 の個別分類における対応関係は次のような 4 つのタイプに分けることができる。対応関係が A_1 から A_2 へと向かう方向を持っているとする。

(1) 対応関係のタイプ 1 は A_1 と A_2 の個別分類コードが 1 対 1 に対応する関係である。このタイプではグループに含まれる対応する分類コードの個数は 1 個である。

表1 商品分類の対応関係における対応関係の数と商品グループの数

(1) 商品グループの数

タイプ 対応関係	1	2	3	4a	4b	total
(a) SITC-R2 to SITC-R1	870	211	64	14	13	1172
(b) SITC-R3 to SITC-R1	659	250	10	0	88	1007
(c) SITC-R3 to SITC-R2	356	250	22	29	41	1669
(d) SITC-R3 to SITC-R2 *	827	345	16	75	56	1322
(e) HS1988 to SITC-R3	2218	0	893	1	0	3112
(f) HS1996 to SITC-R3	2126	0	930	0	0	3056
(g) HS2002 to SITC-R3	2067	0	972	0	0	3039

(2) 対応関係の数

タイプ 対応関係	1	2	3	4a	4b	total
(a) SITC-R2 to SITC-R1	870	797	145	76	118	2006
(b) SITC-R3 to SITC-R1	659	763	20	0	4277	5719
(c) SITC-R3 to SITC-R2	356	923	49	186	3805	5319
(d) SITC-R3 to SITC-R2 *	827	1105	32	471	2345	4789
(e) HS1988 to SITC-R3	2126	0	2988	4	0	5118
(f) HS1996 to SITC-R3	2218	0	2799	0	0	5017
(g) HS2002 to SITC-R3	2067	0	3157	0	0	5224

(出所) 商品分類の SITC-R1 と SITC-R2 の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*、商品分類の 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* から得られた対応関係から著者作成。HS の各改訂版 (HS1988、HS1996、Hs2002) と SITC-R3 の対応関係は UN のホームページの <http://unstats.un.org/unsd/cr/registry/regdnld.asp?Lg=1>、から取り出して著者作成。

(注) *は対応関係の切断モデルを表す。切断の要を加えると (1) の (d) の total は 2,261、(2) の (d) のそれは 5,719 となる。

(2) 対応関係のタイプ 2 は A_1 と A_2 が 1 対多の対応関係である。グループに含まれる対応する分類コードの個数は A_2 に含まれる分類コードの個数に等しい。

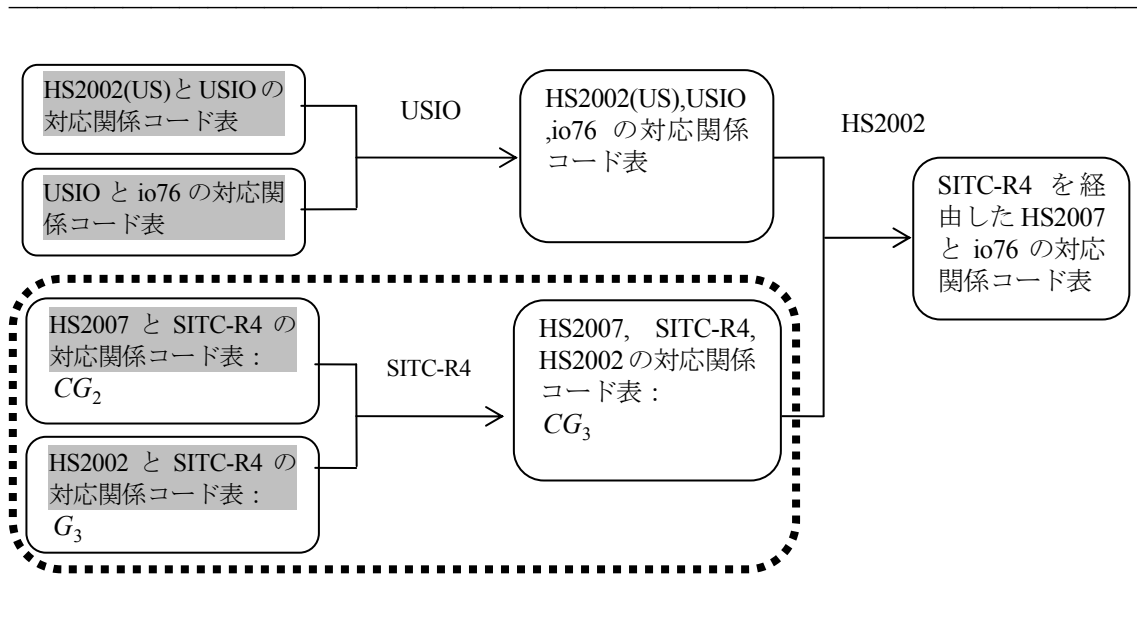
(3) 対応関係のタイプ 3 は A_1 と A_2 が多対 1 の対応関係であり、タイプ 2 とは逆にグループに含まれる分類コードの個数は A_1 に含まれる分類コードの個数に等しい。

(4) 対応関係のタイプ 4 は A_1 と A_2 が多対多の対応関係である。このタイプのグループに含まれる分類コードの個数について特に決まったパターンはない。タイプ 4 はさらにタイプ 4a と 4b とに分けることができる。

1.1 対応関係コード表の作成

本章では Comtrade データの報告国日本を利用して、SITC-R2 から SITC-R1 への貿易デ

図3 HS2007 から io76 へ変換するのに必要な対応関係コード表の一覧



(出所) HS2007 と HS2002 の対応関係コード表および HS2002 と io76 の対応関係コード表に基づき著者作成。

(注) 既存の対応関係コード表は影の付いている対応関係コード表である。枠の外に記載されている USIO、SITC-R4 および HS2002 は対応関係コード表の連結において基準となる分類である。

一タの変換を試みている。変換に必要な商品分類の SITC 改訂第 1 版 (SITC-R1) と SITC 改訂第 2 版 (SITC-R2) の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。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。SITC-R1 と SITC-R2 の対応関係コード表の商品グループは SITC-R1 を旧分類の A_1 、SITC-R2 を新分類の A_2 とするとき、 A_1 と A_2 の FCD として $fcd(A_1, A_2) \rightarrow G_1$ により求められる。表 1 は商品分類の対応関係における対応関係の数と商品グループの数がそれぞれ (1) と (2) で示されている。(a) の SITC-R2 to SITC-R1 によれば、870 個の基本項目コードがタイプ 1 の関係にある。タイプ 2 の関係では、SITC-R1 では 211 個あり、この 212 個の分類コードが SITC-R2 の 797 個のそれと対応する。タイプ 3 では SITC-R2 の 64 個の分類コードが SITC-R1 の 145 個のそれと対応する。タイプ 4 では、SITC-R1 の残りの分類コードが SITC-R2 の残りの分類コードと複数コード対複数コードの関係にある。

この変換は最新年次 (この時点では 2009 年まで利用可能) までを SITC-R1 へと変換できる。その場合には SITC 改訂第 3 版 (SITC-R3) から SITC-R1 の対応関係コード表、HS

の各改訂版から SITC-R1 への対応関係コード表が必要となる。商品分類の SITC-R2 と SITC 改訂第 3 版 (SITC-R3) の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の商品グループは SITC-R2 を旧分類の A_1 、SITC-R3 を新分類の A_2 とするとき、 A_1 と A_2 の FCD として $fcd(A_2, A_3) \rightarrow G_2$ により求められる。表 1 にある (1) の (b) に SITC-R3 to SITC-R2 の商品グループ数、同じく (2) の (b) にその対応関係の数がそれぞれ示されている。商品分類が HS の各改訂版については SITC_R3 との対応関係コード表は対応関係のタイプ 1 あるいはタイプ 3 にまとめられている。

1.2 対応関係コード表の連結の例

商品分類の SITC-R1 と SITC-R3 の対応関係コード表は UN 統計局発行の出版物として存在しないが、対応関係コード表を連結することで作成される。SITC-R1 を分類の A_1 、SITC-R2 を A_2 、SITC-R3 を SITC-R1 を A_3 とする。前述した処理過程において [5] の $k=3$ のときに得られる CG_2 が連結された A_1, A_2, A_3 の対応関係コード表である。この対応関係コード表から A_1 と A_3 の対応関係を取り出し、重複しているものを取り除き、再度グループ化したのが A_1 と A_3 のグループ化された対応関係コード表である。表 1 にある (1) の (c) に SITC-R3 to SITC-R1 の商品グループ数、(2) の (c) にその対応関係の数がそれぞれ示されている。

対応関係コード表の連結を利用した例に、貿易データを HS の 2007 年改訂版 (HS2007) から国際産業連関表の 76 部門分類 (io76) へ変換した内田・野田 [2010] がある。図 3 の下段の破線で囲まれた箇所を例として対応関係の連結の状態を示す。ここで、SITC-改訂第 4 版 (SITC-R4)、米国産業連関表の部門分類 (USIO)、HS の 2002 年度改訂版 (HS2002) としている。対応関係の連結において HS2007 を A_1 、SITC-R4 を A_2 、HS2002 を A_3 、USIO を A_4 、io76 を A_5 とする。 $fcd(A_1, A_2 : \phi) \rightarrow G_2$ となり、 G_2 を CG_2 とする。 $fcd(A_2, A_3 : \phi) \rightarrow G_3$ であり、共通に存在する A_2 を基準として $fcd(CG_2, G_3 : A_2) \rightarrow CG_3$ とする。 CG_3 は $A_1 \cdots A_3$ の連結された共通したグループである。同じようにして CG_5 は $A_1 \cdots A_5$ の連結された共通したグループである。この対応関係コード表から A_1 と A_5 の対応関係を取り出し、重複しているものを取り除いたとき、再度グループ化したものがグループ化された HS2007 と io76 のグループ化された対応関係コード表である。

1.3 対応関係の切断とサブグループ

閉じた対応関係の集まりである 1 つのグループからいくつかの対応関係を取り除くとグ

グループがさらに2つ以上のグループに分かれるとき、この対応関係の取り除きによってグループが「切断」されたといい、そのときに取り除いた対応関係を「切断の要素」という。このとき得られたグループを元のグループに対するサブグループという。サブグループがグループから切断の要素を取り除いた対応関係コード表に対して再度グループ化をすることで得られるということは、切断の仕方によってサブグループの内容や個数が決まるということである。このことは切断というのは対応関係コード表のグループ化に対する1つのモデルであると考えることができる。切断をしない対応関係のモデルを対応関係の基本モデル、切断によりサブグループ化された対応関係を対応関係の切断モデルという。

表1にある(1)の(d)にSITC-R3 to SITC-R2の商品グループ数、(2)の(d)にその対応関係の数がそれぞれ示されており、これがSITC-R2とSITC-R3の対応関係における切断モデルである。同表の(c)がSITC-R2とSITC-R3の対応関係における基本モデルである。前者の切断モデルについては野田・山本[1995]が参考になる。この切断モデル作成の目的は大きなグループに属する類似なサブグループを分離することである。それと同時に後述する配分ウェイト行列の推計において対応関係のタイプ4bを減らし、グループおよびサブグループのサイズを小さくすることにある。表1の(2)において(c)と(d)の4bの割合を計算すれば基本モデルは71.5%であるのに対して切断モデルは49.0%になっている。

2. 商品分類を変換するための配分ウェイト行列の推計

商品分類の改訂に伴って作成される新旧それぞれの商品分類である分類*A*と*B*において、分類*A*から*B*への対応関係をもとに前者の取引額を変換して後者の取引額を推計するとき、その変換のフィルターの役割を果たすのが配分ウェイト行列である。配分ウェイト行列は商品分類の改訂前後のそれぞれの取引額を利用して推計されるがそのために以下で説明するような仮説が必要となる。この仮説を変換可能であるための仮説といい、グループ化された分類*A*から*B*への配分ウェイト行列をもとに前者の取引額を後者のそれに変換するための配分構造が定式化される。

商品グループ内における分類*A*から*B*の方向に対する対応関係が存在するとき、商品分類の改訂前後のそれぞれの期間にこの両分類による取引額が存在し、取引額の構造は多変量定常確率過程に従っているものとする²。すなわち、分類*A*における取引額の構成比はその期間において*A*の取引額の構造から得られた標本として解釈する。同じように分類*B*についても*B*の取引額の構造から得られた標本とする。しかもその期間に両分類によるそれぞれの標本が同時に得られるものとする。

実際には同一期間からは両分類の取引額の構成比に相当する標本は得られず、両分類の

図4 分類Aから分類Bの方向に対する取引額XからYへの変換とその構造

年		1962	……	1975	1976	……	1987
分類A : SITC-R2	$\begin{pmatrix} a_1 \\ \vdots \\ a_n \end{pmatrix}$				$X^* = \begin{pmatrix} x_1^* \\ \vdots \\ x_n^* \end{pmatrix} = \begin{pmatrix} x_{11}^* & \cdots & x_{1h}^* \\ \vdots & & \vdots \\ x_{n1}^* & \cdots & x_{nh}^* \end{pmatrix}$		
分類B : SITC-R1	$\begin{pmatrix} b_1 \\ \vdots \\ b_m \end{pmatrix}$	$Y^* = \begin{pmatrix} y_1^* \\ \vdots \\ y_m^* \end{pmatrix} = \begin{pmatrix} y_{11}^* & \cdots & y_{1k}^* \\ \vdots & & \vdots \\ y_{m1}^* & \cdots & y_{mk}^* \end{pmatrix}$			$\hat{Y}^* = \begin{pmatrix} \hat{y}_1^* \\ \vdots \\ \hat{y}_m^* \end{pmatrix} = \tilde{W}X^*$		

(出所) 著者作成

(注) 日本の商品分類を例としており、SITC-R2を分類A、SITC-R1を分類Bと想定している。影の部分が実際に存在するデータであり、…は欠損値である。分類Aのn個ある個別分類コードを $a_1 \cdots a_n$ 、その取引額は X^* で表わされ、 $x_j^* = (x_{j,1976}^* \cdots x_{j,1987}^*)$ である。分類Bのm個ある分類コードを $b_1 \cdots b_m$ 、その取引額は Y^* で表わされ、 $y_i^* = (y_{i,1962}^* \cdots y_{i,1975}^*)$ である。 \tilde{W} は推計された配分ウエイト行列、 \hat{Y}^* は推計値を表わす。

取引額の構成比は商品分類における改訂前後の期間からそれぞれ得られる。この仮説は改訂前の期間から得られた標本が改訂後の期間から得られた標本と同一の無作為抽出から求められるという相当に強い仮説といえる。ここで重要なことは、この仮説は改訂前の分類Bの構造がそのまま改訂後にも同じように維持されていることであり、しかも得られる標本は取引額そのものではなくその構造を示す構成比としていることである。貿易データのような経済データにはトレンドや各種の周期を内蔵しており、取引額そのものは一般的には大きく変動する傾向にあることによる。

図2はComtradeデータの報告国日本における商品分類の改訂前後を示したものである。この図において分類Aは1976年から1987年までを対象とするSITC-R2、分類Bは1962年から1975年までを対象とするSITC-R1をそれぞれ表わしており、影の付いているところが実際に得られる貿易データである。この仮説において実際には存在していない1976年から1987年までの期間において分類Bの取引額の構成比が同時に得られることを可能にしている。したがって、1976年から1987年までについては分類AとBの取引額の構成比は同時に得られることになる。さらに、得られた両分類の取引額の構成比は時系列データの特性を考慮せずに、両者それぞれの分類が持つ構造から得られた無作為抽出による標本としている。

図2において、商品グループ内における分類Aのn個ある個別分類コードの $a_1 \cdots a_n$ のそれぞれに統計値であるn個の取引額 $x_1^* \cdots x_n^*$ が対応している。右上に*で表わされている

表2 商品グループ内の分類AからBの向けた変換の配分構造

分類A 分類B	a_1	a_2	...	a_j	...	a_n	Total
b_1	$x_1\omega_{11}$	$x_2\omega_{12}$		$x_j\omega_{1j}$		$x_n\omega_{1n}$	y_1
\vdots	\vdots	\vdots		\vdots		\vdots	\vdots
b_i	$x_1\omega_{i1}$	$x_2\omega_{i2}$		$x_j\omega_{ij}$		$x_n\omega_{in}$	y_i
\vdots	\vdots	\vdots		\vdots		\vdots	\vdots
b_m	$x_1\omega_{m1}$	$x_2\omega_{m2}$		$x_j\omega_{mj}$		$x_n\omega_{mn}$	y_m
Total	x_1	x_2		x_j		x_n	1

(出所) 野田容助編『東アジア諸国・地域の貿易指数—作成から応用までの基礎的課題—』(SDS No.88 アジア経済研究所 2005) の第2章の表2を引用

(注) ω_{ij} は商品グループ内における分類Aの a_j から分類Bの b_i への方向に対する配分ウエイトである。表2において i, j 要素は取引額を表示しているため、この表は配分額表とも呼ばれる。配分額表において分類AのTotalとなる $x_1 \dots x_n$ は分類Aの周辺和、分類BのTotalとなる $y_1 \dots y_m$ は分類Bの周辺和と呼ばれる。

変数は取引額を表わしている。 $j=1 \dots n$ に対して、 x_j^* は年次データに相当する h 個の標本から構成されるベクトルとして表わされ、 $x_j^* = (x_{j1}^* \dots x_{jh}^*)'$ である。図3において、 h は1976年から1987年までの期間を一連番号で表わしているため12となる。 $(x_1^* \dots x_n^*)' = X^*$ とすれば、

$$\begin{pmatrix} x_1^{*1} \\ \vdots \\ x_n^{*1} \end{pmatrix} = \begin{pmatrix} x_{11}^* & \dots & x_{1h}^* \\ \vdots & & \vdots \\ x_{n1}^* & \dots & x_{nh}^* \end{pmatrix} = X^*$$

となる。したがって、分類Aのある期間の取引額の行列は $n \times k$ 行列の X^* で表わされる。

同一商品グループ内の分類Bの m 個ある個別分類コードの $b_1 \dots b_m$ のそれぞれに対する統計値を $y_1^* \dots y_m^*$ とする。年次データの y_i^* は k 個の標本から構成されるベクトルで表わされ、 $i=1 \dots m$ に対して $y_i^* = (y_{i1}^* \dots y_{ik}^*)'$ である。図2において、 k は1962年から1975年までの期間を一連番号で表わしているため14となる。 $(y_1^* \dots y_m^*)' = Y^*$ とすれば、分類Bの取引額の行列は $m \times k$ 行列の Y^* で表わされる。取引額の行列 X^* と Y^* の標本数である年次データの個数をそれぞれ h と k としているのは実際においても両者の個数が一致することはほとんどないからである。

このままでは同一年の取引額は分類AとBとに同時に配分されているという仮説の意味する同一年の分類Aの取引額がすべて配分ウエイトによって分類Bの取引額へ変換される条件が与えられていない。分類AからBへの変換において年次を表わす j に対して前者の取引額である $(x_{1j}^* \dots x_{nj}^*)$ の合計は配分されて後者の取引額 $(y_{1j}^* \dots y_{mj}^*)$ のそれぞれ合計と一致しなければならない。しかも、これがすべての年において成り立つわけであるから、

煩雑さを避けるために $h=k$ とすれば、取引額の年毎の和は一致し、 $(x_{\bullet 1}^* \cdots x_{\bullet k}^*) = (y_{\bullet 1}^* \cdots y_{\bullet k}^*)$ となる。ここで、 \bullet は対象とするすべての和を表わし、 $y_{\bullet j}^* = y_{1j}^* + \cdots + y_{mj}^*$ とする。すなわち、 $j=1 \cdots k$ に対して $x_{\bullet j}^* = y_{\bullet j}^*$ となる。 l_m をすべての要素が 1 である m 次ベクトル、 $D(x)$ をベクトル x を対角要素とする対角行列とすれば、

$$(2-1) \quad l_n' X^* = l_m' Y^*$$

という関係が成立つ。分類 A と B のそれぞれの取引額行列に対してその構成比を取り、 X と Y とする。 X^* については、 $X = X^* D(l_m' X^*)^{-1}$ 、としてその構成比を要素とする行列 X が作成できる。 Y^* についても同じように、 $Y = Y^* D(l_n' Y^*)^{-1}$ として Y を作成できる。 $l_n' X = l_m' Y = l_k'$ となり、(2-1) 式の条件を満たすことができる。構成比を用いることにより必ずしもではないが、貿易データが持つ長期トレンドや周期を含めた経済変動から生ずる経済変動固有の一部の変動を取り除くことができる。

商品グループ内における分類 A の $a_1 \cdots a_n$ から分類 B の b_i に向けた変換の配分構造と配分ウエイトを表として示したのが表 2 である。この表は配分額行列あるいは取引額表と呼ばれる。商品グループ内における分類 A から B の方向に対する変換のために ω_{ij} は分類コード a_j から b_i への方向に対する配分ウエイトとする。 ω_{ij} を要素とする行列を W とする。商品分類の対応関係では商品グループ内に分類 A の a_j と分類 B の b_i それぞれの個別分類コード間に対応関係がない状態が存在するのが一般的である。このような状態は配分ウエイト行列において 0 となる要素が存在する状態であり、要素として $\omega_{ij} = 0$ を含む一般的な配分ウエイト行列を W_g とする。 a_j の $j=1 \cdots n$ に対して、 $x_j = (\omega_{1j} + \cdots + \omega_{mj})x_j$ である。これから、 $\omega_{1j} + \cdots + \omega_{mj} = 1$ となり、一般的な配分ウエイト行列はウエイト条件として、

$$(2-2) \quad l_m' W_g = l_n'$$

となる。表 2 における b_i の $i=1 \cdots m$ に対して y_i は、 $y_i = x_1 \omega_{i1} + \cdots + x_n \omega_{in}$ と表すことができる。この配分ウエイト行列に対して構造式は、

$$(2-3) \quad Y = W_g X + U$$

と表すことができる。 U は誤差であり Y と同じ構造を持つものとする。以上のことから、本章の目的は商品グループ内に存在する分類 A から分類 B への方向に対する対応関係とそれぞれの分類に対応する構成比で構成された取引額行列 X と Y が得られたとき、(2-2) 式と (2-3) 式から配分ウエイト行列 W_g を推計することになる³。

仮説により分類 B の取引額の構成比 Y が得られているのに配分ウエイト行列により改めて Y を推計する必要があるかという疑問が生ずる。確かに一般的な配分ウエイト行列として $W_g = Y$ とすれば (2-2) の条件を満たしている。分類 B の取引額は X の年ごとの総額を利用すれば、 $\hat{Y}^* = Y D(l_n' X^*)$ 、として計算することができる。後述するように Y は配分ウエイト行列の特殊な推計方法として説明される。しかしこの推計値は取引額 X^* との対応関係にかかわる情報は利用していない。本節では取引額の構成比の X と Y が対応関係にどのようにかかわっているかを知るための一般的な配分ウエイト行列の構造を推計すること

が興味の対象である。したがって仮定によって得られた Y のみを配分ウエイトの推計値として利用すると同時に、 X と Y を利用した配分ウエイト行列 W も推計値として利用する。 W が推計されれば、(2-3) 式により既存の X から分類 B には存在しなかった \hat{Y} を計算することができ、分類 B にもとづく時系列データの取引額 (Y, \hat{Y}) の利用が可能となる。

2.1 等号制約条件付き最小 2 乗法

ウエイトの等号制約条件付き最小 2 乗法の中で配分ウエイト行列を行列のまま直接的に解くゼロ制約を考慮した等号条件付最小 2 乗法は野田 [2003] で紹介されている。分類 A から B の方向に対する変換のための対応関係において商品グループ内に対応関係がない状態、すなわち 0 となる配分ウエイトが存在するのが一般的である。2 次元の自然数から構成される集合を、 $S = \{(i, j) | i = 1 \cdots m, j = 1 \cdots n\}$ とし、 $(i, j) \in S$ に対して $\omega_{ij} = 0$ となる任意の部分集合を S_2 とする。すなわち配分ウエイト行列 W における 0 となる要素は、 $S_2 = \{(i, j) | \omega_{ij} = 0\}$ 、と表わすことができる。

最小 2 乗法の定義には一般化分散や全変動等を最小にするいくつかの方法があるが本章では後者を採用する。 $\lambda' = (\lambda_1 \cdots \lambda_n)$ を n 次ベクトル、 $i = 1 \cdots m$ および $j = 1 \cdots n$ に対する ξ_{ij} を mn 個のスカラ、 $e_i(m)$ を i 番目が 1 でそれ以外は 0 を要素とする m 次ベクトルとすれば、配分ウエイトのゼロ制約持つウエイト制約条件付き最小 2 乗法の解はラグランジェ関数を、

$$(2-4) \quad \ell = \text{tr}(UU') - (l_m'W - l_n')2\lambda - \sum_{(i,j) \in S_2} e_i(m)'We_j(n)2\xi_{ij}$$

としたとき $\partial \ell / \partial W = 0$ から得られる W の解であり、これをの \tilde{W} とする。また $\partial \ell / \partial \lambda = 0$ から $l_m' \tilde{W} = l_n'$ が得られウエイト条件を満たしていることが確かめられる。ゼロ制約を構成する行列を、

$$(2-5) \quad B = \sum_{(i,j) \in S_2} e_i(m)e_j(n)'\xi_{ij}$$

とする。 B は $(i, j) \in S_2$ に対して $\xi_{ij} \neq 0$ 、それ以外は 0 となる行列である。 W のゼロ制約 $\omega_{ij} = 0$ の位置と B における $\xi_{ij} \neq 0$ の位置は一致している。また、 $\partial \ell / \partial \xi_{ij} = 0$ から、 $(i, j) \in S_2$ に対して $e_i(m)'\tilde{W}e_j(n) = 0$ が得られる。 \tilde{W} は、

$$(2-6) \quad \tilde{W} = \tilde{W}_0 + (I_m - l_m l_n' / m)B(XX')^{-1}$$

と求められる。 \tilde{W}_0 は (2-4) 式においてゼロ制約のないウエイト条件付きの最小 2 乗法の解である。 $YX'(XX')^{-1} = G$ 、 $mI_m - l_m l_m' = A$ とすれば、

$$(2-7) \quad \tilde{W}_0 = l_m l_n' / m + AG / m$$

として得られる確定した値である。したがって、(2-6) 式において \tilde{W}_0 は確定した値なので

行列 B が求められれば \tilde{W} が得られる。 \tilde{W} をゼロ制約とウエイト条件による等号制約付き最小 2 乗法 (Least square with zero restriction and Weight condition into Matrix form: wm 方式) による推定量とする。

2.2 分類間の独立性を仮定した同一配分パターン方式

商品グループ内における分類 A と B の取引額は商品分類の改訂前後においてそれぞれ得られた無作為標本であるが、この標本は同一期間から得られた無作為標本と同一であるという仮説のもとでは、分類 A と B の 2 次元分割表として表わすことができる。表 2 はこの 2 次元分割表を示したものであり、分類 B における取引額の構成比は、 $i=1\cdots m$ に対して $y_i = (y_{i1} \cdots y_{ik})'$ 、分類 A の取引額の構成比は $j=1\cdots n$ に対して $x_j = (x_{j1} \cdots x_{jk})'$ として共にベクトルで表わされる。これをスカラー表示で表わすため、構造式を表わす (2-3) 式において、 $U=0$ とおき、両辺に右から l_k を乗じて k で除せば、 $(\bar{y}_1 \cdots \bar{y}_m)' = W(\bar{x}_1 \cdots \bar{x}_n)' = WD(\bar{x})l_n$ が得られる。 \bar{y}_i は y_i の平均である。 $\bar{y}' = (\bar{y}_1 \cdots \bar{y}_m)$ とする。 \bar{x} についても同様である。適当な数を Q とするとき、この式の右辺の $WD(\bar{x})$ に Q を乗じて、

$$(2-8) \quad V = WD(\bar{x})Q$$

とすれば、(2-8) 式 2 次元分割表であり \bar{x} で表わされた配分額行列となる。 V の列および行における周辺和はそれぞれ $l_m'V = \bar{x}'Q$ と $Vl_n = \bar{y}Q$ となり、 V の総和は、 $l_m'Vl_n = \bar{x}'l_nQ = l_m'\bar{y}Q = Q$ である。したがって、 V の要素が得られる同時確率を p_{ij} としてその実現値を v_{ij} とすれば、

$$(2-9) \quad v_{ij} = p_{ij}Q$$

と表わされる。

分類 A と B が互いに独立と仮定される分割表に対する配分額行列の推計が配分ウエイト行列における分類間の独立性を仮定した同一配分パターン方式である。最初は、分類 A と B の商品グループ内において両分類の個別分類コード間のすべてに対応関係があるとす。同時確率が独立であるときには (2-9) 式において、 $p_{ij} = p_{i\cdot}p_{\cdot j}$ となるため、 $v_{ij} = p_{\cdot j}p_{i\cdot}Q$ と表わすことができる。 p_{ij} を要素とする同時確率の行列を P とすれば、 $P = D(Pl_n)l_m l_n' D(l_m'P)$ 、となる。この P を利用して v_{ij} を要素とする行列 V を表示すれば、 $V = D(Pl_n)l_m l_n' D(l_m'P)Q$ が得られる。 V が与えられているときには配分ウエイト行列の W を計算することができ、

$$(2-10) \quad W = V\{D(\bar{x})Q\}^{-1} = (p_{1\cdot} \cdots p_{m\cdot})l_n'$$

となり、分類 B における周辺分布のみから求めることができる。しかも、分類 A に相当する列に対しては同じように B の周辺確率が適用される。このことが分類間の独立性を仮定した推計方法に対して同一配分パターンと呼ばれる根拠となっている。分類間の独立性を仮定した同一配分パターンによる推計方法は X のみに依存し Y とは無関係となる。

一般的な配分ウェイト行列 W_g の要素が 0 以外のときに 1 に置き換える関数を $a(\cdot)$ とする。 $a(\cdot)$ は次のように解釈され、

$$a(W_g) = \sum_{(i,j) \in \{S-S_2\}} e_i(m) e_j(n)'$$

となる。配分ウェイト行列作成のためのつぎの処理は $a(W_g)$ を利用して対応関係のないところを調整することである。 $W_2 = D(\bar{y})a(W)$ とおき、ウェイトの条件を満たすように作り直せば、

$$(2-11) \quad W_i = W_2 \cdot D(l_m' W_2)^{-1}$$

として得られる。 W_i は分類間の独立性を仮定した同一配分パターンを持つ配分ウェイト行列の推計量であり、これを i 方式による推定量とする。

2.3 初期条件なしの最大エントロピー法

取引額を考慮せずに対応関係の配分構造のみから推計する方法の 1 つに単純均等配分法がある。単純均等配分法はエントロピー最適化の特殊解として解釈できる。分割表を近似する確率分布モデルがいくつか想定されるとき、実現値を生成する真の確率分布に対してモデルによって既定された確率分布の近似は Kullback-Leibler 情報量 (K-L 情報量) によって評価することができる。K-L 情報量の符号を逆転させた値は負のエントロピーであり、この値が大きい程、同じことであるが、K-L 情報量が小さいほど近似の程度が良いとして評価される。この方法は最大エントロピー法と言われる。

ある条件のもとで配分額行列の初期値 $V^{(0)}$ は知られているものとして、推計したい取引額行列を V とする。最初は、商品グループ内の対応関係がすべて存在するとする。知られていない取引額行列に対して、 $V^{(0)} = l_m l_n'$ 、 $\bar{x} = l_n$ とする。配分額行列は $V = W D(l_n) = W$ となり、配分ウェイト行列に一致する。最大エントロピー法を利用するためのラグランジェ関数は $V = W$ とすれば、 $i = 1 \dots m$ および $j = 1 \dots n$ に対して、

$$(2-12) \quad \ell_3 = \sum_i^m \sum_j^n \omega_{ij} (\log \omega_{ij} - 1) + \sum_j^n \eta_j \left(1 - \sum_i^m \omega_{ij} \right)$$

となる。(2-12) 式における右辺の第 2 項目目は配分ウェイト行列におけるウェイト条件である。(2-12) 式を ω_{ij} で偏微分した結果を 0 とおけば、 $W = l_m l_n' / m$ となる。

一般的な配分ウェイト行列において、商品グループの対応関係に $\omega_{ij} = 0$ が存在するときの処理は $a(W)$ を利用して対応関係のないところを 0 に置き換えて調整することである。 $W_3 = D(l_m) a(W) / m$ とおき、ウェイトの条件を満たすように作り直す。単純均等配分の配分ウェイト行列は、

$$(2-13) \quad W_s = W_3 \cdot D(l_m' W_3)^{-1}$$

として得られる。これを s 方式による推定量とする。

3. 貿易データにおける SITC-R1 への変換

貿易データを推計された配分ウェイト行列にもとづいて新商品分類から旧商品分類へと変換するためには整合性の保証された貿易データを必要とする。貿易データについての整合性の評価および補正についての基本的な考えは野田 [2009] においてまとめられ、本章において貿易データの整合性に直接関わるのは商品分類である。貿易データの変換において本章では商品分類として階層的に構成された桁レベル分類コードではなく、下位の階層の分類コードを持たないものを最も詳細なる分類コード (the most detail classification code: *mdcc*) を採用している。貿易統計で使用される SITC の各改訂系列は商品総額のもとに、1桁レベルから5桁レベルまでの各層に分けられた階層的な商品分類コードから構成されている。本書の附表2に3桁レベル分類コードとその名称が示されている。HSの系列は商品総額のもとに2,4,6桁レベルの各層の商品分類コードから構成されている。Comtradeデータでは各桁レベル分類コードの取引額を合計しても必ずしも商品総額に一致するとは限らない。この階層的に構成された商品分類コードの中で取引額が存在しておりしかも下位レベルの階層の分類コードを持たないものが *mdcc* である。取引額を考慮すれば、*mdcc* は報告国、輸出入区分別に決まり、*mdcc* にもとづく分類コードの対応関係は対応関係の具体的切断モデルである。新商品分類から旧商品分類への変換は商品分類において *mdcc* の分類コードにもとづく整合性のある貿易データを対象にした変換である。

本節では Comtrade データから得られた報告国日本の輸出貿易データを利用して SITC-R2 で編集されている 1976 年から 1987 年までの貿易額を同期間の SITC-R1 への変換を試み、Comtrade データが採用している変換方法と本章で紹介した推計方法の違いによる貿易額の比較および検討をおこなう。配分ウェイトの推定方法は以下に示す 5 つの方法を採用している。

- (1) 取引額を利用せずに対応関係の配分構造のみを考慮した単純均等配分の方法である *s* 方式。(2-13) 式を参照。
- (2) 取引額と配分構造の両者を考慮するとともに分類間の独立を仮定した配分構造同一パターンの方法, である *i* 方式。(2-11) 式を参照。
- (3) *i* 方式の変形となる *u* 方式。
- (4) 取引額と配分構造の両者を考慮するとともに制約条件付き最小 2 乗の方法, である *wm* 方式。(2-7) 式を参照。
- (5) それに加えて、Comtrade データから得られた SITC-R1 系列の 1976 年から 1987 年までを *un_c* 方式とする。

本節はこの 5 種類の異なった方法により推計された SITC-R1 における貿易額の特性を比較検討する。(3) において、*u* 方式というのは *i* 方式で求められた配分ウェイト行列において列ごとに見て最大値を 1、それ以外を 0 と置き換える推計方法である。*u* 方式は UN が

表3 対応関係の数の多い順に並べられた商品グループ


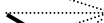



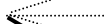


G_i	j	t	s_2	s_1	m	G_i	j	t	s_2	s_1	m
(1) SITC-R1 と SITC-R2 の対応関係						0522	1	4b	15	6	25
0734	1	4b	13	8	23	0070	1	4b	10	8	21
0645	1	4b	5	3	15	0287	1	4b	6	5	20
0412	1	4b	4	7	12	0109	1	4b	4	7	18
0730	1	4b	6	2	12	0543	1	4b	14	2	16
0018	1	4a	6	5	10	0616	1	4b	12	3	16
1100	1	4a	8	2	9	0196	1	4b	7	2	14
0057	1	4b	4	3	8	0624	1	4b	11	3	14
0437	1	4b	4	4	8	0215	1	4b	7	4	12
0580	1	4b	4	2	8	0257	1	4b	10	2	12
0904	1	4b	4	3	8	0570	1	4b	9	3	12
0539	1	4a	3	5	7	(3) SITC-R2 と SITC-R3 の対応関係					
0733	1	4a	4	4	7	0888	1	4b	1014	525	2850
:						0604	1	4b	112	44	391
0242	1	2	1	4	4	0901	1	4b	36	54	270
(2) SITC-R1 と SITC-R3 の対応関係						0367	1	4b	34	11	45
0513	1	4b	1327	439	3225	0422	1	4b	23	17	44
0306	1	4b	33	17	71	0056	1	4a	21	18	38
0603	1	4b	36	7	56	0783	1	4b	27	6	35
0271	1	4b	38	11	49	0830	1	4a	19	10	28
0031	1	4b	21	11	36	0661	1	4a	22	6	27
0539	1	4b	20	4	28	0762	1	4a	20	5	24
0499	1	4a	22	4	25	0650	1	4b	17	2	24

(出所) 著者作成

(注) SITC-R1 と SITC-R2 の対応関係では m が 8 以上のものと 0242 を表示、SITC-R2 と SITC-R3、SITC-R1 と SITC-R3 のそれぞれの対応関係では m が 12 と 24 以上を表示している。影の付いている商品グループの 0057 と 0242 は表 4 と図 5 で参照されている。

採用している配分構造を回避する変換方法に準拠している。(4)において wv 方式を採用しているとき、等号による制約条件を用いているため商品分類の改訂年前後の貿易構造によっては商品グループ内の配分ウエイトは必ずしも正の値を取るとは限らない。配分ウエイト負の値が含まれているときには、すべての配分ウエイトが正の値となるように再配分をする必要がある。この再配分による調整を「負の調整」という。負の値の調整は基本的には2段階の処理でおこなわれる。最初は wm 方式により推計された配分ウエイト行列において絶対値が最大となる負の推定値を 0 と置き、取引額の X と Y を変更せずに同じ処理を繰り返し負の推定値がなくなったときの配分ウエイト行列を一時的な解とする。次は 0 に置き換えられた配分ウエイト行列に小さな値を与え、この置き換えた配分ウエイト行列を初期値として最大エントロピー法により最終推計とする。また、商品分類における対応

図5 SITC-R2 から SITC-R1 への方向に対する商品グループの対応関係

SITC-R2 の名称	SITC-R2	→	SITC-R1	SITC-R1 の名称
(1) 商品グループ : G_{0057} (対応関係のタイプ 4b)				
Bananas, fresh or dried	0573		0513	Bananas incl plantains fresh
Pineapples, fresh, dried	05795		05201	Tropical fruit, dried
Dates, fresh or dried	05796		05195	Tropical fruit nes fresh
Avocado, mango, guava, etc	05797			
(2) 商品グループ : G_{0242} (対応関係のタイプ 2)				
Mineral substances, nes	27899		27692	Earth colours, etc
			27694	Meerschaum, amber, jet
			27696	Naturalarsenic sulphides
			27699	Mineral substances, nes

(出所) UN 統計局発行の *Standard International Trade Classification Revision 2* と、UN 作成による SITC-R2 と SITC-R1 の対応関係コード表にもとづき著者作成

(注) 対応関係コード表において波線は出所におけるの前者、実線は出所の後者から得られる対応関係を表わしている。

関係のタイプの中で対応関係のタイプ 1 とタイプ 3 に属する分類コードの配分ウエイトは無条件に 1 であるので推計はおこなわず、配分ウエイトの推計が必要とされるのは対応関係のタイプが 2,4a,4b の場合である。

表 3 に SITC-R1 と SITC-R2、SITC-R2 と SITC-R3、SITC-R1 と SITC-R3 の対応関係の中で対応関係のタイプが 4a,4b となる商品グループのみが示されている。この表では分類 A と B の対応関係において、 G_i はグループの分類番号、 j はサブグループの一連番号、 t は対応関係のタイプを表しており、分類 A の個数を s_2 、 B の個数を s_1 、対応関係の個数を m で表している。 wm 方式では推計のとき最大の問題は s_1 と s_2 が大きくなるのに伴って m が大きくなり、推計のためのプログラムが異常停止を起こしてしまうことである。SITC-R2 と SITC-R3 では商品グループが 0888 において s_1 と s_2 が 525 と 1,014 であり、 m は 2,850、SITC-R1 と SITC-R3 では商品グループが 0635 のにおいて m が大きくなっている。このような商品グループに対しては wm 方式ではこのままの状態では推計プログラムの容量を超えてしまうため直接的には推計できない。本章では上位桁レベル分類コードに置き換え計算可能な商品グループに集計して推計している。商品グループの大きさがある程度を超えたときにはそのときは対応関係を切断し、計算可能な小さな商品サブグループとそうではない商品サブグループに分割する必要がある。サブグループであってもそのサイズが大きいものが存在するからである。計算不可能な商品サブグループについてはこれ以上の無理な切断はおこなわず、3 桁レベルの分類コードの対応関係置き換えて計算可能な状態に直

表4 推計された配分ウェイト行列の例

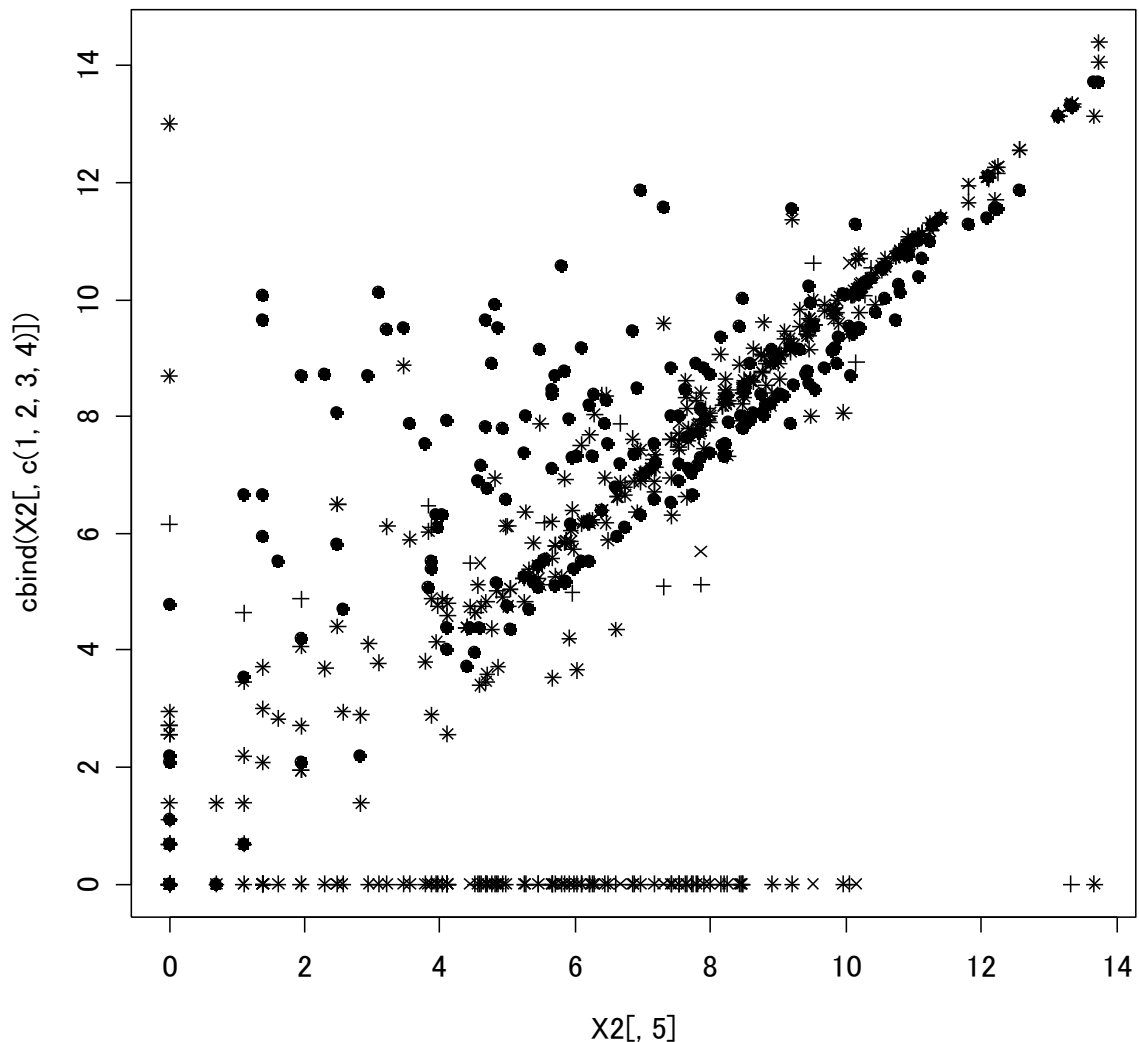
SITC-R1: <i>i</i>	SITC-R2: <i>j</i>	ω_{ij}	<i>wm</i>	<i>s</i>	<i>un_c</i>	<i>u</i>	<i>i</i>
(1) 商品グループ: G_{0057} (タイプ 4b)							
0513 1	0573 1	ω_{11}	1.000000	0.50	1	1	0.997567
05201 3	0573 1	ω_{31}	0.000000	0.50	0	0	0.002433
05195 2	05794 2	ω_{22}	1.000000	0.50	1	1	0.969803
05201 3	05794 2	ω_{32}	0.000000	0.50	0	0	0.030197
05195 2	05795 3	ω_{23}	0.536576	0.50	1	1	0.969803
05201 3	05795 3	ω_{33}	0.463424	0.50	0	0	0.030197
05195 2	05796 4	ω_{24}	1.000000	0.50	1	1	0.969803
05201 3	05796 4	ω_{34}	0.000000	0.50	0	0	0.030197
(2) 商品グループ: G_{0242} (タイプ 2)							
27692 1	27899 1	ω_{11}	0.008149	0.25	0	0	0.008149
27694 2	27899 1	ω_{21}	0.002435	0.25	0	0	0.002435
27696 3	27899 1	ω_{31}	0.002578	0.25	0	0	0.002578
27699 4	27899 1	ω_{41}	0.986838	0.25	1	1	0.986838

(出所) 著者作成

(注) 商品グループは図2における G_{0057} と G_{0242} を対象にしている。配分ウェイト ω_{ij} において *i* は SITC-R1 の分類コード、*j* は SITC-R2 の分類コードをそれぞれ表わしている。*wm, s, un-c, u, i* は推計方法の表記を表し、*wm* は制約条件付き最小2乗法、*s* は単純均等配分の方法、*un-c* はUNのComtradeデータによる推計方法、*u* はUN準拠による方法、*i* は分類間の独立性を仮定した配分構造同一パターンの方法である。

して推計する。分類コードの3桁レベルで得られた配分ウェイトの推計値を4桁レベルへ再配分する。この方法としては1個の3桁レベルの分類コードを複数個の4桁レベルの分類コードに配分することから対応関係のタイプ2の方法で処理できる。しかし、本章ではデータ処理を簡単にするため単純均等配分の方法を適用している。商品グループの*m*が大きくなったときの推計方法については今後も引き続き検討課題として残されている。

貿易データの変換方法として、Comtradeデータの変換方法と本章との違いを具体例として示す。図5にSITC-R2からSITC-R1の方向に対する商品グループ内における商品分類コードの対応関係が示されている。図5の(1)で示されている商品グループ0057、すなわち G_{0057} はSITC-R2の4個の商品分類コード0573,05795,05796,05797とSITC-R1の3個の商品分類コード0513,05201,05195が1つの商品グループを形成している。そのグループの中でSITC-R2の0573が2つに配分され、1つはSITC-R1の0513、もう1つは同05201と対応している。この対応関係に基づいた変換を実現するために、Comtradeデータが採用している変換方法では実線で示した0513のみに対応して、05201は無視されている。それに対して本書の方法は取引額を考慮して配分ウェイトを推計し、このウェイトにより破線で示

図6 報告国日本における SITC-R1 へ変換された *mdcc* の散布図 (1)

(出所) 著者作成。

(注) 年度は1987年である。●は *s* 方式、*は *vm* 方式、+は *un_c* 方式、×は *u* 方式を表している。

されている2つに配分させているところに違いがある。SITC-R2の他の分類コードについても同様である。また、図5の(2)の商品グループ G_{0242} ではSITC-R2の27899が4つに配分される例である。Comtradeデータが採用している変換方法では実線で示した27699のみに対応して、それ以外は無視されるのに対して、本書の方法は取引額を考慮して配分ウェイトを推計し、このウェイトにより4つに配分させている。

図5で示された商品グループの0057と0242は表3において影を付けて示されている。

表5 変換された SITC-R1 における絶対誤差の総和に対する構成比

y	s	$un-c$	u	i
1976	0.993757	0.002420	0.002172	0.001651
1977	0.995157	0.001864	0.001685	0.001294
1978	0.994854	0.002127	0.001688	0.001330
1979	0.939933	0.023526	0.020419	0.016122
1980	0.992172	0.003112	0.002684	0.002032
1981	0.991197	0.004273	0.002553	0.001976
1982	0.992775	0.003295	0.002203	0.001726
1983	0.993866	0.002851	0.001869	0.001414
1984	0.991678	0.003936	0.002433	0.001952
1985	0.992902	0.003396	0.002084	0.001618
1986	0.924556	0.050584	0.013895	0.010965
1987	0.993540	0.003419	0.001697	0.001344

(出所) 著者作成

(注) 制約条件付き最小 2 乗法の wm を基準とした時の対誤差の挿話に対する構成比である。 y は年、 $s, un-c, u, i$ は推計方法の表記を表し、それぞれ表 4 で示された $s, un-c, u, i$ の推計方法である。

この両者の配分ウェイト行列を推計した結果が表4に示されている。Comtradeデータでは対応しなかった配分構造から構成された配分ウェイトのいくつかについては wm 方式と i 方式には0ではない推計値が見られる。 $un-c$ 方式は i 方式における推計値の最大値を1、それ以外を0と置き換えた推計方法である。表4において、商品グループ0057を例にする。このグループは対応関係が4bである。 i 方式ではSITC-R2の0573は2つに配分され、SITC-R1の0513には配分ウェイトとして0.997567、05201には0.002433として配分される。したがって、 $un-c$ 方式では0513には配分ウェイトとして1、05201には0として配分される。この結果は u 方式と一致している。 wm 方式では0513には1、05201には0として配分され、 u 方式と一致している。しかし、 s 方式では u 方式とは一致していない。商品グループ0242については対応関係はタイプ2であるため、 i 方式と wm 方式の推計値は一致する。 $un-c$ 方式は u 方式に一致しているが、 s 方式では u 方式とは一致していない。表4で示された結果では $un-c$ 方式はすべて u 方式に一致している。しかも、 wm 方式に同じ方法を適用しても u 方式に一致しているが、 s 方式では u 方式とは一致していない。

変換後の個別分類コードごとの評価は推計された取引額の中で、分類間の独立を仮定した同一配分方法である i 方式を基準としたときにすべての $mdcc$ に対する取引額の $s, un-c, u, wm$ 方式の散布状態を利用する。日本の 1978 年における輸出データの散布状態は図 6 に示されている。 i 方式の結果を x 軸、 $s, un-c, u, wm$ のそれぞれの方式による結果を y 軸として散布図を作成すると多くの点が右上に集中するので、図 6 では対数変換したものを採用している。図において、●は s 方式、*は wm 方式、+は un_c 方式、×は u 方式を表している。この図では i 方式に一致しているものは対角線上に位置し、 i 方式より大きければ対角線よ

り上に、逆に小さければ対角線より下に位置する。 i 方式は存在するが $s, un-c, u, wm$ 方式のいずれかが存在しないときは x 軸上に、逆のときは y 軸上に位置している。結果として上位に位置するのは s 方式が多く、それに対して下方に位置するものは $un-c, u, wm$ 方式である。図 6 によれば単純均等配分方法である s 方式は誤差が大きいことになる。

変換後の総合的な評価は絶対誤差の総和に対する構成比を利用する。 $mdcc$ 分類コードの個数を N とし、 wm 方式を基準としたときの $mdcc$ 分類コードに対する絶対誤差を ae とする。例えば s 方式については、

$$ae_s = \sum_{i=1}^N |s_i - wm_i|$$

とする。ここで、 s_i と wm_i はそれぞれ s 方式と wm 方式で推計された $mdcc$ 分類コードにおける取引額の推計値とする。その他についても同様である。表 5 は絶対誤差の構成比を示したものである。誤差の構成比 e は、 $ae_s / (ae_s + ae_{un-c} + ae_u + ae_i)$ として表される。その他についても同様に示す。

表 5 によれば誤差に占める割合は圧倒的に s 方式が大きいことがわかる。すなわち、 wm 方式を基準としたとき $i, u, un-c, s$ 方式の順に近似している状態を示している。変換後の $mdcc$ コードにおける総合評価の結果から貿易データ X^*, Y^* の誤差の程度や推計プログラムの難易度を考慮すれば i 方式が現在のところ最良の推計方法と判断される。

おわりに

貿易データを長期時系列で利用するには商品分類の改訂に伴って同一商品においてもカバリッジに変更が生じていることがあり、改訂年を含む連続年次において取引金額や数量は接続という意味において必ずしも整合性のある状態とはいえない。同一商品分類をもとにした貿易データの長期時系列利用のためには、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。本稿はこの配分ウェイトを推計し、新商品分類で編集されている貿易データを旧商品分類で編集されているそれへの変換のための一連の処理過程とそのため必要な各種方法論をまとめたものである。変換後の $mdcc$ コードにおける総合評価の結果から貿易データ X^*, Y^* の誤差の程度や推計プログラムの難易度を考慮すれば i 方式が現在のところ最良の推計方法と判断される

Comtrade データが採用している新商品分類コードから旧商品分類コードへの変換方法は配分構造が生じている場合には商品分類の例示品目を利用して関連の多い分類コードの 1 つに対応させるという方法であるが、基本的には i 方式における特殊な方法である。しかし、この方法は報告国ごとに選択されるのではなくすべての報告国に対して一律に同一の対応表が適用されている。それに対して本章で採用した方法は対応関係コード表をその

まま利用してその配分ウェイトを新旧商品分類に対するそれぞれの報告国から得られた取引額を利用して推計しているところに違いがある。しかし、本章の方法は異なった年に対しては同一の配分ウェイト行列を適用しているところに問題は残されている。

1 野田 [2010] においていくつかの誤りや説明の不完全な箇所があるため、野田 [2012] において再考している。

2 確率過程 $\{y_t\}$ が正常状態であるとは、すべての t に対して平均が $E(y_t) = \mu$ 、分散が $E(y_t - \mu)^2 = \gamma(0)$ 、自己共分散が $E\{(y_t - \mu)(y_{t-\tau} - \mu)\} = \gamma(\tau)$ となることである。一般的に時刻 t における確率過程の平均は時刻を t とするとき、 $\mu_t = E(y_t)$ で表わされる。これは確率変数 y_t の可能なすべての実現値についての平均であり、時刻 t において m 系列の観測値が得られるとき、平均値の推定値は $\hat{\mu}_t = (y_t^{(1)} + \dots + y_t^{(m)})/m$ となる。しかし、時系列データにおいて観測されるのは単一系列のみであり、このままの状態では $\hat{\mu}_t$ を計算することは不可能である。この場合には $\hat{\mu}_t = \mu$ であれば観測値の時間的平均により μ を推計することが可能となる。すなわち、 y_t の平均値が時刻 t に依存しないことである。 $\{y_t\}$ がすべての t に対して $E(y_t) = \mu$ 、 $E(y_t - \mu)^2 = \gamma(0)$ 、 $E\{(y_t - \mu)(y_{t-\tau} - \mu)\} = \gamma(\tau)$ となる条件が満たされるとき、平均や分散、共分散、自己共分散が求められ、確率過程の $\{y_t\}$ は定常状態にあるという。正確に言えば弱定常状態である。多変量確率過程の定常状態は単一変量の拡大で表わされ、平均ベクトルと共分散行列時刻 t に依存せず、自己共分散行列が2時点間の差 τ にのみ依存することである。

3 分類 A から分類 B の方向に対する配分ウェイト行列 W_g が得られ、両者の取引額が存在するとき、逆の方向である分類 B から分類 A の方向の配分ウェイト行列も簡単に計算できる。表2において、分類 A から分類 B の方向の配分額が示されており、配分額を表わす行列は、 $1 \dots k$ の年次 k に対して、 $W_g D(x_{1k} \dots x_{nk})$ である。これを利用して同じ年度 k における分類 B から分類 A の方向の配分ウェイト行列は、 $W_g^* = D(y_{1k} \dots y_{mk})^{-1} W_g D(x_1 \dots x_n)$ として表わされる。

参考文献

- 佐藤英人 [1995] 「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 (SDS) No.67 アジア経済研究所)
- 野田容助 [2002] 「対応関係におけるグループ化とその連結」(『世界貿易マトリクスの作成と評価』調査研究報告書 (開発研究部 2001-III-12) アジア経済研究所)
- [2007] 「商品分類統一のための配分ウェイト行列の推計と変換」(野田容助・黒子正人・編『貿易関連指数と貿易構造』統計資料シリーズ (SDS) No.91 アジア経済研究所)
- [2009] 「UN Comtrade貿易データにおける整合性の評価と補正」(野田容助・黒子正人・吉野久生 編『貿易指数と貿易構造の変化』統計資料シリーズ (SDS) No.93 アジア経済研究所)
- [2010] 「商品分類における対応関係のグループ化と分類の接続」(野田容助・黒子正人 編『貿易指数の作成と応用：貿易構造の変化と国際比較』調査研究報告書 開発研究センター2009-II-03 アジア経済研究所)

- [2012] 「貿易データ変換のための対応関係コード表のグループ化と連結」 (野田容助・黒子正人編『 』統計資料シリーズ No.96 アジア経済研究所) (出版予定)
- 野田容助・山本泰子 [1995] 「体系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変換の試み—」 (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 (SDS) No.67 アジア経済研究所)
- 古河俊一・野田容助 共著 [1998] 『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 (SDS) No.80 アジア経済研究所
- Hideto Sato [1983] , *Fundamental Concept of Social/Regional Summary Data and Inference in Their Database*, Doctoral Thesis, University of Tokyo

