

第 1 部

第1章

UN Comtrade Database貿易データによる貿易指数作成

— 貿易指数作成の準備段階としての基礎的課題とデータ処理方法 —

野田容助

要約

UN Comtrade Database貿易データを基礎として作成されるアジア経済研究所の世界貿易統計データシステム (AID-XT) の概要とそれを基にした貿易関連指数の準備段階としての貿易データに関する基礎的課題とそのために必要とされるいくつかの具体的なデータ処理方法を紹介する。

キーワード

UN Comtrade Database貿易データ、AID-XT 基礎データ、貿易データにおける整合性の評価と補正、報告国の世界合計

はじめに

世界貿易マトリクスは国・地域や商品分類などの分類カテゴリーに関して共通の分類基準を採用している貿易データを必要とする。共通分類基準に従って編集されており一般的に広く利用されている貿易データの1つは International Trade Statistics Section, TSB, UN Statistics Division 作成による UN Comtrade Database 貿易データで

ある。アジア経済研究所が整理し、維持・管理している世界貿易統計データシステム:AID-XT (Ajiken Indicators of Developing economies: eXtended for Trade statistics) もこの UN Comtrade Database 貿易データを基礎としている。

本調査研究報告書別冊はUN Comtrade Database貿易データに基づいて作成されるアジア経済研究所の貿易価格および数量指数、産業内貿易指数 (IIT指数) 顕示的比較優位指数 (RCA指数) について報告国を東アジア諸国・地域である韓国、中国、台湾、香港、フィリピン、マレーシア、シンガポール、タイ、インドネシア、日本および米国の11カ国として表で表わし、図示している。長期時系列データを作成するためにUN Comtrade Database貿易データのすべての年をカバーする1962年から2003年までを対象年とし、指数分類を木下・山田方式による産業の20部門分類を採用している。紙面の関係上本書で紹介するのはその一部であるが、残りの部分はCD-ROMにて利用可能である。

本章はUN Comtrade Database貿易データに基づいて作成されるアジア経済研究所の世界貿易統計データシステム (AID-XT) の概要とそれを基にした貿易価格および数量指数、産業内貿易指数 (IIT指数)、顕示的比較優位指数 (RC

A指数)の準備段階としての貿易データに関する基礎的課題とそのための具体的なデータ処理方法を紹介することを目的としている。

第1節はUN Comtrade Database 貿易データの特性とAID-XT基礎データについてデータを利用するさいの基礎事項を紹介する。第2節ではアジア経済研究所が試みている商品分類における詳細分類コードの取引額をすべて合計すると商品総額に一致することを整合性の評価基準として、可能な限り整合性を保つように補正する方法と具体例を示す。第3節は商品分類のSITC-R1から木下・山田方式による産業20分類への変換方法の紹介である。第4節において商品分類の4桁レベル分類コードにおける報告国の世界合計あるいは地域合計を作成するための処理過程と作成方法を示す。

1. UNC-Oデータと新AID-XT 基礎データ

アジア経済研究所が整理し、維持・管理している世界貿易統計データシステム：AID-XT (Ajiken Indicators of Developing economies: eXtended for Trade statistics) は旧AID-XTと新AID-XTの2種類が存在する。旧AID-XT基礎データはUN貿易統計、OECD貿易統計、台湾貿易統計から構成されており、それぞれの作成機関の違いによるデータ固有の特性をアジア統一コードを使用して共通に利用できるようにしている。UN貿易統計はInternational Trade Statistics Section,

TSB, UN Statistics Division作成によるUN Comtrade Database貿易データであり、総務省統計局統計基準部国際統計課がUNから毎年購入し、当研究所がこのデータを整備し、維持・管理することになっていた。このデータには商品分類がUN作成によるSITCの体系および関税協力理事会 (Customs Co-operation Council : CCC) 作成によるHSの体系の違いに関わらず、商品総額も含めてすべての桁レベルの商品分類コードが

存在する。OECD貿易統計はアジア経済研究所がOECDから直接購入する貿易統計データであり、このデータも同じような商品分類コードから構成されている。

台湾貿易統計はアジア経済研究所が台湾財政部関税總局統計室 (Statistical Department Directorate General of Customs Ministry of Finance, The Republic of China) から直接購入する貿易統計データであり、当研究所の独自の方法によりUN貿易統計に準拠した内容および形式に変換している。台湾のAID-XT基礎データ作成については野田の「台湾のAID-XT基礎データ作成と評価」を参照すること。

新AID-XT基礎データはUN統計局が2003年から開始したon-line検索によるUN Comtrade貿易統計データから得られたUN貿易統計と台湾貿易統計から構成される。旧AID-XT基礎データがOECD加盟国のデータとしてOECD貿易統計を採用していたのに対して新AID-XT基礎データは台湾以外の国についてはUN貿易統計に一元化したところに特徴がある。また、再輸出を含めて輸出と定義しているようにon-line検索によるUN Comtradeの概念の変更あるいは新規項目の追加にともなって旧および新AID-XT基礎データはそれぞれの分類カテゴリーあるいは統計値に違いが生じていることに注意する必要がある。したがって、旧および新の混在した利用はせずに、別系列として利用をする必要がある。

貿易統計データは一般的には貿易統計固有の項目である7個の分類カテゴリーと3個の統計値から構成されている。on-line検索により得られるUN Comtrade Databaseの貿易データを対象とすれば、分類カテゴリーの項目は報告国 (reporter code : その省略形はrc) 、輸出入区分 (trade flow code またはdirection of trade : d) 、商品分類 (classification またはsystem of commodity classification : sc) 、商品分類コード (commodity code : c) 、相手国 (partner code : pc) 、年

(year : y)、数量単位 (quantity unit : qu) であり、後者の項目は取引金額 (value : v) と2種類の数量であるkg表示の重量数量 (netweight (kg) : qw) と取引数量 (Supplementary Quantity : q) である。

1.1 UNC-Oデータ

アジア経済研究所ではon-line検索によって得られたすべての分類カテゴリーと統計値を含む報告国ごとに作成されたUN Comtrade Databaseの貿易統計データをUN Comtrade Databaseのオリジナル貿易統計データといい、UNC-Oデータと表わすことにしている。商品分類コードは商品総額とすべての桁レベル分類コードを含んでいるのに対して相手国は世界合計と個別相手国のみから構成され、地域合計や経済ブロック計は含まれていない。UN Comtrade Database貿易データの特徴は各国作成の元データをもとに商品分類を新規改訂版から旧改訂版に変換して同一分類によるデータ系列を作成していることである。例えば、報告国が日本の場合には元データは1976年から1987年についての商品分類は新改訂版はSITC-R2により編集されている。このSITC-R2による分類を旧改訂版であるSITC-R1へ変換している。結果として1962年から2003年までのSITC-R1による貿易データを作成している。同じようにして、1976年から2004年まではSITC-R2による貿易データが存在する。すなわち、それぞれの報告国は年度に違いはあるもののSITCおよびHSの各改訂版ごとの貿易データが揃っている。

UNC-Oデータの例が表1に示されている。分類カテゴリーと統計値の順番はon-line検索により貿易統計データを取り出すときの指定に依存しており、アジア経済研究所では旧AID-XT基礎データのフォーマットに準拠した順番を採用している^(注1)。表1からわかるとおりUNC-Oデータはcsv形式で保存されており分類カテゴリー

の中で、報告国、輸出入区分、相手国、年、数量単位と統計値の金額と2種類の数量は整数による数値表示、分類カテゴリーの商品分類と商品分類コードは文字表示のためダブルクォテーション(")によって囲まれて表示されている。UN Comtrade Databaseの貿易データの2004年度版では統計値の欠損値はN/A (Not Available data) であるが、同2005年度版ではそれはNULLとなっている。NULLはデータが存在していないという意味であり、..として、の間に何も存在していない状態で表わされている。表1において、影を付けたような形式で表現され、カンマの間に囲まれた状態、またはカンマのみで表わされたカンマと終了記号とで囲まれた状態で表わされる。したがって、表1からわかるようにUNC-Oデータの個々のデータの長さは一定ではない。商品分類コードの桁レベル分類コードはダブルクォテーションを取り除いた文字列の個数により確かめることができる。しかし、商品総額は"TOTAL"として表わされているため特別な処理が必要となる。

アジア経済研究所では on-line 検索による UN Comtrade Database の貿易データを当研究所の統計サーバである iders12 の/ts/org/un/2005_IE/へ商品分類体系、報告国別に text ファイルとして保存しており、商品分類体系と報告国の国コードを組み合わせてファイル名を付けている。例えば、商品分類体系が SITC-R1 である日本の場合には s1.392 をファイル名としている。しかし、on-line 検索の作業場の都合から同一1つのファイルにすべての貿易データを保存させるのは難しいので、ある適当な容量ごとのファイルを作成し、一連番号でファイルを連結させている。さらに、保存されたデータ形式は csv 形式であるためそれらの情報をファイル名に表わして、s1.392_1.csv がファイル名である。表2に商品分類が SITC-R1 であるファイルの一覧の一部が示されている。この表から報告国の日本の 392 は

表1 on-line検索により得られたUN Comtrade Databaseオリジナル貿易統計データ (UNC-O) の例

Reporter Code, Trade Flow Code, Classification, Commodity Code, Partner Code, Year, Value, Netweight (kg), Quantity Unit Code, Supplementary Quantity

```

36, 1, "S1", "0", 0, 1963, 96398464, N/A, 0, N/A
36, 1, "S1", "00", 0, 1963, 1700572, N/A, 0, N/A
36, 1, "S1", "075", 0, 1963, 1603139, 0, 1, 0
36, 1, "S1", "0751", 0, 1963, 660553, 0, 1, 0
36, 1, "S1", "0752", 0, 1963, 942586, 0, 1, 0
36, 1, "S1", "07521", 0, 1963, 134259, 0, 1, 0
36, 1, "S1", "07529", 0, 1963, 808327, 0, 1, 0
36, 1, "S1", "08", 0, 1963, 2148442, N/A, 0, N/A
36, 1, "S1", "081", 0, 1963, 2148442, 0, 1, 0
36, 1, "S1", "08", 0, 1963, 2148442, , 0,

```

(出所) on-line 検索による UN Comtrade Database の貿易統計データ

(注) 報告国オーストラリア (aus:036) の例である。フォーマットはcsv (comma separated value) 形式で表示されている。

表2 on-line 検索により得られた UN Comtrade Database の SITC-R1 系列のデータ保存状況

/ts/org/un/2005_IE/s1.004_1.csv	/ts/org/un/2005_IE/s1.414_1.csv
/ts/org/un/2005_IE/s1.008_1.csv	:
:	/ts/org/un/2005_IE/s1.841_1.csv
/ts/org/un/2005_IE/s1.392_1.csv	/ts/org/un/2005_IE/s1.841_2.csv
/ts/org/un/2005_IE/s1.392_2.csv	/ts/org/un/2005_IE/s1.842_1.csv
/ts/org/un/2005_IE/s1.392_3.csv	/ts/org/un/2005_IE/s1.842_2.csv
/ts/org/un/2005_IE/s1.392_4.csv	/ts/org/un/2005_IE/s1.842_3.csv
/ts/org/un/2005_IE/s1.398_1.csv	/ts/org/un/2005_IE/s1.842_4.csv
/ts/org/un/2005_IE/s1.400_1.csv	/ts/org/un/2005_IE/s1.850_1.csv
/ts/org/un/2005_IE/s1.404_1.csv	/ts/org/un/2005_IE/s1.854_1.csv
/ts/org/un/2005_IE/s1.410_1.csv	:
/ts/org/un/2005_IE/s1.410_2.csv	/ts/org/un/2005_IE/s1.894_1.csv

(出所) 著者作成

(注) アジア経済研究所における UNC-O のデータ保有状況であり、/ts/org/un/2005_IE/はディレクトリー、s1.004_1.csvはファイル名である。

4 個、韓国の 410 は 2 個、米国は国コードが別れていて 1980 年以前に使用されてる 841 が 2 個、それ以降に使用されている 842 が 4 個、US 領バージンアイランドの 850 が 1 個の合計して 7 個のファイルから構成されていることがわかる。したがって、データ処理にさいしては同一商品分類体系における報告国のすべてをカバーするために部分集合として作成されている報告国のファイルをすべて指定することが必要である。例えば、日本における SITC-R1 系列のファ

イルを得るには UN 国コードの 392 利用して、
`ls /ts/org/un/2005_IE/s1*392* > a`
 として得ることができる。a は任意の一時的な Unix のファイルである。

1.2 新AID-XT基礎データ

貿易統計で使用される貿易商品分類は商品分類体系としては基本的には UN 作成の標準国際商品貿易分類 (SITC: Standard International Trade

表3 *mdcc*と個別相手国をもとに作成された貿易マトリクスの取引額表

<i>C</i>	<i>P</i>	P_1	P_j	P_n	<i>Error of P</i>	<i>World</i>
C_1		x_{11}	x_{1j}	x_{1n}	$e_p(1)$	x_{1W}
:		:	:	:	:	
C_i		x_{i1}	x_{ij}	x_{in}	$e_p(i)$	x_{iW}
:		:	:	:	:	
C_m		x_{m1}	x_{mj}	x_{mn}	$e_p(m)$	x_{mW}
<i>error of C</i>		$e_c(1)$	$e_c(j)$	$e_c(n)$	$e_{c,p}$	$e_c(\bullet) + e_{c,p}$
<i>Total</i>		x_{T1}	x_{Tj}	x_{Tn}	$e_p(\bullet) + e_{c,p}$	x_{TW}

(出所) 野田容助「世界貿易マトリクス作成における整合性の評価と補正」(『改訂版世界貿易マトリクス—国際産業連関表24部門分類にもとづいて—』SDS No.84 改訂版)の表1を引用

(注) 影の部分は実際に得られるデータである。完全に整合性の取れた貿易マトリクスの取引表では*error of P*および*error of C*の部分がすべて0で表現される。

表4 相手国*P_j*が欠損値であるときの貿易マトリクスの取引額表

<i>C</i>	<i>P</i>	P_1	P_j	P_n	<i>Error of P</i>	<i>World</i>
C_1		x_{11}	0	x_{1n}	x_{1j}	x_{1W}
:		:	:	:	:	
C_m		x_{m1}	0	x_{mn}	x_{mj}	x_{mW}
<i>Error of C</i>		0	0	0	0	0
<i>Total</i>		x_{T1}	0	x_{Tn}	x_{Tj}	x_{TW}

(出所) 表1にもとづき著者作成

(注) P_j が欠損値であるので影の部分はすべて0に置き換えられるが、替わって*error of P*にその分が誤差として表示される。

Classification) 系列と関税協力理事会が作成する国際統一商品分類あるいは統一システム(HS:Harmonized Commodity Description and Coding System) 系列が存在する(注2)。SITCの商品分類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商品分類コードから構成されている。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。アジア経済研究所では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0でなく、しかも下位レベルの階層の分類コードを持たないものを詳細分類コード(*mdcc*: the most detailed classification code)と呼んでいる(注2)。*mdcc*分類コードの特徴は世界合計を含む個別相手国を固定すれば整合性が保障されているときには桁レベル分類コードとは

異なっており、これらに対応する取引額を合計すると商品総額に一致することである。

次節で紹介するようにアジア経済研究所では商品分類コードにおける詳細分類コード*mdcc*の取引額をすべて合計すると商品総額に一致することを整合性の評価基準としている。この基準によれば桁レベルの分類コードによるUNC-OデータはSITC系列では3桁レベル分類コードまでなら整合性は確保できるものの、4,5桁レベル分類コードから構成される基本分類(item)は必ずしも整合性のあるものとはいえない。HS系列についても2,4桁レベル分類コードでは整合性があっても6桁レベル分類コードのsub-headingでは必ずしも整合性のあるものとはいえない。

AID-XT基礎データはUNC-Oデータを可能な

限り整合性のあるように補正した $mdcc$ により置き換えた貿易データである。補正についての説明は次節で触れている。ここで注意することはAID-XT基礎データはUNC-Oデータを加工したり修正したりするのではなく、上位桁レベル分類コードに置き換えて補正していることである。

旧AID-XT基礎データはUNから直接購入するUN Comtrade Database貿易データ、新AID-XT基礎データはon-line検索によって得られるUN Comtrade Database貿易データを基礎としているため、データの特性がそのままそれぞれのAID-XT基礎データに直接反映している。説明は省略するが両UN Comtrade Database貿易データは分類カテゴリーおよび統計値とも明確に定義が異なっているため、両AID-XTとも別系列として取り扱う必要がある。また、今後はon-line検索によるUN Comtrade Database貿易データによるデータの追加、更新しかおこなわれなくなるため新AID-XT基礎データが利用の対象となる。

2. 貿易データの整合性評価と補正

アジア経済研究所ではコードにおける詳細分類コード $mdcc$ の取引額をすべて合計すると商品総額に一致することを整合性の評価基準としている。すなわち、サムチェックにもとづく評価方法である。貿易統計データにおける取引額にもとづく整合性の評価および補正については野田・深尾の「貿易マトリクス作成における整合性の評価—新および旧AID-XT基礎データにもとづいて—」に説明されている。これによると、表3に示されているように完全に整合性の取れた貿易マトリクスの取引額表が存在するとき、 $j=1\cdots n$ に対して $e_c(j)=0$ であり $i=1\cdots m$ に対して $e_p(i)=0$ 、 $e_{c,p}=0$ である。したがって、記号 \bullet をすべての要素の合計とすれば、相手国による誤差は $e_p(\bullet)+e_{c,p}=0$ 、商品分類コードによる誤差は $e_c(\bullet)+e_{c,p}=0$ 、総合誤差は $e=0$

となる。

(1) 完全に整合性の取れた貿易マトリクスにおいて本来存在すべき相手国の P_j がすべて欠損値であるとする。すなわち、表4において P_j の列の要素がすべて0となるものとする。整合性のある貿易マトリクスなので $error\ of\ P$ の列に P_j のすべての要素が現れる。これを誤差で表現すれば $e_{c,p}=0$ 、 $e_p(\bullet)+e_{c,p}=x_{Tj}$ 、 $e_c(\bullet)+e_{c,p}=0$ 、 $e=x_{Tj}$ となる。 P_j が欠損値であることは貿易統計の作成機関のデータが正しいとしたときに相手国コードの中でアジア研統一国コードに対応していないものが存在することを意味する。

(2) 完全に整合性の取れた貿易マトリクスの取引額表において本来存在すべき $mdcc$ の C_i がすべて欠損値であることは貿易統計の作成機関のデータが正しいとしたときに $mdcc$ に検討すべきものが存在することを意味する。

(3) 完全に整合性の取れた貿易マトリクスの取引額表において本来存在すべき相手国の P_j と $mdcc$ の C_i のすべて欠損値であることは貿易統計の作成機関のデータが正しいとしたときに相手国の対応付けに間違いがあり同時に $mdcc$ に検討すべきものが存在することを意味する。

この結果を逆に利用することにより、整合性の評価表から貿易統計データの状態を知ることができる。場合によっては補正も可能となる。旧AID-XT基礎データにおいて(2)に対応する商品分類コードについてはここ数年にかけて不整合の箇所を補正あるいは調整してきている。しかし、(1)あるいは(3)に示されるような相手国、数量単位あるいは数量についてはようやく評価

の方法論がまとまった段階であるため整合性の評価は今後の課題として残されている。

2.1 不整合における補正の方法

商品分類コードの整合性に欠ける状態は下位レベル分類コードの取引額合計が対応する上位

レベル分類コードの取引額と一致しないものが存在することである。一般に商品分類コードの桁レベルにおける整合性の評価により上位桁レベル程整合性が高いことが知られている。したがって、階層構造を持つ桁レベル分類コードにおいて、 k 桁レベル分類コードの取引額と k 桁レベル分類コードが同一であるすべての $k+1$ 桁レベル分類コードの取引額合計の差が大きいときには $k+1$ 桁レベルの分類コードを使わずにその上位の桁レベルである k 桁レベルのそれを使用して整合性を高めることが必要になってくる。このような方法により整合性を高める処理を $k+1$ 桁レベル分類コードに対する補正あるいは k 桁レベル分類コードによる補正という。SITC系列において k 桁レベルで表わされた商品分類コードの取引額を $x(i_1, \dots, i_k)$ 、 k 桁レベルで合計された $k+1$ 桁レベル分類コードの取引額合計を $x(i_1, \dots, i_k, \bullet)$ とする。 k 桁レベル分類コードによる補正は絶対誤差と相対誤差の2種類によって補正基準が決められる。絶対誤差による補正は絶対誤差を、

$$(2-1) \quad \alpha_k = |x(i_1, \dots, i_k) - x(i_1, \dots, i_k, \bullet)|$$

として、その補正基準を α_k^* とする。補正は、

$$(2-2) \quad \alpha_k \geq \alpha_k^*$$

であるとき、(2-2)式を満たす k 桁レベルの分類コードは $k+1$ 桁レベル分類コードが存在していてもそれを無視して k 桁レベル分類コードを採用することにする。(2-2)式を満たさないときには $k+1$ 桁レベル分類コードを採用する。ここで、 $\alpha_k^* \geq 0$ を k 桁レベルにおける絶対補正係数という。また、相対誤差による補正基準は相対誤差を、

$$(2-3) \quad \beta_k = \alpha_k / x(i_1, \dots, i_k)$$

として相対誤差の補正基準を β_k^* とすれば、

$$(2-4) \quad \beta_k \geq \beta_k^*$$

となるとき、(2-4)式を満たす k 桁レベルの分類コードは $k+1$ 桁レベルの分類コードが存在していてもそれを無視して k 桁レベル分類コードを採用することにする。(2-4)式を満たさな

いときには $k+1$ 桁レベル分類コードを採用する。 β_k^* を相対補正係数といい、 $0 \leq \beta_k^* \leq 1$ である。

補正係数が $\alpha_k^* = 0$ のときは β_k^* も0となり、すべての $k+1$ 桁レベル分類コードを k 桁レベルのそれへ置き換えることを意味する。一方では、 $\beta_k^* = 1$ のときは丸めの誤差により必ずしも正しくはないが、一般に貿易統計データでは、

$$x(i_1, \dots, i_k) - x(i_1, \dots, i_k, \bullet) > 0$$

であるので(2-4)式において $x(i_1, \dots, i_k) < 0$ となる。このようなことは起こり得ないので $k+1$ 桁レベル分類コードから k 桁レベル分類コードへの置き換えはないことになる。このようにして α_k^* および β_k^* の値を適当に選択することにより $k+1$ 桁レベル分類コードに対する補正をするときと同じように補正をしないときにも利用できる。商品分類体系がHS系列でも同じようなことが生ずる^(注3)。

2.2 不整合補正の具体例

前掲した野田・深尾によればon-line検索によるUN Comtradeのmdccにおいて米国の1981年に不整合の状態が知られている。この不整合を検討し、その補正を試みる。補正に先立って補正する桁レベルを確かめるため、桁レベル分類における整合性の評価が必要である。桁レベル分類における整合率評価表はon-line検索によるUN Comtradeの桁レベル分類コードの貿易統計データをもとにする。表5は米国の1981年における桁レベル分類コードの整合率表である。この表のdigitに桁レベルが示されている。

表5において米国の桁レベルが1のときには総合誤差 e は小さく整合性があると見なされる。利用された商品分類コードの個数は d_1 で表わされているように10である。同じようにして2桁レベル、3桁レベルでも不整合の状態ではないことが確かめられる。ところが4桁レベルになると総合誤差における絶対誤差が12564となり、商品総額を基準とする相対誤差は表では桁落ちになっ

表5 on-line 検索の UN comtrade Database 貿易データ、米国（1981）桁レベル分類コードの整合性

<i>digit</i>	v_{TW}	e	d_1	d_2	d_3	d_4	d_5	d_6	$e_c + e_{c,p}$	$e_p + e_{c,p}$			
(import)													
1	273352196	1	0.0000	10	0	0	0	0	1	0.0000	-1	-0.0000	
2	273352196	1	0.0000	0	66	0	0	0	-1	-0.0000	-1	-0.0000	
3	273352196	1	0.0000	0	0	236	0	0	0	0.0000	-1	-0.0000	
4	273352196	12564	0.0000	0	0	0	734	0	0	0.0000	-1	-0.0000	
<i>mdcc</i>	273352196	60012	0.0002	0	0	0	358	1112	0	60012	0.0002	-1	-0.0000
(export & re_export)													
1	233666101	5	0.0000	10	0	0	0	0	5	0.0000	3	0.0000	
1	233666101	5	0.0000	0	67	0	0	0	5	0.0000	3	0.0000	
1	233666101	5	0.0000	0	0	236	0	0	5	0.0000	3	0.0000	
1	233666101	588749	0.0025	0	0	0	706	0	0	588750	0.0025	3	0.0000
<i>mdcc</i>	233666101	1975495	0.0085	0	0	0	375	984	0	1975495	0.0085	3	0.0000

(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—』における第1章の表3から一部引用

(注) 米国の商品分類はSITC系列、 v_{TW} の単位は1,000US\$である。

表6 on-line 検索による UN Comtrade Database 米国、1981年の桁レベル分類コードの補正過程

<i>Corr</i>	v_{TW}	e	d_1	d_2	d_3	d_4	d_5	d_6	$e_c + e_{c,p}$	$e_p + e_{c,p}$			
(import)													
2	273352196	1	0.0000	0	66	0	0	0	-1	-0.0000	-1	-0.0000	
2←3	273352196	1	0.0000	0	0	236	0	0	0	0.0000	-1	-0.0000	
3←4	273352196	2	0.0000	0	0	1	732	0	0	0.0000	-1	-0.0000	
4←5	273352196	17	0.0000	0	0	1	357	1111	0	17	0.0000	-1	-0.0000
(export & re_export)													
2	233666101	5	0.0000	0	67	0	0	0	5	0.0000	3	0.0000	
2←3	233666101	5	0.0000	0	0	236	0	0	5	0.0000	3	0.0000	
3←4	233666101	5	0.0000	0	0	2	702	0	0	5	0.0000	3	0.0000
4←5	233666101	5	0.0000	0	0	2	383	946	0	5	0.0000	3	0.0000

(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—』における第1章の表6から一部引用

(注) *corr*は補正の桁レベルを表わしており、 $k \leftarrow k+1$ は k 桁レベル分類コードにより $k+1$ 桁レベル分類コードの補正をおこなうことを意味する。整合性の評価表は補正後のものを対象としている。記号←のないものは桁レベルの整合性評価である。単位は1,000US\$である。

ていて見えないが0.00005となる。これを誤差と見なすかそうでないかは誤差の基準をどのように設定するかで決まる。本節ではすべての桁レベル k に対して絶対補正係数 α_k を100、相対補正係数 β_k を0.0001と設定してそのうちのどちらかを満たせば誤差としている。この誤差基準によれば米国の1981年では4桁レベル分類コードで整合性がなくなったことになる。すなわち、3桁レベル分類コードでは整合性があり4桁レベルのそれでは整合性がないので補正は4桁レベル分類コードに対しておこなわれる。*mdcc*においても当然総合誤差が観測されることになるが、4桁レベルと比較して絶対誤差および相対誤差

がほぼ同じ値を持てば5桁レベル分類コードから4桁レベルのそれへの補正は生じない。しかし、表5から*mdcc*における総合誤差は大きくなっているため5桁レベル分類コードによる補正も必要であることになる。

桁レベルの補正は上位桁レベルから下位桁レベルの方向に処理される。表6に米国の1981年における桁レベルの補正過程が整合率表で表示されている。誤差基準によって補正がおこなわれるのは3桁レベル分類コードによって4桁レベル分類コードを置き換えることであつたので最初の処理は3桁レベル分類コードによる補正である。補正の結果に対する整合性の評価は表6に示

されており、この補正は*corr*として3←4で表わされる。輸入では4桁レベル分類コードの個数は表5の d_4 から734個あることが知られているが、補正により7223と7224の2個の個別分類コードが取り除かれたため、表6の d_4 ではそれが732個となり、3桁レベル分類コードの722が追加されたので表5において0であった d_3 は表6において1となっている。この補正により4桁レベル分類コードの整合性は表5で示された総合誤差の絶対誤差の12564から表6の*corr*における3←4の総合誤差の絶対誤差の2へと減少しており、4桁レベル分類コードの整合性を確認できる。

輸入についての次の処理は補正された4桁レベル分類コードにより5桁レベル分類コードを補正することである。具体的な置き換えは省略し結果だけを紹介する。5桁レベル分類コードにおいて表5の*mdcc*では d_5 は1112であるのに補正結果の表6では*corr*の4←5で示されている d_5 は1111であり、1個の5桁レベル分類コードが取り除かれたことがわかり、必然的に1個の4桁レベル分類コードが追加されたことになる。4桁レベル分類コードにおいて表5の*mdcc*では d_4 は358であるのに補正結果の表6では*corr*の4←5で示されている d_4 は357である。これは、最初の処理において2個の個別分類コードが取り除かれ、2番目の処理において1個が追加されたため結果的に1個の減少があったことから得られる。この2回の補正により*mdcc*の整合性は表5で示されている総合誤差の絶対誤差の60012から表6の*corr*の4←5で示されている総合誤差の絶対誤差の17へと減少しており、*mdcc*の整合性が保証されていることを確認できる。再輸出を含む輸出も同じように処理され補正された*mdcc*については整合性が保証される。

2.3 新AID-XT基礎データの課題

新 AID-XT 基礎データすなわち補正された UNC-O データの *mdcc* は野田編の『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—』における第4部の表1「東アジア諸国・地域および米国における補正された新 AID-XT 基礎データの整合率評価」に示されている。この中にはまだ on-line による UN Comtrade の抱える問題点がそのまま現れてきており、不整合の状態が明らかにされていないものや、明確な補正方法が未だに見つかっていないもの、補正が完了していないものも含まれている。同書の第4部、表1によれば東アジア諸国・地域および米国を対象としたとき商品分類について今後検討する必要があるいくつかの問題点が存在する。

(1) 1999年においてHSの2桁レベル分類コードで補正されている報告国は中国、韓国、日本、マレーシア、シンガポール、インドネシア、タイである。

(2) 中国は1995年の輸出入、韓国は1999年以外にも88年から98年まで輸出入の一部、シンガポールは96と97年の輸出、フィリッピンは2003年に(1)に含まれる年が存在する。

(3) 中国の1995年の輸出入、インドネシアの1995年輸出では2桁レベル分類コードで補正してもなお大きな誤差が残っている。

(4) 日本の1992年は当該国の貿易統計がHSであるにもかかわらず、それが存在せずSITC-R3とSITC-R2による編集のみ存在する。

(5) フィリピンの1982年輸出に3桁レベル分類コードによる補正が多い。

(6) タイは1988年の輸出、2002年の輸出入は存在しない。

(7) インドネシアの1962年の輸出は存在しているが同年の輸入、1963年から1966年までの輸出入は存在しない。

(8) マレーシアについては1962年と1963年は報告国としてマレーシアが存在せずにマレーシア半島、サバ州に分かれており、相手国もマレーシア半島、サバ州およびサラワク州に分かれている。

また、中国の1984年から86年まではSITC-R1による2桁レベル分類の編集であるため、作成された貿易関連指数の精度についてはこの期間については保証されていないため利用しないほうがいい。また、中国とインドネシアの1995年の誤差はHS98コードによる委託加工等の特殊な対応付けによる誤差である。

3. 異なる分類に対する変換

貿易統計を共通の概念で長期時系列的に分析をするときには商品分類の改訂前後のどちらかの同一商品分類体系へ統一することが必要になる。商品分類体系統一化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エイトを推計し、この配分ウエイトでそれぞれの分類コードに対応する取引金額および数量を再配分することで可能となる。同じような状態は本書が採用している商品分類コードを産業分類あるいは別の分類体系へ変換するときにも発生する。

アジア経済研究所ではこれまで配分ウエイトによる貿易統計の変換の方法として商品グループ内における配分ウエイトの構造を定式化し、

(1) 商品分類体系における対応関係から得られる配分構造が均等に配分されるという仮定の下で推計される均等配分の方法、(2) 配分構造に対する分類コードの取引額をウエイトとして配分されるという配分ウエイトの方法、という基本的には2つの方法が試みられてきている。商品グループ内における配分ウエイトの推計方法については野田の「商品分類の対応関係における配分ウエイトの推計方法」において説明されている。

アジア経済研究所の変換表は実際の貿易データの変換において取引額を考慮せずに商品分類体系における対応関係から得られる配分構造をもとにした均等配分の方法が広く利用されていた。アジア経済研究所が1970年後半、SITC-R2

からSITC-R1への変換モデルとして採用していたのは基本的にはこの均等配分の方法である。SITC-R2は1960年以降の対外貿易の構造変化を考慮して分類の有用性を増大させるための修正あるいは技術の進歩によって分類を拡張するために作成されたSITC-R1の改訂版である。そのため、SITC-R2はSITC-R1を拡張するための修正をのぞいては基本的にはSITC-R1の構造を保っている。SITC-R2における分類の構造からSITC-R2からSITC-R1への変換は配分構造を含んでいるとはいえ基本的には統合型の対応表が基本となる。野田・深尾の「同一商品分類に変換された貿易額の比較—配分ウエイトにおける推計方法の違いを中心に—」によれば取引額を考慮しない均等配分方式は誤差の程度が大きく、利用については問題がありと指摘している。

本書では商品分類SITC各改訂版から「木下・山田方式」による産業分類20部門対応表を用いた産業分類への変換であり、両分類に対応するそれぞれの統計値が存在しないため統計値に基づく配分構造は考慮できない。序章の付表1に示されている木下・山田方式による変換表は基本項目だけでなく上位レベル分類コードまでも変換式で表わされ、控除される分類コードにはが利用されている。しかし、配分構造を持たず、配分ウエイトがすべて1と設定されているところに特徴がある。基本的には入力をSITC系列の2桁レベル分類コード、出力を2桁で表わされた産業分類の20部門として変換表を作成し、SITC系列の下位レベル分類コードで対応関係を調整した変換表である。

表7は序章の付表1からSITC-R1変換表のみを取り出してSITC-R1順に並べ替えたものである。この方法の特徴は産業分類に対してSITC-R1の2桁レベル分類コードで対応させ、配分構造は持っていないが下位レベル分類コードが上位レベルとは異なる産業に属することを可能にさせ、できるだけ少ない分類コードを利用してモデル化していることである。SITC-R1の4,5,9

表7 木下・山田方式によるSITC-R1から産業分類の20部門への変換表 (KY20)

20 部門 SITC	20 部門 SITC	20 部門 SITC	20 部門 SITC	20 部門 SITC
01 +00	03 -081	02 +27	07 +63	04 +8414
03 +01	03 +09	01 +2711	08 +64	05 -8414
03 +02	03 +11	02 -2711	04 +65	05 -8416
03 +03	03 +12	02 -2732	04 -6518	11 +8416
01 +04	03 +21	12 +2732	12 +6518	05 -842
01 -0422	01 +22	02 -2763	04 -654	06 +842
03 +0422	01 +23	03 +2763	05 +654	06 +85
01 -046	01 -2311	02 -2766	04 -6557	19 +86
03 +046	09 +2311	13 +2766	05 +6557	10 +8623
01 -047	01 -2312	02 +28	04 -6566	19 -8623
03 +047	10 +2312	02 -282	05 +6566	10 +8624
01 -048	01 -2313	13 +282	04 -6578	19 -8624
03 +048	09 +2313	02 -284	20 +6578	19 -863
01 +05	01 -2314	14 +284	12 +66	20 +863
01 -052	09 +2314	01 +29	12 -667	20 +89
03 +052	01 +24	01 -2929	20 +667	19 +891
01 -053	01 -243	03 +2929	13 +67	20 -891
03 +053	07 +243	02 +32	14 +68	17 +8911
01 -0546	08 +25	02 -3215	15 +69	19 -8911
03 +0546	01 +26	11 +3215	16 +71	08 +892
01 -055	01 -2612	02 -3218	16 -7114	20 -892
03 +055	04 +2612	11 +3218	18 +7114	18 +8941
03 +06	01 -2613	02 +331	17 +72	20 -8941
01 +07	04 +2613	11 +332	17 -7261	15 +8943
01 -0713	01 -2622	11 +34	19 +7261	20 -8943
03 +0713	04 +2622	02 +3411	17 -7295	10 +8993
01 -0722	01 -2626	11 -3411	19 +7295	20 -8993
03 +0722	04 +2626	21 +35	18 +73	19 +8996
01 -0723	01 -2627	03 +4	13 +7358	20 -8996
03 +0723	04 +2627	03 -4311	18 -7358	21 +9
01 -073	01 -2628	10 +4311	15 +81	16 +951
03 +073	04 +2628	03 -4312	15 -8124	21 -951
01 -074	01 -2629	10 +4312	17 +8124	15 +961
03 +074	04 +2629	03 -4313	07 +82	21 -961
01 -075	01 -266	10 +4313	06 +83	
03 +075	10 +266	10 +5	05 +84	
03 +08	01 -267	06 +61	05 -8413	
01 +081	04 +267	09 +62	06 +8413	

(出所) 木下宗七・山田光男「国別・商品別デフレータの推計と若干の吟味—国連貿易統計による—」(名古屋大学経済学部附属経済構造研究センター『調査と資料』第97号, 1993)の付表1にもとづき著作作成

は1桁レベル分類コードが使用されている。

木下・山田方式では上位レベルと下位レベルのそれぞれの分類コードの包含関係が重要であるため、作成された変換表の形式的な評価は下位レベル分類コードをすべて合計すると0なることである。表7の影の部分はSITC-R1の04を上位桁レベルとしてもつ集まりであり、産業分類の01と対応している。下位レベル分類コードの0422,046,047,048はともに産業分類の03に属しているため、これらを産業の01から取り除いてい

る。影の部分のSITC-R1をすべて加えると下位部分が消えてなくなり、04のみが残り、過不足なく変換されることになる。表7においてすべての下位レベル分類コードを合計すると0となることを確かめることができる。

4. 報告国における世界合計の作成

UN Comtrade Database貿易データあるいはUNC-Oデータにおいて個別報告国のデータはファ

イルごとに保存されているが報告国の世界合計は存在しない。もちろん、報告国における特定の地域合計も存在しない。そのため、貿易マトリクス作成やRCA指数のように報告国の世界合計を必要とするときにはすべての報告国を加算して世界合計を作成する必要がある。本節において商品分類コードの4桁レベル分類コードにおける報告国の世界合計を作成するためのデータ処理過程を示す。

[1]表2に示されているすべての報告国に対して商品分類コードの4桁レベル分類コードを検索のキーとしてUN国コードの3桁分類コードを出力ファイル名として貿易データを出力する。出力は表8に示されている。その出力の分類カテゴリは輸出入区分、相手国、数量単位のすべてを対象としている。統計値の範囲は金額と2種類の数量である。本処理においてすべての報告国で使用されている商品分類コードの4桁レベル分類コードと相手国の国コード一覧表も同時に作成する。

[2] 処理[1]で得られた商品分類コードの4桁レベル分類コードと相手国の国コード一覧表をもとに同じく処理[1]で作成された貿易データにおける年、輸出入区分別にすべての報告国の金額の合計値を計算する。

上記処理過程はアジア経済研究所の統計サーバであるiders12のディレクトリーの/ts/dat/no day/rca/rc_t/ws_u/において処理される。処理過程の[1]はプログラムwsu_p7.cによって報告国ごとの出力ファイルが作成される。処理過程[2]では表8で示されたすべての報告国の金額を合計することを目的としており、プログラムwsu_p11.cを利用する。報告国の世界合計を計算するための実行ファイルはwsu_p11_e.exであり、以下に示されている。

```

#*-----*
#|          wsu_p11_e.ex          |
#*-----*
#! /bin/ksh

echo "[          remove out/*          ]"

```

```

rm out/*
cp wsu_p7.e0 wsu_p7.e2
ls /ts/org/un/2005_1E/s1* > /tmp/$$
while read FN
do wsu_p7_e.ex $FN
done < /tmp/$$
rm /tmp/$$
echo "[          wsu_p11          ]"
ls out/* > wsu_p11.inp
rm wsu_p11.out2
LDR_CNTRL=MAXDATA=0x80000000          ¥
wsu_p11.x 4 wsu_p7.e2 wsu_p11.inp    ¥
wsu_p11.out2
#

```

4.1 プログラムwsu_p7.cの概要

プログラムのwsu_p7.cは処理過程の [1] を処理する機能を持つ。入力データとしてon-line検索によって得られたUN Comtrade Databaseの貿易統計データの中で商品分類がSITC-R1のみを利用するため、同iders12のディレクトリーの/ts/org/un/2005_1E/s1/に保存されているすべてのファイルを参照することになる。実行ファイルのwsu_p11_e.exにおいて処理過程 [1] で述べられているようにすべての報告国をファイルとして読み込む必要がある。そのため、

```

ls /ts/org/un/2005_1E/s1/* > /tmp/$$

```

として一時的なファイル作成する。/tmp/\$\$は影で示した部分の処理を実行するために作成される。wsu_p7.cはwsu_p7_e.exにおいて実行され、商品分類コードの桁レベル指定をパラメータとして与えなければならない。以下の例では4桁レベル分類コードとしている。

```

#*-----*
#|          wsu_p7_e.ex          |
#*-----*
#! /bin/ksh

rm /tmp/$$
echo "[          wsu_p7          ]"
echo "[          $1          ]"
wsu_p7.x 4 $1 wsu_p7.out wsu_p7.e
echo "[          sort          ]"
sort -A wsu_p7.e2 wsu_p7.e -o dou_p5.inp
echo "[          dou_p5          ]"
dou_p5.x dou_p5.inp wsu_p7.e2
#

```


表8 wsu_p11_e.exにより出力された商品分類コードの4桁レベル分類コードを持つファイル一覧

out/004	out/040	out/064	out/858
out/008	out/044	out/068	out/862
out/012	out/048	out/070	out/868
out/020	out/050	out/072	out/876
out/024	out/051	out/076	out/886
out/028	out/052	out/084	out/887
out/031	out/056	out/090	out/890
out/032	out/058	:	out/891
out/036	out/060	out/854	out/894

(出所) 著者作成

(注) out/はディレクトリー、3桁の数字はUN Comtrade Databaseで使用されている3桁の国コードである。

サブプログラムdig_cは商品分類コードの4桁レベル分類コードのとき1、それ以外のとき0を出力として返す機能を持っている。さらに、このサブプログラムは4桁レベル分類コードの一覧表を作成する機能も持ち、外部プログラムのmktblを利用して配列cc_pとその個数ncを作成している。一覧表の出力はサブプログラムのOut_Cであり、

```

/*-----**
|                Out_C                |
/*-----**
int  Out_C (char *c[], int m) {
int  i;
    for (i=0; i<m; i++)
        fprintf(fe_p, "c %s\n", cc_p[i]);
return(0); }

```

と示される。

相手国の一覧表の作成はサブプログラムのp3において、配列gの初期値として、

```
for (i=0; i<1000; i++) g[i]=-1;
```

とすべての要素に-1を与えた後、

```
g[atoi(c)]=1;
```

として相手国に対する国コードの存在を1で定義する。さらに、サブプログラムのOut_PCにおいて一覧表を出力している。

```

/*-----**
|                Out_PC                |
/*-----**
int  Out_PC (int g[]) {
int  i,k=-1;
    for (i=0; i<1000; i++) {
        if (g[i]==1) {
            k++;

```

```

        fprintf(fe_p, "p %03d\n", i);
    }
}
return(k); }

```

相手国の国コードと商品分類コードの一覧表はwsu_p7_e.exに示されているように同一の出力ファイルのwsu_p7.e2に前者はc、後者はpを先頭に付けて出力される。この一覧表は報告国ごとに作成される。したがって、すべての報告国で使用されている国コードと商品分類コードを得るには原理的にはすべての報告国から作成された出力ファイルを集めて重複しているものを取り除くことで可能となる。しかし、本章の処理は報告国ごとの処理を実行するたびごとに当該報告国の出力結果と前報告国の出力結果から共通の国コードと商品分類コードを取り出すという方法を繰り返して処理するという方法を採用している。

4.2 プログラムwsu_p9.cの概要

処理過程[2]では表1で示されたすべての報告国の金額を合計することを目的としており、プログラムwsu_p9.cを利用する。このプログラムは処理過程[1]で得られた商品分類コードと相手国の国コードのそれぞれの一連番号を配列の要素とするマトリクスとして展開し、さらに輸出入区分を配列に組み込んだ3次元の配列で定義してすべての報告国の金額をこの配列の中で

合計するという機能を持っている。プログラムのメモリーの関係から年については配列に組み込むのが不可能であるため、この処理は年毎におこななければならない。3次元に配列を限定しても配列が大きくなるとプログラムが異常停止することがあり、これを避けるために配列の大きさは実行時にサブプログラムsum_mainの中で指定するようにしている。npおよびncがそれぞれ相手国の国コードおよび商品分類コードの個数を表わしている。

```

/*-----**
|               sum_main               |
/*-----**/
int  sum_main(int nc,int np) {
double x3[3][nc+2][np+2],*x2[3][nc+1],
      **x[3];
int   g2[np+2], i, j, k;
  for (i=0; i<3; i++) {
    for (j=0; j<nc+1; j++)
      x2[i][j]=x3[i][j];
    x[i]=x2[i];
  }
  ini2_v(x, g2, nc, np);
  INpD(x, nc);
  OutD(x, g2, nc, np);
return(0); }

```

wsu_p9.c の特徴はサブプログラム INpD で表 1 で示されている報告国で表わされたファイル名を読み込み、その報告国のファイルごとに INpD2 により貿易統計データを読むという 2 段階の方法を利用していることである。

```

/*-----**
|               INpD                   |
/*-----**/
int  INpD (double **x[], int nc) {
  while (NULL!=fgets(a, sizeof(a)+1, fi_p))
    INpD2(x, nc);
  fclose(fi_p);
return(0); }

/*-----**
|               INpD2                  |
/*-----**/
int  INpD2 (double **x[], int nc) {
  a[7]='¥0';
  printf(" ..... %s¥n", a);
  fj_p=fopen(a, "r");
  while (NULL!=fgets(&inp, sizeof(inp)+1,
    , fj_p)) sum_p(x, nc);
  fclose(fj_p);
return(0); }

```

3次元の配列にすべての報告国の金額が合計されるため、最後の報告国の処理をした後の結果を出力すれば報告国の世界合計が計算される。繰り返しになるが、この処理は年毎に処理されなければならない、wsu_p9.cの実行ファイルは省略するが1962年から2003年までの指定が必要となる。

4.3 プログラムwsu_p11.cの概要

プログラムwsu_p9.cの改良版がwsu_p11.cである。wsu_p9.cはプログラムのメモリーの関係からプログラムの中で年についての配列を組み込むのが困難であったが、改良版はハード的にメモリーを最大に利用できるようにしているため年の配列の組み込みが可能となっている。そのため、実行時に繰り返し年を指定をしなくても処理ができるようになっている。前述したwsu_p11_e.exは改良版を実行ファイルで利用している。

実行ファイルにはメモリーを最大に利用できるように、

```
LDR_CNTRL=MAXDATA=0x80000000
```

を指定することが必要である。

4.4 報告国における地域合計の作成

報告国の地域合計を計算するプログラムがwsu_p13.cとwsu_p14.cである。前者のプログラムはwsu_p11_e.exのwsu_p7.cに対応する機能を持ち合計する報告国を指定し、出力となる報告国合計に対して任意の3桁コードを指定する。後者のプログラムはwsu_p11_e.exのwsu_p11.cに対応しており、前者で得られた出力データを分類カテゴリーに対して合計する機能を持っている。例えばアジアNIEsの報告国合計を作成したいとする。アジアNIEsは香港(344)、韓国(410)、シンガポール(702)、台湾(490)なので地域合計のための報告国の指定ファイルとしてwsu

_p13.aniesとすれば、

900
344
410
490
702

として新規にファイルを作成する。ここで、先頭の900はアジアNIEs地域コードを指定しており、以下に344,410,490,702とアジアNIEsの国・地域を続ける。必ずしも昇順に並んでないといけないわけではないが、並んでいるほうが入力ミス等の検証が楽である。プログラム名をwsu_p14_e.exとして、

```
##*-----*  
#|   wsu_p14. ex   |  
##*-----*  
#! /bin/ksh
```

```
echo "[      wsu_p13      ]"  
ls out/* > wsu_p13. inp  
wsu_p13. x wsu_p13. anies wsu_p13. inp      ¥  
wsu_p14. inp
```

```
echo "[      wsu_p14      ]"  
rm wsu_p14. anies  
LDR_CNTRL=MAXDATA=0x80000000      ¥  
wsu_p14. x 4 wsu_p7. e2 wsu_p14. inp      ¥  
wsu_p14. anies  
#
```

を作成する。このファイルを実行モードに変換するため、

```
chmod u+x wsu_p14_e. ex
```

を実行し、続いてwsu_p14_e. exを実行すれば、結果は同ディレクトリーの下にout/900として出力される。

この地域合計のプログラムは報告国における世界合計の一般化されたプログラムである。このプログラムを利用して報告国の世界合計や個別報告国の抜き出しも可能である。報告国世界を作成するにはwsu_p14_e.exにおいて、地域合計のためのファイルをwsu_p13.worldとして、

000
とする。さらに、wsu_p14_e.exの中のaniesをすべてworldに置き換える。結果は同ディレクトリーの下にout/000として出力される。個別報告国として日本を選んだときには、日本を指定する

ファイルをwsu_p13.jとして、

901
392

とする。この例では日本の新たな国コードを901としている。wsu_p14_e.exの中のaniesをすべてjに置き換えて実行すれば、結果は同ディレクトリーの下にout/901として出力される。

おわりに

本章においてUN Comtrade Database貿易データを基礎として作成されるアジア経済研究所の世界貿易統計データシステム (AID-XT) の概要とそれを基にした貿易関連指数の準備段階としての貿易データに関する基礎的課題とそのためが必要とされるいくつかの具体的なデータ処理方法を紹介した。

貿易データの整合性と補正についてはアジア経済研究所において商品分類コードのみを対象とした方法論が具体化され実施されているが、相手国および数量単位や数量に関しても必要であり現在検討中の課題である。特に長期時系列で数量を利用するためには数量単位の取り扱い方を検討する必要がある。

報告国の世界合計を作成するデータ処理についてはCにより書かれたプログラムの主要部分が具体的に示されており、その実行ファイルはwsu_p11_e.exである。また、報告国の地域合計の実行ファイルであるwsu_p14_e.exはより一般的な処理方法を表わしており、報告国の世界合計や個別国の計算あるいは抽出も可能になっている。本章で紹介した報告国の世界合計あるいは地域合計はUN Comtrade Database貿易データでは直接に得られないものであるため、各種の貿易マトリクスあるいは貿易関連指数において利用されることが期待される。

(注1) United Nation Statistics Devision (UN統計局)

のホームページ：<http://unstats.un.org/unsd/comtrade/>を開き、Customize Your Download FormatにあるAvailable Columnsの中から順番にSelected Columnsへ登録することにより、後者に登録された上から順番にフォーマットでは左からの項目に順番が確定される。

(注2) 関税協力理事会 (Customs Co-operation Council) に替わって現在はワーキングネームとして世界税関機構 (World Customs Organization) が一般的に利用されている。

(注3) 詳細分類 (*mdcc*) についての概念、抽出方法、補正の仕方は野田の「商品分類における詳細分類コードの抽出」が参考になる。

(注4) 表7の産業分類とSITC-R1の対応表において産業分類の21部門に対応するSITC-R1の分類コードの35,9,951,961が存在する。しかし、本書で作成された貿易指関連数には21部門を対象外としているので本書ではこれを除外して産業分類を20部門としている。そのため本書の第2部資料編、表1-1,1-2では対象としている20部門のみが示されている。

【参考文献】

[1] 木下宗七・山田光男共著「国別・商品別デフレータの推計と若干の吟味—国連貿易統計による—」(名古屋大学経済学部附属経済構造研究センター編『調査と資料』第97号 1993)

[2] 野田容助「商品分類における詳細分類コードの抽出」(野田容助編『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 2002)

[3] ——「世界貿易マトリクス作成における整合性の評価と補正」(野田容助編『改訂版世界貿易マトリクス—国際産業連関表24部門分類にもとづいて—』統計資料シリーズ No.84改訂版 アジア経済研究所 2003)

[4] ——「台湾のAID-XT基礎データの作成と評価」(野田容助編『貿易指数の作成と応用—長期時系列貿易データの推計と分析に向けて—』調査研究報告書(開発研究センター 2003-IV-20) アジア経済研究所 2004)

[5] 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—』統計資料シリーズ No.88 アジア経済研究所 2005

[6] 野田容助・黒子正人「貿易指数の作成と応用に向けた諸課題」(野田容助・黒子正人編『長期時系列における貿易データと貿易指数の作成と応用』調査研究報告書(開発研究センター 2005-II-04) アジア経済研究所 2006)

[7] 野田容助・深尾京司「貿易マトリクス作成における整合性の評価—新および旧AID-XT基礎データにもとづいて—」(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—』統計資料シリーズ No.88 アジア経済研究所 2005)

[8] 野田容助・深尾京司「同一商品分類に変換された貿易額の比較—配分ウェイトにおける推計方法の違いを考慮して—」(野田容助・黒子正人編『長期時系列における貿易データと貿易指数の作成と応用』調査研究報告書(開発研究センター2005-II-04) アジア経済研究所 2006)

