

## まえがき

貿易統計で使用される国際比較のための商品分類体系は基本的には2つの貿易商品体系から構成されている。その1つは国連 (UN) が作成する標準国際貿易商品分類 (SITC: Standard International Trade Classification) であり、この分類体系は SITC (オリジナル版) から同改訂版 (改訂第1版とも言われ、SITC-R1 と省略する)、同改訂第2版 (SITC-R2)、同改訂第3版 (SITC-R3) まで存在する。もう1つは、関税協力理事会 (CCC: Customs Co-operation Council) が作成する関税協力理事会品目表 (CCCN: CCC Nomenclature) の系列である。この系列は初版の品目表にあたる 1955 年関税品目表 (BTN: Brussel Tariff Nomenclature) から CCCN を経て現在の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) へと変わってきている。HS の分類体系も初版である 1988 年版の HS88 (HS-R1 ともいう) から 96 年改訂の HS96 (HS-R2) へと改訂されてきている。この2つの分類体系は CCCN と SITC-R2、HS と SITC-R3 というように比較的粗い状態ではあるが、対応関係は整っている。

しかし、同一の SITC 分類体系であってもその改訂された前後の接続という点においては必ずしも一貫性があって整合性のある対応関係にあるというわけではなく、貿易統計の時系列的なデータ利用にあたっては十分なる注意が必要になる。SITC-R1 を改訂した SITC-R2 は最も詳細な個別分類コードである基本項目数ではかなりの改訂がおこなわれ、また、上位レベルにおいてもいくつかの変更や修正等がおこなわれて

はいるが、基本的には SITC-R1 と同じ商品分類体系の構造を保っている。しかし、SITC-R2 を改訂した SITC-R3 については分類の概念において SITC-R2 とは大きく異なる商品分類体系をもっている。

このように分類体系の異なる商品分類から編集されている貿易統計を時系列で利用するための試みとして、野田・山本は SITC-R2 と SITC-R3 の対応関係について商品分類の基本項目の対応関係ではなく、複数の基本項目をまとめてグループ化することによって、商品グループ間で対応させるという考え方を提案している。さらに、この商品グループに存在する対応関係の中で比較的關係がないと仮定される部分を取り除くと商品グループがいくつかの商品サブグループに分割されるとき、商品グループが切断されたという。切断の仕方によって商品グループおよび商品サブグループが決まるので切断の仕方は対応関係コード表をグループ化するモデルと考えることができる。すなわち、商品分類のグループ化および切断は体系の異なった SITC-R2 と SITC-R3 の基本項目間で関連する対応関係をすべてつなげて商品グループにまとめ、さらに、切断によってサブグループ化することにより対応関係をモデル化することである。

この対応関係のモデル化は商品分類の改訂時点における前後の変換のさいに生じる貿易統計の不整合をできるだけ小さくすることを目的として作成される。しかし SITC-R2 と SITC-R3 の対応関係に見られるように対応関係が複雑になってくるとこの方法ではどうしても商品グループが大きくなりすぎて実際の利用にあたっては

該当する商品グループが特定化できなくなってしまうという問題を抱えている。そのため、個別分類コードの集まりとしての商品グループを取り出すには対応している個別分類コードをある条件の下で配分するウエイトの推計が必要になってくる。

最近、コンピュータの処理能力の飛躍的な向上により蓄積された大量のデータから知識を獲得するデータマイニングの方法が開発され実用化されてきている。本書はニューラル・ネットワークによるデータマイニングの方法を商品分類の対応関係における配分ウエイトの推計に適用して個別分類コードによる貿易統計の直接的な変換を試みている。変換の対象としたのは OECD が編集した貿易統計の日本および UN が編集した貿易統計の韓国の 1962 年から 99 年までデータである。

本来、大量データを基礎とするニューラル・ネットワークの手法に対して限られた年しか存在しない年次データの貿易統計をデータとして利用しているため、相当に厳しい仮説の設定が必要である。商品分類は SITC の 3 桁レベルとして、輸出入区分を考慮せずに同一のものとする。さらに、長期トレンドや周期を含めた経済変動から生ずる経済データ固有の変動を取り除くため、取引金額を直接使用せず商品グループ間に占める割合をデータとしている。今回の試みの結果としていくつかの改善しなければならない課題を抱えているとはいえ、商品分類の改定に伴う貿易統計の変換にニューラル・ネットワークがうまく適用できたのではないかと考えられる。データ数の制約との関係もあるが、厳しい仮説をいかに緩めていくかも含めてさらに検討されなければならない課題は今後に残されている。

本書は第 1 部および第 2 部から構成され、前者は論文からなる 4 つの章、後者は各章で共通

して使用される 6 つの表と 1 つの図を含む。第 1 章の「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—」は本書の概要にあたる。本章では本書の目的である日本および韓国の商品分類 SITC3 桁レベルにおける改訂第 1 版から第 3 版までの各系列の推計のための方法論の概要が紹介される。紙面の制限から SITC3 桁レベルごとにすべての相手国の貿易統計を掲載できず、特定商品グループのみを紹介することしかできなかったが、CD-ROM によるデータの提供は可能である。本章で用いたそれぞれの方法論については第 2 章から第 4 章までにおいて具体化される。

第 2 章の「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—」は配分ウエイトの具体的な推計方法であるニューラル・ネットワークによるデータマイニングの方法の紹介である。この方法は「検証すべき事前の仮説なしの盲目的なデータ検索」として否定的に解釈されてきてはいるが、その有効性が紹介されるにつれて最近では情報処理的な方法論としてあらゆる分野で利用され始めてきている。もちろん商品分類の改訂にともなう貿易統計の変換については初めての試みである。

第 3 章は「商品分類の改定に伴う貿易統計の整合性評価」であり、得られた SITC 各系列に対して商品分類の改訂時点で分類の不整合による断層状態が生じているかどうかの検討をおこなっている。断層が改訂による不整合が原因であるかどうかは必ずしも判断できないが、その可能性があるとの立場で検証している。不整合の存在を断層状態、すなわち変化点として捉えているため、時系列による構造変化の方法論として回帰モデルとベイズ型の事後分布を利用する。変化点として認識された商品グループについてはその原因が対応関係によるものかあるい

は配分ウエイトの推計によるものかの判断は今のところしていない。データの利用にあたって該当した商品グループの取り扱いに注意するようという範囲に留めている。

第4章は「貿易統計データベースにおける国コードの利用」である。商品分類の改訂と同じく貿易統計の長期時系列利用にあたっての問題となるところは国・関税地域の変遷である。本節では当研究所で所蔵している UN 作成の貿易統計、OECD 作成の貿易統計および台湾作成の貿易統計について当該機関が作成している国・関税地域コードと当研究所が国際比較のために作成した「アジア統一国コード」との関係を明らかにしている。また、貿易統計データベースにおける国・関税地域のコード化およびその検索方法についても検討している。

本書の後半の第2部では各章で共通して使用される表および図がまとめられている。表1は商品分類の1,2,3桁レベルにおける分類コードを SITC の各改訂版に対して横並びで見ることができるよう一覧表としてまとめたものである。同一分類コードであっても各改訂版によりその内容が違うのを確認できる。表2は商品分類を1,2桁レベルの分類コードをもとにしてグループ化された対応関係である。最初に SITC-R1 と SITC-R2 の対応関係が示され、続いて SITC-R2 と SITC-R3 の対応関係が示される。表3は商品分類を3桁レベルの分類コードでグループ化された対応関係を示す。さらに、この個別分類コードには第2章の成果であるニューラル・ネットワークにより推計された配分ウエイトが割振られている。日本および韓国のそれぞれに対して、対応関係の SITC-R1 から SITC-R2 への方向を持つ配分ウエイト、逆に SITC-R2 から SITC-R1 への方向を持つ配分ウエイトが同時に示される。対応関係が SITC-R2 と SITC-R3 に

についても同じように相互の方向を持つ配分ウエイトが示される。

表4および表5は第1章で作成された日本および韓国の SITC 改訂版ごとの時系列データにおける改訂年の構造変化の評価表である。前者は商品分類の改訂を考慮せずに分類コードを共通に使用したときの評価表である。この結果は SITC の各改訂版を横並びにした表1と組み合わせて使用される。商品分類の改訂を考慮していないので、この表4では共通した3桁分類の系列が1個のみが対象となる。表5はニューラル・ネットワークにより推計された配分ウエイトと等配分した配分ウエイトを用いて作成されたそれぞれの SITC の改訂版に対して改訂年に構造変化を生じているかを評価している表である。

表6は国および関税地域の変遷を表わす表であり、アジア経済研究所が所有する世界貿易統計データの1962年から97年まで間に使用されたすべての報告国および相手国の国コードおよび名称を含んでいる。国コードはアジア統一国コードを中心として、その原データである UN 貿易統計、OECD 貿易統計、台湾貿易統計のそれぞれで使用されている国コードとの関係を年毎の変化として捉えている。表7は原データで使用されている国コードの索引である。表6ではアジア統一国コードの名称で並べられているが、表7ではアジア統一国コードも含めてそれぞれの国コードがコード順に並べられている。

最後は図1であり、第2章で作成されたニューラル・ネットワークの評価基準であるエラーのグラフである。ニューラル・ネットワークの推計に使用した学習データでその精度を調べるときに利用されるグラフであり、ニューラル・ネットワークと等配分したときのそれぞれのエラーグラフが示されている。