

# 博士学位論文審査要旨

2016年1月25日

論文題目： Constrained Analysis of Asymmetric dissimilarity data  
by using both MDS and Clustering  
非対称非類似度データに対する MDS とクラスタリング法を用いた制  
約付き解析法

学位申請者： 谷岡 健資

審査委員：

主査： 文化情報学研究科 教授 宿久 洋

副査： 文化情報学研究科 教授 川崎 廣吉

副査： 大阪大学大学院人間科学研究科 教授 足立 浩平

要 旨：

本論文は、非対称非類似度データに対してクラスタリング法と非対称多次元尺度構成法の同時分析法である制約付き非対称多次元尺度構成法に関する研究を行ったものである。

第1章では、制約付き非対称多次元尺度構成法を提案する必要性について述べ、第2章では本論文で共通に用いる表記方法について定義している。第3章では、既存手法である4種類の「非対称多次元尺度構成法」および「クラスタリング法と対称多次元尺度構成法の同時分析法」であるCDSについてまとめている。第4章では、展開法に基づく制約付き非対称多次元尺度構成法である、制約付き展開法および制約付きスライドベクターモデルを提案している。さらに、各手法に対して目的関数の分解による特徴付けを行い、欠損値を含む非対称多次元尺度構成法の定式化が可能であることを示している。また、各手法のパラメータを推定するためのマジョライジング関数を導出し、その更新式を示している。最後に、制約付きスライドベクターモデルが展開法、スライドベクターモデル、制約付き展開法、制約付きスライドベクターモデルの一般化となっていることを示している。第5章では、分解に基づく制約付き非対称多次元尺度構成法である、制約付き丘陵モデルおよび制約付き半径モデルを提案しており、それぞれの手法に関して目的関数の分解による手法の解釈を示している。また、手法ごとに既存手法との関係を示し、提案手法が既存手法の一般化となっていることを明らかにしている。第6章では、4つの提案手法とタンデムクラスタリングの結果を数値シミュレーションにより比較し、提案手法の優位性を示している。また、第7章では実データに既存の非対称多次元尺度構成法と提案手法を適用した結果を示し、提案手法の解析結果が解釈容易であることを明らかにしている。

本論文により、大規模な非対称非類似度データに内在する非対称性を容易に解釈する手法が提案され、従来無視されることの多かった情報を解釈する可能性が示された。また、現状における制約付き非対称多次元尺度構成法の課題と展望についてもまとめられている。

よって、本論文は、博士（文化情報学）（同志社大学）の学位論文として十分な価値を有するものと認められる。

## 総合試験結果の要旨

2016年1月25日

論文題目： Constrained Analysis of Asymmetric dissimilarity data  
by using both MDS and Clustering  
非対称非類似度データに対する MDS とクラスタリング法を用いた制  
約付き解析法

学位申請者： 谷岡 健資

審査委員：

主査： 文化情報学研究科 教授 宿久 洋

副査： 文化情報学研究科 教授 川崎 廣吉

副査： 大阪大学大学院人間科学研究科 教授 足立 浩平

要 旨：

学位申請者は2012年度4月より本学大学院研究科博士課程後期課程に在学しており、国内会議および国際会議での研究発表を通じて研究活動を積極的に行い、それらの内容が3本の国際会議 Proceedings, 2本の統計教育関連の論文が論文誌に掲載されている。また、英語の語学試験にも合格していることから語学（英語）について十分な能力を有していると認定されている。

博士論文の主たる内容は、Behaviormetrika にて掲載予定である内容と関連があるものとなっている。

2016年1月14日木曜日 12:00 から 13:30 まで、約1時間の公聴会と30分の審査会において、種々の質疑応答の結果により博士（文化情報学）の学位を有するに十分な学力を有することを確認した。

よって、総合試験の結果は合格であると認める。

# 博士学位論文要旨

論文題目： Constrained Analysis of Asymmetric dissimilarity data  
by using both MDS and Clustering  
非対称非類似度データに対する MDS とクラスタリング法を  
用いた制約付き解析法

氏名： 谷岡 健資

## 要旨：

本論文は、対象間の近さと遠さの関係が必ずしも対称ではない大規模非対称非類似度データが得られた際、クラスター間の非対称な関係を容易に解釈することを目的としたクラスタリング法と非対称多次元尺度構成法の同時分析法の提案を行い、既存の非対称多次元尺度構成法や提案した手法間の関係を目的関数の分解や実行可能領域の観点からまとめ、その有効性を数値シミュレーションおよび実データ解析を用いて示したものである。

第1章では序章として、非対称非類似度データについて実データを通し説明を行った後、非対称非類似度データから対象間の非対称な関係を明らかにすることを目的とした非対称多次元尺度構成法について説明し、既存の記述的モデルに基づく非対称多次元尺度構成法が修正距離モデル、非距離モデル、拡張距離モデルに分類されることについて述べ、解釈容易性の観点より本論文では拡張距離モデルに着目することについて述べる。その後、既存の非対称多次元尺度構成法では対象間の非対称性を低次元空間で表現することが困難になる場合があるということについて述べ、提案手法である非対称多次元尺度構成法とクラスタリング法の同時分析の必要性について言及する。

第2章では、本論文で統一的に用いる諸定義やその表記方法について3つの観点よりまとめている。2.1節では非対称多次元尺度構成法の入力に関連する諸概念について述べ、2.2節では非対称多次元尺度構成法の出力に関連する諸概念について定義している。2.3節では非対称多次元尺度構成法の目的関数に関連する諸概念について定義している。

第3章では目的関数の分解に用いる諸性質、本提案手法の元となる非対称多次元尺度構成法の先行研究、およびクラスタリング法と対称多次元尺度構成法の同時分析法である Cluster difference scaling (以下 CDS) (Heiser, 1993) について述べる。本章は4節からなっており、3.1節では後述する分解型非対称多次元尺度構成法の目的関数を対称部と歪対称部に分解する際に必要となる性質である、非対称非類似度データの対称部と歪対称部への分解、およびその分解とフロベニウスノルムの関係、歪対称行列間の関係について示す。3.2節では第4章で扱う展開法型の非対称多次元尺度構成法の説明のため、展開法 (Coombs, 1950; Gower, 1977) およびスライドベクターモデル (De Leeuw and Heiser, 1982; Zieltman and Heiser, 1993) のそれぞれに関するモデル、目的関数、通常の対称多次元尺度構成法との関係について示している。具体的には、ある部分に欠損値を含む対称非類似度データに対する多次元尺度構成法として展開法を定式化する。同様に 3.3節でも目的関数を対称部と歪対称部に分解可能な丘陵モデル (Borg and Groenen, 2005) および半径モデル (Okada and Imaizumi, 1987) のそれぞれに関するモデル、目的関数について定義する。また、1節で導入した分解に関する諸性質を用いて2つの手法の目的関数それぞれが対称部と歪対称部に分解可能であることを示し、丘陵モデルと半径モデルの目的関数が非対称非類似度データの対称部に対する対称多次元尺度構成法と歪対称部に対するモデルの目的関数の同時分析になっていることを示す。最後に 3.4節では CDS のモデル、目的関数について定義し、その目的関数がクラスター間のばらつき、クラスター内のばらつき、クラス

ター重心に対する多次元尺度構成法, 同質性, と解釈可能な 4 つの項に分解可能であることを紹介する.

第 4 章では, 展開法型の非対称多次元尺度構成法とクラスタリング法の同時分析法について述べる. 4.1 節では, 展開法とスライドベクターモデルの関係性について説明する. また, 本章でクラスタリング法と展開法の同時分析法および, クラスタリング法とスライドベクターモデルの同時分析法を提案する必要性について述べる. 4.2 節では提案手法である制約付き展開法のモデル, 2 つの制約付き展開法に関する目的関数について述べた後, それらが同一の目的関数であることを示し, 制約付き展開法が非対称非類似度データに対するクラスタリング法の項とクラスター重心に対する展開法の項に分解可能であることを示す. 最後に交互最小自乗法に基づくアルゴリズムについて述べ, SMACOF (Borg and Groenen, 2005) に基づき Majorizing 関数の導出を行い, 各クラスターの座標行列を推定するための更新式, および指示行列の更新式について説明する. 4.3 節では, 提案手法である制約付きスライドベクターモデルについても制約付き展開法と同様の形式で説明を行う. 4.4 節では制約付きスライドベクターモデルのパラメータがある条件を満たす場合, 制約付き展開法や展開法, スライドベクターモデルと実行可能領域が同一となり, 制約付きスライドベクターモデルがこれらの手法の一般化となっていることについて言及する.

第 5 章では, 目的関数を対称部と歪対称部に分解可能な非対称多次元尺度構成法とクラスタリング法の同時分析法について述べる. 5.1 節では制約付き丘陵モデルおよび制約付き半径モデルを提案する必要性, 解釈方法について述べ, 第 5 章の構成について説明する. 5.2 節では, 提案手法である制約付き丘陵モデル, 2 種類の制約付き丘陵モデルの目的関数について述べた後, それらが同一の目的関数であることを示す. 次に制約付き丘陵モデルが従来の丘陵モデルと同様, その目的関数を対称部と歪対称部に分解可能であることを示し, そこから対称部が非対称非類似度行列の対称部に対する CDS の目的関数であることを明らかにする. また, 制約付き丘陵モデルに対する目的関数の歪対称部についての解釈方法についても述べ, クラスタ間間の非対称性をスローベクトルとクラスター重心の座標行列間の関係によって表現していることを明らかにする. さらに, 目的関数の対称部, 歪対称部それぞれに関して 4 つの項に分解可能であることを示し, 計 8 つに分解した項に基づき目的関数の解釈可能であることを示す. 最後に, 交互最小自乗法に基づくアルゴリズムについて述べ, 各パラメータの更新方法について述べる. 5.3 節では制約付き半径モデルについて述べ, 制約付き半径モデルを制約付き丘陵モデルと同様の形式で説明する. 目的関数の歪対称部については, クラスタ間間の歪対称部をクラスタ間間の半径の長さの差によって表現するための目的関数であることを明らかにする. 加えて, 制約付き半径モデルでは半径の長さの推定に不定性があることを示し, 半径の長さの推定問題を半径モデルと同様, 無制約最適化問題と同値であることを示す. 最後に他の方法と同様, 交互最小自乗法に基づくアルゴリズムについて述べ, 各パラメータの更新方法について述べる. 特に, クラスタの座標行列が CDS と全く同様の方法で解けることを明らかにする. 5.4 節では制約付き丘陵モデルのパラメータがある値の場合, 丘陵モデルと同一であることを示し, 制約付き半径モデルについても同様にある条件下で半径モデルと同一となることを示す.

第 6 章では, 第 4 章, 第 5 章で提案した各手法の有効性をクラスタリング結果の観点から数値シミュレーションにより明らかにする. 具体的には真のクラスタリング構造を持った人工非対称非類似度データを生成し, Adjusted rand index (Hubert and Arabie, 1985) に基づき「提案手法のクラスタリング結果」と「非対称多次元尺度構成法と 2 相クラスタリング法のタンデムクラスタリング法のクラスタリング結果」を比較する. 本章は 3 節からなっており, それぞれの提案手法を要因ごとに検証している.

第 7 章では日本茶に対するブランドスイッチング行列に制約付きスライドベクターモデル, 制約付き丘陵モデル, 制約付き半径モデルを適用し, 視覚化の観点より考察を行う. ブランドスイ

ツチング行列は MACROMILL, INC により提供頂いたスキャンパネルデータより作成したものを  
を用いる。第 6 章と同様、それぞれの提案手法ごとに結果を比較、考察している。

第 8 章では、本論文が非対称非類似度データに対する解釈容易性の向上を試みることを目的と  
した、非対称多次元尺度構成法とクラスタリング法の同時分析法に関する研究であるという総括  
について述べる。最後は、将来の展望に関して 5 つ挙げる。一つめはクラスタリング法と非対称  
多次元尺度構成法の同時分析法について、非類似度に対する重みの推定方法を改良することでク  
ラスタリング結果が向上する可能性があることについて言及し、二つめでは、クラスター数およ  
びハイパーパラメータの決定方法についてクロスバリデーションが有用となりうることについ  
て述べる。三つめは、クラスター内の非対称性を提案手法では考慮していないため、考慮方法と  
してクラスター内のバラつきを表現するためのペナルティ項を加えることを示唆し、四つめでは、  
各対象に対して外部情報が与えられている際、その外部情報の考慮方法について説明する。最後  
の五つめでは、クラスタリング法と非対称多次元尺度構成法の同時分析法を適用する際のモデル  
選択の方法について述べる。