論文の内容の要旨

論文題目　Extensions of Bayesian Multidimensional Scaling
（ベイズ推定による多次元尺度構成法の拡張）

氏名　岡田 謙介

１．研究の背景（第１章）

本博士論文ではベイズ推定による多次元尺度構成法（multidimensional scaling）モデルを取りあげ、これについてのいくつかの方法論的拡張とその評価、および実データへの適用を行った。多次元尺度構成法は非類似度・類似度のデータを入力として低次元空間での対象の布置を得る統計手法であり、Torgerson (1952)によって確立された。他の多くの統計手法と同様、この方法は当初最小二乗法の考え方に基づくものであった。その後 Kruskal の非計量多次元尺度構成法をはじめ多くの方法論的拡張ならびに応用研究が行われてきたが、現在でも依然として最小二乗推定に基づくモデルが主流である。しかし最小二乗推定には、（とくに計量的モデルにおいて）誤差を含むデータに対して頑健でない、拡張性に乏しいといった欠点がある。最尤推定を用いた多次元尺度構成も提案されているが、局所最適解の問題もあり、主要な統計パッケージには依然実装されていない。

近年、Oh & Raftery (2001)によってベイズ推定による多次元尺度構成法が提案された。この方法は観測誤差を測定モデルに内包しており、シミュレーション研究においても既存の方法より二乗誤差を小さくでき、拡張性の面でも優れているなどよい性質をもっている。そこで本研究では、このベイズ推定による多次元尺度構成法を拡張して、Minkowski 距離を扱うモデル（第 2 章）、多次元展開法モデル（第 3 章）、および確認的モデル（第 4 章）

を提案し、それぞれ数値シミュレーションによる手法の評価および実データへの適用を行った。また、この方法を応用研究者が手軽に実行するためのフリーソフトウェアを開発し、インターネット上で公開した（第 5 章）。

２．Minkowski 距離を用いたベイズ推定による多次元尺度構成法（第 2 章）

本章では、ベイズ推定による多次元尺度構成法を拡張し、Minkowski 距離を扱うことを可能にした。心理学研究では、古くから多次元尺度構成法を用いた心理的空間のモデリングが行われてきた。多次元尺度構成法では多くの場合ユークリッド距離を用いて対象間距離が測られるが、ユークリッド距離とは Minkowski 距離の特殊な場合（Minkowski 指数=2 の場合）であり、ほかにも同指数 1 の場合の市街地距離などが心理学的見地から重要とされる。多次元尺度構成法でデータの最適な Minkowski 指数を探る方法としては、Kruskal (1964)以来長らく Minkowski 指数の値を網羅的に設定して分析を繰り返し、最もストレス（観測非類似度とモデル距離との二乗誤差を標準化した指標）が小さくなる Minkowski 指数の値を求めることが行われてきた。しかし、この最もよく知られた応用のひとつである Ekman (1954)の色の類似度データの既存の分析では、ストレスが正しく最小化できていない可能性があった。そこで、本章ではまず古典的な多次元尺度構成法モデルに新しい最適化方法を適用してこのデータを再分析し、やはり既存の研究でストレスが最小化できていなかったこと、新しい最適化法を用いれば理論的に導かれる Minkowski 指数 1 の場合と∞の場合のストレスの等値性が得られることを示した。

しかし、依然こうしたアプローチにはいくつかの問題が残る。たとえば、ストレスの定義には様々なものが知られているので、あるストレス（多くの場合ストレス-1 と呼ばれる指標）を最小化する Minkowski 指数をもって真の Minkowski 指数（の推定値）と考えてよいのかという問題がある。また、そもそも異なる Minkowski 指数が設定された異なったモデルのよさを、ストレスで評価できるのかという問題もある。

近年、Lee (2008)は Minkowski 指数を様々な値に固定して何度も分析を繰り返すのではなく、これをパラメータとして扱ってベイズ推測を行う方法を提案した。こうしたアプローチは先行研究の問題点を改善するため有望と考えられるが、彼の方法は一様分布を多く使うなど特殊なモデルを用いており、またシミュレーションによる手法の評価も行われていなかった。そこで、この方法を用いて真のモデルからデータを発生させ Minkowski 指数を推定するシミュレーション研究を行ったところ、推定値に大きなバイアスがあることがわかった。

そこで､本研究では Minkowski 指数をパラメータとして扱うという考え方は継承しつつ、既存のベイズ推定による多次元尺度構成法モデルをシンプルに Minkowski 距離へと拡張することにより、安定した推定のできる方法を提案した。提案手法では同様のシミュレーションにより正しい Minkowski 指数を復元することができた。また、Ekman のデータにおいて本手法を用いた分析を行ったところ、Minkowski 指数=1.7 付近の値が事後推定値として

得られた。この結果はストレスの最小化を用いた場合の推定値とほぼ一致するものであった。

３．ベイズ推定による多次元展開法（第３章）

本章では、多次元展開法の問題に対してベイズ推定によるモデルを構築した。展開法は典型的には複数の評定者が複数の対象について評定を行ったデータに対して、評定者と対象の双方を同じ低次元空間上に布置する統計手法である。この展開法のモデルは、（評定者＋対象）×（評定者＋対象）という大きな非類似度データにおいて、（評定者×評定者）の部分のデータおよび（対象×対象）の部分のデータが欠損した場合の、欠測を含む多次元尺度構成法としてとらえることができる。こうした欠測は、データ拡大法（data augmentation）のアルゴリズムによってマルコフ連鎖モンテカルロ（Markov chain Monte Carlo, MCMC）法を用いたベイズ推定で扱うことができる。したがって、本研究ではベイズ推定による多次元尺度構成法モデルにデータ拡大法を組み合わせることにより、ベイズ推定による多次元展開法を提案した。多次元展開法にベイズ推定を用いたのは本研究が初めてである。

手法の評価としては、Green & Carmone (1970)の提案に基づいて「A」「M」を構成する 33 点のうち、「A」に含まれる点と「M」に含まれる点の間の距離だけが観測されている（「A」の点どうし、「M」の点どうしの距離は欠測とする）状況から布置を復元するシミュレーション研究を行い、正しく元の布置が復元できることを確認した。つづいて実データの解析を行ったが、ここでは既存の多次元展開法アルゴリズム・ソフトウェアである ALSCAL, PREFSCAL と提案手法との比較を行った。各種ストレスおよび相関係数に基づく評価から、提案手法は ALSCAL よりも優れ、PREFSCAL と同等または僅かに優れた成績を収めた。また ALSCAL および PREFSCAL では推定誤差を扱えないため、推定誤差も得られる点においても本手法が優越しているということができる。

４．ベイズ推定による確認的多次元尺度構成法（第４章）

統計手法は探索的方法と確認的方法とに大別される。両者は両輪のように密接な関係にあり、データから有意義な情報を取り出すためには双方の発展が不可欠である。心理学研究において多次元尺度構成法と並んでよく使われる多変量解析法である因子分析・共分散構造分析においては、確認的方法が探索的方法と並んでよく発展している。しかし、多次元尺度構成法では、確率モデルを用いた確認的方法が相対的に未発展であった。ある種の制約付き推定を行う方法を確認的多次元尺度構成法と呼ぶ場合はあったが、このアプローチは単純な制約・重み付きの多次元尺度構成法にすぎず、推測統計学的なモデルの評価・比較などができなかった。そこで本章では、ベイズ推定を利用して距離行列に構造がある場合の確認的多次元尺度構成法を提案した。この方法では、あらかじめ距離行列への不等式制約が事前仮説としてある場合、MCMC 法における繰り返し毎に距離行列で不等式制約

が成立するか否かを検討し、制約しない場合はその繰り返しを棄却する。こうして制約が満たされる場合の MCMC 標本のみを用いて事後分布を構成することにより、条件付き事後推定を行うことができる。またこの方法では制約を導入しないモデルと導入したモデルでのベイズ比（Bayes factor）を計算することができ、これを用いて導入した制約を定量的に評価・確認することができる。

本手法を用いて、まずシミュレーション研究により正しいモデルを正しいと評価できること、および誤ったモデルを誤りと評価できることを確認した。続いて多特性・多評定者（multitrait multi-informant）データへの応用を行った。

５．ソフトウェアの開発（第５章）

本章では、ベイズ推定による多次元尺度構成法の分析を行うためのプログラムの開発と公開を行った。統計的手法は方法論であり、実際に心理学研究のためにデータ解析を行う研究者によって利用され、新たな知見を得られてこそ意味がある。例えば共分散構造分析の今日の発展を支える要因として、Amos などの使いやすいソフトウェアの果たした役割は大きいと考えられる。これまでベイズ推定による多次元尺度構成法を実行できるフリーのプログラムは存在しなかったため、フリーの統計環境 R 上で動作するベイズ推定による多次元尺度構成法プログラムを公開した。本プログラムの関数仕様は Appendix B に示した。

６．まとめと議論（第６章）

本研究では、ベイズ推定を用いた多次元尺度構成法を方法論的に拡張する 3 つの研究、およびソフトウェアの開発と公開を行った。本研究では Minkowski 距離への拡張や確認的モデルの推定などにおいてベイズ推定の特徴を生かした多次元尺度構成法の拡張を行うことができた。

MCMC 法を用いたベイズ推定は、しばしばその時間的コストが欠点として指摘される。しかし、本研究で行った分析に要する時間は短いものでは数十秒を要するのみである。また、ムーアの法則として知られるように計算機の性能は指数関数的成長を過去数十年にわたって続けており、今後も計算機性能の躍進が続くことを考えると、将来的にベイズ推定の重要性はさらに増していくと考えられる。