審査の結果の要旨

論文題目： Multivariate Linear Mixed Models with Application to Small Area Estimation
題目和訳： 多変量線形混合モデルと小地域推定への応用
氏名： 伊藤　翼

本論文は，小地域推定問題に対して多変量線形混合モデル（Multivariate Linear Mixed Model）を用いた手法の開発に取り組んでいる。線形混合モデルとは，通常の線形回帰モデルに確率変数による変量効果を加えたモデルのことで，小地域推定の枠組みでは地域の差違が変量効果として組み込まれる。その代表的なモデルとして Fay-Herriot モデルと枝分かれ誤差回帰モデル（Nested Error Regression）モデルがあり，前者は集計データを解析するために用いられ，後者は個票データが利用できるときに使われている。両モデルはこの分野の基本モデルとして非常に広く用いられているが，従来の研究の多くは一変量データを解析するモデルに関するもので，例えば家計調査の支出項目「教育費」と「教養娯楽費」など，２変量以上のデータを扱うモデルについての小地域推定問題はあまり研究されてこなかった。その理由の一つとして，変量効果と誤差項の分布にそれぞれ異なる未知の共分散行列が入るため，それらの推定量に基づいた統計的推測手法に関して漸近２次近似の議論を解析的に行うことは大変複雑な計算を伴う必要があることがあげられる。本論文は，多変量 Fay-Herriot モデルと多変量枝分かれ誤差回帰モデルを扱い，従来の研究成果を多変量モデルへ拡張し，多変量小地域推定の推測方法を確立している。特に，多変量の信頼区間の一つとしてマハラノビスの距離に基づいた信頼領域の構成を行い，漸近展開によるバートレット型の補正を通して，名目信頼係数に漸近２次のオーダーまで正確な信頼領域の構成を行っている。具体的には以下のような章立てで構成される。

第１章　序
第２章　多変量 Fay-Herriot モデル
第３章　多変量 Fay-Herriot モデルにおける小地域推定の平均２乗誤差行列の頑健な推定
第４章　多変量枝分かれ誤差回帰モデル

本論文で扱う小地域推定の背景，各章のモチベーションと概要を第１章で与えた上で，第２章～第４章で研究成果をまとめている。

小地域推定の分野で集計データの解析のために利用されるモデルが Fay-Herriot モデルと呼ばれるもので，先行研究の多くが一変量データを解析するためのモデルを扱ってきたのに対して，第２章では多変量の Fay-Herriot モデルを扱い，多次元の小地域データを解析するための統計的推測手法に関する研究成果をまとめている。例えば家計調査に基づいて支出項目の「教育費」と「教養娯楽費」の地域別平均値を推定することを考えるとき，家計調査により与えられる数値は直接推定値と呼ばれるが，地域のデータ数が多くないためバラツキが大きくなる傾向にある。そこで，線形混合モデルを用いることによって周辺地域のデータを利用する工夫がなされ推定精度を高めることができる。線形混合モデルの枠組みでは推定問題は，回帰係数に基づいた多次元の母数効果と多次元の変量効果の和を予測する問題として捉えることができる。共変量の共分散行列が既知のときには，地域別平均の最良線形不偏予測量(BLUP)が与えられ，共分散行列が未知の場合には，共分散行列の推定量を最良線形不偏予測量に代入することによって経験最良線形不偏予測量(EBLUP)が得られる。このとき，経験最良線形不偏予測量の予測誤差を見積もる必要が生ずる。そのために，第２章では，変量効果ベクトルと誤差ベクトルに多変量正規分布を仮定して，平均２乗誤差行列とその漸近２次不偏推定量の導出及び信頼領域の導出を行っている。

まず，経験最良線形不偏予測量の平均２乗誤差行列を計算すると３つの項の和に分解で

きることがわかる。第１項は，観測値を与えたときの変量効果の条件付き共分散行列に対応する項，第２項は，共分散行列を既知としたときの一般化最小２乗推定量で回帰係数ベクトルを推定したときに生ずる推定誤差に対応する項であり，これらは明示的に与えられる。第１項と第２項の和は，変量効果の共分散行列を既知としたときの最良線形不偏予測量の平均２乗誤差行列に対応していることがわかる。第３項は，共分散行列を推定することから生ずる誤差に対応しており，共分散行列の推定量を代入した経験最良線形不偏予測量と共分散行列を既知とした最良線形不偏推定量の差の平均２乗誤差行列として与えられる。この項については，共分散行列の推定量に適当な条件を仮定した上で地域の個数を大きくする漸近理論を考えて，２次漸近近似の値を求めている。こうして求めた経験最良線形不偏予測量の平均２乗誤差行列の漸近２次近似の式に，共分散行列の推定量を代入することによって平均２乗誤差行列の推定式が得られることになる。しかし，第２項と第３項が２次のオーダーであるのに対して第１項が定数オーダーになるため，その推定式は２次のバイアスをもってしまう。そこで，第１項に共分散行列の推定量を代入したものの期待値について漸近展開を行うことにより，第１項の２次漸近不偏推定量を導出し，最終的に経験最良線形不偏予測量の平均２乗誤差行列の２次漸近不偏推定量を与えている。

次に，経験最良線形不偏予測量を中心に予測領域の構成を行っている。1 次元のときの信頼区間と異なり，多次元の場合の構成方法は一通りとは限らない。ここでは，経験最良線形不偏予測量の平均２乗誤差行列の２次漸近不偏推定量の第１項と第２項の和を重みとして組み込んだマハラノビスの距離を考え，その分布の漸近展開を特性関数の漸近展開を用いて導出している。この方法は多変量解析において用いられるものであるが，小地域推定の文脈で用いたのは初めてである。１次元の場合と比べ，多くの項を評価する必要性が生じ計算は極めて煩雑になるが，最終的にきれいな形にまとめることができている。この漸近展開からバートレット補正のための補正項を求め，マハラノビスの距離を修正することによって，信頼領域の被覆確率が名目的な信頼係数の値に漸近２次のオーダーまで一致するような信頼領域を求めている。

信頼区間の問題点として，変量効果の分散の推定値が小さくなるにつれて信頼区間の幅が長くなってしまうことが知られているが，ここで構成した信頼領域についても同様な問題があることがわかる。すなわち，共分散行列の推定値の最小固有値が小さくなるにつれて信頼領域の面積が増大してしまうことになる。通常のモーメント推定量は正の確率で負の値を取ってしまい，またそれをゼロで打ち切った推定量も正の確率でゼロの値を取ってしまうため利用することができない。そこで，正定値であり，しかも漸近的に２次不偏になるような共分散行列の推定量の導出を行っている。これは，極めて独創的な方法であり，この推定量を用いたときのバートレット補正項の近似値を計算し最終的な信頼領域の構成を行っている。さらに，任意の２つの地域の平均ベクトルの差に関する信頼領域の構成についても同様な結果を導出している。

提案手法の数値的な性質については第２章の最後にまとめられている。標本誤差が地域によって異なる状況をつくるため地域別誤差共分散行列の値として２つの場合を設定してシミュレーション実験を行っている。一つは地域別にあまり差がない場合，もう一つは一つの地域だけ大きな誤差共分散行列をもつ場合である。いずれの場合も，経験最良線形不偏予測量の平均２乗誤差行列は直接推定量のものよりもかなり小さくなっており，特に，共分散行列の相関係数が大きくなるにつれて改善度は大きくなる傾向にある。１次元のモデルから導かれる予測量との比較においても，相関係数とともに改善度が増していくことを明らかにしている。また平均２乗誤差行列の漸近２次不偏推定量についても相対バイアスの値は小さく抑えられている。提案された信頼領域の被覆確率については，相関係数が小さいときには名目的信頼係数の値に近いものになるが，相関係数が大きくなるにつれて信頼係数の値よりやや大きな値をもってしまう傾向にあることを指摘している。この点は今後の課題として取り組むことになるであろう。最後に，家計調査における「教育費」と「教養娯楽費」の地域別平均の推定への応用が与えられている。

第２章は多変量 Fay-Herriot モデルの変量効果ベクトルと誤差ベクトルに多変量正規分布を仮定して議論しているのに対して，第３章では分布の正規性を外した議論を展開している。最良線形不偏予測量及び経験最良線形不偏予測量は正規性の仮定の有無に関わらず同じ形をしていることに注意する。経験最良線形不偏予測量の平均２乗誤差行列が４つの項に分解され，第１項と第２項は明示的に与えられ，第３項と第４項については漸近２次近似が計算されている。正規性のもとでは３つの項に分解されたのに対して，正規性を外したときには，第３項と第４項が変量効果及び誤差項の分布の尖度パラメータに基づいて表現されることが示されている。正規性のもとでは，正規分布の尖度が３になることを代入すると，第３項は第２章で求めた近似式に一致し，第４項はゼロになることがわかるので，一般的な近似式を与えていることになる。

第３章では，次に，経験最良線形不偏予測量の平均２乗誤差行列の２次漸近不偏推定量の導出を行っている。平均２乗誤差行列の第１項のみ定数オーダーなので，その第１項に共分散行列の推定量を代入したものの期待値は２次のバイアスをもつことになる。そこで２次漸近近似を行いバイアス補正を行うことにより，経験最良線形不偏予測量の平均２乗誤差行列の２次漸近不偏推定量が得られる。これは５つの項から構成されるが，変量効果の分布の尖度パラメータに依存しないことがわかる。このことは，変量効果の分布に対する頑健性を示している。この結果は，Lahiri and Rao (1995, JASA)の多次元への拡張であるが，彼らの結果は誤差項に正規分布を仮定しているのに対して，第３章の結果は誤差項に正規性を仮定しないので，２次漸近近似のための計算は多次元化と非正規化とから極めて複雑になっている。平均２乗誤差行列の推定量の挙動はシミュレーション実験を通して調べられている。変量効果と誤差項の分布に標準化された t-分布とカイ２乗分布を仮定したときには，正規性のもとでの結果に比べ改善しているものの，両者の差はさほど大きくない。これに対して，標準化された対数正規分布を仮定したときには，正規性のもとでの結果に比べ大きく改善していることが示されている。

第４章では，多変量枝分かれ誤差回帰モデルが扱われる。第２章と第３章が集計データを解析するためのモデルを扱ってきたのに対して，多変量の個票データが利用可能なときには多変量枝分かれ誤差回帰モデルを用いることによって地域別平均値などの推定を行うことができる。このモデルの特徴は，変量効果の分布と誤差項の分布にそれぞれ異なる未知の共分散行列があるという点であり，それぞれ群間共分散行列，群内共分散行列と呼ばれる。１変量の枝分かれ誤差回帰モデルにおいては最尤法を含め様々な推定方法が提案されているが，多変量の場合には最尤推定量を数値的に求めることは困難で，これまであまり研究されてこなかった。第４章では，モーメント法を用いた明示的な推定量を提案し，その一致性や２次漸近不偏性を示している。経験最良線形不偏予測量の平均２乗誤差行列の２次近似を行うとともに，その２次漸近不偏推定量の導出を行っている。また地域平均の線形結合の推定やその信頼区間の構成を考え，漸近２次の意味で正確な手法を求めている。提案手法の被覆確率が名目的な信頼係数に近い値を与えていることを，正規分布，t-分布，カイ２乗分布からのシミュレーション実験を通して確認している。これは，提案手法の補正項がうまく働いていることを示すとともに分布に関する頑健性を有していることを示している。

論文の評価

本論文は，多変量の小地域推定問題について解析的な結果を与えている優れた論文である。市区町村や更に細かい地域の特性値を推定しようとすると，データ数が少なくなるためその地域のデータだけで推定したのでは大きな推定誤差が生じてしまう。これを小地域推定問題という。この問題を解決するために全領域で共通な母数として埋め込まれる回帰係数と地域毎に異なる変量効果とを組み入れた線形混合モデルが使われる。線形混合モデルから導かれる経験最良線形不偏予測量は，小地域毎の標本平均を全データに基づいた安

定した推定値の方向へ縮小することにより，推定値の安定化と推定精度の改善が図られている。これまでの多くの研究は１変量データを扱うための線形混合モデルに関するものである。その理由の一つとして，変量効果の分布と誤差項の分布にそれぞれ異なる共分散行列が入っているため，それらの推定量に基づいた統計手法の漸近２次近似の議論を解析的に行うには漸近展開から生ずる多くの複雑な項を評価する必要があることがあげられる。そこで多変量小地域推定問題に関する先行研究では，解析的に扱うことを避け，共分散行列に事前分布を想定した階層的ベイズモデルを利用して数値的に解く方法が提案されてきた。これに対して，本論文の特徴は，経験ベイズ的アプローチを用いている点で，漸近展開の各項を解析的に正確に評価して２次漸近近似の項を明示的に導出していることである。多変量 Fay-Herriot モデルや多変量枝分かれ誤差回帰モデルにおいて，経験最良線形不偏推定量の平均２乗誤差行列の２次近似とその漸近２次不偏推定量の導出を解析的に行うことは，漸近展開の項の多さと複雑さとから大変な労作業になっているが，最終的には明示的な表現式を導出しており，この分野への貢献として評価できるものである。

特に，マハラノビスの距離に基づいた信頼領域に関する研究はこの分野では初めての試みであり，マハラノビスの距離による統計量の分布の漸近展開を特性関数の漸近展開から導きバートレット型の補正項を明示的に与えている点は優れた研究成果として高く評価できる内容である。この導出の過程で，変量効果の分布の共分散行列の推定量の固有値がゼロもしくはゼロに近くなるにつれて補正項が大きくなり信頼領域の面積が大きくなってしまうという欠点に注目し，通常のモーメント推定量の打ち切り手法では不十分であるため，正定値でしかも漸近２次不偏性をもつ推定量を提案してこの問題の解決を図っている点も評価に値する。

以上，説明してきたように，本論文は，多次元の小地域推定を行うための多変量 Fay-Herriot モデル及び多変量枝分かれ誤差回帰モデルに関して，伊藤翼氏自身の独創的で優れた研究成果をまとめたものである。その中で提案された手法と理論結果は小地域推定などの応用分野において利用価値の高いものであり，この分野における貴重な貢献であると評価できる。この論文の発表会は，平成３１年1月２９日に行われ，その後開かれた論文審査委員会において，全員一致で伊藤翼氏が博士（経済学）の学位を授与されるにふさわしいという結論に達した。

審査委員
久保川達也
大森裕浩
下津克己
入江薫
菅澤翔之助