

科学研究費助成事業 研究成果報告書

平成 27 年 6 月 26 日現在

機関番号：82616

研究種目：基盤研究(C) (一般)

研究期間：2011～2014

課題番号：23530968

研究課題名(和文)計量心理モデルにおける統計的推定の高速化と大規模データへの応用

研究課題名(英文)Parallelization of statistical estimations for psychometric models and its application to large scale data

研究代表者

大津 起夫(Otsu, Tatsuo)

独立行政法人大学入試センター・研究開発部・教授

研究者番号：10203829

交付決定額(研究期間全体)：(直接経費) 3,900,000円

研究成果の概要(和文)：心理学をはじめとする行動諸科学の研究においては、多変量統計データの統計的分析が、研究を進める上で重要であり、適用領域の拡大とともに分析対象となるデータの規模も急速に大型化してきている。本研究においては、近年進展の著しいコンピュータの並列処理機能(マルチコア化)を前提に、これまで心理統計および心理測定分野で開発されてきた統計モデル(特に、大規模な非線形の潜在変数モデル)の推定法の高速化をおこなった。また、大規模なテストデータを対象としてモデルの推定を行い、特に試験問題の潜在構造と難易度比較についての分析例をした。非線形因子分析モデルの推定においては、6コアPC上で、3倍超の高速化を実現した。

研究成果の概要(英文)：The role of multivariate statistical analyses has increased in psychology and related social and behavioral sciences. Extensions of their application area led to enlargement of target data and computational sizes. We developed parallelization for speeding up estimations in statistical methods for psychometric models, especially nonlinear latent variable models. We applied the developed method to analyses of latent structures in educational testing. The parallelization of non-linear factor analysis on six cores PC make the speed of estimation more than three times faster compared to serial computation.

研究分野：計量心理学

キーワード：潜在変数 並列計算 テスト理論 因子分析 スプライン テスト等化

1. 研究開始当初の背景

心理学をはじめとする行動諸科学および社会科学の研究においては、多変量統計データの統計的分析が、研究を進める上で重要な役割を果たしており、適用領域の拡大とともに分析対象となるデータの規模も急速に大型化してきている。本研究においては、近年進展の著しいコンピュータの並列処理機能（マルチコア化）を前提に、これまで心理統計および心理測定分野で開発されてきた統計モデル（特に、大規模な非線形の潜在変数モデル等）の高速な推定法の検討と実装を行う。計算量の多寡は、ある意味時間をかければ解決される問題ではあるが、現実的な応用を考えると、長時間の計算が適用範囲を極端に狭めることは否めない。また、1980年代以降に開発された、ブートストラップやクロスバリデーションなどモデル評価の手法が、それ自体多くの計算量を必要とする場合が少なくない。

2. 研究の目的

近年利用コストの低減した並列計算を利用して、心理学で多用される統計モデルの統計的推定方法の高速化を実現し、さらに大規模なテストデータを対象としてモデルの推定を行い、特に試験問題の潜在構造と難易度比較についての分析例を示す。

潜在変数を伴う統計モデルについては、パラメータの信頼区間などの導出も解析的な理論を用いることが困難である場合が多く、ブートストラップなどの手法に頼らざるを得ない。これらの問題を解消するため、近年進展の著しいマルチコア（計算素子）を利用した、並列計算用の推定プログラムを作成し、推定の高速化を実現する。

3. 研究の方法

(1)統計モデルの推定に有効な並列化の実現方法（OpenMP などの各種の並列化プログラムツール）について検討し、主に利用する実装方法を決定する（開発言語は C++ を想定）。特に、開発したコードの保守の容易さと並列化による効果について検討する。(2)基礎統計量(平均,分散,分位点,多重集計など)の導出の高速化方法について検討し、ツールの決定および実装を行う。(3)非線形因子分析の尤度計算の並列化の検討と実装を行う。(4)大規模多次元尺度構成法の高速推定法について検討する。(5)尤度関数の数値最適化の高速化方法について検討と実装を行う。(6)実装した方法を大規模テストデータへ適用し有効性を検討する。

4. 研究成果

主要な成果は、記号処理プログラムにおいて並列文字列探索機能を組み込んだ同一文字列探索プログラムと、OpenMP による非線形因子分析の推定の高速化とその応用、および

線形計算の高速化による多次元尺度構成法の計算の高速化であるが、ここでは後者の応用例について示す(大津 2014, 行動計量学会)。

センターにおける科目選択形態別の

得点分布--非線形因子分析による比較--.

(1)データの背景：平成 24 年度以後の大学入試センター試験は、それ以前と教科科目は大きくは異なるもの、受験生が選択可能な科目の組み合わせは、平成 23 年以前とは大きく変わった。ここでは、新しい科目選択システムのもとでの、試験問題の難易度と受験者群の特徴について検討する。

平成 18 年から平成 23 年度までのセンター試験においては、地理歴史については一つの試験時間において 6 科目から 1 科目のみを受験することが可能であり、公民についても一つの試験時間において 3 科目から 1 科目のみを受験できた。このため、受験者は、地理歴史から 2 科目を選択することや、公民から 2 科目を選択することはできなかった。また、理科については平成 23 年度までは 3 つの試験時間が設定されており、それぞれの試験時間において 2 つの理科の科目からいずれかを選択することが可能であった。

平成 24 年度センター試験においては、公民に「倫理、政治・経済」科目が新設され、さらに地理歴史 6 科目、公民 4 科目合計 10 科目から 2 科目を選択することが可能となった。ただし、「世界史 A」と「世界史 B」などのように内容の重複する科目（科目名に重複する部分があるもの）の組み合わせは許容されない。また、理科については、6 科目から任意の 2 科目が選択可能になった。地理歴史、公民および理科のそれぞれにおいて 2 科目を受験する者は、時間割上 2 コマの受験時間を用いて試験室から途中退出することなしに 2 科目続けて受験を行う。いずれの領域においても、問題冊子は最初の試験時間の冒頭にすべて配布されるが、解答用紙はそれぞれの時間の冒頭に 1 科目配布され、その時間の終了時に回収された。

平成 24 年センター試験の時間割のもとでは、2 科目目に受験する者は問題冊子の内容を 1 科目目の試験時間中から知ることが可能である。このため 1 科目受験者と 2 科目受験者の間での公平性が保たれるか否か危惧された。そこで、前半のコマにて受験した科目（第 1 解答科目）と後半のコマで受験した科目（第 2 解答科目）について、取扱いに差を付ける対応がとられた。

具体的には、平成 24 年度の国公立大学を含む多くの大学の大学入試において、センター試験の利用法に制約が課せられた。これは、地理歴史、公民（または理科）から 1 科目のみを試験科目として利用する場合には、第 1 解答科目を用いるというものである。この制約のもとでは、第 1 解答科目への解答時間を削って、第 2 解答科目への解答に用いることの利点は多くはないと推測される。

ここでは、平成 24 年に実施されたセンター試験におけるいくつかの科目において、第 1 解答科目として受験した者と、第 2 解答科目として受験した者の得点差について検討を加える。分析にあたっては、大津(2011)と類似の枠組みで、欠測構造に MAR(Missing At Random, Rubin,1976)を仮定する非線形因子分析を適用した。また、地歴・公民科目と理科科目の得点は受験者層によってかなり違う様子を示しているため、それぞれについて分析を行ったが、ここでは理科について結果の概要を示す。

(2)分析の枠組み:平成 20 年度から平成 22 年度までの科目得点については、大津(2011)において、欠測値に対応した非線形因子分析を用いた分析の結果が報告されている。ここでは、

上記の分析と同様の手法を用いるが、平成 24 年度よりセンター試験に導入された「第 1 解答科目」、「第 2 解答科目」の区分を理科の各科目について考慮する。具体的には、理科の科目選択において、2 科目選択者の第 2 解答科目を他の受験者(2 科目受験者の第 1 解答科目および 1 科目受験者)の科目得点とは別の変数とみなすことにする。これにより、第 2 解答科目として受験された科目の難易度を、他の場合とは区別して推定する。以下では、「2 科目受験者の第 1 解答科目および 1 科目受験者」の科目得点を、「第 1 科目得点」と呼び、「2 科目受験者の第 2 解答科目」による得点を「第 2 科目得点」と呼ぶことにする。

分析対象は、本試験において、数学 I・数学 A、英語(リスニングを含む)をともに受験した者(約 38 万人)を対象とし、数学 I・数学 A、数学 II・数学 B、理科の 6 科目(物理 I、化学 I、生物 I、理科総合 A、理科総合 B)、および英語(リスニングを含む)の科目得点について検討する。「数学 II・数学 B」は選択問題を含むため、厳密にはこれらの選択パターンごとの分析を行うべきではあるが複雑になるため、ここでは選択問題の別は無視し科目得点のみを用いる。理科の 6 科目については、「第 1 科目得点」と「第 2 科目得点」を別の変数とみなすので、15 の変数を分析対象とする。

(3)分析モデル:ここでは、いささか大胆ではあるが、分析対象となる 15 個の変数に共通する能力があり、それが 1 次元の潜在変数によって表されるとする因子分析モデルを仮定する。また、観測変数の条件付平均(潜在変数の値を固定した場合の観測変数の平均)が、区分多項式(スプライン関数)(de Boor,2001)によって表されるモデルを用いる。この際、独自因子は潜在変数の値によらず、一定の分散に従うと仮定する。しかしながら、試験の得点を分析する場合には、満点付近での独自因子による分散は、より平均的な得点における場合よりも小さいと思われる。そこでそれぞれの変数(素点) X_j を次の式(逆正弦変換)によって Y_j に変換し、

因子分析モデルの適用がより妥当となる変数を構成する。

$$Y_j = f_j(X_j) = \sin^{-1} \sqrt{X_j / X_j \text{ の満点}} \quad (1)$$

非線形の因子分析モデルは次のようなものである。

$$Y_j = \mu_j + \epsilon_j, (j=1, \dots, J) \quad (2)$$

ここで μ_j は標準正規分布に従い直接には観測されることのない確率変数

(潜在変数)であり、ここでは分析対象とする各科目に共通な能力を表すものと仮定する。また ϵ_j は および他の $\epsilon_k, (k \neq j)$ と互いに独立な確率変数(独自因子)であり、分散 σ_j^2 、平均ゼロの正規分布に従うものとする。また μ_j はそれぞれの変数に対応した区分多項式(スプライン)とする。

ここで潜在変数の次元は 1 次元とし、正規分布を 50 個の等分位区間に分割し、それぞれの区間を平均 $\mu_k, (k=1,50)$ によって代表させることにより離散的に近似した。また各 μ_j は、離散化された μ_j の最小値 μ_{j1} 、最大値 μ_{j50} および $\mu_{j25} = 0$ (中位数)の 3 点を節点とする 2 次多項式による B-スプラインによって推定した。さらに、分析対象者のすべてが科目 j を受験したら得られるであろう科目の平均点の推定値を、零点と満点の境界を考慮し次の式で求める。

$$E_j = \mu_j [g_j(Y_j) = f_j^{-1}(\min(\max(0, Y_j), 1/2))] \quad (3)$$

ここで μ_j を上述のように 50 個の離散値で近似するが、さらにそれぞれの μ_j も同様に 50 個の離散値で近似して確率分布の推定を行う。このようにして得られた各科目の平均点の推定値を「補正済み平均」と呼ぶことにする。

(4)結果:平成 24 年理科科目の比較結果の科目別平均点をみると(詳細情報はここでは省略する)、第 1 科目得点と第 2 科目得点との関係は、さまざまな場合があるが、補正済み平均をみると、理科科目のいずれについても、第 1 科目得点が第 2 科目得点を上回っている。ここでは欠測メカニズムに MAR を仮定しているため、補正済み平均は、観測されたデータに基づいて、その科目得点と他の科目得点との関係を推測したうえで、分析対象者全員が該当科目を解答したら得られるであろう得点の平均を予測するものである。第 1 科目得点と第 2 科目得点の差の大小は科目によってあるものの、第 1 科目得点の方が一貫して大きい。これらを検討すると、第 2 解答科目として解答することが、科目得点に特に有利になるという傾向は、少なくともこの分析結果からは認められない。

(5)計算技術についての考慮:本報告の統計的な推定においては、モデルの特性上非常に多くの計算が必要になる。線形の因子分析とは異なり非線形の構造を含むため、簡潔な十

分統計量が存在せず、モデルの尤度を求めるためには毎回全てのデータを参照する必要がある。また、現在のところ数値微分を利用した準ニュートン法を利用しているため尤度計算の回数はかなり多くなるので、計算の高速化は大規模なデータを分析する上で実用上重要になる。ここでの計算には 64bit Linux 上の R-2.15 を用いたが、計算の高速化を図るために考慮した点について示す。

計算には CRAN で配布されているバイナリ版ではなく、インテル社製の C++ および Fortran コンパイラ (Intel Composer XE 2013) を用いて、R システムを再構築して利用した。また、コンパイルにあたっては、インテル社製の CPU 用に最適化された数値計算ライブラリ (MKL) を、BLAS (ベクトル和、行列積などの基礎的な線形計算ライブラリ) および LAPACK (LU 分解、固有値分解などのライブラリ) の代わりにリンクして用いた。

非線形因子分析の尤度計算の部分は C 言語によって作成し、R のパッケージとして組み込む方式をとった。さらに尤度計算を高速化するために OpenMP (マルチコア用の共有メモリ並列計算支援のためのコンパイラ仕様) を用いて並列計算が実行されるようにした。OpenMP の利用により、6 コアの CPU (Xeon E5-1660) でシングルスレッドに比べておおよそ 1/3 に計算時間の短縮が可能となった。シングルスレッドでは、本稿の計算例で 4 時間超の経過時間を必要としたが、OpenMP を用いた計算コードでは約 1 時間 15 分であった。

64 ビット Linux (CentOS 6.4) 上でインテル社製コンパイラ (Intel Composer XE 2013) を用いて R-2.15 システムをビルドする場合の指定法は、後記の著者サイトに示した。インテルのサイトに示されている R システムのビルドの例では `-fp-model precise` の指定はなされていないが、この指定がないと R システムのビルド後の `make check` においてエラーが生じる。

5. 主な発表論文等

(研究代表者、研究分担者及び連携研究者には下線)

〔雑誌論文〕(計 1 件)

大津起夫 (2011) 大学入試センター試験における科目別得点の非線形因子分析による比較. 大学入試センター研究紀要, 40, 1-23. (査読有)

〔学会発表〕(計 5 件)

大津起夫 (2014) センターにおける科目選択形態別の得点分布 -- 非線形因子分析による比較 --. 日本行動計量学会第 42 回大会, 2014-09-05, 東北大学, 宮城県仙台市

大津起夫 (2014) 標準化英語試験とセンター試験英語科目得点との関係分析. 日本テスト学会第 12 回大会, 2014-08-30, 帝京大学, 東京都八王子市

Otsu, T. (2014). Multidimensional IRT analysis of large scale data in NCT (poster).

The 9th Intern. Conf. of the International Test Commission (ITC2014). 2014-07-03, San Sebastian, (Spain).

大津起夫 (2013). 記号処理言語を用いた問題テキスト処理プログラム. 日本テスト学会第 11 回大会, 2013-08-28, 九州大学, 福岡県福岡市.

Otsu, T. (2012) Conditional likelihood of stratified categorical data for analysis on a large scale examination. JCS-CLADAG 2012 (日伊分類学会), 2012-09-04, Anacapri, (Italy).

〔図書〕(計 0 件)

〔産業財産権〕
出願状況 (計 0 件)

名称：
発明者：
権利者：
種類：
番号：
出願年月日：
国内外の別：

取得状況 (計 0 件)

名称：
発明者：
権利者：
種類：
番号：
取得年月日：
国内外の別：

〔その他〕
ホームページ等

<http://www.dnc.ac.jp/~otsu/R215withMKL.html>

6. 研究組織

(1) 研究代表者

大津 起夫 (Otsu Tatsuo)

大学入試センター・研究開発部・教授

研究者番号：10203829

(2) 研究分担者

宮埜 壽夫 (Miyano Hisao)

大学入試センター・研究開発部・教授

研究者番号：90200196

(3) 連携研究者

()

研究者番号：