

平成21年 5月 27日現在

研究種目：若手研究（B）

研究期間：2006～2008

課題番号：18700271

研究課題名（和文） 高次元離散分布からの効率的なサンプリング手法に関する研究

研究課題名（英文） Efficient sampling methods from discrete sampling space of higher dimensions

研究代表者

青木 敏（AOKI SATOSHI）

鹿児島大学・理学部・准教授

研究者番号：90332618

研究成果の概要：

高次元の離散分布からの効率的なサンプリング手法として、マルコフ連鎖を利用したモンテカルロ法を実行するために、以下の3つのテーマに関する研究を行なった。（1）必要となる連結な推移基底（マルコフ基底）の性質解明。（2）純粹数学（代数学）の分野で既存の結果を利用した応用研究。（3）実験計画法の問題に対する定式化と応用研究。

交付額

（金額単位：円）

	直接経費	間接経費	合計
2004年度			
2005年度			
2006年度	1,100,000	0	1,100,000
2007年度	1,100,000	0	1,100,000
2008年度	1,100,000	330,000	1,430,000
総計	3,300,000	330,000	3,630,000

研究分野：数理統計学

科研費の分科・細目：情報学・統計科学

キーワード：マルコフ連鎖・モンテカルロ法、グレブナー基底、マルコフ基底、分割表、トーリックイデアル、実験計画法、Segre-Veronese 型配置、階層モデル

1. 研究開始当初の背景

マルコフ連鎖・モンテカルロ法は、離散データに対する条件付き検定の有意確率を計算する、という問題に対し、有効な方法のひとつであるが、特に高次元の問題になると、連鎖な推移基底（マルコフ基底）の計算が困難となることが知られていた。実際、計算機の発達と普及にとともに、大規模データの解析

が必要となる応用の場面は増えており、高次元の構造をもつ実社会のさまざまな問題について、マルコフ基底の計算とその性質解明は、重要な問題である。これまでに、研究代表者は、東京大学の竹村彰通教授との共同研究によって、マルコフ基底の中でも極小な構造をもつものについて、その性質をある程度解明していた。一方で、極小性以外のマルコフ基底の構造解明や、得られた理論を用いた

実際のマルコフ基底計算のアルゴリズムの開発については、十分な結果が得られていなかった。

2. 研究の目的

(1) マルコフ基底のうち、極小性に加えて不変性(分割表に対する、軸の入れ替えや、水準の入れ替えを、一般化した群作用に関する不変性として定式化した)をもつものの性質解明を行う。得られた結果を用いて、実際のマルコフ基底の計算の高速化を目指すと共に、得られた基底をデータベースとして保存することを念頭に置いた、起動の代表元による簡明な表記法についても考察する。

(2) 実問題への適用としては、極小マルコフ基底を含む集合であるグレブナー基底を利用したマルコフ連鎖・モンテカルロ法も念頭に置き、グレブナー基底の性質を利用したアルゴリズムの開発と実問題への適用を行う。

(3) 別の実問題への適用として、多因子の一部実施計画によって計数データが得られる場合についても、条件付き検定問題を定式化する。また、従来の分割表に関するマルコフ基底の結果との対応を明らかにしながら、さまざまな実験計画法の問題に対してマルコフ基底を求める。

3. 研究の方法

これまで通り、東京大学の竹村彰通教授との共同研究を中心に、研究を遂行した。また、代数学の研究者である、大阪大学の日比孝之教授、立教大学の杉英史准教授とも、共同研究を行った。得られた研究成果は、国内外の研究集会、学会において報告を行い、また、論文としてまとめて投稿した。

4. 研究成果

(1) まず、マルコフ基底の極小性の概念を拡張して、マルコフ基底の要素(二項式として表わされる)の各項が、いずれかのマルコフ基底の要素の項として必ず含まれるか否かの特徴付けを行い、必須項(indispensable monomial)の定式化を行った。この結果により、すべての極小マルコフ基底には含まれない基底成分(dispensable binomial)についても、その項の必須性からさらなる分類が可能となった。また、マルコフ基底の中でも不変な構造を持つもの(不変マルコフ基底)について、その極小性、一意極小性の特徴付け、極小不変マルコフ基底の構造、不変群の定義

とその構造解明を行った。不変群の構造は、モデルから自然に想定される「自然な」群が、実際に最大の不変群となるか否かを示すことが困難であるが、応用上重要である、多元分割表における独立モデルや、遺伝子型データにおける Hadry-Weingerg モデルを含む、いくつかの具体的な問題について、証明を行った。また、計算代数の分野で近年注目されている、数独の問題についても、不変群の構造を解明した。

(2) 代数学の分野で Segre-Veronese 型として知られる問題について、そのグレブナー基底に関する性質を利用した、グレブナー基底によるサンプリング手法を考案した。考案した手法は、二項式の項に現れる変数の添え字の入れ替えを考えることにより、グレブナー基底の元を、効率的にサンプリングするものである。また、マルコフ基底でなく、(集合として、マルコフ基底よりも大きい)グレブナー基底を直接推移基底として利用することにより、極小マルコフ基底を用いるよりも収束の面で望ましい性質をもつことも、本研究の意義のひとつである。応用面では、センター試験の科目選択データに対するさまざまな階層的なモデルや、遺伝子データ(アレル頻度データ、haplotype 頻度データ、diplotype 頻度データ)に対するさまざまな平衡仮説のモデルが、Segre-Veronese 型として表記できることを示し、これらの実データに対するマルコフ連鎖・モンテカルロ法による仮説の検証を行った。

(3) 計測値が計数値であるような多因子要因計画について、特に計画が regular とよばれるものについて、条件付き検定問題を定式化した。また、2 水準、3 水準のそれぞれの設定に対して、regular な実験計画と従来の高次分割表との対応を明らかにし、マルコフ基底の導出、マルコフ連鎖・モンテカルロ法の実行を行った。実験計画により得られるデータは、因子の効果に関しては階層モデルを想定していても、それを高次の分割表に対応させた場合には、対応する階層モデルが存在しない例があることを示し、従来の分割表の枠組みでは得られない、新たなモデルに対するマルコフ基底が必要となることを示した。特に、水準数が 3 の場合に、従来の分割表の枠組みでは得られなかった、3 要素からなる集合を結ぶ基底要素が、必須基底成分として現れる、という、興味深い性質を得た。提案する手法は、実験計画法により得られる計数データに対する解析手法として、新しい視点を与えるものである。また、regular な実験計画について、因子の主効果、交互作用の交絡関係は、多項式環のイデアル所属問題として定式化できる点にも注目し、本研究と、当

該分野に対する計算代数の従来 of 研究成果との関連を明らかにした。

(4) *regular* でない一部実施実験計画に対して、母数の推定可能性の観点から示唆される、新たなクラスを提案し、その代数的、幾何的な性質を調べた。提案するクラスは、どの *regular* な一部実施計画の部分集合にもならない、と特徴付けることができ、これは、幾何的には、 $GF(2)$ におけるアフィン部分空間の次元が縮退していないことに対応することを示した。また、提案するクラスの性質として、従来の *D*-最適基準との関連を明らかにするために、行列式の値に関する基本的な性質を証明し、実験回数が 99 以下の知られている *D*-最適計画 (主効果モデル) と、提案するクラスの関連を調べた。その結果、実験回数に関する周期性、つまり、実験回数を 8 で割った余りが 5, 6, 7 となる場合に限り、*D*-最適計画が提案するクラスに含まれる、という興味深い事実を発見し、行列式の値が既知の極限值 (bound) を達成する場合には、常にこの性質が成り立つことを証明した。

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 9 件)

- (1) S. Aoki and A. Takemura (2009). Some characterizations of affinely full-dimensional factorial designs, *Journal of Statistical Planning and Inference*, accepted. 査読あり
- (2) T. Sei, S. Aoki and A. Takemura (2009). Perturbation method for determining group of invariance of hierarchical models, *Advances in Applied Mathematics*, accepted. 査読あり
- (3) S. Aoki and A. Takemura (2008). Markov basis for designed experiments with three-level factors, in *Algebraic and Geometric Methods in Statistics, Mathematical explorations in contemporary statistics, (dedicated to Professor Giovanni Pistone on the occasion of his sixty-fifth birthday)*, edited by P. Gibilisco *et al.*, 220—233. 査読あり
- (4) S. Aoki, T. Hibi, H. Ohsugi and A. Takemura (2008). Groebner bases of nested configurations, *Journal of Algebra*, **320**, 2583—2593. 査読あり
- (5) S. Aoki, T. Hibi, H. Ohsugi and A. Takemura (2008). Markov basis and Groebner basis of Segre-Veronese

configuration for testing independence in group-wise selections, *Annals of the Institute of Statistical Mathematics*, accepted. 査読あり

- (6) S. Aoki, A. Takemura and R. Yoshida (2008). Indispensable monomials of toric ideals and Markov bases. *Journal of Symbolic Computation*, **43**, 490—507. 査読あり
- (7) S. Aoki and A. Takemura (2008). The largest group of invariance for Markov bases and toric ideals, *Journal of Symbolic Computation*, **43**, 342—38. 査読あり
- (8) S. Aoki and A. Takemura (2008). Minimal invariant Markov basis for sampling contingency tables with fixed marginals, *Annals of the Institute of Statistical Mathematics*, **60**, 229—256. 査読あり
- (9) S. Aoki and A. Takemura (2007). Statistics and Groebner bases – the origin and development of computational algebraic statistics, *Sugaku*, **59**, 283—302. 査読あり

[学会発表] (計 9 件)

- (1) S. Aoki and A. Takemura. Some characterization of affinely full-dimensional factorial design, Computational Algebraic Statistics, Theories and Applications (CASTA2008), Kyoto, 11-12 December, 2008.
- (2) 青木敏、竹村彰通。regular でない一部実施計画のあるクラスの性質。日本計算機統計学会第 22 回シンポジウム、臨床研究情報センター、神戸市、2008 年 11 月 6 日。
- (3) 青木敏、竹村彰通。3 水準の一部実施計画に対する MCMC 法による要因効果の検定。2008 年度統計関連学会連合大会、慶應義塾大学、2008 年 9 月 8 日。
- (4) 青木敏、竹村彰通。一部実施計画における要因の推定可能性と指示関数で表現される交絡の関係。2008 年度統計関連学会連合大会、慶應義塾大学、2008 年 9 月 8 日。
- (5) S. Aoki and A. Takemura. Markov basis for design of experiments with three-level factors. Mathematical explorations in contemporary statistics, (dedicated to Professor Giovanni Pistone on the occasion of his sixty-fifth birthday), Genova, 19-20 May, 2008.
- (6) 青木敏、竹村彰通。一部実施計画に対する MCMC 法による要因効果の検定。2007

年度統計関連学会連合大会、神戸大学、
2007年9月8日。

- (7) 青木敏、日比孝之、大杉英史、竹村彰通。
グループごとに制約がある選択問題の条件付検定手法と Segre-Veronese 型配置のグレブナー基底。2007年度統計関連学会連合大会、神戸大学、2007年9月8日。
- (8) 青木敏、宮川雅巳。2値入出力系に対する多因子要因効果の検定。2007年度統計関連学会連合大会、神戸大学、2007年9月8日。
- (9) 青木敏、日比孝之、大杉英史、竹村彰通。
グループごとに制約がある選択問題の条件付検定手法と Segre-Veronese 型配置のグレブナー基底。日本数学会 2007年度年会、埼玉大学、2007年3月29日。

6. 研究組織

(1) 研究代表者

青木 敏 (AOKI SATOSHI)
鹿児島大学・理学部・准教授
研究者番号：90332618

(2) 研究分担者

(3) 連携研究者