

平成 10 年度研究報告会要旨

と き：1999 年 3 月 16 日 午前 10 時～午後 4 時 45 分
17 日 午前 10 時～午後 5 時 45 分
ところ：統計数理研究所 講堂

プログラム

3 月 16 日 (火)

午前 (10 時～12 時 15 分)

あいさつ

所長 清水 良一

【統計科学情報センター】

球面上の配置の問題 (II)

種村 正美

多変量解析における信頼性

馬場 康維

Estimation for a Scale Parameter with Known Coefficient of Variation

金藤 浩司

外乱の非ガウス成分に対する系確率密度の応答

岡崎 卓

金融機関の統計的監査

山下 智志

Application of OR Techniques in Finance (客員, Univ. of Otago) Premachandra, I. M.

【予測制御研究系】

統計モデルと知識発見

北川源四郎

株式市場の共和分解析

川崎 能典

脊椎動物の系統進化

長谷川政美

午前 (13 時 15 分～16 時 45 分)

GPS データからの時空間情報の抽出

樋口 知之

情報量規準の差における選択バイアスと多重比較法

下平 英寿

予測と非線形モデルの最近の発展

尾崎 統

不完全情報下における制御系設計に関する研究

宮里 義彦

Web Decomposition の成果

佐藤 整尚

Prior Model Selection Criterion, RCC

田邊 國士

自己双対線形計画問題と内点法

水野 眞治

選挙の票割の変遷

上田 澄江

ノンパラメトリック統計学と J 言語

鈴木義一郎

MZ ファミリーを用いた二次錐計画法に対する主双対内点法

土谷 隆

拡張アンサンブルの応用

伊庭 幸人

周波数帯相互作用モデル / Interactive Frequency Band Model

石黒真木夫

2次元化に基づく非定常信号のスペクトル解析と

アンチエイリアスフィルタリング

対称錐の平均演算と情報幾何について

瀧澤 由美
(客員, 大阪大学) 小原 敦美

3月17日(水)

午前(10時~12時15分)

【領域統計研究系】

事前情報が曖昧な場合の Bayes 検定の限界

臨床試験における国際統計ガイドライン

文化計量のための基礎研究

A Social Quantum Theory for the Analysis of Public Opinion Survey Data

Nested case-control 研究での Mantel-Haenszel タイプ推定量 (客員, 宮崎大学) 藤井 良宜

柳本 武美
佐藤 俊哉
村上 征勝
吉野 諒三
藤井 良宜

【調査実験解析研究系】

方法論としての電子調査法に関する考察

目黒区の住民意識調査から

インターネットの学術利用調査

確率過程における情報量規準の構成

(客員, 横浜市立大学) 川浦 康至
前田 忠彦
丸山 直昌
内田 雅之

午後(13時15分~17時45分)

世帯主年齢別データのコウホート分析

子どもの体験活動等に関する調査結果について

電子調査法とその周辺 — いくつかの実験調査を例として —

統計数理研究所での10年間の研究総括

時空間解析に関する二・三の話題

「第10次日本人の国民性調査」の企画について

中村 隆
土屋 隆裕
大隅 昇
高木 廣文
柏木 宣久
坂元 慶行

【統計基礎研究系】

二変量マルコフ試行列における待ち時間問題

正規確率場の最大値の分布 — tube の方法と Euler 標数の方法 —

マルチンゲールに対するエントロピー法

確率分布関数のフラクタル近似

単調正則変動関数の分解とその応用

主成分ベクトルのロバストな推定量のあるクラスについて

平野 勝臣
栗木 哲
西山 陽一
松縄 規
志村 隆彰
紙屋 英彦

【統計計算開発センター】

余震活動から周辺地域の大地震を予測できるか?

非線形時系列解析

密行列計算に適した汎用計算機について

TIMSAC72 for workstation について

視覚的操作による時系列解析支援システム

半無限計画法

尾形 良彦
田村 義保
泰地真弘人
荒畑恵美子
中野 純司
伊藤 聡

統計科学情報センター

球面上の配置の問題 (II)

種村 正美

球面上にできる限り均等に分布する点配置の問題を考える。従来から知られている球面上のミニマクス最適問題と最小エネルギー問題について論じた後、これらと原理的に全く異なる、最適配置を得るための「球面調節法」とよぶ新しい方法を提案した。

まず、球面調節法は点の個数が少ない場合、ポアソン初期配置から大局的最適配置を生成できること、点の個数によっては、ミニマクス解や最小エネルギー解に一致する配置が得られることが示された。また、球面調節法は、広範囲の点の個数に対して局所的最適配置を生成するためにも有用であることを示した。

また、われわれの新しい方法は、注意深く選ばれた良い配置を初期条件に用いると、その配置を改善したより良い配置が得られることが判明した。その一例として、反発相互作用の下でマルコフ連鎖モンテカルロ法によって作られた平衡配置が、球面調節法で改善されることが示された。

参 考 文 献

種村正美 (1998). 球面上の最適配置の問題, 統計数理, 46, 359-381.

多変量解析における信頼性

馬場 康 維

二つの話題について報告した。

1) 数量化Ⅲ類の結果は分析対象のクロス集計表の頻度分布に依存している。したがって、頻度分布に何らかの仮定を置くことによって、Ⅲ類による結果の信頼性についての情報が得られる。数量化スコアの信頼領域を求め、信頼領域が重なった点を一つのクラスターとみなすことにより、変数あるいは個体のクラスター分析を行うことができる。

2) ランキングデータの構造をグラフィカルに発見する方法の一つである順位グラフはランキングの頻度表から描ける。上記と同様に、アイテムを表現する点の信頼性領域が導出できる。これを利用して、アイテムのクラスター分析を行うことができる。

Estimation for a Scale Parameter with Known Coefficient of Variation

金 藤 浩 司

実数直線上で定義された分布の位置母数の損失関数としては、二乗損失関数が一般的である。しかし、二乗損失関数は、正の領域で定義された分布において、尺度母数の推定量の損失関数としては適切ではない。そこで、Wasan (1970) が提案した損失関数が、損失関数の性質と正の

領域で定義された分布の確率変数とカイ2乗分布との関係が存在する意味で適切であることを示した。同時に、真の母数の近傍では、Wasan (1970) が提案した損失関数は、二乗損失関数に一致する。また、この損失関数に基づいて、母変動係数が既知の場合、尺度母数の不偏推定量の族の中で、最適な推定量を求めた。

参 考 文 献

Wasan, M. T. (1970). Parametric estimation, McGraw-Hill, New York.

外乱の非ガウス成分に対する系確率密度の応答

岡 崎 卓

1. はじめに

外乱を受けて変動しながら発展する系の確率密度は、外乱の非ガウス性によってどのように左右されるであろうか。外乱が Rayleigh 過程に従う場合など外乱が特定されていれば、系の応答確率密度は外乱変数と系変数の結合系に対する Fokker-Planck (FP) 方程式により、あるいはシミュレーションにより調べることができるが、特定の外乱と応答確率密度の関係ではなく、外乱のもつ非ガウス性と応答確率の一般的関係についてはこれを探る系統的方法は未だに開発されていない。本報告では、ガウス外乱に限定せずに系変数の確率密度を定め得る一般化 Fokker-Planck (GFP) 方程式を拠り所として、外乱の skewness や flatness など非ガウス性を象徴する統計量で応答確率密度を表現する試みについて述べる。

2. 指数分解型 e -GFP 方程式の解

外乱 W を受けつつ、確率微分方程式 $(d/dt)U = M(U) + W(t)$ に従って変動する変数 U の確率密度 $f(u, t)$ を定める GFP 方程式に、指数演算子分解近似による簡約と補正を施した e -GFP 方程式は次の形をもつ。

$$\begin{aligned} \frac{\partial}{\partial t} f(U, t) = & -\frac{\partial}{\partial U} M(U) f(U, t) + \left(\frac{\partial}{\partial U}\right)^2 \lambda_A(t) f(U, t) \\ & + \left(\frac{\partial}{\partial U}\right)^2 \int dV \lambda_B(V, t) f(U - V, t) \end{aligned}$$

ここに λ_A は local な拡散係数、 λ_B は non-local な拡散係数であって、ガウス外乱、あるいは時間的相関を持たない外乱に対しては λ_B は消滅する。従って外乱の非ガウス性の効果は λ_B に集約されているわけである。

定常線形過程に限定すれば、この方程式の解を具体的に表現できる：

$$f(U) = \int dV f_{\text{Gauss}}(V) \cdot f_{\text{mod}}(U - V).$$

ここに $f_{\text{Gauss}}(U)$ はガウス分布の確率密度関数、 $f_{\text{mod}}(U)$ は λ_B を与えれば確定する関数であって、非ガウス外乱を受ける線形系の応答確率密度は、ガウス分布に外乱の非ガウス成分に対応

した変調関数 f_{mod} による変調を施したものとなる。

3. 変調関数と応答確率密度の一般表現

変調関数 f_{mod} を定めるには、拡散係数 λ_B を構成する多時間統計量

$$\langle W(a) \delta(W(b) - w) W(c) \rangle \quad (a \leq b \leq c)$$

を外乱 $W(t)$, $t \in R^+$ の特性に即して計算することが鍵となる。外乱の特性汎関数 $\widehat{G}[\phi] = \widehat{G}[\phi(\cdot)]$ とその cumulant 表現

$$\begin{aligned} \widehat{G}[\phi(\cdot)] = \exp & \left[\frac{1}{2} \int h(t_1, t_2) i\phi(t_1) i\phi(t_2) dt_1 dt_2 \right. \\ & \left. + \frac{1}{6} \int h(t_1, t_2, t_3) i\phi(t_1) i\phi(t_2) i\phi(t_3) dt_1 dt_2 dt_3 + \dots \right] \end{aligned}$$

を活用してその計算を実行すれば、変調関数 f_{mod} を外乱の 3 次や 4 次の cumulant と 1 時間確率密度 $g(w, t)$ で表現することができ、外乱の非ガウス性を表わすパラメータと応答確率密度との具体的関係が得られる。

金融機関の統計的監査

山下 智志

金融機関の破綻が現実的になってきた今日、銀行のリスク許容が実体経済に与える影響がますます増加している。そこで、政府としては銀行のリスクに対する安全管理を監督する必要性が生じている。本発表では、こうした問題に対する政策として、銀行のリスク許容が適当であることを、金融当局が監査する方法について論議した。

確率的に発生する事象に対する評価は、従来の会計的な監査ではとらえることができない。そこで、統計的手法を使った取り組みが必要となる。特に、銀行が提出するリスク計算プロセスが適当であるかどうかを、一定の基準で評価し、指導する必要がある。しかし、計算プロセスには、多くのステップが存在するため、論理的にはその優劣が証明しづらい。

本発表では、新たな評価基準として、分布の乖離をローレンツ曲線と 45 度線の間面積 (ジニ係数) を採用した。このローレンツ曲線は金融分野では q-q プロットといい、一部で利用されていた。そのためこの評価基準は、純粋に統計的な見地からいえば正確性に欠けるが、実務家が理解しやすく、金融当局の評価を得ている。

また、発表では評価基準のほか、確率的リスク会計のバックボーンについて簡単に説明し、現在統計学者が実業界から求められている貢献について紹介した。

Application of OR Techniques in Finance

(Visiting Professor) University of Otago I. M. Premachandra

Portfolio managers face the following two types of problems when investing their funds.

They face a diverse mix of conflicting investment objectives (constraints or goals) such as risk and return optimizing goals, liability constraints of the institution, legal restrictions, cash flow restrictions and achieving the performance targets of the portfolio manager. It is shown (Premachandra and Powell (1998)) that the Goal Programming technique can effectively be used in this case to find the investment proportions while trying to achieve these goals simultaneously. On the other hand, they might face cash outflows and even possible job loss if their short-term performance lags behind the competition. These problems are severe in countries where the small number of predominantly resource based companies has meant that a few unfortunate stock picks may lead to sharp divergences among managed fund performance. Therefore, ranking of the most efficiently performed funds is quite important to the fund manager. The article (Premachandra et al. (1998)) argues that a spreadsheet-based DEA (data envelopment analysis) model can be used in this case for ranking of the DMUs (decision making units). The DEA model which is formulated on a Excel spreadsheet (Premachandra (1998)) is simulated using @RISK add-in and the average efficiency measure and its standard deviation obtained from the simulation are used to define a new ranking measure. It was shown by using non-parametric tests that the DEA ranking is consistent with the traditional penalty function ranking.

However, the original DEA model proposed by Charnes et al. in 1978 allows each DMU to decide its factor weights freely in order to demonstrate its own value system. Sometimes, DMUs can abuse this freedom by assigning bigger weights for its favorable inputs and outputs while ignoring the unfavorable ones in order to appear as efficient thus producing nonsensical results. We propose (Premachandra (1999a)) an AHP (analytic hierarchy process) based procedure for controlling this factor weight flexibility which incorporates the decision maker's value judgement about the relative importance of input output factors into the DEA model. Furthermore, Zhu (1998) has shown that there is a consistency between the DEA ranking and the ranking based on PCA (principal component analysis) but his approach is not valid in some cases. We modify (Premachandra (1999b)) his approach so that the new approach gives significant results even in extreme cases where majority of the DMUs are efficient.

References

- Premachandra, I. M. (1998). A simplified guide to Excel, Postgraduate Institute of Management, Colombo, Sri Lanka.
- Premachandra, I. M. (1999a). Controlling factor weights in data envelopment analysis by incorporating decision maker's value judgement: An approach based on AHP (under review).
- Premachandra, I. M. (1999b). A note on DEA vs. principal component analysis: An improvement to Joe Zhu's approach (under review).
- Premachandra, I. M. and Powell, J. G. (1998). Accommodating diverse institutional investment objectives and constraints using non-linear goal programming, *European Journal of Operational Research*, **105**, 447-456.
- Premachandra, I. M., Powell, J. G. and Shi, J. (1998). Measuring the relative efficiency of fund management strategies in New Zealand using a spreadsheet-based stochastic data envelopment analysis model, *Omega*, **26**(2), 319-331.
- Zhu, J. (1998). Data envelopment analysis vs. principal component analysis: An illustrative study of economic performance of Chinese cities, *European Journal of Operational Research*, **111**, 50-61.

予測制御研究系

統計モデルと知識発見

北川 源四郎

統計モデルを対象・目的に応じて適切に構成することにより、データの持つ本質的な情報を取り出すとともに、知識発見へと導くことができる。多くの場合、従来のモデルでは捉えられない部分は、パラメータの変動や残差の分布の偏りや変動の形に反映される。

パラメータの時間変化を自動的に捉える方法として自己組織型の状態空間モデルを開発した。このモデルでは、本来の状態空間モデルの状態ベクトルを x_n 、パラメータを θ とするとき、拡大した状態を $z_n = (x_n^T, \theta_n^T)^T$ と定義するものである。このとき、この z_n に関する状態空間モデルに非線形平滑化を適用することによって、状態推定と同時にパラメータの時間変化を検出することができる。本報告では、確率的ボラティリティ・モデル

$$\begin{aligned}x_n &= \beta x_{n-1} + \alpha + v_n \\r_n &= x_n + \log w_n^2\end{aligned}$$

$v_n \sim N(0, \tau^2)$, $w_n \sim N(0, \sigma^2)$ の場合について、係数パラメータ α , β および分散パラメータ τ^2 , σ^2 の時間変化を推定した結果を報告した。

また、ピアソン分布族などの非ガウス型ノイズ分布を持つ状態空間モデルを用い、その形状パラメータなどを自己組織型平滑化によって推定することによって、あらかじめ分布形を特定しなくても、自動的に平均値などのパラメータの急激な変化も自動的に捉えることができる。

一方、推定されたモデルの残差の中に新しい知識発見の手掛かりがみつかった例として、地下水位データの解析結果を示した。この例では、観測時系列から適当なモデリングによって、気圧、潮汐、降雨の影響を除去することによって地下水位中の地震の影響が明確となり、地震のマグニチュードや距離と地下水位の変化の関係などが検出できた。

参 考 文 献

Kitagawa, G. (1998). Self-organizing state space model, *J. Amer. Statist. Assoc.*, **93**, 1203-1215.

株式市場の共和分解析

川崎 能典

資産価格、特に株価の時系列特性は定常からはほど遠く、個々の銘柄の価格を将来にわたって予測することは殆ど不可能である。しかしながら、特定の銘柄の組み合わせによる線形結合に注目すれば、スプレッドは少なくとも平均非定常の域を脱し、時には統計的に定常と見なせる場合がある。資産運用の現場でしばしば行われるロング・ショートを用いるペアトレード戦略は、おおむねこのような動機に基づくものであるが、この研究の目的は、ペアトレード戦略の妥当性を東証一部に限って統計的に検証することである。技術的には、まず流動性等によるスクリーニングを経た後で、個々の株式時系列に単位根検定を行い、単位根プロセスと判定された銘柄全ての組み合わせで共和分検定を行う。割高・割安の目安を差で見ると比で見るとはトレーダーの好みの問題であると言えるが、仮に株価比が1より大か小かで判断するとすれ

ば、株価の対数系列に対する共和分モデルで、共和分ベクトルに特定の線形制約がかけられているケースに等しい。一般には市場環境に依るが、少なくとも近年の東証一部において、統計的にそのような共和分関係が妥当性を持つ組み合わせは20から100程度であり、リスク水準を高く維持することは難しいという結論を得た。一方、株価そのものの線形結合から生み出されるスプレッド系列を元に運用を考えることも可能である。この場合は必ずしもマーケットニュートラルな戦略にならない分、株価比で判断する場合よりも共和分関係を有するペア数が増え、リスク枠を効率的に利用することが可能である。マーケットニュートラルリティからの乖離度合いは、近年の市場環境からすると整合的ではあるが、総じて若干ショート気味のポジションとなった。(なお、この研究は日興証券・立木成氏との共同研究である。)

GPS データからの時空間情報の抽出

樋口知之

GPSとはGlobal Positioning Systemの略であり、複数の人工衛星からの相対的な位置情報を総合的に解析することによって、地球上のあらゆる地点の高精度の位置情報を随時提供する目的のために構成されたシステムである。GPSデータは、このGPSに関わるさまざまなデータをさし、一般に理解されているような高精度位置データのみをさすものではない。そもそも人工衛星と地上間の相対的位置は、相互間を伝播するmicrowave帯の電磁波信号から算出されるため、伝播に関連する物理量の変位——propagation delay——は、最終的に相対的位置情報として解釈・変換される相互間の距離以外に、伝播途中の媒体の物理的性質によって生じる。この自然現象によるpropagation delayの主たる要因の一つは、地上受信点と人工衛星間に存在する大気中の水蒸気量である。この水蒸気によるpropagation delayは、当然のことながら観測点の気象状況に依存し、結果として、ある地点の位置を高精度に計測するという測地学的見地からは、時間・空間に強く依存したノイズとして取り扱われる。

もし、propagation delayから大気中の水蒸気量を適切に推定することができたならば、これは新しい気象観測方法を我々は手に入れたことになる。日本においては、建設省国土地理院によって1,000点近くものGPS観測点がこの数年間で敷設されている。この観測点が構成するネットワーク——GPS arrayと呼んでいる——上で水蒸気量の推定が可能となったならば、これは気象庁のもつAMeDAS (Automated Meteorological Data Acquisition System) という呼称で一般に知られている降雨量観測ネットワークの空間分解能と比較して、同等あるいはそれ以上のものとなる。このような木目細かいGPSネットワークは世界でも例が無い。さらに重要なことは、AMeDASでは実際の降雨量を計測しているだけだが、一方GPSではprecipitable water vapor (PWV)を推定することが原理的にはできるため、天気予報のためのシミュレーションモデル中の物理量との整合性が高く、結果として予測力の高い天気予報システムの開発を支援することができる。本年度は、GPS日データの上下変動成分中に顕著に見られる、たぶんPWVによる影響である時間・空間に強く依存した成分、いいかえれば季節変動成分を推定し、その影響を合理的に除去する準備的解析を行った。

情報量規準の差における選択バイアスと多重比較法

下 平 英 寿

情報量規準もしくは対数尤度のバラツキを現実のデータにおいて評価すると、モデルの良さの差が有意とは言えない場合が多い。データが十分に多ければ AIC や MDL などの情報量規準を最小にするモデルだけが他に比べて有意に良いと言えるのだが、そうはならない場合が多いのである。モデル選択の一致性に関する議論はデータが無限に多くなった場合のものであり、現実のデータは有限である。またデータが多くなるにつれて、より詳細なモデルの比較をしたり、同時に比較するモデルの数が増えたりすることも、現実のモデル選択における不確実性を増す原因となっていて、データマイニングでは特に注意が必要である。

情報量規準の小さい順にモデルを並べ、1 番のものからの規準の差の有意性を検定することによってモデルの信頼集合を構成する方法が、遺伝データから生物の系統樹を推定する問題で広く用いられている (Kishino and Hasegawa (1989))。しかし 1 番のモデルの規準の値は選択バイアスにより小さくなる傾向があり、この効果を考慮しないと規準の差が有意と誤って判断される確率が増える。下平 (1993, 1999), Shimodaira (1998) は多重比較法 (multiple comparisons) をモデル選択に適用することによって、この選択バイアスを考慮してモデル信頼集合を構成した。

Shimodaira and Hasegawa (1999) では、本方法を哺乳類の分子系統樹の問題に応用した。6 種の哺乳類について、ミトコンドリア DNA にコードされたアミノ酸シーケンスをデータとする統計解析を行い、系統樹の信頼集合を計算した。これにより、分子生物学の分野で主張されていた幾つかの互いに矛盾する仮説が、じつは信頼性の評価に誤りがあり、正しい計算の結果どちらも信頼集合に含まれることが分かった。

参 考 文 献

- Kishino, H. and Hasegawa, M. (1989). Evaluation of the maximum likelihood estimate of the evolutionary tree topologies from DNA sequence data, and the branching order in Hominoidea, *Journal of Molecular Evolution*, **29**, 170-179.
- 下平英寿 (1993). モデルの信頼集合と地図によるモデル探索, 統計数理, **41**(2), 131-147.
- Shimodaira, H. (1998). An application of multiple comparison techniques to model selection, *Ann. Inst. Statist. Math.*, **50**, 1-13.
- 下平英寿 (1999). モデル選択理論の新展開, 統計数理, **47**(1), 3-27.
- Shimodaira, H. and Hasegawa, M. (1999). Multiple comparisons of log-likelihoods with applications to phylogenetic inference, *Molecular Biology and Evolution*, **16**, 1114-1116.

この他、論文やプレプリントについては、ホームページ <http://www.ism.ac.jp/~shimo/> を参照するか、shimo@ism.ac.jp まで EMAIL を下さい。

予測と非線形モデルの最近の発展

尾 崎 統

最近の非線形モデルの発展

非線形モデルに関連するいくつかの研究分野で最近大きな発展があった。今年度は非線形時

系列モデリングにおける以下のようなトピックスで研究を進めた。

1) ExpAR モデルの拡張

AR モデルの係数を振幅 $x(t-1)$ だけでなく、速度、 $x(t-1) - x(t-2)$ 、や加速度、 $x(t-1) - 2x(t-2) + x(t-3)$ に従属する形にすることにより非線形性の機能を拡大させることを試みた。また施招雲博士らと共同で RBF-AR モデル、ノンパラメトリック AR モデルの可能性と限界などを追及した。

2) ジャンプ拡散過程とヴォラティリティ

ジャンプ拡散過程の推定とその応用、特にファイナンス時系列のフィルターと予測を飯野光徳氏と共同で行った。Hawkes 型の Self Exciting Point Process のようなクラスター型の点過程とヴォラティリティな時系列を結び付ける新領域に発展する可能性が出てきた。

3) 火力発電プラントの非線形 ARX 制御

PID 制御下にある火力発電プラントを多入力多出力非線形システムと見て外生変数つき多次元 AR モデルを用いてシステム同定し非線形制御系を設計する方法を研究した。実プラントでの実験データの他 40 次元の微分方程式からなるシミュレーターによって生成した人工データにも応用して有効性をチェックした。H. Peng 外国人研究員と日本ベレーの豊田、織田両氏らとの共同研究。

4) 気象データのアシミレーション

気象の数値予報の精度をあげる為には全地球上の観測データと大規模な気象力学系モデルを結び付けるデータアシミレーションが最も重要な仕事の一つとなる。データアシミレーションの効率をよくする為に NIWA の Zheng 博士らと共同で局所線形フィルターをこれに応用する計画に着手した。

5) サポートベクターマシン (SVM) とダイナミックシステム

SVM を非線形力学系のロバスト推定に利用することに取り組んだ。

不完全情報下における制御系設計に関する研究

宮里 義彦

制御のためのモデルの設定と同定から制御手法までを総括的に含む統合化制御系設計理論の構築を試みている。その一環として、モデリングと制御の接点を扱う適応制御の基礎理論の研究や、実用化のための様々な制約を取り除いた適応制御系の設計の研究を行っている。

この数年間は外乱や非線形成分、次数や相対次数に依存しない適応制御系の構成法の研究に関わってきた。その中でこれまでに、可変構造制御系、ハイゲインフィードバックやバックステッピング法などにより、外乱・非線形成分・次数、そして部分的に相対次数についても従来の制約を緩和した適応制御手法を導出することができた。今年度はさらに相対次数に3次の幅の不確定性がある場合や、対象の高調波利得の符号に不確定性がある場合でも対処できる適応制御系の構成法、具体的にはモデル規範形適応制御系、適応サーボ系、次数が未知の場合に近似的なモデルマッチングを達成する非線形モデル規範形適応制御系の構成手法について研究を行い、参考文献にあげるような結果を得た。

これらと平行して、バックステッピング法に代わる適応系の安定化手法についても考察を進めた。ハイゲインオブザーバを用いた構造の簡単な適応制御系の構成手法を求め、そのロバスト特性についても解析を行い、参考文献にあげるような結果を得た。

また従来の適応制御理論が漸近安定性に主眼を置くのに対して、制御性能をもより定量的に考慮する観点から、適応制御の過程を最適制御問題や H_∞ 制御問題としても定式化することを試みた。モデル規範形適応制御を含む一般的な形式の適応制御問題について、安定解析に用いるリアプノフ関数の一部と Hamilton-Jacobi (Isaacs) 方程式の解を同一視することで、特定の評価関数に対して最適 (または準最適) な 3 つの型の適応制御系の構成法を導出した。

参 考 文 献

- Kawabe, T., Isobe, O., Watanabe, Y., Hanba, S. and Miyasato, Y. (1998). New semiactive suspension controller design using quasi-linearization and frequency shaping, *Control Engineering Practice*, **6**, 1183-1191.
- 宮里義彦 (1998). 次数と相対次数に依存しない非線形適応制御系の構成法 (相対次数が 3 次の範囲で不確定な場合), システム制御情報, **11**(6), 330-340.
- Miyasato, Y. (1998a). A design method of universal adaptive servo controller, *Proceedings of the 37th IEEE Conference on Decision and Control*, Vol. 2, 2294-2299.
- Miyasato, Y. (1998b). A simple redesign of model reference adaptive control system and its robustness, *Proceedings of the 37th IEEE Conference on Decision and Control*, Vol. 3, 2880-2885.
- Miyasato, Y. (1998c). A design method of universal model reference adaptive controller, *Proceedings of 1998 American Control Conference*, Vol. 1, 609-613.
- Miyasato, Y. (1999). Model reference adaptive control for a class of nonlinear systems with unknown degrees and uncertain relative degrees, *Proceedings of 1999 American Control Conference*, Vol. 1, 571-575.
- 宮里義彦 (1999). 次数と相対次数に依存しない適応サーボ系の一構成法, 計測自動制御学会論文集, **35**(3), 450-452.
- 大森浩充, 宮里義彦 (1998). 適応制御の潮流, システム/制御/情報, **42**(6), 287-299.

Web Decomp の成果

佐藤 整 尚

WWW 上で簡単に時系列解析のできるシステム “Web Decomp” を運用してきた。このシステムはインターネットブラウザのみで手元のデータを最新手法で解析できるもので、実質的な計算はすべて統数研の計算機で行われる。今年度は新しい Version に作り変え、より使いやすいものとなった。この約 1 年間の総アクセス数も 19732 にのぼり、前年度に比べて、大幅に増加した。これは、新しい Version を導入したこととともに、サーバーを 2 台に増やしたことが功を奏したものと思われる。前回の調査では企業等の実務家からのアクセスが多かったが、今年は、大学関係からのアクセスも増えてきた。実際の研究等への利用も進んでおり、ローカルでの利用を求める声もいくつかあった。来年度以降はこのような要請に応じて、ローカルな環境でも使えるようにプログラムを作成するつもりである。

自己双対線形計画問題と内点法

水野 眞 治

線形計画問題を解く内点法は、Karmarkar (1984) により提案されて以来、活発に研究されて

きた。線形計画問題を実際に内点法で解く場合に重要な問題の一つに、初期点をいかに求めるかといったことがある。これにはいくつかのアプローチがあり、大きな係数を使った人工問題を作る方法、初期点をまず求めてから問題を解く Phase I-Phase II 法、勝手な正の初期点を使うインフィージブル内点法などがある。ここでは、それらとは異なり、自己双対システムを使う方法を研究した。この方法は、Ye et al. (1994) により初めて提案したが、特徴として、勝手な初期内点を使うことができる、反復回数が多項式オーダーで理論的におさえられている、問題の実行不能性を判定できる、実行可能な場合には解を計算できる、big M と呼ばれる大きな係数を必要としないといったことがある。

本研究では、二つの異なった自己双対システムを使った内点法を調べ、その違いと共通点について研究した。二つの方法とは、上記の Ye et al. (1994) で提案した方法と最近 Nesterov et al. (1999) により提案された方法である。これらの方法は、センターパスを逆方向に追跡するといったように、一見全く異なるように見える。水野 (1998) と Mizuno and Todd (1998) では、これら二つの方法の関連を詳細に調べあげた。その結果、一部のパス追跡法では、ある条件の元で2つの方法により生成される点列が全く等しくなることを明らかにした。このようなことは、ごく限られた方法のみで見られ、その他のパス追跡法では、異なった点列が生成される。

参 考 文 献

- Karmarkar, N. (1984). A new polynomial-time algorithm for linear programming, *Combinatorica*, 4, 373-395.
- 水野眞治 (1998). 自己双対システムを使った内点法, 京都大学数理解析研究録, 1068, 171-181.
- Mizuno, S. and Todd, M. J. (1998). On two homogeneous self-dual systems for linear programming and its extensions, Research Memo., No. 687, The Institute of Statistical Mathematics, Tokyo.
- Nesterov, Y. E., Todd, M. J. and Ye, Y. (1999). Infeasible-start primal-dual methods and infeasibility detectors for nonlinear programming problems, *Math. Programming* (to appear).
- Ye, Y., Todd, M. J. and Mizuno, S. (1994). An $O(\sqrt{n}L)$ -iteration homogeneous and self-dual linear programming algorithm, *Math. Oper. Res.*, 19, 53-67.

選挙の票割の変遷

上 田 澄 江

日本では各政党はそれぞれの選挙区における候補者の公認を、車を道路に駐車するように割り振っているようにみえる。この路上駐車の問題、すなわち、一次元パッキングモデルを参議院選挙全国区データ (1947~1980) と衆議院選挙データ (1949~1993) に適用した。また、参院選全国区における相転位、すなわち、生態学研究にしばしば適用されるブローケン・スティック・モデルからランダム・パッキング・モデルへの移行期を観察した。

政党が各候補者に組織票を分配する方法が路上駐車の問題であるとは、道路の長さが1政党が獲得した票数に相当し、1台の車は1候補者に相当し、1台の車の前からすぐ後ろの車の前までの長さは1候補者の獲得した票数に相当する。これは各候補者がいかに地盤を獲得するかを説明するものである。

各政党は支持者数に比例して選出される候補者数を最大にするように努力をしている。政党にとって最も有効な方法とは、選出されるために必要な最小の得票を全候補者が得ることである。このような状況下において、ランダム・シークエンシャル・パッキングの手続きは政党が

候補者を公認するときのよい手法であると思われる。公明党と共産党の各候補者の得票は非常に効果的に分布していることがわかる。

衆院選は 1993 年までは中選挙区制で遂行されていたが、1996 年以降は小選挙区制と比例代表制が導入され、また、参院選全国区は 1980 年を最後に比例代表制によるドント方式が採用された。選挙制度改革に賛成した自民党・社会党にとって、綿密な地盤割りを要しないドント式は有利であった筈である。無党派層が多数をしめる今日、第 2、第 3 の相転位が同じ尺度で観察できないのは残念に思われる。

拡張アンサンブルの応用

伊庭 幸人

マルコフ連鎖モンテカルロ法の新しい流れとして、与えられた分布族を合併あるいは拡張したものを不変分布とするマルコフ連鎖を用いる手法群 (拡張アンサンブル法) がある (伊庭 (1998))。これによって、与えられた分布をそのまま扱うよりもマルコフ連鎖の混合が飛躍的に速くなる場合があることが報告されている。代表的な手法としてはマルチカノニカル法、交換モンテカルロ法 (Metropolis-Coupled Chain) などがある。われわれは、格子タンパクモデルに、マルチカノニカル法を拡張した拡張アンサンブル法を適用した。この手法のポイントは、ある割合で自己回避条件を破ることを許し、予備的なシミュレーションでその割合がちょうどよくなるように自己交差に対する罰金の大きさを学習させることにある。いままでの実験によって、この方法は分布からのサンプリングと最適化の両方について非常に有力なことが示された (Iba et al. (1998), Chikenji et al. (1999))。

参 考 文 献

- Chikenji, G., Kikuchi, M. and Iba, Y. (1999). Multi-self-overlap ensemble for protein folding: ground state search and thermodynamics, *Physical Review Letters*, **83**, 1886-1889.
- 伊庭幸人 (1998). 拡張アンサンブル — MCMC の新しい流れ —, 統計学会予稿集 1998 年度, 中央大学, 東京.
- Iba, Y., Chikenji, G. and Kikuchi, M. (1998). Simulation of lattice polymers with multi-self-overlap ensemble, *J. Phys. Soc. Japan*, **67**, 3327-3330.

対称錘の平均演算と情報幾何について

(客員) 大阪大学 小原 敦 美

2 つの正数に対して算術, 幾何, 調和平均が定義され, この 3 者間に成立する大小関係は様々な不等式の基礎となっているが, その拡張として 2 つの正定値対称行列 X, Y にも

$$\begin{aligned} \text{算術平均 } X_a Y &:= (X + Y)/2, \\ \text{幾何平均 } X_g Y &:= X^{1/2}(X^{-1/2} Y X^{-1/2})^{1/2} X^{1/2}, \\ \text{調和平均 } X_h Y &:= 2(X^{-1} + Y^{-1})^{-1} \end{aligned}$$

が定義され正定値性の順序に関して同様な大小関係が成立する (Kubo (1980)).

本研究ではこれらの平均演算が、正定値対称行列上の情報幾何 (Ohara et al. (1996)) から自然に定義される ∇ , Levi-Civita, ∇^* の各接続に対して, X, Y を結ぶそれぞれの測地線 ($t \in [0, 1]$)

$$\begin{aligned} \nabla\text{-測地線} & X + t(Y - X), \\ \text{Levi-Civita 測地線} & \exp \{t \log (X^{-1/2} Y X^{-1/2})\}, \\ \nabla^*\text{-測地線} & \{X^{-1} + t(Y^{-1} - X^{-1})\}^{-1} \end{aligned}$$

の中点となることを示し, 平均を生成する作用素単調関数 (Kubo and Ando (1980)) まで立ち戻ってその関係を調べた。

また, 上記の平均演算と情報幾何の関係はさらに一般に対称錘 (Faraut and Korányi (1994)) まで拡張可能であることも示すことができる。

参 考 文 献

- Faraut, J. and Korányi, A. (1994). *Analysis on Symmetric Cones*, Oxford Press, New York.
 Kubo, F. and Ando, T. (1980). Means of positive linear operators, *Math. Ann.*, **246**, 205-224.
 Ohara, A., Suda, N. and Amari, S. (1996). Dualistic differential geometry of positive definite matrices and its applications to related problems, *Linear Algebra Appl.*, **2247**, 31-53.

領域統計研究系

事前情報が曖昧な場合の Bayes 検定の限界

柳 本 武 美

統計的検定に対する批判は、帰納的推論、更には証拠に基づく推論そのものへの批判から決定理論・モデル選択等に至るまで多くある。しかし実証的で検証的な推論の標準的な形式の骨子を成している。

近年根強い批判の1つとして、Lindley パラドックスに始まる Bayes 検定からの検定が一致性を持たないとの主張がある。当初は病理的な事実とみなされたようであるが、Bayes 推測に対する期待の高まりと情報・通信 (特に圧縮) 理論の進展の中でこの批判は重要性を増している。統計的仮説との対比で、データの蒐集を考慮に入れて考察すると、Bayes 検定がより有用であるのは

- 1) 帰無仮説が本当に正しいことが信じ得る
- 2) 帰無仮説からの観測誤差に比べてたわずかなずれも有意義である
- 3) 大量のデータが (解析者にとって) 短期間に得られる

を満たす場合であると結論される。直観的にみても、Bayes 推論が有効なときはデータの情報量に比べて事前分布の情報が無視できない場合である。大量の良質な情報が手に入るときに敢えて情報量の少ない事前分布 (曖昧な事前情報) に強く依存する推論を行うことはない。

以上の結果は現在統計数理に投稿中である。また関連した論文を参考文献にあげる。

参 考 文 献

- 柳本武美 (1991). 統計的検定における帰無仮説の理解, 応用統計学, 20(2), 97-108.
柳本武美 (1994). 統計的推測における手法の妥当性, 統計数理, 42(2), 215-224.

臨床試験における国際統計ガイドライン

佐藤俊哉

日米欧三極医薬品承認審査調和国際会議 (ICH) では, 新医薬品承認審査の迅速化, 共通化を目指して, 新医薬品承認申請のための臨床試験に関する様々なガイドラインを作成している。筆者は東京理科大学吉村氏とともに, その一つである “Statistical Principles for Clinical Trials” の作成に携わってきた。この ICH ガイドラインは 1998 年 2 月に日米欧で承認されているが, 日本語訳「臨床試験のための統計的原則」が 1998 年 11 月 30 日に厚生省医薬安全局審査管理課長から各都道府県に通知された。今後は, このガイドラインに沿った臨床試験を実施しなければならないことになった。

これに先立ち, 統計数理第 46 巻 1 号の特集「ヘルスサイエンスと統計科学」の中で, 「臨床試験のための統計的原則」のミニ特集を企画した。ICH ガイドラインの作成に携わった吉村氏には, ガイドライン作成の際に他国の委員と議論になった点を中心にまとめてもらい, 旧統計解析ガイドライン作成のメンバーであった椿・藤田氏には, これまでのわが国の臨床試験について, コントローラー委員会の活動を中心にまとめてもらった。この 2 編の論文に対し, 著者相互に議論するほか, アカデミックな視点から大橋氏, 審査を行う側の視点から山口・安藤・秋田氏, メーカーの視点から佐々木・酒井氏の議論を加え, 再度吉村, 椿・藤田による回答を掲載した。

臨床試験の現場にかかわっている生物統計家による白熱した議論が掲載されているので, ぜひ一読をお勧めする。

参 考 文 献

- 厚生省医薬安全局審査管理課長 (1998). 「臨床試験のための統計的原則」について (平成 10 年 11 月 30 日医薬審第 1047 号). <http://www.nihs.go.jp/dig/ich/ichindex.htm> より pdf ファイルとしてダウンロード可。

A Social Quantum Theory for the Analysis of Public Opinion Survey Data

吉野諒三

本年度は, 「米国西海岸の日系人調査 (科学研究費補助金・基盤研究 A [2])」, 「第 6 回世界青少年の国際比較調査 (総務庁青少年対策本部, 十一か国比較)」, 「日米青少年の親子の意識調査」, 「第 10 回日本人の国民性調査」に関する米国オムニバス調査, 日本オムニバス調査などの多様な比較実験調査を実践的に遂行した。これらの調査は, それぞれの回答分布の分析も重要ではあるが, 住民票や選挙人名簿などの既存のリストが存在しない場合の特殊なサンプリング

の工夫,あるいは国によって異なるサンプリング法で収集したデータの比較などの問題解釈を要し,国際比較調査データを実践的に取り扱う際の重要な情報,知見を与えた。

理論面では,統計数理研究所の既存のパネル調査データ(同じサンプルに,ある期間を置いて同じ質問をする)を2次分析し,力学系の理論の視点から「一時収束点」などの挙動の考察を推進し,「社会調査データ解析の為の社会量子理論」の構築を試みた(Yoshino(1998))。特に,「波動尺度」を構成し,戦後から今日までの国政選挙の投票率の推移の説明,将来の変化予測等の応用を試み,理論を有効に発展,応用させることが出来る感触を得た。

参 考 文 献

- Yoshino, R. (1998). A social quantum theory for the analysis of public opinion survey data, *Behavior-metrika*, 25(2), 111-132.
統計数理研究所国民性国際調査委員会 編 (1998). 「国民性七か国調査」, 出光書店, 東京。

Nested case-control 研究での Mantel-Haenszel タイプ推定量

(客員)宮崎大学 藤 井 良 宜

疫学研究において,明確に定義されたコホート内で,ケースコントロール研究を行う研究のことをネステッドケースコントロール研究と呼んでいる。このデザインでは,一般のケースコントロール研究にくらべて,コホートのメンバーが明確であるばかりでなく,ある種のコホートメンバーに関する情報が既に得られているという点が特徴である。Langholz and Borgan (1995) は,この研究デザインにおける counter-matching と呼ばれる新しいコントロールのサンプル方法を提案した。このサンプリング方法は,研究対象の曝露あるいは曝露の代用変数がある場合に,この変数によって層別してサンプリングする方法で,「ケースが得られた層では,コントロールの数を一つ減らす」というこれまでの層別サンプリングとは異なる方法が取られている。このサンプリング方法における推定方法としては,部分尤度法が提案されているが,ここでは Mantel-Haenszel タイプの推定関数を用いた推定方法を提案した。この方法は,不偏な推定関数に基づく方法であり,ある種のロバスト性を持つとともに,部分尤度法と比べて,それほど効率が減少しないことを示した。

参 考 文 献

- Langholz, B. and Borgan, Ø. (1995). Counter-matching: A stratified nested case-control sampling method, *Biometrika*, 82, 69-79.

調査実験解析研究系**方法論としての電子調査法に関する考察**

(客員) 横浜市立大学 川 浦 康 至

1. 今年度の仕事

電子調査としてインターネット上のウェブ調査を取り上げた。具体的には、重点領域研究(代表者:大隅昇)の共同研究者としてウェブ調査の統計的検討とともに、調査法としての側面についても考察した。それらの成果のうち、筆者が直接報告したのは川浦(1998a, 1998b)である。

2. 主な知見**(1) 実験的ウェブ調査の実施**

各種の世論調査で用いられた設問をウェブ調査の質問項目に盛り込み、両者の回答結果を比較することで、ウェブ調査の特性を抽出しようとした。対人ネットワーク密度に関する設問では、ウェブ調査による回答の方が低密度であった。こうした背景として、ネットワーク利用による影響や利用者の特性が考えられた。

(2) ウェブ調査の構成要素による検討

ウェブ調査とは狭義に考えると、質問紙法の一つに位置付けられる。しかし現実には回答者の選定など、それ以外の側面まで混同して議論されるケースが多く、いわば混乱状態にある。ウェブ調査を意義のある調査法とするためには、調査の構成要素(対象者、調査票、実施場面など)に注目した検討が欠かせないことが明らかになった。

参 考 文 献

- 川浦康至(1998a)。調査空間としてのインターネット, 日本行動計量学会第 26 回大会発表論文抄録集, 275-276。
川浦康至(1998b)。電子ネットワークにおける自己表出, 日本社会心理学会第 39 回大会発表論文集, p. 11。

目黒区の住民意識調査から

前 田 忠 彦

1997 年度と 1998 年度に目黒区住民(成人男女)を対象として実施した、人間関係観や満足感をテーマとする意識調査の概要を報告した。

上記調査は、1998 年 2 月に行った同一テーマの全国調査の予備調査及び補足的吟味調査として実施したものであり、日本人の人間関係観や日常生活に対する満足感を中心とする意識の動向を探る資料とすることが目的であった。このため、本研究所が行っている「日本人の国民性調査」との比較を図るべく同調査でも用いられている質問項目を用いた他、今後の意識動向を探るためのオリジナルの質問項目を含めて調査票を構成した。

調査実施要領は以下の通りであった。

調査対象母集団 調査年度の 10 月 1 日現在で 20 歳以上の東京都目黒区全域に在住す

る男女。

調査方法 1997年度は留置法，1998年度は郵送法。

調査時期 調査年度の10月～11月。

標本抽出方法 等間隔抽出を用いた2段抽出法。詳細は下記の通り。

	1997年度調査	1998年度調査
利用台帳	有権者名簿	住民基本台帳
抽出地点数	61	72
地点当り標本数	5	約7
計画総標本数	305	500
回収標本数	150	247
不能率	50.8%	50.6%

上記標本について、母集団年齢構成との比較、および年齢別調査不能率の検討を行ったところ、1) 計画標本段階では母集団の年齢構成からのズレが少ないが、回収標本においてはズレが大きく、特に男性標本でこの傾向が顕著であること、2) このズレの原因は、若年（特に男性）層における回収率の低さを反映したものであること、3) ズレの程度は1998年度調査の方が顕著であり、これは郵送法という調査方法の特質によるものである可能性があること、などが指摘された。

また、同時期に行われたオムニバス調査等との重複質問計17項目に関して、上記調査との回答分布の比較を行ったところ（オムニバス調査は全国調査であるが、そのうち「大都市」のサンプルのみを集計した）、両者の結果の差が無視しうる程度の項目と、無視し得ない差を生じる項目とがあり、それぞれの項目の特質を反映していることが推察された。

なお、本研究は平成9年度10年度科学研究費（奨励研究；課題番号09780221）の補助を受けて行った研究の一部である。

インターネットの学術利用調査

丸山直昌

インターネットは、当初学術研究ネットワークとして発展して来たが、近年は商用的にも広く利用されるようになり、今や現代社会の基盤をなすと言って良いほどの存在となっている。このため最近はその利用の形態や価値、功罪について多方面から議論されているが、元々の学術利用に関する議論はかえって下火になっているようにさえ感じられる。そこで、昨年筆者が関係する社団法人日本ネットワークインフォメーションセンターでは、インターネットの学術利用の実体を探るためのアンケート調査を実施した。調査票を日本学術会議に登録されている全学術研究団体と、文部省科学研究補助金採択課題から抽出した約1000課題のプロジェクトリーダーへ送付し、約32%の回答を得た。この調査により、インターネットの学術研究分野での利用が未だに充分では無いことが伺えるが、その原因の推定についてもこの調査は有効な情報を与えていると思われる。調査結果の分析と次回調査への展望を概観した。

確率過程における情報量規準の構成

内 田 雅 之

赤池情報量規準 (AIC) が導出されて以来、竹内情報量規準 (TIC) や一般化情報量規準 (GIC) など、さまざまな状況においてモデル選択が行なえるような情報量規準を構成する研究が盛んに行なわれている。本年度の研究の要旨は以下の通りである。

幾何強ミキシング条件を満たす連続時間確率過程モデルにおける情報量規準を確率解析を使って導出した。特に、従来の情報量規準の論拠である平均値不偏性だけでなく、2次漸近中央値不偏性に基づく別の情報量規準 (MUIC) の構成も行なった。

参 考 文 献

- Uchida, M. and Yoshida, N. (1998). Information criteria in model selection for stochastic processes, Research Memo., No. 709, The Institute of Statistical Mathematics, Tokyo.

世帯主年齢別データのコウホート分析

中 村 隆

コウホート分析は、継続調査データの分析法であり、年齢×調査時点別に集計されたデータ表から年齢・時代・コウホート (世代) 効果を分離することにより、社会の変動についての知見を得るための方法である。この方法は、本来は個人の年齢やコウホートを想定しており、総務庁の「家計調査」などで得られる世帯単位の集計データに適用するには問題がある。

家計調査データは消費動向を捉える情報源として重要であり、コウホート分析を適用する意義は大きい。そのため世帯主年齢別データであるという問題点があっても、適当な方法がないためそのままコウホート分析を適用することがほとんどである。この場合、20歳未満の年齢効果は当然得られず、また年齢効果やコウホート効果の解釈には十分気を付ける必要がある。

第2の接近法として、もし調査世帯の個票データを処理できるのであれば、ある種の回帰モデルを用いて個人単位の消費量を推計し、その推計値に対してコウホート分析を適用することができる。ただし、膨大な作業が必要であり、実際には限られた調査年について少数の項目の推計値しかなかかなか得られない。

そこで、第3の中間的接近法として、揃えやすい世帯単位の集計データと、世帯主年齢別世帯員年齢構成データを組み合わせるコウホートモデルを考えた。統計モデルとしては、通常のコウホートモデルのデザイン行列に年齢構成データ行列を右からかけたものを新たにデザイン行列とすればよい。

この第3の接近法によって、たとえば「米」の分析結果では、第2の接近法によれば得られる年齢効果の10歳台の消費量のピークが得られている。もちろん、これは第1の接近法では得られないものであり、簡便性と有用性をもつ接近法として有望である。

なお本研究は、統計数理研究所共同研究 (10-共研 A-89) により実施されたものである。

電子調査法とその周辺 — いくつかの実験調査を例として —

大 隅 昇

改善の兆しの見えない調査環境の急速な変化の中にあつて、新たな調査法として登場した電子調査法が益々注目を集めている。昨年度に引き続いて今年度も電子調査法に関連した研究を進めたが、この一年間の動向には良きにつけ悪しきにつけ著しいものがあつた。とくに、インターネット環境を背景として登場した Web 調査（インターネット・サーベイ、インターネット・リサーチ）と呼称される新しい調査法が抱える諸問題を、その方法論、従来の調査法との関連、とくに標本調査設計（標本抽出、調査票作成、実査方法等）に関わる諸事項を組織的に研究した。とくに、市場調査の分野への Web 調査の普及・浸透は著しく、面接法・郵送法調査等との比較研究も広く行われるようになってきた。この他、消費者動向、ネットワーク・コミュニケーション等の実務・応用分野では、多様かつ大量のデータベース化蓄積データ、電子メール・WWW ホームページ等を活用した電子調査関連の応用研究が積極的に行われている。

こうした社会情勢を念頭に、また昨年度の研究成果を踏まえて、Web 調査を用いた際の調査環境のあり方、調査法としての適用可能性や利用の限界を、実査を通じて具体的に取得した調査データを用いて検証することに重点をおいた。また国内外の関連した先行的研究や調査報告の問題点を明らかにするとともに、実査を行うための調査設計・企画を初めとする一連の調査過程をなるべく総合的に追跡し分析するように努めた。これらの状況を勘案して Web 調査を可能な限り客観的に評価し検証するため、以下の方針で研究を進めた。

- (1) 昨年度に実施した 12 回の Web 調査で取得のデータを用いてより詳細な解析を行い、これらを基に現状の Web 調査の類型化を試みたこと。
- (2) 社会の要請に応えるべく研究成果を一般に公開し、学会・諸機関における公開セミナー、シンポジウムを組織的に行ったこと。こうした情報公開を通じて、Web 調査に対する社会の期待や関心の対象がどのような事にあるかの情報収集を行ったこと。
- (3) さらに Web 調査を調査法としていかに確立すべきかを、具体的な実査を通じて検証したこと。
- (4) 実査に際しては、昨年度の経験を生かして、調査協力機関を複数（4 機関）とする、各機関での実査時期を揃える、同一の調査票を用いる、各サイトで複数回の調査を行う、通常の調査方式（面接方式オムニバス他）を併用する等、独自の調査計画を立案し実施したこと。
- (5) また、調査機関を複数とすることで、より顕在化するであろう調査環境の間の類似や差異の客観的評価を試みたこと。
- (6) Web 調査を用いる利点の一つとされてきた自由回答取得のあり方の研究とその統計解析法、解析ソフトの開発を併せて行ったこと。

この一連の実験調査は国内で初めての組織的な共同研究であつたが、これにより現時点における国内の Web 調査の実態を実証的に明らかにすることができた。また、これらの調査結果はインターネットを通じて公開される。

なお、本研究の一部は、研究課題名「調査環境の変化に対応した新たな調査法の研究」（研究課題番号 09206117；平成 9 年度・10 年度）として文部省科学研究費補助金特定領域研究（領域番号：114，領域略称名：マイクロ統計データ）による助成を受けた。

参 考 文 献

- Ohsumi, N. and Yoshimura, O. (1999). The online survey in Japan: An evaluation of emerging methodologies (Invited paper), Meetings of "Improving the Effectiveness of Data Collection through Innovative Technology", in the 52nd Session of the International Statistical Institute (ISI99), August 10-18, 1999 in Helsinki.
- Yoshimura, O., Ohsumi, N., Kawaura, Y. et al. (1998). Some experimental trials of electronic surveys on internet environments, *Advances in Data Science and Classification*, 663-668, Springer, Heidelberg.

時空間解析に関する二・三の話題

柏 木 宣 久

大規模なデータを解析しなければならない状況が増えると考え、典型的な大規模データ解析である時空間解析へと発展し得る問題を特に選び、その解決に取り組んできた。本報告では、そうした問題の内、領域内の複数の井戸で観測された水頭の時空間データから領域内の地下帯水層を推定する時空間逆解析、東京湾の複数の地点で観測された水質の時空間データから水質の動向を推定する時空間平滑化および時空間季節変動調整、および地球環境のリモートセンシングに現れる時空間解析へと発展し得る問題等について述べた。

「第 10 次 日本人の国民性調査」の企画について

坂 元 慶 行

「日本人の国民性」の統計的研究には、(1)日本人の意識動向の解明、(2)調査法の研究、(3)解析法の研究の三つの目的がある。

「第 10 次 全国調査」の第一の目的は、(国民性調査の質問で可能な範囲で) 20 世紀後半期の意識動向を総括することである。そこで、過去に 1 回しか調査されたことのない質問でも、現在と比較することによって興味ある調査結果が得られる可能性があれば、できるだけ調査に付すことにした。

第二の目的は、調査法の研究、特に、調査方式の調査結果に与える影響について検討することである。今日のように、統計調査環境が悪化し回答の質の低下が憂慮されている状況の下で意識の動向を的確に把握するためには、つぎのような課題について研究する必要がある。一つは、これまで当研究所で実施してきた様々な調査結果を活用するために、従来の調査結果との継続性を有する調査方式を見い出すことであり、もう一つは、様々な専門調査機関の調査結果を活用するために、それらの調査結果を解読する方法について研究することである。このため、一方で、同時期に複数の調査方式で調査を行ない、調査方式の違いによる回答の揺らぎを測定する、と同時に、他方で、各調査方式を固定し、継続して調査することによって各調査方式の下での意識の動きを追跡することとした。調査方式に関しても質問量に関してもやや不完全ながら、前々回 (1988 年) と前回 (1993 年) の 2 回は、この構想にそって調査を設計した。今回 3 時点目の調査を同じ方式で行なうこととし、ようやく可能になる 3 時点の時系列変化に基いてその系列間比較を行ない、所期の目的の達成に努めたい。

統計基礎研究系

二変量マルコフ試行列における待ち時間問題

平野 勝 臣

本年度の研究

1. m 次 $\{0, 1\}$ -値マルコフ系列において、長さ k の成功連がはじめて起こるまでに、長さ ℓ ($m \leq \ell < k$) の成功連が起こる回数の分布はオーダー $(k - \ell)$ の幾何分布に従うことが知られている。これを negative version と呼ぶならば、これに対応する positive version についてまとめた (Aki and Hirano (1998))。
2. 2次元 $\{0, 1\}^2$ -値マルコフ系列において sooner and later waiting time の問題についての結果をまとめた (Aki and Hirano (1999))。

ここでは Aki and Hirano (1999) の結果について述べる。

二変量マルコフ試行列における待ち時間問題

二変量 $\{0, 1\}$ -値マルコフ系列, すなわち $\{0, 1\}^2$ -値マルコフ系列において, 第一成分で長さ k の 1 の連が起こる事象と, 第二成分で長さ r の 1 の連が起こる事象を考える。このとき, 2つの事象のどちらかが先に起こるまでの待ち時間 (試行数) 分布 (sooner waiting time problem) と両方が初めて起こるまでの待ち時間 (試行数) 分布 (later waiting time problem) を調べた。

具体的及び技術的なことは直接論文を参照されたい。ここではこの問題が極めて一般的な待ち時間問題であることを注意しておくことにとどめる。 (S, \mathcal{S}) を任意の標本空間とし, 独立で同一の分布に従う S -値確率変数列 Z_1, Z_2, \dots とする。二つの事象 A と $B \in \mathcal{S}$ が連続して起こることに興味をもつ。 A と B は必ずしも排反とは限らない。 $i = n - k + 1, \dots, n$ に対して $Z_i \in A$ であるとき, n 番目の試行で長さ k の A -run が起こったということにする。長さ k の A -run と長さ r の B -run についての sooner and later waiting time の問題を考えるためには $\{0, 1\}^2$ -値二変量確率ベクトル $\begin{pmatrix} I_A(Z_i) \\ I_B(Z_i) \end{pmatrix}$, $i = 1, 2, \dots$ を観測すればよい。ここに $I_A(\cdot)$ は A の indicator function である。考察している問題は独立試行列の場合にも十分興味深いものである。 A と B が排反で余事象であれば $\{0, 1\}$ -値確率変数列 $I_A(Z_1), I_A(Z_2), \dots$ を観測すれば十分であり, この場合は多くの研究に帰着される。

参 考 文 献

- Aki, S. and Hirano, K. (1998). Number of success runs of specified length until some stopping times and generalized binomial distributions of order k , Research Memo., No. 670, The Institute of Statistical Mathematics, Tokyo.
- Aki, S. and Hirano, K. (1999). Sooner and later waiting time problems for runs in Markov dependent bivariate trials, *Ann. Inst. Statist. Math.*, **51**(1), 17-29.

正規確率場の最大値の分布 — tube の方法と Euler 標数の方法 —

栗 木 哲

推測統計の基本的な道具である尤度比検定においては、その検定統計量の帰無仮説の下での分布(帰無分布)の上側裾確率の計算が応用上重要である。正則な統計モデルにおいてはこの帰無分布が少なくとも漸近的にはカイ 2 乗分布であることはよく知られている。しかしながら母数空間に特異点を含むような確率モデルにおいて、その特異点を帰無仮説とする尤度比検定を考えると、帰無分布の漸近カイ 2 乗性はもはや成り立たない。このような場合には、ある種の確率場の最大値の分布として特徴付けられる分布が統計量の帰無分布としてしばしば現れる。

こういった問題意識を出発点として、本研究では最初にこれらの確率場の最大値の分布を積分幾何学的手法で導出するという研究を展開した。とくに、Sun (1993) の与えた分布の上側裾確率の漸近展開の誤差評価を行った。また、確率場が滑らかな閉多様体を添字集合とする正規確率場である場合には、Sun の tube 法と、Worsley (1995) らの Euler 標数法と呼ばれる 2 つの異なる導出原理が、数学的には同等で正しい結果を導くことを証明した(栗木・竹村(1999))。

また、この一般論を個々の検定問題に適用することを行った。その結果、従来は解析的扱いが困難とされていた幾つかの検定統計量の分布の上側裾確率を、陽な漸近展開の形で与えることに成功した。具体的には、最大特異値を特殊な場合として含むような多重線形形式の最大値の分布、Johnson and Graybill (1972) の検定を特殊な場合として含む多元配置における交互作用の検定統計量、多次元正規性の検定に関する Malkovich and Afifi (1973) の検定統計量などの上側裾確率を導出した(Kuriki and Takemura (1998))。

参 考 文 献

- Johnson, D. E. and Graybill, F. A. (1972). An analysis of a two-way model with interaction and no replication, *J. Amer. Statist. Assoc.*, **67**, 862-868.
- Kuriki, S. and Takemura, A. (1998). Tail probabilities of the maxima of multilinear forms and their applications, Discussion Paper CIRJE-F-4, Faculty of Economics, University of Tokyo.
- 栗木 哲, 竹村彰通 (1999). 正規確率場の最大値の分布 — tube の方法と Euler 標数の方法 —, *統計数理*, **47**(1), 201-221.
- Malkovich, J. F. and Afifi, A. A. (1973). On tests for multivariate normality, *J. Amer. Statist. Assoc.*, **68**, 176-179.
- Sun, J. (1993). Tail probabilities of the maxima of Gaussian random fields, *Ann. Probab.*, **21**, 34-71.
- Worsley, K. J. (1995). Boundary corrections for the expected Euler characteristic of excursion sets of random fields, with an application to astrophysics, *Adv. Appl. Probab.*, **27**, 943-959.

マルチンゲールに対するエントロピー法

西 山 陽 一

$\theta \rightarrow \gamma(\theta)$ はパラメータ $\theta \in \Theta$ をもつ deterministic process であるとする。我々は関数 $\theta \rightarrow \gamma(\theta)$ の最大点 θ_0 を推定することに興味があるとする。もしも $\theta \rightarrow \gamma(\theta)$ がある stochastic process $\theta \rightarrow \Gamma_n(\theta)$ によって“よく”近似されるならば、自然な推定量は后者の最大点 $\hat{\theta}_n := \operatorname{argmax}_\theta \Gamma_n(\theta)$ である。すなわち、criterion process $\theta \rightarrow \Gamma_n(\theta)$ に関する M -推定量である。そのような M -推定の手続きのうち、 (Θ, d) を一般の距離空間にしたようなものは、最近になっ

て研究されてきた。一般論によると、

$$\gamma(\theta) - \gamma(\theta_0) \leq -d(\theta, \theta_0) \quad (\theta_0 \text{ の近傍において})$$

$$E \sup_{d(\theta, \theta_0) \leq \delta} |(\Gamma_n - \gamma)(\theta) - (\Gamma_n - \gamma)(\theta_0)| \leq \varphi_n(\delta), \quad \forall \delta > 0$$

という2条件がチェック出来れば、 $\varphi_n(r_n^{-1}) \leq r_n^{-2}$ をみたす定数列 $r_n > 0$ に対して

$$\lim_{K \rightarrow \infty} \limsup_{n \rightarrow \infty} P(r_n d(\hat{\theta}_n, \theta_0) > K) = 0$$

が成立する。

このアプローチのポイントは、上記の2番目の不等式を示すところである。そこで、筆者はそのような「最大不等式」を、連続マルチンゲール、離散時間マルチンゲール、点過程に関する確率積分、の3種類の状況下でエントロピーの言葉で証明し、その応用をそれぞれ得た。その一例として、連続マルチンゲールに対する結果を用いた拡散過程のノンパラメトリック推定問題について述べる。確率微分方程式

$$dX_t^n = \theta(X_t^n) dt + n^{-1/2} dW_t, \quad X_0^n = x_0 \in \mathbb{R}$$

における未知関数 $\theta(\cdot)$ を推定する問題を考える。常微分方程式

$$dx_t = \theta(x_t) dt, \quad x_0 \in \mathbb{R}$$

の解を $x = (x_t^{\theta})_{t \in [0, 1]}$ とし、擬距離

$$d(\theta, \vartheta) = \sqrt{\int_0^1 |\theta(x_t^{\theta_0}) - \vartheta(x_t^{\theta_0})|^2 dt}$$

を導入する。そこで、最尤推定量のこの擬距離に関する収束率を、上述の一般論を用いて導出することができる。

参 考 文 献

- Nishiyama, Y. (1998). Entropy methods for martingales, Ph. D. Thesis, University of Utrecht.
 Nishiyama, Y. (1999). A maximal inequality for continuous martingales and M -estimation in a Gaussian white noise model, *Ann. Statist.* (to appear).

確率分布関数のフラクタル近似

松 縄 規

連続型分布の分布関数を近似するために、Stieltjes 変換の中のある有理関数部分に、絶対収束性と自己相似性を持つ無限級数展開を利用し、従来は評価が難しかったいくつかの確率分布関数の近似を精度よく行えることを検証した。

この過程で、次の型の Stieltjes 変換 $\int_0^{\infty} (x+t)^{-1} g(t) dt$ ($x+t \neq 0$) は統計理論の中で屢々重要な役割を演じることが分かった。量 $(x+t)^{-1}$ は一般に次の様な有限和表現が可能である。 $x+t \neq 0$, $x \neq 0$ とし、 $\{a_i\}$ ($i = 0, 1, \dots, n-1$; $n \geq 2$, $a_0 \equiv 0$) を $x \neq a_i$ ($i = 1, \dots, n-1$) なる条件を満たす n 個の実数とする。この時

$$\frac{1}{x+t} = \frac{1}{x} + \sum_{i=1}^{n-1} (-1)^i \frac{(t-a_0)(t-a_1)\cdots(t-a_{i-1})}{x(x+a_1)(x+a_2)\cdots(x+a_i)} \\ + (-1)^n \frac{t(t-a_1)(t-a_2)\cdots(t-a_{n-1})}{x(x+a_1)(x+a_2)\cdots(x+a_{n-1})} \cdot \frac{1}{x+t}$$

と表せる。一般にこの表現の中にフラクタル性とカオス性が同居しているが、上記の積分範囲において $x > 0$, $t > 0$ の時カオスは生じない。特に $a_i = i$ ($i = 0, 1, \dots, n-1$), かつ $n > t+1$ である時, 上記の関係式は

$$\frac{1}{x+t} = \frac{1}{x} + \sum_{i=1}^{n-1} (-1)^i \frac{t(t-1)(t-2)\cdots(t-i-1)}{x(x+1)(x+2)\cdots(x+i)} + \frac{\eta_n \Gamma(n)}{x(x+1)(x+2)\cdots(x+n-1)} \cdot \frac{t}{x+t}$$

となる。ここに η_n は $|\eta_n| < 1$ を満たす量である。従って、後者の条件の下で、 $(x+t)^{-1}$ の絶対収束する無限級数表示が可能となる。これらの関係に現れている $(x+t)^{-1}$ の自己相似性を利用し、上記 Stieltjes 変換で $g(t) = t^{-p}e^{-t}$ ($0 < p < 1$) として、平均 p のガンマ分布や関連する正規分布、逆ガウス分布などの確率分布関数や hazard rate の近似が可能であることを確かめた。なお、これらの近似において、上記の条件の内 $x > 0$, $t > 0$ であるが $n \leq t+1$ である領域を考える必要もある。この際の確率評価に係る積分量が無限小になること、すなわち、この領域が近似主領域であることを示すことが一つのポイントになる。

単調正則変動関数の分解とその応用

志村 隆 彰

0 でない指数 $-\alpha$ の正則変動関数 F が可測関数 f により、次のような形で表わされているとする。

$$F(x) = \sum_{k=0}^{n-1} b_k f(a^k x),$$

ここで、 $0 < a < 1$, b_k ($k = 0, 1, \dots, n-1$) は実数である。条件として、さらに $\sup_{x \geq 0} |f(x)/F(x)| < \infty$ と $\sum_{k=0}^{n-1} a^{-k\alpha} b_k > 0$ を加える。問題は正則変動関数 F がこのように表わされているとき、 f が正則変動性をもつかということである。この問に対して、次のような判定条件を得た。方程式

$$\sum_{k=0}^{n-1} a^{-k\alpha} b_k z^k = 0$$

が絶対値 1 の解をもたなければ、上記の式で f は正則変動する。しかし、解に絶対値が 1 のものがあれば、指数 $-\alpha$ の任意の正則変動関数 \tilde{F} に対して、次の式を満たす正値単調関数で正則変動する f_1 と正則変動しない f_2 が存在する。

$$\lim_{x \rightarrow \infty} \sum_{k=0}^{n-1} b_k f_i(a^k x) / \tilde{F}(x) = 1 \quad (i = 1, 2).$$

これは上記の表現において、 F の正則変動性から f の正則変動性は必ずしも導かれないことを意味する。この結果は、独立な正値確率変数 X, Y の積 XY の分布の裾が指数負の正則変動すると仮定した場合に、 X, Y いずれかの分布の裾が正則変動性をもつかという分布の分解問題へ応用できる。前述の結果と対応して、 Y の台が有限等比数列に含まれることを仮定する。こ

のような状況のもとで、確率変数 X が存在して、独立積 XY の分布の裾が正則変動するために Y の分布が満たすべき必要十分条件が得られた。

参 考 文 献

志村隆彰 (1998). Decomposition of distributions with regularly varying tails in the sense of Mellin-Stieltjes convolution (III), 無限分解可能分布に関する諸問題 (3), 統計数理研究所共同研究リポート, No. 112, 69-85.

主成分ベクトルのロバストな推定量のあるクラスについて

紙 屋 英 彦

$N_p(0, \Sigma)$ からの独立な標本 x_1, \dots, x_n が与えられたとする。このとき、第1主成分ベクトル、すなわち Σ の最大根に対応する固有ベクトル γ_1 を推定する問題を考える。

ここでは γ_1 の次のような推定量を提案する： $\gamma \in \mathcal{S}^{p-1}$ (\mathcal{R}^p の単位球面), $x \in \mathcal{R}^p$ に対し、 $z(\gamma, x) = (1/2) \{\|x\|^2 - (\gamma^T x)^2\}$ とおく。このとき $\rho(0) = 0$, $\psi(0) = 1$, $\psi := \partial\rho/\partial z$, を満たす非減少凹関数 ρ に対し、 \mathcal{R}^p 上の分布 G の汎関数 $T_\rho(G)$ を $T_\rho(G) = \arg \min_{\gamma \in \mathcal{S}^{p-1}} L_G(\gamma)$, ただし $L_G(\gamma) = E_G[\rho\{z(\gamma, x - \mu_G)\}]$, $\mu_G = E_G(x)$, で定義する。そして x_1, \dots, x_n の経験分布 \hat{F}_n に対し、 γ_1 の推定量 $\hat{\gamma}_*$ を $\hat{\gamma}_* = T_\rho(\hat{F}_n)$ によって定義する。

ρ の例として、

- (i) $\rho_0(z; \beta, \eta) = -\frac{1 + \exp(-\beta\eta)}{\beta} \log \frac{1 + \exp\{-\beta(z - \eta)\}}{1 + \exp(\beta\eta)}$, $\beta > 0, \eta > 0$;
- (ii) $\rho_1(z) = \lim_{\beta \rightarrow 0} \rho_0(z; \beta, \eta) = \lim_{\eta \rightarrow \infty} \rho_0(z; \beta, \eta) = z$;
- (iii) $\rho_2(z; \eta) = \lim_{\beta \rightarrow \infty} \rho_0(z; \beta, \eta) = \min\{z, \eta\}$, $\eta > 0$,

等が考えられる。 $\hat{\gamma}_*$ は、 $\rho(z) = \rho_0(z; \beta, \eta)$ の場合は、ロバストな主成分ベクトルとして Xu and Yuille (1995) が提案した自己組織化法則に対応し、他方 $\rho(z) = \rho_1(z)$ の場合は、古典的な推定量に一致する。

以上の下で、次のことが示される。まず T_ρ はフィッシャー—一致性を満たす： $T_\rho\{N_p(0, \Sigma)\} = \gamma_1$ 。また影響関数 $IF\{x; T_\rho, N_p(0, \Sigma)\} := (d/de) T_\rho\{(1 - \epsilon)N_p(0, \Sigma) + \epsilon\delta_x\}|_{\epsilon=0+}$, ただし δ_x は $x \in \mathcal{R}^p$ における point mass 1, は次で与えられる：

$$IF\{x; T_\rho, N_p(0, \Sigma)\} = \psi \left\{ \frac{1}{2} \|a_{(2)}(x)\|^2 \right\} a_1(x) \sum_{j=2}^p \frac{\lambda_j a_j(x)}{\lambda_j^* (\lambda_1 - \lambda_j)} \gamma_j,$$

ただし $a_1(x) = \gamma_1^T x$, $a_{(2)}(x) = \{a_2(x), \dots, a_p(x)\}^T = \Gamma^T x$; $\Sigma = \Gamma \Lambda \Gamma^T$, $\Gamma = (\gamma_1, \gamma_2) = (\gamma_1, \gamma_2, \dots, \gamma_p) \in \mathcal{O}(p)$ (p 次の直交群), $\Lambda = \text{diag}(\lambda_1, \dots, \lambda_p)$, $\lambda_1 > \dots > \lambda_p > 0$; $\lambda_j^* = E[\psi\{(1/2)\|a_{(2)}(x)\|^2\} a_j(x)^2]$, $j = 2, \dots, p$.

影響関数のこの明示的な表現により、 $\hat{\gamma}_*$ のロバストネスと効率を調べることが出来る。

まずロバストネスの観点からは、

$$\sup_{z \geq 0} \left\{ z^{1/2} \frac{\partial \rho(z)}{\partial z} \right\} < \infty$$

を満たす ρ を用いることが望ましい。この条件は、 T_ρ のいわゆる gross-error sensitivity が有限となるための条件となっている。 ρ_0 と ρ_2 はこの条件を満たすのに対し、古典的なケースに対応する ρ_1 はこの条件を満たさない。

一方、効率に関しては、上で述べたフィッシャー一致性と影響関数の表現により、 $n^{1/2}(\hat{\gamma}_* - \gamma_1) \rightarrow N_p[0, V\{T_\rho, N_\rho(0, \Sigma)\}]$ in distribution, ただし $V\{T_\rho, N_\rho(0, \Sigma)\} = \lambda_1 \sum_{j=2}^p [\lambda_j^{*2} \lambda_j^2 / \{\lambda_j^{*2} (\lambda_1 - \lambda_j)^2\}] \gamma_j \gamma_j^T$, $\lambda_j^{*2} = E[\{\psi((1/2)\|a_{(2)}(x)\|^2) a_j(x)\}^2]$, $j = 2, \dots, p$, となる。ここで $\hat{\gamma}_* = T_\rho(\hat{F}_n)$ の漸近相対効率を $\text{Eff}(\hat{\gamma}_*) = \det V\{T_{\rho_1}, N_\rho(0, \Sigma)\} / \det V\{T_\rho, N_\rho(0, \Sigma)\}$, ただし $\det V$ は V のゼロでない固有値の積, により定義すると, $\text{Eff}(\hat{\gamma}_*) = \prod_{j=2}^p \lambda_j^{*2} / (\lambda_j \lambda_j^{*2})$ となる。

この報告は Kamiya and Eguchi (1998) に基づいている。

参 考 文 献

- Kamiya, H. and Eguchi, S. (1998). A class of robust principal component vectors, Research Memo., No. 699, The Institute of Statistical Mathematics, Tokyo.
- Xu, L. and Yuille, A. L. (1995). Robust principal component analysis by self-organizing rules based on statistical physics approach, *IEEE Transactions on Neural Networks*, 6, 131-143.

統計計算開発センター

余震活動から周辺地域の大地震を予測できるか？

尾 形 良 彦

一年以上に亘る多数の余震活動データ（発生時刻とマグニチュード）を以下のように解析した。まず、全体区間を通して単一の ETAS モデルを当てはめる。さらに全体区間を任意の時点で 2 つに分けて前半と後半に 2 つの別々の ETAS モデルをあてはめ、AIC の和が最小になるような時点を求める。この時点が余震活動の変化点として有意であるか否かについて次のようにして調べる。最初の全体区間を通しての単一の ETAS モデルの AIC の値に対して後者の最小の AIC の和に対して変化点に対応するペナルティ（総余震数に依存する）で補正したものを比べ、小さい値を取ったほうが適合度が優れており、それが後者の場合は変化点がありであると見なす。

このようにして 3/4 世紀に亘る日本周辺に於ける約 70 例の余震活動を解析し、各余震活動を次のように分類した。すなわち余震活動が順調に変化点無く単一の ETAS モデルで表現できた場合、及び、変化点があり、その後、相対的静穏化がみられた場合である（変化点があり且つ相対的に余震活動が活発化した例はかなり少なかった）。それぞれの場合の本震を基準に時空間的に他の大地震の発生状況を追跡調査をしたところ、余震域から数度以内（1 度 = 約 111 km）の距離の周辺地域に於いて 10 年の時間内で、その本震とほぼ同規模以上の大きな地震が発生する確率に大きな違いが見られた。かくして、余震活動の ETAS モデルに対する相対的静穏化の有無が、その後の周辺部での大地震の発生の確率予測に意味を持つことが示されたものと考えられる。

非線形時系列解析

田村 義保

非線形性をとらえるための一つの手法として非線形確率微分方程式がある。実際のデータは離散時間でしか得ることができないため、データ解析と結び付けるためには離散化が必要である。1980年代半ば頃から、尾崎統教授は局所線形時系列モデルの研究を行っている。尾崎教授の局所線形時系列モデルの研究は非線形現象を示す対象の測定データである非線形時系列から精度高く、対象の本質を推定するためのものである。総合研究大学院大学の院生であった庄司功博士（現筑波大学助教授）による研究により、より高精度の計算が可能になっている。田村は庄司による局所線形化に注目した。どちらの研究も、ノイズ項のマルチプリカティブな形を避けるために、確率変数の非線形変換を用いている。加法的ノイズにはなるが非線形性は消えず離散化を行うためには局所線形化近似が必要になる。問題によっては別の変換を用いるとノイズ項はマルチプリカティブな形のままであるがドリフト項が線形の確率微分方程式を得ることができる。従って、局所線形化を行うことなく離散化が可能である。このような場合はドリフト項が線形になるような変換を用いるのが良いことを主張したい。物理乱数および擬似乱数を用いたシミュレーションではどちらの変換を用いても結果はほぼ同じであった。どちらの変換を用いるべきかの優劣をつけるための研究を継続していく。

報告会では、平成10年度中に行った「覚せい剤乱用者総数推定のための調査研究」、「人口動態統計死亡個票の開示方法についての研究」等の研究にも言及した。

密行列計算に適した汎用計算機について

泰地 真弘人

密な行列であらわされる連立一次方程式の解や、対角化、行列式の値などの計算には、行列の大きさ N の3乗の計算時間を要するため、非常に計算量が多い。しかしこうした問題に対しては、多数の演算器に同一のデータを供給し同一の演算をおこなう変形SIMD方式の演算装置で高速化できる。各演算器は共通のデータ入力ポートと演算器に固有のローカルメモリを持ち、これらのデータを用いて計算を行なう。こうした演算装置は入出力の問題が少ないため、通常のマイクロプロセッサに比べてはるかに高い集積度・性能を達成できる。結果として通常の並列計算機に比べて10倍以上コストパフォーマンスを上げることができるであろう。また、汎用性を高めることによって密行列計算以外の計算、例えば分子動力学計算なども本計算機で高速に実行することができる。このような計算機の基本設計を行なっている。

具体的なCPUのモデルを考え、行列乗算、LU分解などの例を実行させた場合の効率を計算した。LU分解(Crout法, Partial Pivoting)の場合、 $N = 10^4$ 、並列度 $P = 250$ で効率が約50%となる。効率が50%となる N の値はおおよそ並列度 P の平方根に比例する。今1並列度あたりの性能を0.5 Gflops程度とすると、テラフロップスの性能を得るには $P = 2000$ となり、 $N = 3 \times 10^4$ 以上の大きさの行列に対して50%以上の効率が得られる。このとき 3×10^4 元の連立一次方程式を解くには約300秒かかる。こうした結果から、密行列計算用の汎用計算機は十分な実効性能が得られることがわかった。こうした汎用計算機は今後の科学シミュレーションにとって重要な意味を持つであろう。

TIMSAC72 for workstation について

荒 畑 恵美子

TIMSAC72 のプログラムを Workstation で動くようにした。Dimension の大きさをパラメータで指定するようにした。データと結果の入出力関係をわかり易くした。使用例もつけるようにした。Fortran f77 での図をなるべくつけるようにした。

参 考 文 献

- 赤池弘次, 中川東一郎 (1972). 『ダイナミックシステムの統計的解析と制御』, サイエンス社, 東京.
 Akaike, H. and Nakagawa, T. (1988). *Statistical Analysis and Control of Dynamic Systems, Mathematics and Its Applications (Japanese Series)*, KTK Scientific Publishers (Kluwer), Tokyo.

視覚的操作による時系列解析支援システム

中 野 純 司

最近の統計解析パッケージには GUI (グラフィカルユーザインタフェース) を備えているものが多い。ところが多くの場合、それらは旧来のコマンド言語によるインタフェースの拡張として実現されている。この方法は、システムの信頼性を損なわずに比較的容易に実現できるという利点を持つとはいえ、GUI そのものの能力を十分に引き出すことは難しいといえる。そこでわれわれは設計の段階から GUI を中心に考えた時系列解析支援システムを作成してみた。本システムでは解析の過程を視覚化するために、データ、統計量、モデルなどをアイコンで表現し、操作はそのアイコンで指定されるポップアップメニューにより行う。このようなプログラムはオブジェクト指向の考え方によりシステムを設計することにより実現された。さらに、データの変換などもできるだけマウスを利用して行えるようにし、操作ミスを軽減できるようにした。プログラミングは、統計計算部には MS Windows 版 TIMSAC72 の C++ コードを利用し、それ以外の部分は Tcl/Tk 言語を用いた。本システムはプロトタイプではあるが、ほとんどの UNIX システムと MS Windows で稼働する。

半無限計画法

伊 藤 聡

半無限計画問題とは、有限次元空間における無限個の制約条件つき最適化であり、例えば次のように表わされる。

$$\min_x f(x) \quad \text{subject to} \quad g(x, y) \leq 0 \quad \forall y \in Y$$

ここで、添字集合 Y は陽に与えられることもあれば $Y = \{y | \gamma(y) \leq 0, \eta(y) = 0\}$ などとして陰に与えられることもある。このような問題は、近似問題一般、分権システムにおける資源配分、競争状況下での意志決定、多目的最適化、信号処理、制御系設計など工学のあらゆる分野に現

われる。また、無限制約式は適当な条件のもとでは最適値関数を含む式に等価的に変換され、この意味で、半無限計画法は階層的最適化に対する基礎を与えるともいえる。

半無限計画問題の汎用的な解法としては、グリッドや切除平面などによる離散化に基づく解法や局所縮約に基づく解法、あるいはこれらを組み合わせたハイブリッド解法、さらに微分不可能最適化問題に対するバンドル法の適用などがある。研究報告会当日は、数値解法の一例として双対パラメトライゼーション法の概略について説明した。これは基本的に局所縮約に基づいた解法であり、微分可能性を最大限に活用できること、許容性のチェックを極力避けること、添字集合 Y が陰に与えられた時にもそのまま対応できること、そして通常非線形計画法のパッケージが利用できることなどの方針に従って設計されたものである。詳細は Ito et al. (1998) を参照されたい。この他、研究報告会では min-max 問題、一般化半無限計画問題、最適制御問題、大域的最適化問題への応用について簡単に触れた。

参 考 文 献

- 伊藤 聡 (1998). 半無限計画の最適性と双対性, 統計数理, **46**, 345-358.
- Ito, S. and Teo, K. L. (1998). Dual sequential quadratic programming for nonlinear semi-infinite optimization, The International Conference on Nonlinear Programming and Variational Inequalities, Hong Kong, PRC.
- Ito, S., Liu, Y. and Teo, K. L. (1998). A dual parametrization method for convex semi-infinite programming, *Optimization: Techniques and Applications* (eds. K. L. Teo, P. F. Siew, Y. H. Leung, L. S. Jennings and V. Rehbock), Vol. 1, 550-557, Curtin University of Technology, Perth.
- Liu, Y., Teo, K. L. and Ito, S. (1999a). A dual parametrization approach to linear-quadratic semi-infinite programming problems, *Optim. Methods Softw.*, **10**, 471-495.
- Liu, Y., Ito, S. and Teo, K. L. (1999b). A semi-infinite programming approach to continuously constrained LQ optimal control problems, *Proceedings of the 38th Annual Conference*, 1149-1154, The Society of Instrumentation and Control Engineers.