

公開講演会要旨

複雑な系の非線形時系列モデリング： 予測誤差ゼロへの嗜好と志向

統計数理研究所 尾 崎 統

(1998 年 11 月 4 日, 統計数理研究所 講堂)

1. 複雑系と予測問題

最近科学ジャーナリズムに複雑系の科学は「21 世紀を主導する科学」などと盛んに取り上げられるようになってきた。複雑系の科学の特徴の説明によく使われる表現は「従来の研究分野を横断し、研究対象を自然現象、社会現象一般に広げるのが複雑系の科学の特徴であり...」とか「反還元主義 (部分の挙動を観測しても全体の動向がとらえられない)...」うんぬんである。これらの考え方の原形はサイバネティックスを提唱したウィーナーやコンピューターを導入して気象予測に利用しようとしたフォンノイマンやボルツマンのエントロピーを帰納的に使ってシステム同定と予測と制御に役立てた赤池弘次博士らの仕事にも見られる。しかしここで注目すべきはこれら先達にすべて共通しているのは予測の追及という視点であるのに対し複雑系の科学ではシミュレーションは強調されても予測は 2 義的な地位に落ちている点である。予測が科学にとって 2 義的であって本質的ではないという一部の複雑系科学者の言は天体の軌道の予測からニュートン力学の誕生までの過去の科学の発展の歴史を考えると多くの人にとって納得のいくものではない。

複雑系の科学において何故予測の問題を避けなければならないのか？ 我々にとってはそれが非常に難しいということ以外に回避しなければならない積極的理由は思い浮かばない。

2. 統計数理研究所における複雑な系の研究の歴史

複雑な系の予測研究は統計数理研究所において河川流量予測や工業プラントの予測制御に関連して 1950 年代から菅原正己博士、赤池弘次博士らによって始められた。とくに複雑な系の時系列研究は予測誤差を小さくするモデルを探すという素朴なアイデアと深く結び付いている。H. Wold, A. N. Kolmogorov らの時系列予測理論までさかのぼらなくてもカルマンフィルターやその前身のウィーナーフィルターなども予測誤差最小化の考えと密接に結び付いている。赤池博士の AIC の前身である FPE (Final Prediction Error) もその名前の示すとおりまさに予測誤差を最小にするモデルを探すという素朴なアイデアに基づいている。

1960 年代の後半から赤池博士は多次元 AR モデルを使って観測変数を予測誤差系列に変換すること (すなわち白色化) を組織的に活用することによって複雑なシステムのフィードバック解析法を確立した。この方法は以後セメントキルンや火力発電ボイラー制御、原子炉解析、マクロ経済システム解析、脳波解析、生体フィードバック解析など多くの複雑な系の解析に利用されてきた。

このように複雑系研究において他に先んじて多くの画期的仕事を発展させ一時期を築いた統計数理研究所の時系列研究もその後訪れたさらなる予測誤差減少化への発展の機会をうまく捉えたかどうかには疑問が残る。

線形モデルとモデル次数選択規準を主体とした統計アプローチの成功は自然人々の関心をそれら線形モデルに立脚した技術の改良へと向かわせ、非線形モデルよりも線形モデル中心のベイズ型分解モデル、ベイズ型平滑化モデルの可能性の追及、予測誤差を減少させる新しいモデルの研究よりもモデルの良さを評価するモデル選択規準の改良への過度の期待となって現われても不思議ではない。

1980年代に入って世の中が非線形モデル、連続時間力学系モデルによる予測誤差減少化への志向を強める中で統計数理研究所では正攻法の非線形モデル、連続時間力学系モデルの研究への無関心がしばらく続いた。しかしこのような状況下においても予測誤差のさらなる減少化に向けて著者が非線形力学系の研究を続けることが出来たのは研究所の伝統、先輩達によって築かれた寛容な研究環境のおかげである。ExpAR モデルと呼ばれ今では非線形時系列モデルの古典として海外の時系列の教科書にも載るようになったモデルはこのようにして1970年代後半の統計数理研究所において生まれ発展したものである。面白いことに1994年刊行の「統計数理研究所50年の歩み」には研究所における主要研究テーマ発展小史の中の時系列研究成果の中にExpARモデルの名前さえ載っていない。科学研究の評価の歴史にはジグザグ歩行はつきものである。

1970年代統計数理研究所における時系列解析研究はTIMSACパッケージの成功とともに華々しく多くの科学技術研究分野で活躍の舞台を上げた。四半世紀後1990年代に入ってTIMSACは停滞期に入る。世の中の制御工学関係者の間でも30年前に出された線形多次元ARモデルによる予測制御の時代は終わったと言う人が出てくる。

1970年代の統計数理研究所におけるTIMSAC成功の原因が

- 1) 線形多次元ARモデルによる白色化の手法
- 2) 情報量規準によるモデル次数の自動化

であったとすると1990年代のTIMSAC停滞の理由も同じものに起因しているように見える。成功の原因であった同じものが逆に停滞の原因として作用しているのである。

1) 線形モデルへの過信はベイズモデル化や状態空間モデル化への過度の期待を引き起こし、例えば状態空間モデルが予測に関しては高だかARMAモデルの機能しか持ち得ないことに気が付かない。したがってそれは基礎となるARやARMAモデルそのものの非線形モデル化や連続時間非線形力学系モデルの取り込みへの無関心となって現われニューラルネットなどの新興モデルに予測モデルの主座を奪われることになる。

2) 情報量規準への過度の期待は稚拙な予測モデルに対しても安心して使える情報量規準の拡張へと人々の関心を惹き付け現存の稚拙な予測モデルそのものを改良する努力への無関心となって現われる。

TIMSACの過去30年の経験から以下のことがはっきりしている。

1) 応用でもっとも役立っている部分は多変量ARモデル(ノイズ寄与率解析と制御系設計)であること。(記述や、特徴抽出に便利なベイズモデルやノンパラメトリックモデルもいざ具体的に現象を予測し制御をする段階になるとパラメトリック予測モデルに道を譲らざるを得ないのである。)

2) 多変量ARモデルは力学系の特殊な一モデルであること。

この2点を考える時1970年代に統計数理研究所の時系列研究グループの中であって予測に関心を持つ30才になったばかりの筆者が非線形なAR型モデル、非線形な力学系モデルへ時系列予測モデルを拡張することに取り組み始めたのは当時大変勇気のいることであったが、いま考えるとごく自然の成り行きであったように思われる。

3. 力学系に託された仮説の利用

時系列のARモデルを非線形化した時にまず出会う困難がシミュレーションにおける発散の問題である。データに当てはめて得られた非線形モデルを白色ノイズを使ってシミュレートすると必ず発散するのである。この種の経験をしたものが最初に抱く考えは「予測は改良するのに何故シミュレーションは改良どころか線形よりかえって悪くなるのか?」、「本来あるべき非線形時系列モデルとはどのようなものか?」という素朴な疑問である。

時系列予測はダイナミックな現象のモデルであるが、同じダイナミックな現象のモデルとして微分方程式を用いた力学系モデルと確率微分方程式を用いた確率過程モデルが物理学、工学、応用数学などの分野で用いられている。ところがこれらの連続時間ダイナミックモデルにおいては非線形モデルの歴史は古く色々な非線形振動モデルがありその性質も調べられている。

上記の素朴な疑問から発してこれら非線形振動モデルを参考にしながら非線形時系列モデルの古典の一つとなったExpARモデルは1976年の秋統計数理研究所において生まれた。それは船のローリング時系列のARモデルの係数の部分を定数の代わりに振幅従属にしたらどうだろうかという素朴な疑問からはじまった。予測性能が改良することはすぐ確認出来たがそれではその振幅従属係数ARモデルが一体どんな意味を持っているかが疑問となる。多項式のほか、B-スプライン、直交関数、調和関数などいろいろな関数族を試しても得られる係数の関数の形はほとんど同じであることからこれは確実にデータの背後にあるダイナミクスの特徴を係数の変化で捉えていると確信した。問題はシミュレーションにおける発散をいかに止めるかであり、そのために始めはアドホックな思いつきでスタートしうまく行ったExpARモデルであるが、後にその思いつきが深い意味を持つことがわかった。このExpARモデル導出のさいは船舶工学者山内保文博士、織田博行博士らの協力そして統計数理研究所における時系列解析の指導者赤池弘次博士の助言が大きかったと思う。これがその後、船のローリング時系列予測だけでなく機械振動解析、火力発電プラントのボイラー制御、脳波、生体信号解析から金融時系列の解析にまで使われようとは当時は想像すらできなかった。

この奇妙に役に立つExpARモデルの奇妙な性質を調べるうちにこの非線形時系列モデルと非線形力学系モデルの間に単に類似したという程度を超えた緊密な関係が感じとられたため、私は確率微分方程式で表わされる確率力学系モデルとExpARモデルとの間の明示的関係を調べる研究を1979年頃から始めた。これには1979年に京都セミナーハウスで当時京都大学教授山口昌哉氏によって組織されたカオスの研究会での出会いが大きく影響を与えたように思う。この研究会にはダイナミクスに関連する物理、経済、数学、工学など色々な分野から新進気鋭の人達が参加していた。当時ブリゴジンのいるベルギーでの滞在を終えて帰国して間もない京都大学基礎物理学研究所の助手をしておられた相沢洋二現早稲田大学教授が参加されており氏からローレンツカオスについていろいろ教えられた。伊東敬祐神戸大学教授からは後で力武カオスに関する御自身の論文を送って頂いた。ご本人達はもう忘れておられるかも知れない。ロジスティックモデルや自分自身で発見したExpARモデルの人工的カオスには全く関心を持てなかった私も連続時間の非線形力学系モデル、ローレンツカオスと力武カオスには惹き付けられるものを感じこれをなんとかしなければと思い続けた。

その後は非線形力学系の局所線形化法、マルコフ拡散過程の局所線形化、ピアソンシステム

と拡散過程モデルの関係を利用した非ガウス時系列モデル, 連続時間非線形力学系の状態空間モデル化と局所線形フィルター, 局所線形フィルターを組み込んだ力学系モデルの最尤法と続き, 休む間もなく現在まで来た。最初は研究会や国際シンポジウムで話してもポカンとして質問さえしてもらえなかった。あるいは確率過程の推定問題そのものが理解できない数学者からよく伊藤確率積分の話と混同してトンチンカンな質問を受けた。1990年代に入って局所線形化法も, 少しずつ注目を集めるようになった。局所線形化研究に日本の外でいち早く注目を私と共同研究を始めたキューバ神経科学センターのヴァルデス博士, ビスカイ博士, ヒメネス博士らはその仕事で1997年のベルヌイ協会主催の国際会議に招待された。彼等はまた局所線形フィルターの脳神経モデルへの応用で1998年 Paul Rapp 教授らによって組織された「非線形ダイナミクスと脳機能」に関する国際会議にも招待された。総合研究大学院大学博士課程で局所線形化の拡張の研究論文を書いた庄司功現筑波大学助教授は1999年の夏オックスフォード大学で開催が計画されている計算数学に関する国際会議での講演に招待されている。最初の局所線形化の発表から今日国際的に認められる兆しが見えるまでに15年以上経っている。これが遅かったか早かったかは研究に対するスタンスの取り方で認識の分かれるところである。ここで庄司功博士の業績は特筆に値する。彼は局所線形化のようなこの世の中の流行からはずれた研究主題を自らの問題意識に基づき選び, それに新しいアイデアを入れ論文を大量に生産した。しかし論文は博士課程修了間際まで投稿する度にことごとく拒否されるという苦杯をなめる。その後博士論文の仕事は徐々に理解者を得て学術誌に受理され始め7つの論文が国際誌に載り, 博士課程における仕事で国際会議の招待講演に招かれる事態にまでなった。統計数理研究所にも最近には多くの総合研究大学院生が統計科学専攻院生として在籍するようになったが, 庄司功博士のような先輩の活躍を見て志を高く持って研究に向かっていたいただきたいものである。

さて本題に戻って, 最近ではデータマイニングとか自己組織化モデリングとか言って統計的モデリングがあたりかも時代遅れであるかのように言う人がいる。赤池博士は「時系列解析の心構え」と題する論文で“解析者が深い知識と経験に基づく仮説を提案することなしには良いモデルとその利用が実現される見込みはない”と言い切っている。一般的なグニャグニャモデルを使って計算機にやみくもに自己組織化モデリングで見つけてもらうことが出来れば解析者にとってはまことに結構な話しであるが科学の最前線でそんなにうまい話しはない。私が30才のころから取り組んできたことは結局力学系に託された仮説の有効利用の方法ではなかったかと最近になって思う。

4. 複雑系予測モデルの実際

1970年代に船のローリング時系列の非線形振動モデルから始まった統計数理研究所における著者らの非線形時系列モデリング研究は時を同じくして始まったカオス研究の Robert May らの「すべての確率的モデルは決定論的モデルで置き換えられる」の論調に激しく反発し1980年代から1990年代にかけてずっとカオス研究と一線を画すかたちで進んできた。しかしこの期間をとおして私の研究を深いところで最も駆り立てたのは前節でも述べたローレンツカオスと力武カオスであった。ノイズの全くない決定論的力学系モデルが現実の時系列データに当てはまるわけではない。しかしこれらを特殊な場合として含む一般多次元確率非線形力学系をスカラー時系列から同定する方法を導入すれば世の中の多くの複雑なダイナミックな現象の時系列解析, 予測に役立つはずである。こうして多次元確率微分方程式の局所線形化, 非線形確率力学系の局所線形フィルター, それに基づく非線形状態空間モデルの最尤法など自前の最新の武器が生まれた。これらの武器を手にした今, 我々の前には限りない可能性がひらけている。

ニュートンの微分方程式導入以来、物理学、化学、工学、生物学、経済学など科学のいたるところでダイナミクスを記述するモデルとして力学系モデルが利用されてきている。更に 20 世紀になって計測技術の進歩と高速コンピュータの出現によりこれらのダイナミックな現象の力学系モデルを観測データと照合させてさらに精密化し、より現実に近い予測精度の高い力学系モデル、あるいはより大規模により複雑に絡み合った力学系モデルへと発展させる道が開けたのである。仮説としての力学系の尤もらしさを観測データから評価する方法を与える統計学の発展によって、長い間批判の対象から除外され続けて来たアカデミズムの申し子としての力学系の時代は終りを近づけようとしている。複雑系のモデルとして次々と提案されてくるモデル群も例外ではない。ダイナミックな現象のモデルである限り予測の問題から逃げることは出来ず、観測データがある限り統計の評価から逃れられない。きびしい評価とチェックによって力学系モデルはさらに改良発展することが出来るのである。

統計数理研究所においては 1980 年代なかばからこの最新の手法を用いた諸科学における複雑な力学系モデルの応用の試験的試みが続けられてきた。以下にそれらの例を挙げると：

1) 金融ダイナミクス

アメリカの Long Term Capital Bank の金融危機で有名になったノーベル賞経済学者 Scholes と Merton らの失敗の原因は金融のダイナミクスのモデルとして理論的に確率微分方程式を導入したあと統計を使った現実データとの照合を怠ったことにある。庄司功氏の局所線形化の仕事はこれらの理論的に導入された Merton らの金利モデルをより現実のデータに合う確率力学系モデルに改良する道を開いたのである。現在この仕事を多次元力学系を用いたさらに洗練されたものに発展させる試みが続けられている。

2) マクロ経済ダイナミクス

ケインズ理論など多くの経済理論は経済の平衡状態を仮定した理論であり平衡状態からはずれた状況での動きを記述するダイナミックマクロ経済モデルの必要性が言われて久しい。1985 年以来統計数理研究所では Hicks の IS-LM 理論をもとにした総生産と利子率の非線形ダイナミックモデルとその構造時間変化を許した非定常非線形ダイナミック IS-LM モデルが導入されデータ解析に有効に働くことが確かめられている。TIMSAC の赤池弘次博士の線形 AR モデルが日本のマクロ経済解析に応用されるまでに 20 年かかっている。この非線形モデルが世界のマクロ経済データに応用されるまでにはさらに数十年を要するかも知れない。現在ニューヨーク大学の Ramsey 教授らはこの統計数理研究所の非線形マクロ経済モデルの仕事に非常に関心を示している。

3) タンクモデル

菅原正己博士のタンクモデルはスカラー時系列への高次元力学系モデルの当てはめを試行錯誤によって行うものである。これはモデル内部の観測されない変数のダイナミクスをデータから決める手続きを与えるものでもあり、ある意味で高次元の非線形ベイズ状態空間モデルを試行錯誤法でデータに当てはめるものとみなせる。統計数理研究所で開発された局所線形フィルターの利用によってこれら高次元力学系タンクモデルの同定が試行錯誤法でなく最尤推定法によって組織的に行うことが可能になった。この方法が Hydrologist (水文学者) の間で注目を集めていると数理水文学の権威、Waterloo 大学の K. Hipel 教授から聞いた。本格的に利用が始まる日も近いかも知れない。

4) 脳神経力学系モデル

脳神経のダイナミックなモデルがいろいろ議論されているがシミュレーション研究が主でありデータに基づく実証的モデル研究には手がつけられていない状況がしばらく続いていたが、この状況にブレークスルーがもたらされた。キューバ神経科学センターのグループが統計数理

研究所と共同で開発した局所線形フィルターを駆使して Zetterberg モデルと呼ばれるマクロ神経モデルの EEG アルファ波データへの当てはめなど次々と結果を出し始めている。また現在これに触発された海外の脳波の力学系モデル研究者達から我々のもとへ共同研究の申し入れが次々と来ている状況にある。

5) 気象方程式

気象予測は観測データをもとに地球上のメッシュ点上の気象変数の推定問題 (Data Assimilation と呼ばれる) とメッシュ点上の推定変数をもとに気象方程式を計算機シミュレーションして得られた変数の将来の値をもとにした気象予測の二つのジョブによって構成される。両ジョブともそこで本質的役割を果たしているのが気象変数のダイナミックスを記述する気象方程式を構成する大規模の力学系モデルである。過去数十年の数値予報研究の経験からこの理論力学系モデルの限界が見えてきたいまデータの情報を効果的に取り入れたより現実に近い力学系モデルへの改良の動きが興りつつあり、一部気象学者によって統計数理研究所の力学系推定手法を Data Assimilation と Weather Prediction に応用する研究が計画されている。

6) 火力発電プラントボイラー制御

多次元線形 AR モデルによる予測制御法の典型的成功例に挙げられる火力発電プラントのボイラー制御は十数年の稼働経験からその出力需要の大きさに従属するシステム特性の非線形性が一つの大きな問題としてあることが認識されている。発電プラントシステムの場合はその他の複雑系モデリングと違ってプラントの物理的力学系モデルそのものの構成よりはむしろ制御のために有効な機能的モデリングが求められている。統計数理研究所では制御工学の専門家達と協力して線形多次元 AR モデルを特別な場合として含む非線形多次元 ExpAR モデルを利用して非線形システムの制御を効率よく行う手法の研究が進められている。

このほかに力武カオスを例にとり局所線形手法の具体的力学系予測への応用の一端を詳しく示した。さて以上のような応用との接点での研究の進行状況は 21 世紀に向かって統計学研究所の未来像を考えると、歴史的に固定した諸科学の境界を取り払い諸科学のハブとして活躍する統計学研究所の未来の姿を写しているように筆者には見える。複雑系の研究に限っても上記のように統計学研究の果たす役割は限りなく大きい。学問、研究の自由な発展を推進していくには開かれた学問、開かれた研究所、開かれた研究グループ、開かれた研究者であることが大切である。研究グループの中で閉じ籠り隣室の研究者と研究上の論争もしないような閉じたナルシス研究所に未来はない。

5. 前進を止めない人間

人間を取り巻く社会、自然界において不可解で予測の困難なダイナミック複雑系には枚挙にいとまがない。予測への関心、未来を正確に知りたいという欲求は太古の昔から人間を惹き付けてきた。17 世紀のニュートンによる微分方程式と力学系の発見は確かに人間に未来の変化を予測するための強力な道具を提供した。18 世紀ラプラスはこれによって世の中の全ての動きを記述する大規模の力学系を用いて初期値さえ正確に与えれば未来を正確に予測できると主張した。この「ラプラスの魔」と呼ばれる決定論がその後 19 世紀から 20 世紀まで人間の歴史に大きな影響を与えたことはよく知られている。プランクやアインシュタインらの 20 世紀の量子力学の誕生に大きな影響を与えたボルツマンのエントロピーの確率的解釈の誕生は 19 世紀末のこの決定論者達の容赦ない批判攻撃の中でもがきしみながら生まれたことは朝永振一郎博士の著書「物理学とはなんだろうか、下」に詳しく書かれている。この他社会科学においても 19 世紀におけるマルクスの唯物論的歴史観と共産主義社会の予言は決定論的世界観の影響が色濃

く現われたものとしてよく引用される。オーストリー生まれの科学哲学者 Karl Popper は著書「歴史主義の貧困」、「開かれた社会の敵」で決定論的マルクス主義の徹底的批判をおこなっているが、その論調は時系列モデルを現実のデータに合わせ予測誤差を出来るだけ小さくするモデルを探すことを生業としている者にとって興味深く示唆するものが多い。その後に着わされた「開かれた宇宙」ではまさに力学系による予測が議論され決定論のアカウントビリティのなさが詳細な条件のもとで示されている。

複雑な系の時系列予測モデリングも開放システムの考えに通じている。閉鎖システムということは状態変数の未来を規定するものは全て過去の状態によって決まるのに対し、開放システムにおいては状態の未来を規定するシステムのダイナミックスの記述は完全ではありえず常に予測不可能なノイズと呼ばれる外からの影響にさらされていると考える。我々の方法は非線形力学系を駆動するシステムノイズの分散と観測ノイズの分散と状態変数の初期値を同時に推定することによって当の予測問題に伴う予測誤差の原因を初期値の不正確さ、システムダイナミックスの記述の不正確さと観測データの不正確さの3つの不正確さからの原因としてその起因の度合を量的に把握することを許す。この意味で我々の発展させてきた非線形力学系と観測式を組み合わせた非線形状態空間モデルに基づく推定手法はこの Karl Popper の開かれた宇宙、あるいは開かれたシステムの主張を現実の観測データを通して量的に確認するための方法を与えると見ることもできる。

私は1990年の Journal of Royal Statistical Society の M. S. Bartlett 教授の論文“Chance and Chaos”の Discussant としてカオス論の実証的裏付けのなさを指摘し“Chaos' determinism is a revival of the ghost of Laplace's daemon which resisted the surge of Maxwell's and Boltzmann's idea of statistical mechanics in the end of the last century”と決定論的予測モデルへの過度の期待をいませめた。人間の予測誤差ゼロへの嗜好は否定すべくもない。1980年代に気象学界を蔽った数値予報による長期予報の可能性への期待もまさにラプラスの魔への期待そのものである。カオスへの一時的な熱狂も冷め、20世紀も終わろうとしているいま我々が努力すべきは哲学的なあれかこれかの議論より具体的な予測誤差減少への不断の努力であろう。時系列のモデルで言えばランダムウォークモデルは予測への全くの無知を自ら認めたモデルであり決定論的カオスモデルはシステムダイナミックスの記述の完全さを主張する全知全能のラプラスの魔モデルである。非線形時系列解析がめざすものは無知から全知へ向けて予測誤差を少しでも小さくするために前進を止めない人間の知的活動の一翼を担うことである。

参 考 文 献

- 赤池弘次 (1997). 時系列解析の心構え, 『時系列解析の実際』(赤池弘次・北川源四郎 共編), 朝倉書店, 東京.
- Ozaki, T., Jimenez, J. C. and Ozaki, V. H. (1996). Role of the likelihood function in the estimation of chaos models, Manchester University/UMIST Tech. Report, 1996/11.
- Ozaki, T., Valdes-Sosa, P. and Ozaki, V. H. (1997). Reconstructing nonlinear dynamics from time series: With application to epilepsy data analysis, Manchester University/UMIST Tech. Report, 1997/02.
- Shouji, I. (1996). Estimation and inference for continuous time statistical models, Ph. D Thesis, Graduate University of Advanced Studies, Tokyo.
- 統計数理研究所 (1994). 『統計数理研究所 50 年のあゆみ』, 統計数理研究所創立 50 周年記念事業, 刊行物等実行委員会 編.