

小西貞則会員の平成 21 年度科学技術分野の文部科学大臣表彰科学技術賞 (研究部門) の受賞に寄せて

統計数理研究所 北川源四郎

2009 年度の文部科学大臣表彰において研究部門の科学技術賞が小西貞則会員(九州大学数理学研究院教授)に授与された。科学技術賞は、「我が国の科学技術分野において顕著な功績を挙げた人」に授与されるもので、特に研究部門では、「我が国の科学技術の発展等に寄与する可能性の高い独創的な研究または開発を行ったものを対象とする」とされており、小西氏に関しては「高度計算機技術環境のもとでの新たな統計科学の研究」が受賞理由とされている。

小西氏は広島大学理学部助手、統計数理研究所研究員、室長、助教授を経て 1994 年に九州大学大学院の教授となり現在に至っている。この間、日本統計学会理事長、応用統計学会会長、統計科学研究会委員長、日本学術会議連携会員を歴任するとともに、Annals of the Institute of Statistical Mathematics, Bulletin of Informatics and Cybernetics, Journal of the Japan Statistical Association の Editor のほか多くの学術雑誌の Associate Editor を務めている。また、今回の受賞にも関連した業績により、日本統計学会賞(2004 年)、応用統計学会優秀論文賞(2006 年)を受賞したほか 2006 年には American Statistical Association の Fellow になっている。ASA の Fellow に列せられた日本人は 1914 年に制定されて以来 10 名だけである。

近年、科学や技術における研究対象は著しく拡大し、それにもなって科学的研究の方法論も大きく進化しつつある。19 世紀までの学術研究は機械的世界観に基づく物理的世界が対象であったが、19 世紀中盤にはダーウィンの進化論の強烈なインパクトによって、現実世界のあらゆるものが変化し進化することを認識せざるを得なくなった。これに触発された実験や観測に基づく実験科学の成立は、生命科学ばかりでなく、工学、経済学、心理学などの多くの分野を科学的研究の土台に乗せることになり、多くの新しい学問分野が創成された。さらに 20 世紀後半になると社会のグローバル化によってあらゆるものがリンクし、多くの社会的重要な課題は、多数の要因が相互関連した非線形・複雑・超多変量システムとしてとらえることが必要となってきた。さらに 21 世紀の現在、情報通信センサー技術(ICT)の飛躍的進歩によって実世界の大規模・大量観測データが得られるだけでなく、大量の Web 情報が氾濫し巨大なサイバー世界が形成されつつある。現代の科学・技術は明確な物理的構造をもった対象だけでなく、このような茫漠たる対象を視野に入れることが必要になっている。

このような研究対象の変化は必然的に科学的方法論の変化の契機となる。対象の複雑化、システム化に対応して、すでに 20 世紀後半には従来の理論科学、実験科学に加えて第三の科学と呼ばれるシミュレーションを中心とする計算科学が勃興し、従来の解析的方法では接近が困難であった非線形系や複雑なシステムの理解や予測に成果を挙げてきた。さて今日、20 世紀終盤以降の ICT の急速な発展によって、生命科学におけるマイクロアレイデータ、ファイナンスにおける高頻度データ、マーケティングにおける POS データ、環境や防災関連のデータ、全天 CCD データなどのように、あらゆる科学・技術の分野や社会において大規模・大量のデータが獲得され、しかもそれがデータベースとして時々刻々蓄積されつつある。科学的方法論における現在の課題は、これらの大規模・大量データを活用して、複雑な現実の問題を解決するための第 4 の科学すなわちデータ中心科学の確立である。

現代の科学技術研究における著しい特徴は、多くの場合“真のモデル”を前提としえないことにある。真の構造を仮定し、いくつかのパラメータや初期値や境界条件を推定することによって対象を表現できるとする従来の立場ではもはや対応できなくなってしまうのである。明確なモデルがアприオリに存在しない場合、対象に関する理論や知識、様々な経験、大量のデータを統合し、さらには研究の目的をも考慮しながらモデルを作り上げていることが不可欠になっている。情報化時代の科学・技術研究における要諦は従来の数学が必ずしも重視してこなかったモデリングにある。

小西氏の今回の受賞は、このような中で、国際的に急速に研究が進展している複雑な非線形現象解明のための新しいモデリングの理論・方法論を提案し、それが諸科学および産業界が直面する問題の解決に役立つことが高く評価されたものである。

これまでに日本人によって提唱され、自然科学はもとより工学や社会科学などのあらゆる分野で応用され、現象解明と知識発見に多大な貢献をしてきた数理的方法として赤池情報量規準 AIC に基づく統計的モデリングがよく知られている(提唱者の赤池弘次氏は 2006 年に京都賞を受賞)。情報量規準 AIC はパラメトリックモデルの評価のための視点と具体的判断基準を与えた画期的なもので、これによってパラメトリックモデルに基づくモデリングの方法が確立することになった。実際この方法によって、火力発電用ボイラの制御、船舶の自動操舵システムの設計、経済時系列や生体システムの解析など科学の様々な分野の現象解明に大きな役割を果たしてきた。しかし、近年、ゲノムデータ、地球環境データ、現象過程・動作過程の実時間記録データ、形や立体情報データなど多様なデータの中から有益な情報やパターンを効率的に抽出するためのより柔軟なモデリング手法、特に、複雑な非線形構造を捉える非線形モデリング手法の研究が不可欠となってきた。

このような状況の中で、小西氏は情報量規準 AIC の基本的な考え方を継承しつつもその

適用範囲を大きく拡大した一般化情報量規準 **GIC** を導出した。情報量規準 **AIC** が最尤法で推定されたモデルの予測分布の評価を実現したものであるのに対して、一般化情報量規準 **GIC** は統計的汎関数で定義されたあらゆる推定量に対して適用できる。最尤推定量に限らず、**M** 推定量などのロバスト推定量や多くのベイズ推定量が統計的汎関数を用いて定義できる。したがって、**GIC** によって、ロバスト推定量で推定されたモデルやベイズモデルなど、**AIC** では直接評価できなかった広範なモデルを評価・比較することが可能となり、モデリングの自由度が著しく高まることになった。

計算統計学の潮流のなかで、情報量規準自体をブートストラップ法にもとづいて数値的に求めるブートストラップ情報量が提案されている。この **EIC** の適用範囲は極めて広範で、ほぼ無制限といってもよい。しかし、その一方、数値的に計算されるために、その偏りや精度などに注意が必要である。これに関しても統計的汎関数にもとづく **GIC** に関する理論的結果は **EIC** の性質の分析に威力を発揮し、**EIC** のバイアス補正や分散減少、高次のバイアス補正等の方法を導き、**EIC** の実用化にも大きな貢献をした。

さらに、小西氏は事前情報をモデル化に取り込んだベイズ理論によるアプローチによって、ベイズ型モデル評価基準 **GBIC** を導出し、それを非線形モデリングのプロセスに導入した。これらの非線形モデリングの方法は、ゲノムデータ解析、タンパク質立体構造データ、ファイナンスデータ解析、画像情報分析、人間の歩行における動作過程の実時間記録解析などにおいて威力を発揮し、共同研究者とともに大きな成果を挙げつつある。これらの研究は、平成 15 年度（2003 年度）に「数学・物理学・地球科学」分野で採択された九州大学の 21 世紀 COE プログラム「機能数理学の構築と展開」の研究教育拠点形成のための機軸の一つに位置づけられている。また、平成 20 年度グローバル COE「マス・フォア・インダストリ教育研究拠点」の重点研究課題の一つにも位置づけられている。

「忘れられた科学—数学」が喧伝される昨今、新しい科学的方法論の確立に直結し、しかも現実の問題解決にもつながる理論的貢献が数理科学の分野からなされたことは大変喜ばしい。これを契機に、学術や科学技術分野の動向を見極めつつ本質的な貢献をする研究が数学や数理科学の分野においてなされることを期待したい。