

# モデリング変革の 4 半世紀を振り返って

樋口 知之<sup>†</sup>

(受付 2019 年 5 月 27 日；改訂 7 月 5 日；採択 7 月 16 日)

## 要 旨

私が研究所に入ったのは平成元年 4 月、また所長を退任すると同時に研究所を退職するのは、平成が終わる直前の平成 31 年 3 月であったため、平成の 30 年間は私にとって研究者人生そのものである。本稿では、データにもとづく予測や判別といった統計的思考の根幹を成す、統計的モデリングにかかわる技術の大きな進展を、実体験した私の視点で概括する。

キーワード：ベイズモデリング，粒子フィルタ，カーネル法，深層学習。

## 1. はじめに：入り口は線形・ガウスモデリング

私が入所した当時(1989 年)は、その 10 年ほど前から研究が活発化していたベイズモデリング研究の成熟期にあった。線形・ガウス型の事前分布および尤度関数に関する理論は確立し、応用の観点からも、いくつかの課題では研究所の研究成果が標準的方法として認知されるなど、研究進展も著しかった。正確に言えば、統計数理研究所では、滑らかさを表現する事前分布と周辺尤度最大化法を採用した、経験ベイズ法が活発に研究されていた(赤池・北川, 1994, 1995; 石黒 他, 2004; 伊庭 編, 2018)。私は学部生の頃から雑誌「数理科学」の愛読者で、ある時、ベイズモデリングの特集号(赤池, 1981)があり、その柔軟な表現能力に大変な感銘を受けた。その解説原稿のベースとなる国際会議録(Akaike, 1980)のコピーが、地震学を専門とされる松浦充宏東大教授の研究室にあることを知り、研究室に行ってコピーさせてもらい(劣悪な多数回のコピーの連続で、式の細かい部分がかかなり判読しづらくなっていた)すぐさま論文内容の検証をした記憶がある。実はその論文にはタイポ的ミスが複数あり、「世界的権威の赤池先生の論文に間違いを見つけた」と学生の私はかなり昂揚し、当時所長をされておられた赤池先生に不躰にもコンタクトを取ってしまった。もちろん、タイポ的ミスなので結果に間違いは無かったのだが、今と違い取るに足らない些細なタイポミスを知り得ようもなく、研究を効率的にすすめることが難しい、逆に言えば、大変長閑な時代であった。経験ベイズ法を採用した赤池先生の 1980 年の論文は、相当な議論をベイズ統計学の分野で巻き起こし、その当時、ベイズ統計学の第一人者であった英国の Smith 教授らベイズ統計学者とは激しいやりとりがあったと赤池先生から聞いている(赤池 他, 2007)。なお、Smith 教授は現在(2019 年)、2015 年に英国政府の肝入で設立された人工知能とデータサイエンスの国立研究所、アラン・チューリング研究所の所長を務めている。後段にも触れるが、英国の人工知能研究においては統計学の存在感が重厚である。

現象や既存知識のモデリングにおいては、計算上からくる線形・ガウス性の制約は、研究者の自然な思考を妨げる。そのため、その制約を、“人手”でもって解析的に緩める論文が 1980

---

<sup>†</sup> 中央大学 理工学部：〒112-8551 東京都文京区春日 1-13-27

年代は数多く出版されていた。事実、研究所の先生方は、現象の振る舞いを決して過度に抽象化することなく、計算が容易となる空間への非線形写像を“巧みに見だし”，データに潜む新しい知識を次々とあぶり出していた。「数理〇〇学」という当該専門分野の研究者の持つ、先入観やある種の価値観に大きく依存した数理表現でなく，“データに『因子とメカニズム(機序)』を語らせる”ことにこだわっていた研究所の先生方は、間違いなくデータ駆動科学の先導者である。先生方の玄人芸的非線形モデリングのアプローチと平行して、計算機集約的な方法の開発も活発化していた。高い次元の潜在変数がつくる超高次元の分布から、マルコフ連鎖でもって直接的に実現値を得るマルコフ連鎖モンテカルロ法(伊庭 他, 2005)や、その分布を低次元の分布に分解し(いわゆる逐次ベイズフィルタ)、低次元の分布の逐次更新を数値的に得る「非線形フィルタ」(北川, 2005)などである。私は学位を宇宙プラズマ物理の領域で得たため、統計力学の自由エネルギーや転送積分とほぼ同一の発想にもとづく両手法を自然に理解でき、大いに魅惑された(伊庭, 2003)。

## 2. 粒子フィルタ：非線形・非ガウスモデルの計算実現

平成にはいってしばらくして北川先生(第十代所長)は、非線形フィルタの適用範囲を大幅に拡張するモンテカルロフィルタを提案された(Kitagawa, 1993, 1996)。この手法は今日において、粒子フィルタの手法群の中で、原始的ではあるがレガシーとして高く評価されている。別分野においても、全く同じ手法がほぼ同時期に英国の Blake らによって提案され、大きな注目を浴びていた(Isard and Blake, 1998)。このアルゴリズムは、Condensation(CONditional DENsity propagATIOn)と呼ばれ、物体追尾分野(特に軍事分野)で盛んに研究開発された。Blake はマイクロソフトの簡易 3D カメラである Kinect の開発にも参加し、後にアラン・チューリング研究所の初代所長に就任している。

当時、北川先生の近くで研究していた自分は、モンテカルロフィルタの大規模並列コンピュータへの実装容易性に衝撃を受け、さまざまな非線形・非ガウス型の時系列モデルへの適用を試みた。また、世界同時的に多くの研究者が、アルゴリズムの高度化と現実問題への適用をすすめる、21 世紀早々にシュプリングから書籍が出版された(Doucet et al., 2001)。私も著者の一人として参加するチャンスを得、当時新進気鋭の海外の若い研究者と交流できたことは、その後の自分の研究者人生にとって大きな糧となった。この友人の多くは現在、後述する深層学習に研究の軸足を移しているが、そのことは偶然でなく、むしろモデリング技術の発展を考えれば極めて自然なことと考える。

粒子フィルタは、時系列モデルが所与であれば、パラメータや状態変数の推定に関して原理的には万能である。もちろん、モンテカルロ誤差(有限サンプルによる表現限界)からくる、分布の表現能力の喪失(いわゆる退化問題)や、尤度値の不安定性など、数値的課題は避けられないが、問題や経験に即して“人”が対象を自由にモデリングできるようになった点は、モデリング技術の発展において大きな飛躍であった。特に、ロボティクス分野への粒子フィルタのインパクトは大きなものがある。前述したシュプリングの本にも寄稿している Thrun らは、ビッグデータから簡便にロボットを制御する方法論を確立した(Thrun et al., 2007)。彼らは、統計学とロボティクスの“真の結婚”を実現した立役者である。さらに彼らはその技術をもって米国国防高等研究計画局(通称 DARPA)のプロジェクト「グランドチャレンジ(2007 年はアーバンチャレンジ)」にスタンフォード大チームとして参加し、自律型無人自動運転車の技術開発に多大なる貢献をした。後に Thrun はグーグル(Google)に入り、Google Street View を開発している。その後、グーグル・グラスを開発した Google-X の社長を務め、現在、グーグルの創設者である Page と一緒に、空飛ぶ自動車 Kitty Hawk の実現に取り組んでいる。

### 3. カーネル法：非線形モデリング

2000年代にはいって、非線形モデリングの観点からは、別の形で大きな進歩があった（樋口, 2014）。それはカーネル法の登場である（Schölkopf and Smola, 2002; Shawe-Taylor and Cristianini, 2004; 赤穂, 2008; 福水, 2010）。1998年にGoogleが、また2004年にはFacebookが創業されるなど、ビッグデータを新しい情報サービスという価値に転換することに成功した企業が当時、続々と誕生していた。ビッグデータを用いて容易なタスクは、識別関数の構築を通じた判別や分別である。カーネル法は、データ空間で複雑な識別境界面を推定するのではなく、あえて超高次元の特徴ベクトル空間を構成し、そこで豊富に蓄積された線形のモデリング技術を活用する手法である。もちろん、データ間の類似度を規定するカーネル関数は“人”が設定しなければならないが、データ空間での非線形モデリングに求められる巧みの技の多くを、カーネルトリックに押しつけられた点は大きい。パターン認識手法の多くはデータどうしの内積計算を含むため、カーネルトリックによって既存の線形モデリングの諸手法が非線形版に自然に一般化された。カーネル法研究の進展は、グラフ間の類似度を計るさまざまなアイデアを喚起し、その後、分子構造や原子配列のようなグラフ構造による記述が本質的に重要である、バイオインフォマティクスやマテリアルズインフォマティクスといった分野で大きく花開いていく（Schölkopf et al., 2004）。

2000年代はビッグデータの登場により、統計的モデリングの主たる興味である生成モデルの構築から、カーネル法と最適化の活用による、複雑な識別関数（識別モデル）の自動構築に、研究のトレンドがシフトした点は指摘しておきたい。特に、1990年代後半から非常に多くの応用成果を生み出したサポートベクターマシンは、カーネル化技法の恩恵を最も受けたと言える（Boser et al., 1996）。それと同時に、機械学習の言葉で代表される研究者コミュニティが世界的に育っていったのもこの時代である。その理由として、データに基づくさまざまな統計的推測手法やデータマイニングのアルゴリズムが、教師あり学習、教師無し学習、そして強化学習といった、機械学習の3つのタイプで整理され、異分野からの研究者の参入障壁が低くなった点は大きい。

### 4. スパースモデリング：非ガウスモデリング

2000年代半ばから2010年代にかけて非ガウスモデリングにおいても大きな発展が、スパースモデリングの普及（伊庭 編, 2017; 科研費新学術領域研究「スパースモデリング」, 2019）によってもたらされた（樋口, 2014, 2016a）。応用ドメインや計測現場における統計的モデリングの成否は、膨大な説明（属性）変数群の中からタスクの解決に有効な特徴ベクトルの適切な構築にあると断言できる。今、興味ある対象を膨大な説明変数の中から少数個で線形回帰表現する問題を考える。何次の回帰モデルとするのか、さらにどの変数を使うかで、膨大な数の回帰モデルが存在する。モデル数は組み合わせ爆発しており、現実にはAICなどの情報量規準による最適モデルの探索（変数選択）は基本貧弱である。一方スパースモデリングでは、回帰係数（変換行列内の要素）にL1正則化（絶対値誤差の最小化）を加えた上でパラメータ推定を行う（川野他, 2018）。最適化関数の形がL1であることから、L1正則化は非ガウスモデリングと言える。この最適化の結果として、重要な説明変数のセットを自動的に浮きあがらせる。

正則化自体は、概念的には古い、汎化能力を高める一つの方策である。前述した1980年代のベイズモデリングの成熟期には、パラメータに対してL2（二乗誤差の最小化）制約を加えることで正則化を実現していた。一方、L1正則化の採用により目的関数の解析的な性質を病的にするかわりに、得られるパラメータの最適解に癖が出ることを意図するアプローチも以前から実はあった。ここでの癖というのは、値が小さいパラメータ値は“大胆にも”ゼロに自動的

にセットされてしまう性質を指す。これにより、属性変数に関するパラメータの値がゼロとなれば、それは事実上その属性変数は不要であることを意味し、変数選択が最適化により実現できるのである。このアプローチは、病的な最適化関数に対して高速に最適解を求めるアルゴリズムが提案されたことを端緒として、1990 年半ば以降この 20 年間、理論的にもまた応用の広がりの意味でも精力的に研究されてきた (Hastie et al., 2008)。スパースモデリング研究の隆盛を後押しした別要因として、最適化法の計算高速化と汎用的ソフトウェアの実用化などの、最適化手法のコモディティ化がある。これは、スパースモデリングはもちろん、前述のカーネル法の進展に大きな役割を果たした (樋口, 2014)。

スパース性のようなデータ空間の低ランク性を陰に陽に活用するアイデアは、1990 年代後半から 2000 年代にかけて、信号処理の成分分離 (複雑なノイズ成分の除去) に有力な手法を多数導き、事実、独立成分分析 (ICA: Independent Component Analysis) (村田, 2004)、Compressed Sensing (CS) や、Non-negative Matrix Factorization (NMF) (亀岡, 2010)、テンソル分解にもそのアイデアを確認することができる。このように、説明変数の選択といった、限定されたモデリング技術に関しては、スパースモデリングにより非ガウスモデリングが実現されたと言える。なお、画像、音声、自然言語の処理においては、現在その課題は、End-to-End と呼ばれる深層学習の技術によって大幅に自動化されている。一方、カーネル法は、データ空間から高次元の特徴ベクトルを構成する方法なので、情報圧縮の文脈では逆センスの手法である。

## 5. 深層学習：非線形モデリング

2010 年代にはいつの最大の衝撃は、特定のタスクでの深層学習の圧倒的パフォーマンスである。その性能の高さは、入出国時の自動顔画像判別、スマートスピーカーでの音声認識、多言語自動翻訳など、私たちの生活に身近な製品として具現化している。これらの技術は人の働き方や社会の構造そのものにも直接的に影響を与えていることから、深層学習はこの四半世紀の情報科学技術における最大のブレイクスルーであると言わざるを得ない。

過去に AI ブームは、1950 年代後半から 1960 年代前半と、1980 年代の 2 回あった (樋口, 2016b)。初回は、一つの脳神経活動を数式でモデル化し、それらを複数組み合わせコンピュータに処理させた、極初歩的なニューラルネットワーク (NN: Neural Network) の提案により巻き起こった。二回目では、「第五世代コンピューティング」がまずブームを誘起し、1980 年代半ばあたりから、ニューロ・ファジー家電製品の開発・販売に代表されるような、第二次ニューロブームがそれに続いた。その後、計算に莫大な時間がかかることや、試行錯誤による構造やパラメータチューニングの面倒さ、また理論構築困難 (精度保証など) により、NN 研究への興味が失われ、AI にとって冬の時代が長く続く。筆者は第二次 AI ブームの時代からベイズ統計学の立場で AI にかかわってきたので、AI 研究とのつきあいも 30 年近くになる。特に、「第五世代コンピューティング」のあとの旧通商産業省にとっては新情報技術の最後となる大型プロジェクト、「リアルワールドコンピューティング (RWC) 計画」(大津, 1993) の理論・アルゴリズム基盤領域 (前期 5 年) へ参加した。RWC が終結した直後からビッグデータの時代が到来する。

従来の教師あり機械学習は、まずは問題毎に元データから“適切な”特徴ベクトルを構成し、この特徴ベクトルを入力として、諸々の目的関数を最大 (最小) 化するように、学習器内の膨大なパラメータの値を決定・推測する。ここで学習器とは、サポートベクターマシンや、ブースティング、ノンパラメトリックベイズなどを指す。パラメータ学習アルゴリズムは学習器毎に異なるが、一度、データが与えられればあとの学習プロセスはマシンにお任せである。一方、特徴ベクトルの構成法はほぼ人間の知的作業に委ねられており、実はこの構成法が予測・判別性能といったパフォーマンスをほとんど決めていると言っても過言でない。特に一般物体認

識では、画素値ベクトルから特徴ベクトルを構成する、さまざまな手法の研究が長年にわたって行われてきた。深層学習では、この特徴ベクトルの構成作業をあえて省略し、画素値ベクトルを直接 NN の入力とする。

深層学習で使われるニューラルネットワークは、層数が大幅に増えた以外、第二次ニューロブームの時のものと違いは無く、そのパラメータ学習アルゴリズムも、Back propagation を基本とする以前のものと大差はない。深層学習では層数が大幅に増えた結果、パラメータ数も爆発的に増え、学習アルゴリズムもさほど賢くなっているわけでもないので、必然と計算リソースはこれまでとは桁違いに必要となる(樋口, 2018)。ただし、上述したように、それまでの特徴ベクトル構成法は機械学習の『匠の技』と言え、“機械”学習にもかかわらず、人間の判断が最も性能向上に大切という羊頭狗肉の面もあったが、その問題を特定領域では基本的に解決した点は画期的である。

深層学習では、通常、最適化関数は L2 であるため、誤差にガウス性を仮定している。よって、非線形・ガウスモデリング技術は、深層学習により、少なくとも予測・判別性能の観点からはほぼ極みに達したと言える。これまでの深層学習の大きな成功事例は、入力データが画像、音声、テキストなどの、データ構造が簡単なものに集中しているが、その制約を逆手にとって、さまざまな入力データをあえて画像などに変換すれば(Poplin et al., 2018)、特徴ベクトルの選択問題を回避可能とも言える。

非線形モデリングの自動化に革命的インパクトを与えた深層学習にも、課題は数多く残っている(樋口, 2016a, 2016b, 2018)。まず、さまざまな泥臭い工夫に、統一的理論(ガイドライン)がない。第二次ニューロブームの時には、層数やノード数の適切な選択問題や学習係数の客観的決定法が、情報量規準の観点などから真剣に研究された。もはや深層学習においては、これらは完全に(当座?)議論の対象外となっている。よって、精度保証のような理論構築は極めて困難なままである。まさに、深層学習の問題毎の最適化が“黒魔術”とも呼ばれる所以である。もう一つの弱点は、推定結果が(ほぼ)完全にブラックボックス化している点である。この NN の本質的な弱点は、第二次ニューロブームの時も、統計学を中心として研究者から、また NN を実問題に適用した産業界からも、大きな不満としてさんざん指摘されてきた。今の応用開発現場では、予測や判別性能が足りなければ、層数をさらに積み増したり、パラメータ学習に使う計算時間を 1, 2 週間延ばすなど、とても系統的とは言えない職人技が横行している。今後、深層学習がさらに大きな飛躍を遂げるためには、精度保証のための理論研究がすすむことが必須であろう。

## 6. 今後の統計的モデリング

非ガウス・非線形モデリングを自動化する試みは、この四半世紀に大きな飛躍を遂げた。スマートセンサ、スマホ、バイオメトリクス、インターネット調査、SNS(Social Network Service)からの即時的集計などなど、データ取得の環境はますます“ビッグデータの”になっていく。膨大なデータ量は統計的モデリングの自動化に大きく寄与する一方、精度(信頼度)はサンプル毎にバラバラで、欠測や異常値の混入も頻発、また属性変数の次元も巨大といった、量と質のアンバランスが極端なデータの解析がさらに求められる。このような状況では、まず、目的に直結する属性変数の選択(言い換えれば、特徴ベクトルの構成)や、サンプルの信頼性の評価により、真に必要なデータセットの作成、つまり「スマートデータ」の生成こそが大切になる。統計学は、外部に作成してもらった「スマートデータ」の丁寧で精緻な分析のみを担うのか、それともビッグデータからスマートデータを構成するプロセスへ踏み込むのか、正念場に來ている。

一般社会においては、ビッグデータ × AI(人工知能)により大きな存在感を示す米中のブラッ

トフォーマー(7 Sisters と通常呼ばれる)に対する風当たりも強くなってきている。今後は、モデリング技術の向上を計る上で、モデルの“人”による解釈可能性、帰納法で構築された意思決定システムの説明責任、データのバイアスと社会的偏見の分離など、人に寄り添った視点が極めて重要になってくる(内閣府, 2018)。プライバシーや Fairness を SNS 業界でどう位置づけ、その適切な管理とモニターを実現する技術が今まさに求められている。個人的には、その解決は永久に解決に至らぬ「永遠の難問」と考えるが、しかしながら、下に示すような具体的な技術課題は一般社会に喫緊に対応策を示さねばならない。

- Accountability (結果の説明責任力)
- Interpretability (結果の解釈可能性)
- Reproducibility (結果の再現性)
- Transparency (手続きの透明性)

これらの課題解決には、以下の問題点

- Data imbalance (どんなにビッグデータになろうとも、データそのものがバイアスのある母集団からの情報収集とならざるを得ない現実)
- Sampling strategy (ビッグデータからの統計的判断に利用するデータのサンプリング方策)
- Generative model (データ生成を模擬する確率モデル)

を深く認識することが肝要であり、それらはすべて Foundation of Statistics につながる、統計学の原理の本質的理解そのものと言える。

## 参 考 文 献

- 赤穂昭太郎 (2008). 『カーネル多変量解析—非線形データ解析の新しい展開(シリーズ確率と情報の科学)』, 岩波書店, 東京.
- Akaike, H. (1980). Likelihood and the Bayes procedure, *Bayesian Statistics* (eds. J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith), 143–166, University Press, Valencia.
- 赤池弘次 (1981). モデルによってデータを測る, 数理科学, **19**(3), 7–10.
- 赤池弘次, 北川源四郎 編 (1994). 『時系列解析の実際 I』, 朝倉書店, 東京.
- 赤池弘次, 北川源四郎 編 (1995). 『時系列解析の実際 II』, 朝倉書店, 東京.
- 赤池弘次, 樋口知之, 川崎能典, 石黒真木夫 (2007). 『赤池弘次: 統計科学を語る—1993 年. 駆け出し研究者によるインタビュー記録—』, 統計数理研究所, 東京.
- Boser, B. E., Guyon, I. M. and Vapnik, V. N. (1996). A training algorithm for optimal margin classifier, *Proceedings of the Fifth Annual ACM Workshop on Computational Learning Theory*, **5**, DOI:10.1145/130385.130401.
- Doucet, A., de Freitas, N. and Gordon, N. (eds.) (2001). *Sequential Monte Carlo Methods in Practice*, Springer-Verlag, New York.
- 福水健次 (2010). 『カーネル法入門—正定値カーネルによるデータ解析(シリーズ多変量データの統計科学)』, 朝倉書店, 東京.
- Hastie, T., Tibshirani, R. and Friedman, J. (2008). *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*, Second Edition, Springer Series in Statistics, Springer, New York.
- 樋口知之 (2014). 統計数理の誕生とその広がり, 横幹, **8**(1), 14–21.
- 樋口知之 (2016a). スモールデータ, ビッグデータ, そしてスマートデータ—人工知能ブームの中での統計学—, 統計, **67**(1), 9–14.
- 樋口知之 (2016b). 人工知能はみよみまねマシンの究極形, 情報管理, **59**(5), 331–335.
- 樋口知之 (2018). データ関連の数理技術の変遷: 深層学習は春秋戦国時代の秦となるのか?, 人工知能,

33(2), 116–123.

- 伊庭幸人 (2003). 『ベイズ統計と統計物理』, 岩波講座 物理の世界 物理と情報 3, 岩波書店, 東京.
- 伊庭幸人 編 (2017). スパースモデリングと多変量データ解析, 『岩波データサイエンスシリーズ Vol.5』, 岩波書店, 東京.
- 伊庭幸人 編 (2018). 『ベイズモデリングの世界』, 岩波書店, 東京.
- 伊庭幸人, 種村正美, 大森裕浩, 和合肇, 佐藤整尚, 高橋明彦 (2005). 『計算統計 II マルコフ連鎖モンテカルロ法とその周辺』, 統計科学のフロンティア 12, 岩波書店, 東京.
- Isard, M. and Blake, A. (1998). CONDENSATION — Conditional density propagation for visual tracking, *International Journal of Computer Vision*, **29**(1), 5–28.
- 石黒真木夫, 松本隆, 乾敏郎, 田辺国土 (2004). 『階層ベイズモデルとその周辺—時系列・画像・認知への応用』, 統計科学のフロンティア 4, 岩波書店, 東京.
- 科研費新学術領域研究「スパースモデリング」(2019). <http://sparse-modeling.jp/> (2019年7月5日).
- 亀岡弘和 (2010). 非負値行列因子分解入門～音響信号処理を題材として, <http://www.kecl.ntt.co.jp/people/kameoka.hirokazu/publications/Kameoka2010Subspace07.pdf> (2019年7月5日).
- 川野秀一, 松井秀俊, 廣瀬慧 (2018). 『スパース推定法による統計モデリング』, 統計学 One Point 6, 共立出版, 東京.
- Kitagawa, G. (1993). Monte Carlo filtering and smoothing method for non-Gaussian nonlinear state space model, Research Memo., No.462, The Institute of Statistical Mathematics, Tokyo.
- Kitagawa, G. (1996). Monte Carlo filter and smoother for non-Gaussian nonlinear state space models, *Journal of computational and graphical statistics*, **5**(1), 1–25.
- 北川源四郎 (2005). 『時系列解析入門』, 岩波書店, 東京.
- 村田昇 (2004). 『入門 独立成分分析』, 東京電機大学出版局, 東京.
- 内閣府 (2018). 人間中心の AI 社会原則(案), [https://www8.cao.go.jp/cstp/ai\\_gensoku.pdf](https://www8.cao.go.jp/cstp/ai_gensoku.pdf) (2019年7月5日).
- 大津展之 (1993). リアルワールドコンピューティング研究計画 1. 情報処理の新たなパラダイムを目指して, 情報処理, **34**(12), 1423–1428.
- Poplin, R., Chang, P.-C., Alexander, D., Schwartz, S., Colthurst, T., Ku, A., Newburger, D., Dijamco, J., Nguyen, N., Afshar, P. T., Gross, S. S., Dorfman, L., McLean, C. Y. and DePristo, M. A. (2018). A universal SNP and small-indel variant caller using deep neural networks, *Nature Biotechnology*, **36**, 983–987.
- Schölkopf, B. and Smola, A. J. (2002). *Learning with Kernels*, The MIT Press, Cambridge.
- Schölkopf, B., Tsuda, K. and Vert, J.-P. (eds.) (2004). *Kernel Methods in Computational Biology*, The MIT Press, Cambridge.
- Shawe-Taylor, J. and Cristianini, N. (2004). *Kernel Methods for Pattern Analysis*, Cambridge University Press, Cambridge.
- Thrun, S., Burgard, W. and Fox, D. (2007). 『確率ロボティクス』(上田隆一 訳), 毎日コミュニケーションズ, 東京.

## Quarter Century of a Paradigm Shift in a Statistical Modeling Technology

Tomoyuki Higuchi

Faculty of Science and Engineering, Chuo University

I entered the Institute of Statistical Mathematics in April 1989, the beginning year of the Heisei period. Just before the end of the Heisei period, in March, 2019, I stepped down from the director and at the same time retired from the institute. Therefore, the “Heisei 30 years” are the life of a researcher for me. In this article, I summarize the major advances in statistical modeling technologies that form the basis of statistical thinking such as data-based prediction and discrimination from my point of view.