

まえがき

1950年代に人工知能という言葉が登場した当初、人間の脳に近い情報処理として目指していたのは、帰納と演繹に基づく命題論理の記号処理による知能の実現であった。コンピュータといえば電卓が巨大化したものとしての位置付けでしかなく、人間が指示した計算をただただ黙々と実行する装置に過ぎなかった当時としては、その人工知能の第一歩は革新的なものであった。しかしながらこの試みは、「察する」、「経験を生かす」、「気付く」という我々があたりまえに行っている営みの実現が困難であることが認識され、大きな壁に突きあたる。

このような状況と前後して、人工的知能創出のためのアプローチの一つである人工ニューラルネットワークは、20世紀半ばに登場したパーセプトロン (**perceptron**)、階層構造をもつネオコグニトロン (**neocognitron**)、フィードバック結合を有する再帰的ニューラルネットワーク (**recurrent neural network**)、多層構造をもつ人工ニューラルネットワークのデータからの学習を実現する誤差逆伝播法 (**back propagation**) の提案を通して、様々な発展を遂げてきた。近年は、畳み込みニューラルネットワーク (**convolutional neural network**) をはじめとする多層ニューラルネットワーク (**multilayer neural network**) が登場し、現在の深層学習 (**deep learning**) [3] の成功へとつながっている。そして、ビッグデータサイエンスという新しい研究分野の登場とともに、そのキーアプリケーションとしての重要性はますます高まりつつある。同時に、数学的裏付けの重要性がより認識されるようになりつつある。

データサイエンスの数学的基盤はいうまでもなく統計学である。大量のデータから重要な情報を抽出し、大量ではあるが不完全なデータから欠損した情報を予測し、新しい知識を発見するような技術は、統計学がもともと目指す方向であった。不完全ではあるが大量のデータが容易に得られる環境が急速に整いつつある

近年の状況は、過去のどの時代にもないほどの統計学に対する期待の高まりへとつながっている。20世紀後半にベイズ統計に基づき予測・推定機能を実現するマルコフ確率場 (Markov random field), ベイジアンネットワーク (Bayesian network) が登場し、人工ニューラルネットワークの学習機能を最尤推定の立場から理解するために有力なボルツマンマシン (Boltzmann machine) が提案され、確率的グラフィカルモデル (probabilistic graphical model) [5, 11, 12] による統計的機械学習理論 (statistical machine learning theory) として、この統計学に基づくアプローチは体系化されつつある。

世界的には **Neural Information Processing Systems (NIPS)** という名称の国際会議が1987年から発足し、統計的機械学習理論の新しい発想を常に生み出し続けてきた。たとえば、深層学習は2009年のバンクーバー/ウィスラーで開催されたNIPSが出发点となっている。国内における統計的機械学習理論の現在につながる取り組みは、1998年に発足した情報論的学習理論 (**Information-based Induction Sciences: IBIS**) という電子情報通信学会の第2種研究会が出发点の一つとして位置付けられる。現在は、情報論的学習理論と機械学習研究会 (**Information-based Induction Sciences and Machine Learning: IBIS-ML**) という電子情報通信学会の第1種研究会としてさらなる拡大を遂げている。2002年に文部科学省科学研究費補助金特定領域研究「確率的情報処理への統計力学的アプローチ (Statistical-mechanical Approach to Probabilistic Information Processing: SMAPIP)」が発足し、2006年に同じ文部科学省科学研究費補助金特定領域研究「情報統計力学の深化と展開 (Deepening and Expansion of Statistical Mechanical Informatics: DEX-SMI)」がこれに続いた。当時、これらの研究プロジェクトが目指したものは、「日本に來れば他では得られない理論や技術を得ることができる」と世界の研究者に思ってもらえるような研究分野の創出であった。日本では1950年代から材料科学の理論的基盤の整備を目的として、統計力学を基礎とする計算技法で世界に先駆けた多くの革新的計算技法の提案がなされていた [4, 6, 9]。これらの計算技法を統計学に基づく大規模データからの予測のための数学的枠組みと融合させることで、その思いの一部を実現できるのではないかというのが上記プロジェクトの最初の発想であった。そしてその思いの拠り所が西森秀稔氏の2つの著書 [7, 8] であった。現在、上記2つの

プロジェクトは2013年に発足した文部科学省科学研究費補助金新学術領域研究「スパースモデリングの深化と高次元データ駆動科学の創成 (High-Dimensional Data-Driven Science through Deepening of Sparse Modeling)」へと進化を遂げ、これまで想定しなかった研究分野への展開とともに新たな研究のステージへと踏み出しつつある。

本書では、確率的グラフィカルモデルの統計的機械学習理論について、画像処理とパターン認識に応用例を絞りつつ概説することから始める。特にパターン認識では、クラス分類問題という視点において多値ロジスティック回帰モデル (multi-valued logistic regression) と制約ボルツマンマシン (restricted Boltzmann machine) という2つの確率的グラフィカルモデルを通して、深層学習の基礎となる数理を紹介する。その上で、グラフ構造の疎 (sparse) 性という深層学習とは真逆の性質をもとに急速に展開しつつあるスパースモデリング (sparse modeling) [1, 2, 10] という新しい研究領域の最近の理論的基盤の深化の様子を、連続最適化問題という視点から概説する。

本書の執筆者の安田宗樹氏、大関真之氏、片岡駿氏は、いずれも上述の3つのプロジェクトを学生時代から現在にいたるまでの間に経験し、直感とセンスだけを頼りにそれぞれの相異なる立場で自らの研究領域を切り拓いてきた。本書はそれらの独特の研究世界観を、次の時代における人工知能の基盤として、学部・大学院学生に紹介するに値する書として結実できたものと考えている。執筆は、第1, 2章を田中和之、第3章を片岡駿、第4章を安田宗樹、第5章を大関真之が担当した。

本書の出版は、東北大学の照井伸彦教授、小谷元子教授をはじめとするクロスセクショナル統計シリーズの編者の先生方からのご提案が出発点となった。また、本書の出版が実現したことは執筆をあたたく見守ってくださった共立出版株式会社編集部の山内千尋氏のおかげと感じている。これらすべての皆様に深く感謝したい。

参考文献

- [1] Elad, M.: *Sparse and Redundant Representations: From Theory to Applications in Signal and Image Processing* Springer-Verlag (2010), 396p. (玉木徹 訳: スパースモデリング— l_1/l_0 ノルム最小化の基礎理論と画像処理への応用—. 共立出版 (2016), 436p.)
- [2] Hastie, T., Tibshirani, R., Wainwright, M.: *Statistical Learning with Sparsity: The Lasso and Generalizations*. Chapman & Hall/CRC (2015), 366p.
- [3] 神島敏弘 編, 麻生英樹・安田宗樹・前田新一・岡野原大輔・岡谷貴之・久保陽太郎・ボレガラダヌシカ 著, 人工知能学会 監修: 深層学習 Deep Learning. 近代科学社 (2015), 267p.
- [4] 菊池良一・毛利哲雄: クラスタ変分法—材料物性論への応用—. 森北出版 (1997), 179p.
- [5] Koller, D., Friedman, N.: *Probabilistic Graphical Models: Principles and Techniques*. MIT Press (2009), 1266p.
- [6] 守田徹: 第2章 フラストレートした磁性体の統計力学 (石原明, 和達三樹 編著: 新しい物性), 共立出版 (1990), 308p.
- [7] 西森秀稔: スピングラス理論と情報統計力学 (新物理学選書). 岩波書店 (1999), 206p.
- [8] Nishimori, H.: *Statistical Physics of Spin Glasses and Information Processing: an Introduction*. Oxford University Press (2001), 255p.
- [9] 小口武彦: 磁性体の統計理論. 裳華房 (1970), 302p.
- [10] Rish, I., Grabarnik, G.: *Sparse Modeling: Theory, Algorithms, and Applications*. Chapman & Hall/CRC (2014), 249p.
- [11] Sucar, L. E., Enrique, L.: *Probabilistic Graphical Models: Principles and Applications*. Springer-Verlag (2015), 277p.
- [12] 鈴木讓・植野真臣 編著, 黒木学・清水昌平・湊真一・石島正和・樺島祥介・田中和之・本村陽一・玉田嘉紀 著: 確率的グラフィカルモデル. 共立出版 (2016), 280p.

本書で使用しているコードの一部は、共立出版ウェブサイトの本書のページ
www.kyoritsu-pub.co.jp/bookdetail/9784320111233
からダウンロードできる。