企画にあたって

足　立　吉　隆*

データには「縦のデータ」と「横のデータ」があり，データセット(数)は縦のデータであり，データの種類(記述子)(数)は横のデータである．データセット数が増えるほど，データサイエンスが一層活用でき，データの種類数が増えるほどデータサイエンスが必要となってくる．データサイエンスは，実験，理論，計算工学に次ぐ第4の研究手法という認識が広がりつつあり，材料工学においても画像認識，パラメータの最適化，順解析，逆解析，データ同化，などへの適用が急速に進んでいる．人類では対応がほぼ不可能な複雑なデータに対して，傾向(順解析モデル)を見出し，データ数が増えるほどベイズ更新によりモデルの精度を上げていく．機械学習で作られた順解析モデルは，究極的な近似式といえる(物理モデルが背景にはない)が，その精度は理論式に戻づく予測を超える場合があり，しかも計算コストが低いため，多量の計算を順解析モデルにより行って傾向を把握することにより，我々が望む方向(例えば特性や特性バランスの最大化)に向うための記述子の最適化を提案してくれる．このような出力データを最大化(あるいは最小化)するための入力データの在り方を解析することを逆解析と呼び，遺伝的アルゴリズムやベイズ的最適化という手法によりシステムに実装される．逆解析は，材料開発あるいは新物質探索の革新的効率化を目指すにあたっては必要不可欠となるであろう．

子供は知識(データ)が少ないにもかかわらず，やがて各素養を備えた立派な大人に成長するが，その過程では周囲(特に親)から報酬とペナルティーを与えられ，成長するべき方向を学んでいく．同様なデータサイエンス手法は強化学習と呼ばれ，それを実装したプログラム「アルファ碁」がプロ棋士に勝利したことは記憶に新しい．このような強化学習を材料工学に適用している例はまだないものと思われるが，材料工学にとって今後注視に値するデータサイエンスであろう．

材料工学におけるデータの多くは，画像やスペクトラムの形で与えられる場合が多いものと思われるが，そこから徹底的に特徴量を抽出することは「横のデータ数」を増やすために大変重要となる課題である．手作業でこの作業を行うことは大変労力を要するが，例えば画像の場合は深層学習をはじめとする最先端画像認識手法(CNN，R-CNN，FCN，機械学習型画像処理など)を活用して対象領域の自動抽出を行った後に，位相幾何学，微分幾何学を使った豊富な特徴量の抽出が行われる．近年では不完全な物体の連結性を幅広いスケールで定量評価できるパーシステントホモロジー群や，連続的な形態変化を曲率に基づいて定量評価する手法や，あるいは銀河における星の分布を定量的に評価する際に用いられる二点相関関数(その後，次元削減が必要な場合が多い)などを材料工学に適用する研究が報告されるようになってきている．

このように目まぐるしく数学，情報工学が発展する中で，材料工学研究者がその要点を理解し，個々の材料工学の課題にうまく適用して，効率よく，精度よく機構解明や材料開発につなげることを願って，今回，数学研究者，情報工学研究者の専門家の先生方に，その基礎と材料工学への適用例について解説いただく機会を設けた．本ミニ特集が，新たな材料工学の一側面を開拓する機会となることを切に願っている．

最後に，材料工学に活用できる最先端の数学，情報工学を分かりやすく解析してくださった執筆者の皆様，そして本企画を強力にご支援してくださった会報委員会委員長の山本剛久先生(名古屋大学)に心より御礼申し上げる．

足立吉隆

★★★★★★★★★★★★★★★★★★★★★★★★★
1990年3月　名古屋大学大学院工学研究科修士課程修了
主な略歴　物質・材料研究機構，鹿児島大学
2017年4月–現職
専門分野：高次元材料情報学
◎鉄鋼材料を主対象に材料組織を定量3D解析し，そこで得た材料ゲノム情報を使って順解析・逆解析モデルを構築する研究に従事している．
★★★★★★★★★★★★★★★★★★★★★★★★★

* 名古屋大学大学院工学研究科材料デザイン工学専攻；教授(〒464–8603 名古屋市千種区不老町)
Preface to “Mathematics and Computer Science to Promote Material, Process, and Measurement Informatics”; Yoshitaka Adachi(Nagoya University, Nagoya)
Keywords: *material informatics, machine learning, 3D4D material science*
2018年12月3日受理[doi:10.2320/materia.58.6]