

2024年8月6日(火)

## 機能発現の3ステップモデル

東京大学・大学院新領域創成科学研究科 複雑理工学専攻

岡田真人

© 2024 Masato Okada

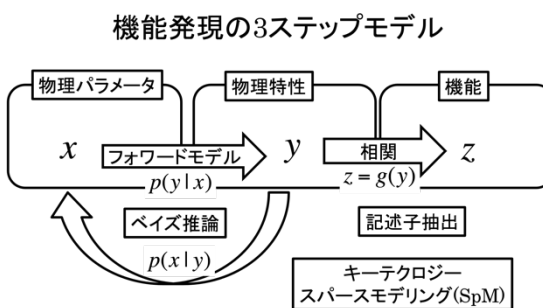
### 1. 機能の数理モデル化を目指して

私は2015年7月からマテリアルズ・インフォマティクス(Materials Informatics)を、急速に我が国に根付かせることを目的とした、国立研究開発法人物質・材料研究機構(NIMS)で行われた国立研究開発法人科学技術振興機構(JST)イノベーションハブ構築支援事業の一環として発足した情報統合型物質・材料開発イニシアティブ(Material Research by Information Integration Initiative:MI2I)に参画した。

### 2. 物質科学と材料科学は違うという教え

これまで私は物質科学関係では、物性物理学の研究の経験しかなかった。最初にMI2Iに参画した時に言われたことは、物質科学と材料科学では目的が全く違うということだった。最初は、その言葉に戸惑ったが、MI2Iで仕事を始めるとすぐに、材料開発に関する学問のことを材料科学と呼ぶからいけないのであって、材料工学と呼ぶべきであるということに気づいた。

考えてみれば当然のことで物性物理学に代表される物質科学では、量子力学/電磁気学/熱統計力学を基礎に、科学的に物質の性質を調べるのが目的であり、まさに物質“科学”と呼ぶべきものである。一方、材料開発の目的は、ある機能をもつ材料を作ることである。つまり材料“科学”ではなく、材料“工学”である。それなのに、科学を工学より一段上のものであるように扱う現在の風潮の顔色をうかがって、材料工学と呼ばずに材料科学と呼ぶ風潮があり、それが二つを別物と考え辛いものとなっている。



Igarashi, Nagata, Kuwatani, Omori, Nakanishi-Ohno, and Okada "Three levels of data-driven science" International meeting on High-dimensional Data-Driven Science (HD3-2015)Journal of Physics: Conference Series, 699 (2016) 012001(2016)

図1 機能発現の3ステップモデル

### 3. 材料開発への数理的アプローチ: Kepler の法則から SpM へ

ここから視野を広げると、材料等の機能の発現のメカニズムの理解は、量子力学/電磁気学/熱統計力学を礎とする物質科学での因果律だけで到達することができないことがあるのではないかとということが容易に想像できる。よく考えると材料での機能の発現は、系の性質で決まる熱統計力学の相転移のような現象ではない。機能は、そのような量子力学/電磁気学/熱統計力学を基礎とする物質科学上の概念とは別の工学的要請から決まるものである。従って、機能は物質科学の延長だけでは理解できないのである。

もちろん機能のメカニズムは理解できれば、それは量子力学/電磁気学/熱統計力学を基礎とする物質科学で説明できる。しかし、そのことが機能のメカニズムから演繹できるわけでないのはこれまで述べたとおりで、そのため、それとは別のアプローチが必要である。このように演繹不可能な状況でも、科学的に現象を扱う手法は、科学史を紐解けば存在する。それが Kepler の法則や前期量子論に代表される定量的現象論である。

その定量的現象論を得るためのアプローチ以下のようなものである。まず、ターゲットしている現象に関するデータを多種多量に獲得する。つぎに、その多種多量のデータから特徴量を定める。例えば Kepler の法則の場合、Tycho Brahe が測定したデータに対して、Kepler は惑星が楕円軌道を描くことから、面積速度、公転周期  $T$ 、公転半径  $R$  を特徴量として採用した。そして例えば、公転周期  $T$  と公転半径  $R$  の間に  $T^2$  と  $R^3$  が比例するという定量的現象論である Kepler 第三法則を求めた。このような多数の定量的現象論が得られた後、Newton の運動方程式と万有引力の法則が提案され、その後、Newton の運動方程式と万有引力の法則からこれら多数の定量的現象論を演繹できることが示されることで、Newton の運動方程式と万有引力の法則の正しさが示唆された。このように、系の数理モデルが決まる前に、その系のデータから定量的現象論が提案されることが科学史を紐解くことで明らかになる。物理学科の教育では、物理数学を用いた数理モデルからの現象の説明に時間を割くので、物理学科を卒業していても、現象論から第一原理が導出されるという、自然法則の展開を理解できてない人が多い。

Kepler の時代よりもデータが大量になり高次元化すると、Kepler が行ったような、天才がその属人的洞察力でデータを現象論的に記述できる特徴量を導出することが難しくなっている。この困難を、データ科学の力で克服するのがスパースモデリング(SpM)である。SpM では、まず(1)考えつく限りの特徴量をヒトが提案する。つぎに(2) 説明したいデータである説明変数が、(1)で提案した特徴量の線形関数でかけることを仮定する。最後に(3) SpM のアルゴリズムにより、説明変数を説明できる特徴量の組みを求める。

#### 4. 機能発現の 3 ステップモデル

データ生成の数理モデルが既知の場合に、威力を発するベイズ推論と、データ生成がわからないときのアプローチのスパースモデリング(SpM)を組み合わせることにより、図 1 の機能発現の 3 ステップモデルが示すように、第一原理だけではアプローチできない材料の機能への数理アプローチの提案できる。さらにこのアプローチは単に材料工学だけのものでないことすぐにわかる。神経科学を例にとると、初期視覚野(V1)の単純型細胞の特性はガボールフィルタで記述できることが知られている。V1 単純型細胞の機能は網膜に入力された画像から直線検出することである。この V1 単純型細胞の機能発現は、図 1 の左のベイズ推論の部分だけからなり、右の SpM は必要ない。一方、高次視覚野でパターン認識の機能を司るとされている側頭葉の神経細胞の数理モデルは提案されていない。その理由は V1 から側頭葉まで至る経路で、どのような視覚情報処理が行われているかが不明であるからである。側頭葉の神経細胞のパターン認識の機能発現の数理モデルは、図 1 の左のベイズ推論の部分ではなく、左の SpM だけである。この時の、図 1 の中間の特徴量は側頭葉の神経細胞の発火パターンであり、その発火パターンを説明変数とし、パターン認識の結果を目的変数とする SpM により、側頭葉のパターン認識の機能を司る神経細胞の発火パターンを探る。この神経科学の例から明らかなように、全ての機能の発現は図 1 の機能発現の 3 ステップモデルで数理モデル化できるというパラダイムを提唱する。

そのような流れもあり、私は、文部科学省科学研究補助金「新学術領域研究」平成 25 年度～29 年度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 Initiative for High-Dimensional Data-Driven Science through Deepening Sparse Modeling (略称 新学術 SpM)を領域代表として申請し、採択された。

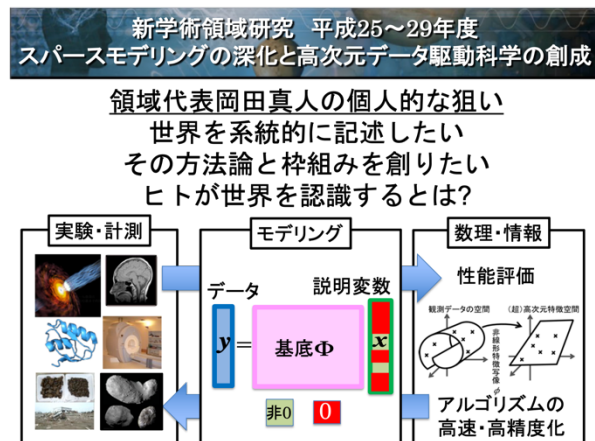


図 2 新学術領域 SpM

<http://sparse-modeling.jp/>

この新学術 SpM では、図 1 の機能発現の 3 ステップモデルの右側の比重が高い中学校の理科第二分野の生物学地学のテーマを、SpM を横串にすることにより、分野横断的に研究する。幸い、この新学術 SpM は大成功に終わり、このことは図 1 の機能発現の 3 ステップモデルのパラダイムが正しいことを示唆することとなった。

#### 5. 機能発現の 3 ステップモデルからマテリアルズ・インフォマティクスの学理

へ

機能発現の 3 ステップモデルにより，材料開発では，物質科学と材料科学は違うのであるという表層的なスローガンではなく，材料開発にも図 1 の機能発現の 3 ステップモデルが適用し，物質科学と材料工学の関係が図 1 の機能発現の 3 ステップモデルの右のベイズ推論と左の記述抽出に記述されていて，相互に助け合うことで材料開発がすすむというマテリアルズ・インフォマティクスの学理の提案に繋がる。

私は MI2I での活動を開始するとそのことにすぐに気づき，第 15 回 NIMS フォーラムでの招待講演[1]でこの機能発現の 3 ステップモデルを発表した。当時はベイズ最適化による第一原理計算の探索時間の短縮がマテリアルズ・インフォマティクスだと信じられており，その人達には，私がマテリアルズ・インフォマティクスの学理である機能発現の 3 ステップモデルの意味は全く理解されなかったようである。そこで我々は，機能発現の 3 ステップモデルのプライオリティを取るために，新学術 SpM の国際会議 International Meeting on “High-Dimensional Data-Driven Science” (HD3-2015)で機能発現の 3 ステップモデルを発表した[2]。現在，心理学/行動科学[3]，地球化学[4]，物質材料科学で機能発現の 3 ステップモデルの有用性が実証されつつある。

#### 参考文献

[1] 第 15 回 NIMS フォーラム 2015 年 10 月 7 日 (水)

<https://www.nims.go.jp/nimsforum/>

マテリアルズ・インフォマティクスとは何か-物質材料科学とデータ駆動科学-

<https://mns.k.u-tokyo.ac.jp/pdf/2015nims.pdf>

[2] Igarashi, Nagata, Kuwatani, Omori, Nakanishi-Ohno and M. Okada, “Three Levels of Data-Driven Science”, *Journal of Physics: Conference Series*, 699, 012001, 2016.

[3] Ichikawa, Kitazono, Nagata, Manda, Shimamura, Sakuta, Okada, Yamaguchi, Kanazawa and Kakigi, “Novel method to classify hemodynamic response obtained using multi-channel fNIRS measurements into two groups: Exploring the combinations of channels”, *Frontiers in Human Neuroscience* 8, 480, 2014.

[4] Kuwatani, Nagata, Okada, Watanabe, Ogawa, Komai and Tsuchiya, “Machine-learning techniques for geochemical discrimination of 2011 Tohoku tsunami deposits”, *Scientific Report*, 4, 7077, 2014.

## 謝辞

この解説は、熊本大学水牧仁一郎教授，東京大学青西亨教授，JASRI 横山優一  
研究員、東大新領域複雑理工学専攻岡田研竝河伴裕氏，森口棕太氏，林悠偉氏に  
チェックいただきました．ここに感謝の意を表明いたします．