自動ハイスループット材料計算法と データ駆動型マテリアルデザイン

福島鉄也

産業技術総合研究所 機能材料コンピュテーショナルデザイン研究センター

ナノマテリアル・ナノデバイスデザイン学 2024年10月28日(月)



新・高機能材料の探索やデザインは難しい…



3 磁性材料のための基盤的シミュレーションコードの開発(AkaiKKR)

4 データリポジトリ、大型施設(スーパーコンピューター「富岳」)との連携



自動網羅計算による磁気物性データの創出とその活用(不規則系)

やりたいこと:広大な材料空間と探索

灰色の点: 未探索の材料群



「O. Isayev, Nature 571 42 (2019)」から引用

新物質探索の難しさ

- ・ 周期表の元素の組み合わせから、無限ともいうべき物質群
- ・「多数」の原子・電子が集まった時に初めて現れる特徴・性質

100個以上の原子が存在する

例) 永久磁石だとNd、Ce、Dy、Fe、Co、B等を混ぜる!



新物質の発見するには 多くの時間とコストが必要





材料探索手法の変換



第一原理計算を用いたデータ駆動型マテリアルデザインにより高効率エネルギー変換材料、省エネルギー材料、環境調和材料、安全安心保障材料、生体調和材料など、これからの社会が必要とする材料やデバイスを効率良く、環境低負荷、省エネルギーで開発する材料探索は社会的急務!

物質の設計(デザイン)、探索

- 目標
 - ▶ 新規機能を持つ新物質の発見
 - ▶ 既存物質の高機能化
- デザインは実現(解明)の逆問題



Computational Materials Design (CMD) workshopのテキストから引用

例) 強磁性体、強誘電体、超伝導体の場合



計算物質科学における数値シミュレーションとスケール





計算機上での仮想実験(電子状態計算) 高精度で物性値を計算可能(汎用かつ非経験的)

例)磁性材料

例)スピントロニクス材料



- ✓ 物性を支配する微視的物理・化学機構の解明
 - ✓ 新規物性・機能を有する未知材料のデザイン
 - ✓ 自動ハイスループット計算を用いた大規模物性データベースの構築

計算機ナノマテリアルデザイン(CMD®)エンジン

https://cmdworkshop.sakura.ne.jp/index.htmlを参照



Fe/MgO/Feを用いた TMR素子の実現

物質科学における方法論の変換

J. Mater. Sci. 54 8361 (2019)を基に福島が編集

MIの重要性は日本においても1980年後期に指摘されていたが…

吉田博、応用物理、第58巻、第9号(1989)

将来的には、人工知能・AIによる新物質の推論予測演算や、コンピューター・ グラフィックス技術との組み合わせによって、新物質開発がより効率的に進め られるものと期待される(本文より抜粋)。

マテリアルズインフォマティクス(MI)の目的

▶ 物質探索

- 新・高機能を持つ未知物質・材料の発見
- 予測モデルは必ずしも構築する必要はない

≻ 物性予測

- 予測モデルを構築し、物質の物性を高精度に予測
- ・ 非線形モデル(Black box 関数)、あるいは線型モデルを利用

> <u>物理法則の発見</u>

- 物性発現機構の解明
- 得られた知見から新たな原理、法則の発見

▶ <u>物性データベースの構築</u>

- 機械学習に資する物性データの創出
- 次世代研究者ためのデータプラットフォームの整備

マテリアルズインフォマティクス(MI)のフロー

MIの応用例:リチウム2次電池正極材料

田中功グループ(京都大学)

- 第一原理計算により2,000個のデータを創出
- Li(Fe_{1-x},Zr_x)(P_{1-2x},Si_{2x})O₄を提案

M. Nishijima, Nature Commun., 5 4553 (2014)

小野寛太グループ(高エネルギー加速器研究機構、大阪大学)

T. Ueno et al., npj Computational Materials, 4 4 (2018)

マテリアルズロボティクス

ー杉太郎グループ(東京工業大学)、安藤康伸ら(産総研)

AI/Robot-driven Materials Research

合成、観測、最適化、データベース構築 をロボットにより自動に行う

https://emira-t.jp/ace/9933/

- ベイズ最適化で成膜の合成条件を最適化し、14回の自動合成により最小の抵抗値を実現
- 本手法により、人的作業に比べ約10倍の効率化に成功

R Shimizu et al, , Applied Phys. Lett. Mater. 8, 111110 (2020).

マテリアルズインフォマティクス(MI)の問題点

現在の状況

バイオ・ケモインフォマティクスの分野に比べて、無機材料の分野におい てはデータ科学との融合は遅れている

理由

元素種・組成の組み合わせから、材料空間は広大であるが、データ数は 極端に少ない(100件以下のデータで機械学習をすることも…)

解決策

- 1. 限れられた少ない教師データから高精度な有効モデルの構築
- 2. 計算と実験データを組み合わせたデータ同化の適用

3. 自動網羅計算による利用価値の高い物性ビッグデータの実現 (第一原理計算 + スーパーコンピュータ「富岳」)

電子系のハミルトニアン(N電子系)

自由度が大きすぎて数値的に解けない(単純な場合を除く)

密度汎関数理論に立脚して何とか計算する

多体問題

有効ポテンシャル中の 一電子問題

コーン・シャム方程式に対する様々な近似や解き方(1)

コーン・シャム方程式に対する様々な近似や解き方(2)

KKRグリーン関数法(全電子第一原理計算手法の一つ)

- 多重散乱理論に立脚し散乱の定常状態を記述する
- 散乱の定常状態 → エネルギーの固有状態
- 他の手法とは異なり、一電子グリーン関数を直接計算する
- 電子状態、磁気特性、伝導特性が高精度で計算可能
- 散乱に関わる問題が扱える:不純物、合金、フォノン・マグノン散乱等

KKRグリーン関数法でできること

- 全電子計算(擬ポテンシャルなし)であるのに高速
- 2次元系、(半無限の)多層膜構造
- オーダーN(N²)計算による大規模系の取り扱い
- グリーン関数を必要とする問題
 - ▶ 不純物問題
 - ▶ 不規則合金
 - ▶ 欠陥を含む系
 - ▶ モデルのパラメーター(磁気相互作用、有効対相互作用等)
 - ▶ 輸送現象(久保公式の利用)
 - ▶ 有限温度における電子・磁気・伝導特性(マグノン、フォノン散乱)

不規則系とは?

置換位置

アモルファス、液体

スピングラス

- 点欠陥
 - ▶ 空孔や不純物
- 置換不規則系
 - ▶ ランダムポテンシャル、合金
- 配置不規則系
 - ▶ アモルファス、液体

• 磁気不規則系

▶ スピングラス、LMD

不規則系では周期性が壊れる

スーパーセル法

- 長所
 - 局所環境効果の考慮
 - カの計算、構造の最適化

短所

- 人工的な相互作用が入る
- 配置平均が難しい
- 低(高)濃度の再現が困難
- 計算コストが高い

不規則系(合金等)では配置平均が重要

2元合金 $A_{1-x}B_x$ の場合 A (B)

個々の原子の配置 (多くの配置が存在) **マクロな性質**を得るには **配置平均**が必要

✓ 有限個の配置のみを考えて平均をとる(スーパーセル法)。
 ✓ 低(高)濃度領域では計算コストが非常に高い。

配置平均を計算する

2元合金 $A_{1-x}B_x$ の場合 (B)A

✓ 上記のように2段階に分けて平均を取るのは厳密に正しい。
 ✓ 原子が無数にある場合は配置平均を取るのは困難。

有効媒質(仮想原子)の考え方

2元合金 $A_{1-x}B_x$ の場合 A (B)

配置平均をとった系

有効媒質系

✓ 配置平均をとった系は有効媒質で置き換えることが可能。
 ✓ 有効媒質系では並進対称性が復活する。
 ✓ 有効媒質系は基本的に非エルミートになる。

有効媒質を用いて平均をとる

2元合金 $A_{1-x}B_x$ の場合 A (B)

✓ 有効媒質系においても上記の平均は厳密に正しい。
 ✓ 有効媒質系の原点にAB原子を置いても周りの環境が変化しないとする → シングルサイト近似

コヒーレントポテンシャル近似(CPA)

(B)2元合金 $A_{1-x}B_x$ の場合 A

 $(1-x)G_A + xG_B = \tilde{G}$

- ✓ CPAでは上記の式を満たすように \tilde{G} を決定する。
- ✓ G_AとG_BはGに依存するのでセルフコンシステント計算を実施する必要がある。

KKR-CPA、KKR-VCA、スーパーセル法の比較

- KKR-CPAは実験傾向を非常によく再現する(参考:スレーター・ポーリング曲線)
- KKR-VCAの場合Mnが増加すると、実験傾向からはずれる
- スーパーセル法は原子配置によって結果が異なる(参考:SQS構造の使用)

スレーター・ポーリング曲線(Slater-Pauling curve)

3d遷移金属合金における1原子当たりの磁気モーメントを,1原子当たりの平均電子数 に対してプロットした曲線

- FeCoやFeNiが大きな磁気モーメントを持つ。
- ・ リジットバンド的にSP曲線を理解できる部分もあるが、複雑な分岐を有する。
- KKR-CPA法は実験結果を、非常に良く再現する。

H. Akai, Hyperfine Interactions 68 3 (1991)

不規則合金・磁性材料の電気抵抗率の温度依存性

H. Shinya et al., Appl. Phys. Lett. 117, 042402 (2020)

計算物質科学界におけるデータリポジトリ

API機能により、ユーザーは容易に材料データ をリポジトリからダウンロードでき、機械学習や スクリーニングを実施可能

一般的な情報科学の国際会議においても
 これらのデータリポジトリは標準的に利活用
 → データ科学者が物質科学に参入しやすい

日本には上記のようなリポジトリが存在しない、どのように進めるべきか?

- 海外と差別化を図り、材料開発にとって利用価値が高いリポジトリを構築
- •「富岳」等の大型計算機による、高品質なマテリアルデータの創出

Materials Project

- 物質材料の第一原理計算結果のデータベース。
 (結晶構造、バンド構造、熱力学量、相図等)を提供する。
- MITの研究グループが運営。
- ・ httpベースのAPIも提供されており、ユーザ独自のスクリーニングが可能。

The Materials Project by the numbers

https://next-gen.materialsproject.org

スーパーコンピュータ「富岳」

CPU数:158,976個 総演算性能:約442ペタフロップス (フロップス:コンピュータが1秒間に処理可能な浮動小数点演算の回数) © RIKEN

スーパーコンピュータ「富岳」と計算物質科学

4期連続4冠

 - 高さ:性能の高さ、裾野の広がり:ユーザーの拡がり (https://www.r-ccs.riken.jp/fugaku/about/) → 性能・使いやすさにこだわったスパコン 	 1. 量子物質の創発と機能のための基礎科学 —「富岳」と最先 端実験の密連携による革新的強相関電子科学
(計算機システムとアフリケーションのCo-design)	2. 次世代二次電池・燃料電池開発によるET革命に向けた計
	算・データ材料科学研究
	3. 省エネルギー次世代半導体デバイス開発のための量子論
	マルチシミュレーション
	4. 大規模計算とデータ駆動手法による高性能永久磁石の開発
	5. 環境適合型機能性化学品
CPU数:158,976個 総演算性能:約442ペタフロップス	6. データ駆動型高分子材料研究を変革するデータ基盤創出
(フロップス: コンピュータが1秒間に処理可能な浮動小数点演算の回数) ©RIKEN	7.「富岳」を活用した革新的光エネルギー変換材料の実現

- ・「富岳」は、2021年11月に4部門で4期連続の世界1位を獲得
- 基礎科学、気象、防災、医療等の幅広い分野で用いられているが、産業分野において も画期的な成果を期待されている(成果創出加速プログラム(物質・材料系))

成果創出加速プログラム(物質・材料系)

 「大規模計算とデータ駆動手法による高性能永久磁石の開発」では、基盤的シミュレーション手法の開発、そして不規則系磁性材料を対象とした大規模物性データベースの 構築を通じ、高性能磁性材料の開発を目指している

大規模計算とデータ駆動手法による高性能永久磁石の開発 (2020.4~2022.3)

物理-化学連携による持続的成長に向けた高機能・長寿命材料の 探索・制御(2023.5~2026.3)

スーパーコンピュータ「富岳」成果創出加速プログラム

永久磁石材料とその用途

http://www.power-technology.com/

Hard Disk Drive : VCM

Hybrid \cdot electric vehicle

Nd-Fe-B (Nd₂Fe₁₄B) 永久磁石

希少金属元素(2015~2025)

K. Binnemans, J. Sustain. Metall., 4 126 (2018)

希少金属を含まない、Nd-Fe-Bを超える 永久磁石材料の探索は非常に重要

永久磁石の研究課題

A) 新規磁石の提案・主相の設計(第一原理計算、データ科学) 高磁化、高磁気異方性、高キュリー温度、省希少元素、耐熱性

B) 材料特性(保磁力)の改良指針の提示(マルチスケールシミュレーション) 磁化反転機構の解明、微細組織の最適化

磁性材料のための基盤的シミュレーションコード: AkaiKKR

自動網羅計算によるユニークかつ利用価値の高い 磁性材料データの創出(AkaiKKR)

- 材料パラメータ並列
- ・ 全電子計算(擬ポテンシャルなし)
- CPAによる不規則性の取り扱い
- 磁化、キュリー温度、磁気異方性
- 電気抵抗率、ゼーベック係数
- 有限温度の磁性・伝導特性
- 磁気エントロピー、スピン凍結温度
- 安定な自動網羅計算

. . .

磁石材料への応用例:磁化とキュリー温度

 $(Sm_{1-\alpha-\beta}La_{\alpha}Ce_{\beta})_{2}(Fe_{1-\gamma-\delta}Co_{\gamma}Ni_{\delta})_{17}N_{3}$

製造コストと性能の観点から最適な、Sm-Fe-N系 永久磁石材料(主相)を発見(合成に成功)

磁化のカラーマップと実施例

特願2023-213937 AIST-ISSP-トヨタ共同研究

磁石材料への応用例:磁気異方性(FP-KKR法)

H. Okumura et al., Solid State Commun. 373-374, 115257 (2023)

Fe、Cu、Niを添加し、 磁化と異方性が高い材料を探索 La、Co添加による異方性の変化 Co少量添加で異方性が増大

AkaiKKRパッケージ・ツールの整備(GitHub: AkaiKKRteam)

	AkaiKK	Rprogram		
AkaiKKRteam / AkaiKKRprogram	ivate			-
<> Code ⊙ Issues 11 11 Pull requests	Actions Projects Security	urity 🗠 Insights 🛞 Settings		61
	🏌 master 👻 🥲 3 branches	s 🛇 O tags	Go to file Add file	• Code -
\wedge	👍 fuku-cmp Merge remote-tr	acking branch 'remotes/origin/develop_testrun'	ca7ff0e 15 days ago	3 54 commits
	akaikkr	util/test_script/exeutil.py iand resultuti.py are	added as show resul	15 days ag
	akaikkr_cnd	util/test_script/exeutil.py iand resultuti.py are	added as show resul	15 days ag
	akaikkr_public	The module structure of util/test_script is cha	anged. spin flipped cal	16 days ag
	akaikkr_sic	fmg related test is added.		16 days ag
	docs	Description of the j mode added		25 days ag
	img	update img		29 days ag
	util	from typing -> from collections import Order	redDict	15 days age
AkaiKK B	LICENCE.md	update LICENCE.md		25 days ag
	 MARK - CONTRACTOR - 10 			

- akaikkr(電子状態、磁気特性)
- akaikkr_cnd(伝導特性)
- akaikkr_sic(強相関系用)
- akaikkr_public(公開用)

	AkaiKK	RPythonUti	I	
AkaiKKRteam / AkaiKKRPythonUtil Public	c			
<> Code ⊙ Issues 3 11 Pull requests 1	⊙ Actions	fiki 🕕 Security 🗠 Insights 🛞 Settings		()
	₽ master - ₽ 1 branch	⊘ 0 tags	Go to file Add file	Code -
\wedge	AkaiKKRteam Merge pull n	request #1 from nim-hrkn/master	e20ec1d 26 days ago	🕲 5 commits
	demo	demo/akaikkr_cnd.small.devel/cnd_test/fi	nite_temperature.py is upd	26 days ago
	library	The contents are added.		26 days ago
	tests	The contents are added.		26 days ago
	🖿 util	The contents are added.		26 days ago
\ltimes	LICENSE.txt	The contents are added.		26 days ago
	D. action of the	The excitence and add		

- cifファイル => AkaiKKRインプット
- 結果の解析(DOS, スペクトル関数等)
- ・ 数値パラメーター、収束等の自動化
- ・ ワークフロー作成(aiida-akaikkr)

AkaiKKRを用いた自動ハイスループット材料計算ツール

- ・ 全電子KKRグリーン関数法に基づいた高速・高精度の自動計算(AkaiKKR)
- 構造・配置・スピン不規則系に対して適用可能(汎用の計算手法では困難)
- 転移温度や伝導率を高速で計算可能(汎用の計算手法では困難)

AkaiKKRと「富岳」を用いた磁気物性データの創出

福島(物性研)、赤井(物性研)、知京(NIMS)、木野(NIMS) プレスリリース:スーパーコンピュータ「富岳」による大規模物性データの自動創出 ~ 不規則系磁性材料におけるビッグデータの実現へ~

	対象となる物理量
\checkmark	磁気モーメント、磁化
\checkmark	強磁性転移温度
\checkmark	残留抵抗

38元素の中から4元素を組み合わせる (e.g, MnFePdW, AlSiCrMn, CrFeCoNi) BCC + FCC = 147,630個

「富岳」を用いて自動網羅計算を実行する

大規模物性データの解析

(<u>T.F.</u> et al., Phys. Rev. Mater. **6** 023802 (2022))

ハイスループット計算の結果:4元HEA

- 147,630個のうち99%以上の系を自動的に計算することに成功
- 富岳を用いた2,000ノード並列計算により一週間もかからず計算可能

頻出パターンマイニング

有名な例 (ビールと紙おむつ)

スーパーでの購買データを対象にの頻出パターン マイニングを適用し、ビールと紙おむつが同時に 買われていることが明らかにした。 ↓

隣同士に配置することで売り上げアップ!

頻出パターンマイニングによる区画特徴特定:磁性の解析

頻出パターンマイニングによる区画特徴特定:磁性の解析

トランザクションデータベースを作成

各トランザクションのアイテム

- A. 物質構成元素
 - element (elm) 1, elm2, elm3, elm4
- B. セルのID
 - A2、B4、C10等
- C. 各構成元素の局所磁化の向き
 - |m₁| > |m₂| > |m₃| > |m₄|に並び替え
 - 閾値m_{th}を定義し|m₁|からの相対磁化で
 強磁性・反強磁性を階層的に定義する

例1) |m₁| > |m₂| > |m₃| > m_{th} > |m₄|で、 m2 < 0、m3 > 0 の場合 FA、FAF、FAFN

例2)|m₁| > |m₂| > m_{th} > |m₃| > |m₄|で、 m2 < 0 の場合 FA、FAN

10×10のセルに区切ってデジタイズする

構築されトランザクションデータベースから 頻出パターンマイニング(飽和集合)を行う

LCM (Linear time Closed itemset Miner)、http://research.nii.ac.jp/~uno/codes.htm

頻出パターンマイニングによる区画特徴特定:磁性の解析

- 高磁化にはMnが必要、FeとCoは高キュリー温度に貢献する
- Mn-Rhのペアは磁化を増強させる
- V、CrはMn、Fe、Coと反強磁性的にカップリングし高磁化は望めない
- この解析結果は既にMRAMの配線材料探索に使用されている

電子状態からの電気抵抗率の理解は難しい…

HEAにおける電気抵抗法則の推定:回帰分析

説明変数 X として、各原子における周期表の特徴量と基礎物理量を用いる (これらの最小値、最大値、平均、標準偏差(合計40個)を利用)

BCC phase

Regression method	Features	R ² _{test}
Linear regression + CV	Х	0.754
Linear regression + CV	X, X ²	0.815
Linear regression + CV	X, X ² , X ³	0.840
Random forest + CV	Х	0.964
k-nearest neighbor method + CV	Х	0.941

FCC phase

Regression method	Features	R ² _{test}
Linear regression + CV	Х	0.766
Linear regression + CV	X, X ²	0.843
Linear regression + CV	X, X ² , X ³	0.868
Random forest + CV	Х	0.959
k-nearest neighbor method + CV	Х	0.939

重要な説明変数の評価

物理では「説明可能」AIが重要

Permutation importance (説明変数のランダム化)

R²の減少から説明変数の重要性を評価する

- ・ 説明変数として周期表の族(group)が重要
- 説明変数に<u>族(group)と周期(period)</u>を用いて 全探索を行い、有効モデルを構築する
- 有効モデルに立脚し電気抵抗法則性を導く

max:最大值、min:最小值、mean:平均、std:標準偏差

BCC phase, k-NN method

R ² _{test}	n	Explanatory variables
0.5845	1	['group_std']
0.773	2	['group_std', 'group_mean']
0.826	3	['group_std', 'group_mean',
0.020	Ũ	'melting_point_mean']
0.898	4	['group_std', 'group_mean',
0.000		'melting_point_mean', 'melting_point_max']
0.914 5	5	['group_std', 'group_mean', 'group_max',
	Ũ	'melting_point_mean', 'melting_point_max']
0.918 6	6	['group_std', 'group_mean', 'group_min',
		'melting_point_mean', 'group_max',
		'melting_point_max']
		['group_mean', 'group_std',
0.927	7	'melting_point_mean', 'group_min', 'group_max',
		'melting_point_max', 'melting_point_std']
	8	['group_mean', 'group_std',
0 928		'melting_point_mean', 'group_min', 'group_max',
0.920 8		'melting_point_max', 'melting_point_std',
		'molar_volume_mean']
0.931		['group_mean', 'group_std',
	9	'melting_point_mean', 'group_min', 'group_max',
		'melting_point_max', 'melting_point_std',
		'molar_volume_mean', 'molar_volume_std']
		['group_mean', 'melting_point_mean',
	7 10	'group_std', 'group_min', 'group_max',
0.937		'molar_volume_mean', 'melting_point_std',
		'atomic_radius_calculated_std',
		'melting_point_max', 'molar_volume_std']

R²の頻出パターンマイニング

説明変数として"group"と"period"のみを用いて全探索を行い、 トランザクションを作る [R²,features]

R²をデジタイズして「R², 説明変数」のトランザクションデータベースを構築する
 R²の連続した非零領域毎に頻出パターンマイニングを適用する

R² > 0.80の領域は、常に説明変数'group_mean'、'group_std'を含む

<u>高エントロピー合金(XYZW)における電気抵抗法則の推定</u>

各4つのパネルは簡易的な周期表 (周期は3~6、族は3~15)を表す

○の大きさは<u>XYZ</u>を固定し、Wに対応する 原子を変化させた時の電気抵抗率を示す

高エントロピー合金(XYZW)における電気抵抗法則の推定

もっと効率よくターゲット材料を探索できないだろうか?

全データを持っているので多目的ベイズ最適化のシミュレーションが可能

ガウス過程は訓練データ点に確率分布を置いていくという考え方で、訓練データ 点では誤差を小さく、訓練データが無い点では誤差を大きくする予測ができる (平均値+標準偏差を予測可能)

ベイズ最適化による探索加速

ベイズ最適化はガウス過程が予測標準偏差が求まることを用いて獲得関数を 通して最適な記述子、ブラックボックス関数の最大(小)値を探索する手法

- 1. 既知の候補点(*x_i*,*y*_{expr,*i*})が存在する。(*i*=1,...,*n*)
- 2. ガウス過程により未探索点 $j o y_{mean,j}$ とその標準偏差 σ_j を予測する
- 3. 最も高い獲得関数スコアの点を評価しy_{expr,k}を得て、既知の候補点セットに入れる。

ベイズ最適化のシミュレーション

新・高機能材料の探索やデザインは難しい…

2 マテリアルデザインとマテリアルズインフォマティクス

3 磁性材料のための基盤的シミュレーションコードの開発(AkaiKKR)

4 データリポジトリ、大型施設(スーパーコンピューター「富岳」)との連携

自動網羅計算による磁気物性データの創出とその活用(不規則系)