

AI for Scienceとデータ駆動科学

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

内容

- 自己紹介
- AIの歴史
 - AI for Game: ゲームソフトの開発と終焉
 - 2016年のAlpha Goの登場
 - Demis HassabisによるAI for Science の提案
- AI for Scienceとデータ駆動科学
- AI for Material science
 - 機能発現の3+1ステップモデル
- AI for ScienceとLLM
- AI for Scienceの国内での動向
- まとめ

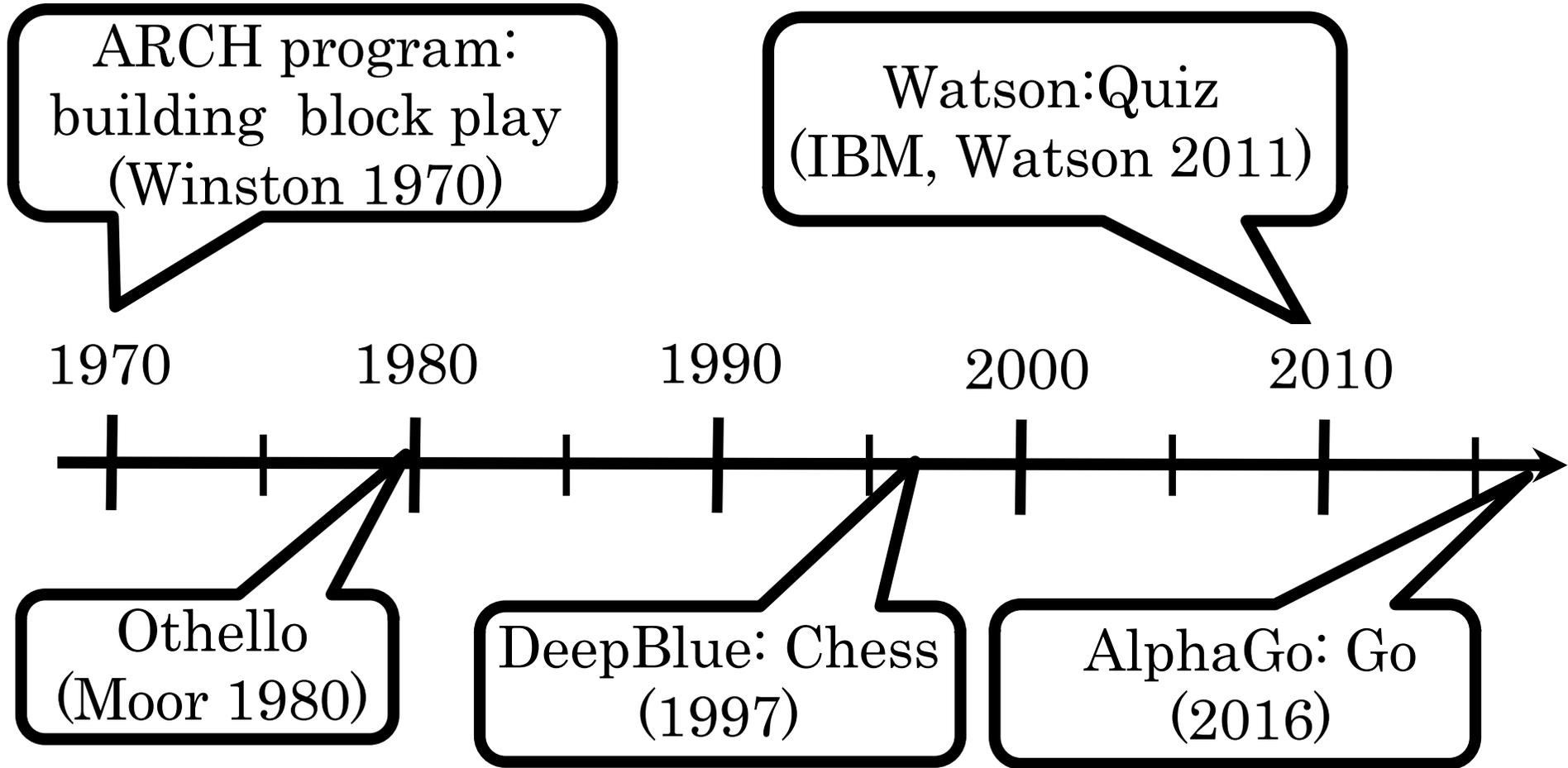
自己紹介(理論物理学)

- 大阪市立大学理学部物理学科 (1981 - 1985)
 - アモルファスシリコンの成長と構造解析
- 大阪大学大学院理学研究科(金森研) (1985 - 1987)
 - 希土類元素の光励起スペクトルの理論
- 三菱電機 北伊丹製作所(量産エンジニア) (1987 - 1989)
 - 化合物半導体(半導体レーザー)のエピタキシャル結晶成長
- 大阪大学大学院基礎工学研究科生物工学福島研 (1989 - 1996)
 - ニューラルネットワーク
 - 福島先生は畳み込み深層ニューラルネットワークの提案者
- JST ERATO 川人学習動態脳プロジェクト (1996 - 2001)
 - 計算論的神経科学
- 理化学研究所 脳科学総合研究センター 甘利チーム (2001 - 04/06)
 - 情報統計力学
 - ベイズ推論, 機械学習, データ駆動科学
- 東京大学・大学院新領域創成科学研究科 複雑理工学専攻
 - データ駆動科学、ベイズ推論、スパースモデリング (2004/07 -)

内容

- 自己紹介
- AIの歴史
 - AI for Game: ゲームソフトの開発と終焉
 - 2016年のAlpha Goの登場
 - Demis HassabisによるAI for Science の提案
- AI for Scienceとデータ駆動科学
- AI for Material science
 - 機能発現の3+1ステップモデル
- AI for ScienceとLLM
- AI for Scienceの国内での動向
- まとめ

AI for Gameの展開と終焉



Masato Okada, "2024 Nobel Prize in Physics and Chemistry: From neural network models to materials engineering" Science and Technology of Advanced Materials: Methods, accepted, (2025)

内容

- 自己紹介
- AIの歴史
 - AI for Game: ゲームソフトの開発と終焉
 - 2016年のAlpha Goの登場
 - Demis HassabisによるAI for Science の提案
- AI for Scienceとデータ駆動科学
- AI for Material science
 - 機能発現の3+1ステップモデル
- AI for ScienceとLLM
- AI for Scienceの国内での動向
- まとめ

2024ノーベル化学賞受賞

Demis HassabisによるAI for Science

- I **don't think** much about **robotics** myself personally.
- What I'm really excited to use this kind of **AI for is science**, and advancing that faster.
- I was giving a talk at **CERN** a few months ago.
- I think it'd be cool if one day an AI was involved in **finding a new particle**.

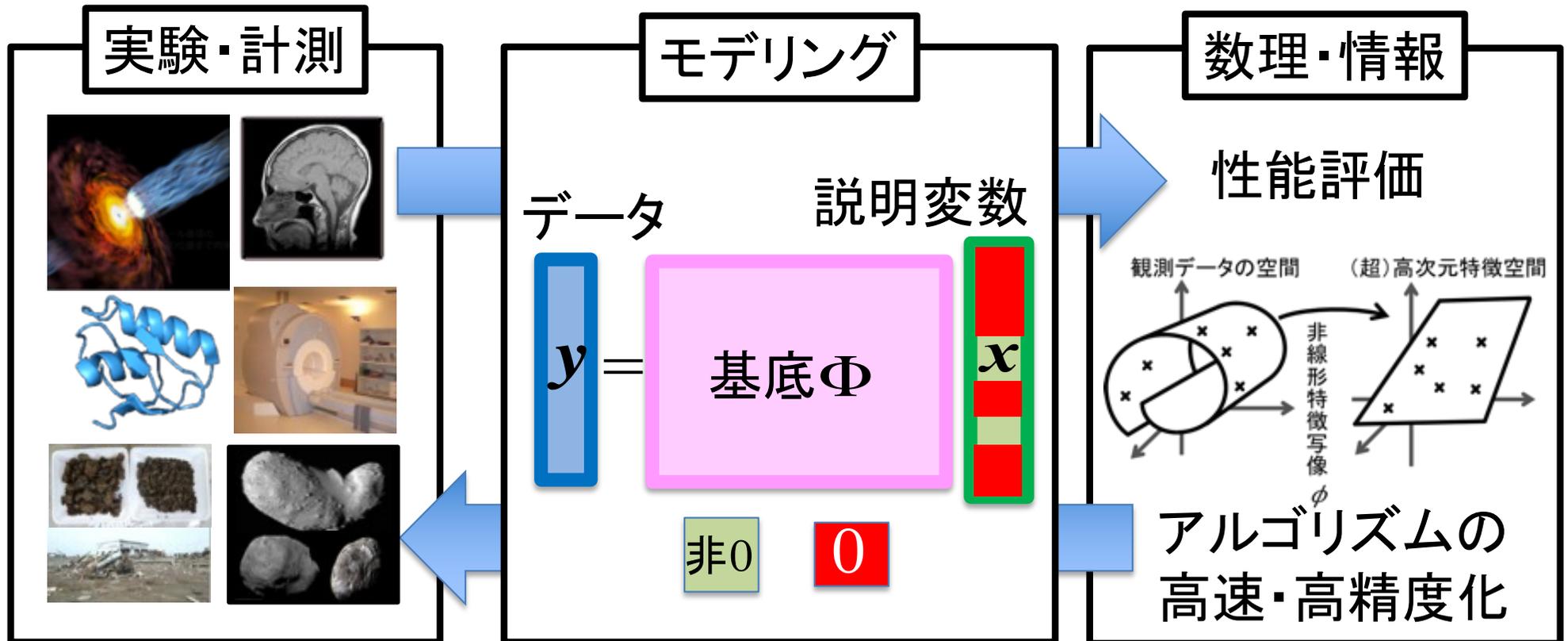
<http://www.theverge.com/2016/3/10/11192774/demis-hassabis-interview-alphago-google-deepmind-ai>

内容

- 自己紹介
- AIの歴史
 - AI for Game: ゲームソフトの開発と終焉
 - 2016年のAlpha Goの登場
 - Demis HassabisによるAI for Science の提案
- AI for Scienceとデータ駆動科学
- AI for Material science
 - 機能発現の3+1ステップモデル
- AI for ScienceとLLM
- AI for Scienceの国内での動向
- まとめ

新学術領域研究 平成25～29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

領域代表岡田真人の個人的な狙い
世界を系統的に記述したい
その方法論と枠組みを創りたい
ヒトが世界を認識するとは？



領域の目的と戦略

目的: 高次元データ駆動科学の創成

大量の**高次元データ**から**仮説(モデル)**を**系統的に導く方法論**を「**生物**」, 「**地学**」分野に確立し, それを**実践するための研究体制のコア**を我が国に形成する.

3つの戦略

1. **スパースモデリング**に重点投資

今後5年で飛躍的發展が確実視される枠組み

2. 分野の壁を取り去り, 知識伝播を飛躍的に加速

分野をまたぐモデルの構造的類似性を明確化

3. 実験家と理論家との有機的協働

仮説の提案／検証ループを効率的に稼働させる体制

データ駆動科学とは

- 機械学習などのAIを使い，各学問分野の問題を解いていくというアプローチ
- 実験/計測/計算データの背後にある潜在的構造の抽出に関して，データが対象とする学問に依存しない普遍的な学問体系
- 同じアルゴリズムがスケールや対象を超えて，有用であることが多いという経験的事実を背景として，その理由を問い，背後にある普遍性から，データ解析自体を学問的対象とする枠組み.
- 全ての実験/計測のデータ解析をデータ駆動学の三つのレベルの鑄型に押し込んで考える

データ駆動科学の三つのレベル (2016)

計算理論(対象の科学, 計測科学)

データ解析の目的とその適切性を議論し, 実行可能な方法の論理(方略)を構築

モデリング(統計学, 理論物理学, 数理科学)

計算理論のレベルの目的, 適切さ, 方略を元に, 系をモデル化し, 計算理論を数学的に表現する

表現・アルゴリズム(統計学, 機械学習, 計算科学)

モデリングの結果得られた計算問題を, 実行するためのアルゴリズムを議論する.

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

連立方程式の応用と データ駆動科学の三つのレベル

データ駆動科学の
三つのレベル

計算理論

表現・モデリング

アルゴリズム

連立方程式とその応用

鶴亀算 食塩水 など

連立方程式への変換

加減法, 代入法

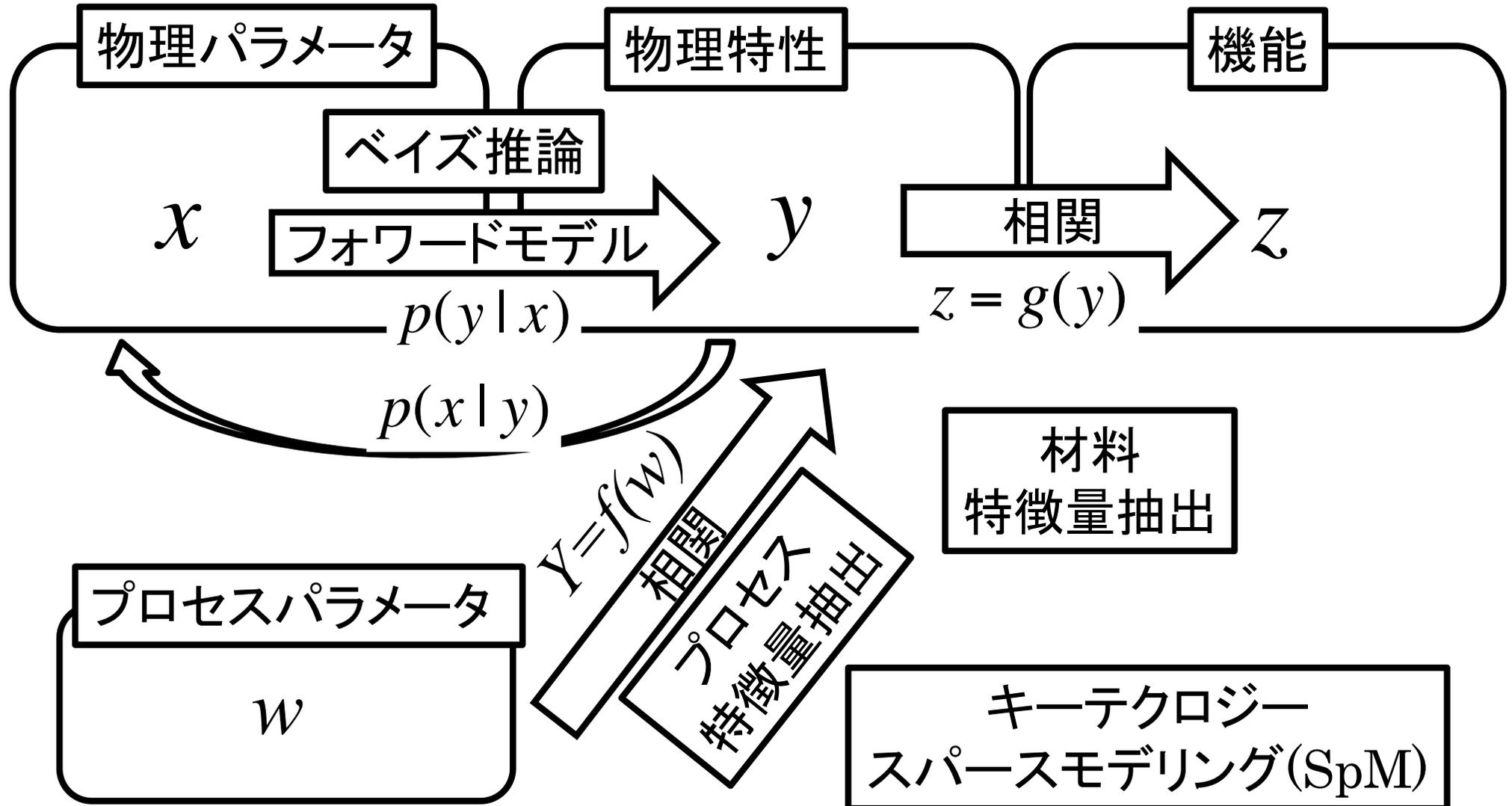
(五十嵐, 竹中, 永田, 岡田, *応用統計学*, 2016)

内容

- 自己紹介
- AIの歴史
 - AI for Game: ゲームソフトの開発と終焉
 - 2016年のAlpha Goの登場
 - Demis HassabisによるAI for Science の提案
- AI for Scienceとデータ駆動科学
- AI for Material science
 - 機能発現の3+1ステップモデル
- AI for ScienceとLLM
- AI for Scienceの国内での動向
- まとめ

AI for Material Science

機能発現3+1モデル (1/2)



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)

AI for Material Science

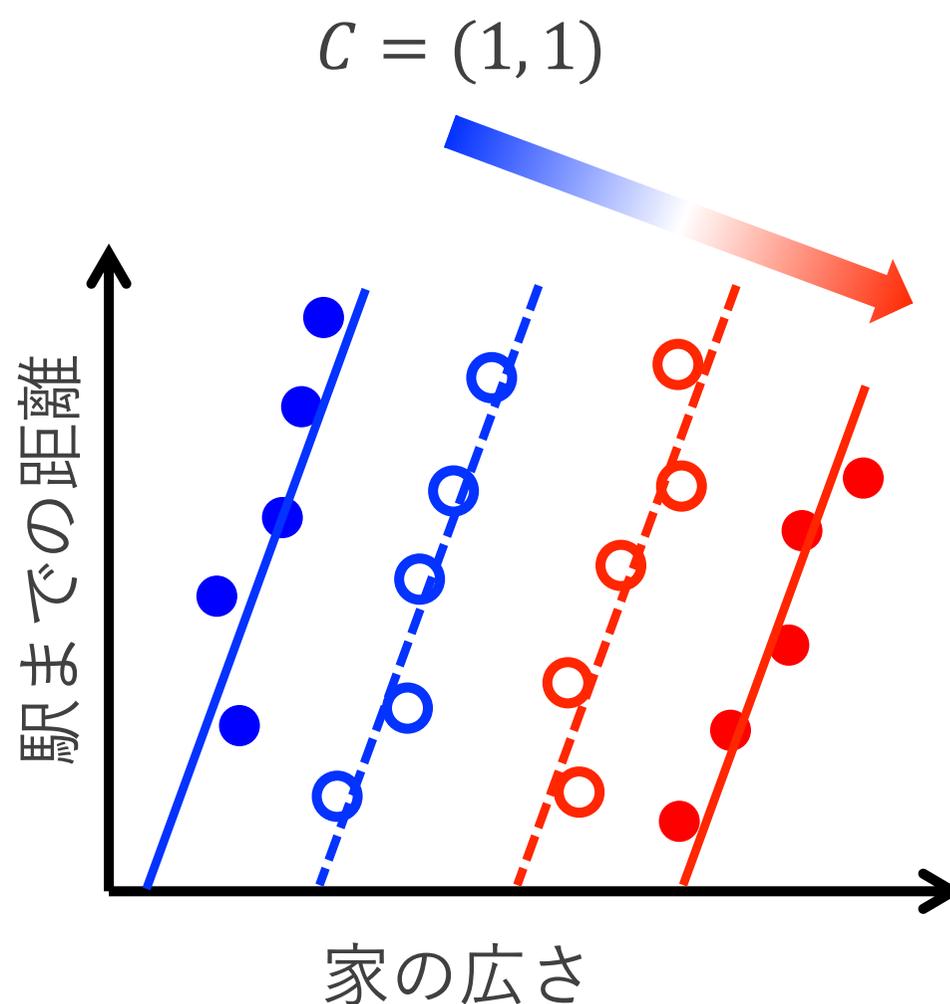
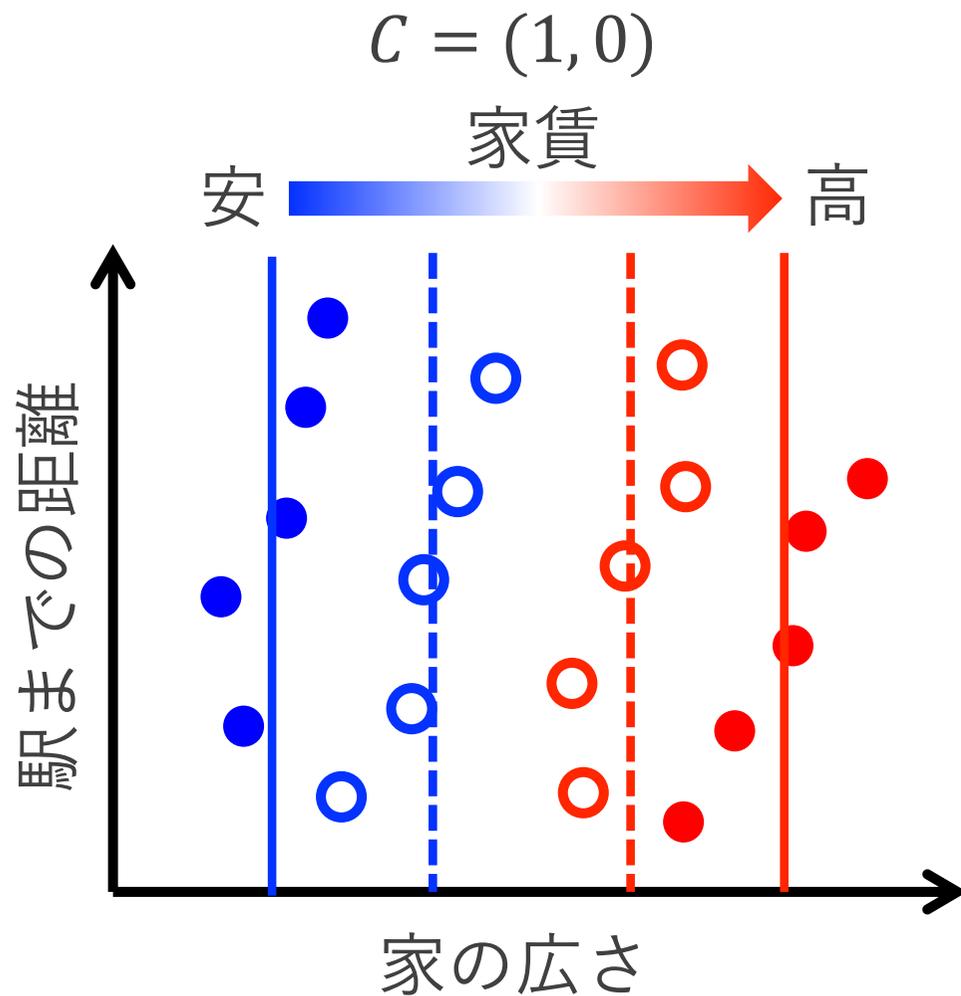
機能発現3+1モデル (2/2)

- 知りたいことを、既知の学問的知見を用いて、モジュール構造で表現
- 各モジュールでのデータ解析は、ベイズ推論とスパースモデリングの二つのみ
- ベイズ推論とSpMに深層ネットワークを組み込むことは可能
- 各モジュールのデータ解析は、データ駆動科学の三つのレベルで階層的に取り扱う

内容

- 自己紹介
- AIの歴史
 - AI for Game: ゲームソフトの開発と終焉
 - 2016年のAlpha Goの登場
 - Demis HassabisによるAI for Science の提案
- AI for Scienceとデータ駆動科学
- AI for Material science
 - 機能発現の3+1ステップモデル
- AI for ScienceとLLM
- AI for Scienceの国内での動向
- まとめ

スパースモデリング (1/2)

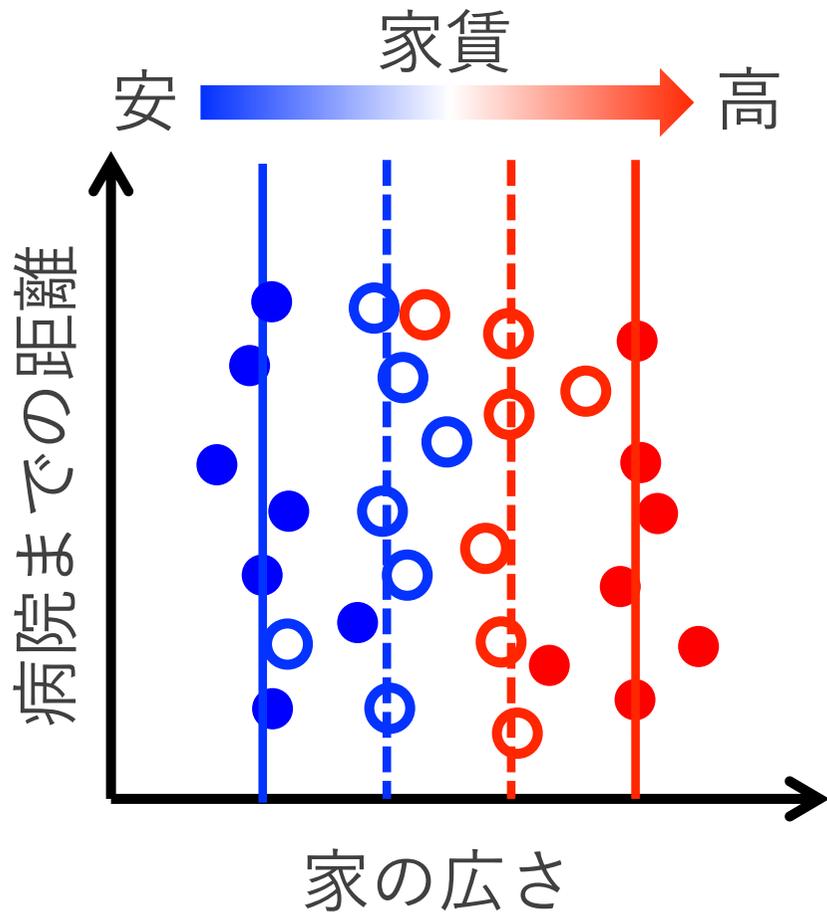


$Error(1, 0) > Error(1, 1)$

- 両方とも予測に必要

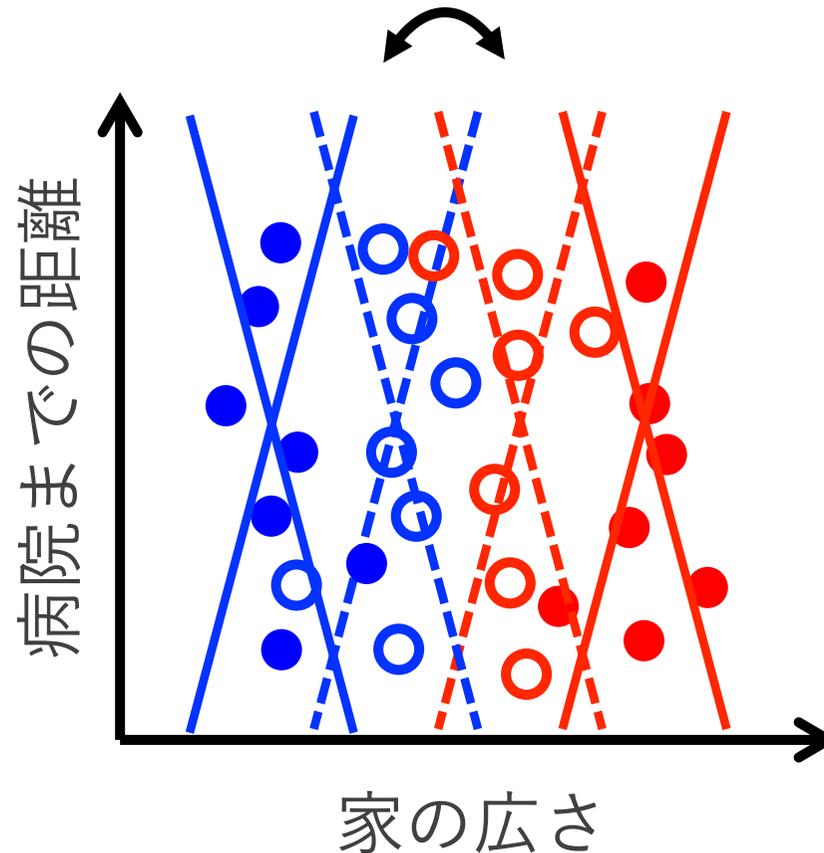
スパースモデリング (2/2)

家の広さのみ $C = (1, 0)$



$Error(1, 0) < Error(1, 1)$

家の広さと病院までの距離
 $C = (1, 1)$



- 家の広さ : 必要
- 病院までの距離 : 不要

スパースモデリングとLLM

- スパースモデリングの鍵は、特徴量をどう選ぶかである。
- 通常は、各研究者が個人の経験に基づいて、特徴量を提案している。
- ここを、LLMを用いて全論文を検索して、関連する特徴量を全て列挙して、スパースモデリングする。
- 特徴量が莫大になると計算量爆発を起こすが、その際は、愚直な全状態探索ではなく、レプリカ交換モンテカルロ法で近似的に全状態探索を行う。

内容

- 自己紹介
- AIの歴史
 - AI for Game: ゲームソフトの開発と終焉
 - 2016年のAlpha Goの登場
 - Demis HassabisによるAI for Science の提案
- AI for Scienceとデータ駆動科学
- AI for Material science
 - 機能発現の3+1ステップモデル
- AI for ScienceとLLM
- AI for Scienceの国内での動向
- まとめ

AI for Scienceの国内での展開

- NIMS(物質・材料研究機構: 文科省): AI for Material engineering
- NICT(情報通信研究機構: 総務省): 協創的AI研究促進ファンド
 - NICT内部のファンド
- DX-GEM(再生可能エネルギー最大導入に向けた電気化学材料研究拠点: 文科省)

内容

- 自己紹介
- AIの歴史
 - AI for Game: ゲームソフトの開発と終焉
 - 2016年のAlpha Goの登場
 - Demis HassabisによるAI for Science の提案
- AI for Scienceとデータ駆動科学
- AI for Material science
 - 機能発現の3+1ステップモデル
- AI for ScienceとLLM
- AI for Scienceの国内での動向
- まとめ

まとめ

- AI for Gameの展開と終焉
- AI for Scienceの登場
 - AIとScienceの双方向的相互作用
- AI for Scienceのロールモデル: AI for Material engineering
- AI for Scienceの今まで
 - 機能発現の3ステップモデル
- AI for Scienceのこれから
 - LLMの機能発現の3ステップモデルへのグランディング