



小特集 プラズマ・インフォマティクス—データ駆動科学のプラズマへの応用

4. トカマクプラズマ運転への応用展開

4. Application Development for Tokamak Plasma Operation

若月琢馬, 横山達也¹⁾, 大山直幸, 山田弘司^{1,2)}

WAKATSUKI Takuma, YOKOYAMA Tatsuya¹⁾, OYAMA Naoyuki and YAMADA Hiroshi^{1,2)}

量子科学技術研究開発機構, ¹⁾東京大学大学院新領域創成科学研究科, ²⁾自然科学研究機構 核融合科学研究所

(原稿受付: 2019年6月13日)

トカマクプラズマ運転への応用について, 異なる源流と思想を持つニューラルネットワークと統計的機械学習によるスパースモデリングそれぞれに基づく具体的な事例を紹介し, この方面への機械学習のアプローチの動向と展開を議論する. ニューラルネットワークの応用例は内部輸送障壁を伴うプラズマのイオン温度勾配の制御, スパースモデリングの応用例はディスラプション予知を課題としたものである. これらは輸送特性の変化や破壊的突発現象という非線形性が際立った端的な例であり, いずれも, 要素還元された物理の時間発展微分方程式で表現するようなモデルによって精確に予測し, 制御することは現状として難しい. このため, 実験およびシミュレーションからのデータと機械学習の手法を用いてその解決をめざすデータ駆動型アプローチが近年注目されている. 世界的に実機への応用は未だ途上であり, 様々な試みがなされる中, ここで紹介する2例はトカマクプラズマ運転の研究開発に対して, 適応的制御と仮説の提供という大きな展開の可能性を持っている.

Keywords:

neural network, deep learning, internal transport barrier, sparse modeling, support vector system, disruption predictor

4.1 トカマクプラズマ運転におけるデータ駆動型アプローチ

本小特集の背景として既に論じられているように, 人口知能や機械学習の物理への応用が盛んになっている[1, 2]. プラズマは生物ほどではないにしろ, 複雑系の典型であり, 要素還元された物理モデルによってシステム全体の挙動を理解, すなわち予測し, 制御することはできるであろうか? 特に本章で議論するトカマクプラズマにあっては, 国際熱核融合実験炉 ITER の実験開始や原型炉建設を見据えた時間軸に沿って, このような対象に対して要素還元的なアプローチと相補的・相乗的にデータ駆動型アプローチによって研究開発の加速を図ることは理に適ったものと言える.

本章では, データ駆動型アプローチの方法論として, 機械学習の中でも異なる源流を持つ, ニューラルネットワークとスパースモデリングの具体的な応用例を紹介する.

強化学習にニューラルネットワークを利用した技術は深層強化学習と呼ばれ, 近年, Alpha GO[3]に代表されるようにゲーム[4]やロボット制御[5]を対象にして目覚ましい発展を遂げている. ここでは, 内部輸送障壁を伴うプラズマのイオン温度勾配を加熱パワーで制御するシステムの, 輸送シミュレーションによるデータを元とする学習を通じた開発が述べられる.

一方, 統計的機械学習の流れにあるスパースモデリング[6]とは, 様々な分野に高次元データに共通して潜在する

スパース (疎) 性を利用し, 効率的に最大限の情報を抽出することをめざすものである. いわゆる「オッカムの剃刀」を合理的に手にする方法論と言える. ここでは, ディラプションの有無という分類を未知のデータに対して精度良く予測して行う分類器の開発が述べられる. そこでは, 特に2クラス分類問題の手法であるサポートベクトルマシン (Support Vector Machine) [7]が利用される.

4.2 イオン温度勾配制御への応用

4.2.1 内部輸送障壁を伴うプラズマの制御

実験やシミュレーションで得られるデータに基づく学習を制御システムに利用する一例として, イオン温度勾配制御への応用が検討されている.

JT-60U では高性能プラズマの定常維持をめざして, 中性粒子ビーム入射 (NBI) を用いた内部輸送障壁 (ITB) を持つプラズマのイオン温度勾配の実時間制御実験が行われた[8]. ITB プラズマではイオン温度勾配の応答が非線形性を示すため, 比例・積分・微分 (PID) 制御を用いながらも, 比例ゲインをプラズマ中心のイオン温度に依存する可変の値とすることで実時間制御を実現したが, 可変ゲインのイオン温度依存性を決定するために, 事前に NBI に対するイオン温度分布の応答特性を調べるための放電を行う必要があった.

しかしながら, ITB プラズマにおけるイオン温度分布の応答特性は, 第一壁のリサイクリング量など実験条件の違

いによって大きく変化するため、幅広い運転条件に適用できるフィードバック制御を実現する可変ゲインを決定するのは困難である。また、長時間放電時にリサイクリング量が時間とともに変化することによって、ITBの特性が動的に変化してしまうような場合にも適切な制御を行うことは難しい。そこで、データ駆動型アプローチとして、統合輸送コードを用いてイオン温度分布の応答特性を様々に変化したプラズマに対する制御シミュレーションのデータを作成し、それらに対して学習を行うことで、幅広い運転条件に適応して適切な制御を行うことができるシステムを構築することをめざした。

4.2.2 深層強化学習のアルゴリズム

ここでは、学習したシステムの動作をJT-60Uの実験データと比較して検証するため、実験と同様に荷電交換分光計測(CXRS)による2つの計測点間のイオン温度差 ΔT_i の制御を通じてイオン温度勾配を制御することを考える。ITBプラズマのMHD不安定性を回避することをめざし、2つの計測点は規格化小半径 $\rho \sim 0.4$ に存在する $q=2$ の有理面を挟んだ2点($\rho \sim 0.3, 0.5$)とした。

イオン温度差 ΔT_i 制御のための最適な加熱NBIパワー P_{NB} はプラズマの輸送特性、特にイオンの熱拡散係数 χ_i の空間分布に依存する。しかし、ITBプラズマの χ_i の時空間発展を実験前に正確に予想することは難しく、適切な P_{NB} の波形を事前に決定することは困難である。一方で、 P_{NB} に対する ΔT_i の応答特性は現在から過去一定の時間の P_{NB} と ΔT_i の計測データから予測できると期待できる。そこで、以下の状態ベクトル \mathbf{s}_j

$$\mathbf{s}_j = [P_{NB}(t_j), P_{NB}(t_{j-1}), \dots, P_{NB}(t_{j-k}), \Delta T_i(t_j), \Delta T_i(t_{j-1}), \dots, \Delta T_i(t_{j-k}), \Delta T_i^{\text{target}}(t_{j+1})] \quad (1)$$

を入力し、最適な P_{NB} を出力できるニューラルネットワークを学習することとした。ここで、 t_j は j 番目の制御ステップの時刻、 $\Delta T_i^{\text{target}}$ は ΔT_i の制御目標である。

学習のためには、サンプルとなる制御シミュレーションのデータを作成する必要がある。NBIの入力パワーは使用するユニット数によって決定され、JT-60Uでは7つの離散値を取る。そのため、100ステップ程度の制御周期分のシミュレーションデータを網羅的に作成しようとした場合、7の100乗回のシミュレーションを行う必要となってしまう、現実的ではない。そこで、制御性能の向上に効果的な学習サンプルを選択的に収集することで、少ない学習サンプルから効率よく制御システムの学習を行う必要がある。

試行錯誤を通じて制御システムを訓練する試みの一つに強化学習がある。強化学習の枠組みでは、学習に用いるサンプルの収集と制御システムの訓練を同時に行い、学習の各時点で制御システムが最適だと予想する解の近傍のサンプルを重点的に収集する。これにより、少ないサンプル数(試行回数)で効率よく制御システムの性能を向上することが可能になる。

強化学習では学習のためにreward(報酬)と呼ばれる量

が定義される。これは、制御システムの各制御ステップにおける制御の“よさ”を表す量であり、この量を最大化するように制御システムは学習される。そこで、 ΔT_i 制御を行うシステムの学習のために、各制御ステップの ΔT_i の制御誤差が小さくなるにつれて大きくなるようなrewardを設定した。ここで、ある制御ステップにおけるrewardを r_j 、制御出力ベクトル(ここでは P_{NB} のみを成分とする1次元ベクトル)を \mathbf{a}_j とすると、強化学習の問題設定は、ある状態ベクトル \mathbf{s}_j が与えられたとき、それ以降の制御ステップで得られるrewardの和 $R_j = \sum_k r_{j+k}$ を最大化するように制御出力ベクトル \mathbf{a}_j を最適化することと定義できる。

強化学習にニューラルネットワーク、特に従来では学習の難しかった多層ニューラルネットワークを利用した技術は深層強化学習と呼ばれ、目覚ましい発展を遂げている。深層強化学習のアルゴリズムは年々開発が進んでおり様々な種類が存在するが、ここでは ΔT_i 制御システムの学習に用いたActor-Critic法と呼ばれるアルゴリズム[9-11]の概要を説明する(図1参照)。Actor-Critic法では状態ベクトル \mathbf{s}_j に対する最適な制御出力ベクトル \mathbf{a}_j の関係を表す方策関数 $\pi(\mathbf{s}_j)$ と、状態ベクトルが \mathbf{s}_j である時点で、制御出力ベクトル \mathbf{a}_j が選択されたときの、その後の報酬和 R_j を計算する行動価値関数 $Q^\pi(\mathbf{s}_j, \mathbf{a}_j)$ を考え、これらをそれぞれニューラルネットワークで近似する。ランダムに初期化された方策関数 $\pi(\mathbf{s}_j)$ に基づいて制御シミュレーションを行うことから始まり、試行錯誤を通じて得られる $\mathbf{s}_j, \mathbf{a}_j, r_j$ の組み合わせから $Q^\pi(\mathbf{s}_j, \mathbf{a}_j)$ の予測誤差を最小化するようにニューラルネットワークを訓練する。そしてそれと同時に、 $Q^\pi(\mathbf{s}_j, \mathbf{a}_j)$ の予測に基づいて、得られる報酬和が多くなる制御出力ベクトル \mathbf{a}_j が選ばれる確率が上がるように方策関数 $\pi(\mathbf{s}_j)$ のニューラルネットワークのパラメータを更新する。この繰り返しにより、最適解周辺の報酬和の予測精度の改善とそれに基づく方策関数の改善を通じ、効率的な制御システムの最適化を達成することができる。以上が使用した深層強化学習アルゴリズムの概要である。

なお、深層強化学習のアルゴリズムの特性上、必要とする制御特性を獲得するためにはrewardの適切な設計が重要である。例えば、ニューラルネットワークの学習を安定に行うためにはrewardの値を0から1の間の値にtanh

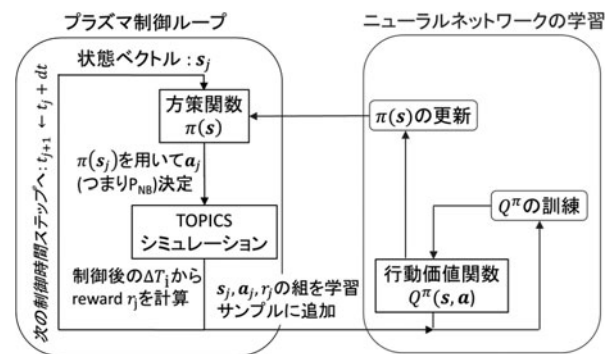


図1 Actor-Critic法を用いたイオン温度差 ΔT_i 制御の深層強化学習のフローチャート。

関数などを介して規格化するのが一般的であるが、規格化のスケールを小さくとりすぎると、初期のランダムな方策では reward が有意に大きくなるケースに遭遇する確率が非常に少なくなってしまい、学習が進まなくなる。一方で規格化のスケールを大きくとりすぎると、大きな制御誤差が残った状態でも reward が頭打ちとなり、それ以上学習が進まなくなる。そこで、最終的に得たい制御成績と学習のためにかける計算時間との兼ね合いから、適切な規格化のスケールを選択する必要があった。

4.2.3 シミュレーションを用いた学習と結果

ここからは、実際に学習したイオン温度差 ΔT_i の制御シミュレーションについて説明する。シミュレーションは統合輸送コード TOPICS[12]を用いて行った。NBIの入射位置やエネルギー、プラズマ形状、プラズマ電流分布、そして電子密度分布はJT-60Uの実験データを使用し、イオン温度分布と電子温度分布の時間発展を計算した。1回のシミュレーションで2.1秒間の温度分布の時間発展を計算し、その間0.02秒周期で P_{NB} 制御を行う。幅広い輸送特性のプラズマに対して学習するため、弱いITBプラズマから強いITBプラズマまで、4段階の勾配強度についての熱輸送モデルをシミュレーション開始時にランダムに採用してシミュレーションを行った。数千回程度のシミュレーション結果を学習することにより、4種類すべての熱輸送モデルに対して適切な ΔT_i 制御ができるようになった。これにより、学習した制御システムでは与えられた状態ベクトル s_j から、ITBの強さに依存して変化する ΔT_i の応答の強さを推定し、その応答特性に適応した適切な制御を行うことができるようになることがわかった。

しかし、一般にデータに対する学習で得られた最適化システムは、学習に使ったサンプルに対して高い性能を示したとしても、学習サンプルとは異なる新規のデータに対して望ましい結果を得られない、いわゆる過学習の可能性がある。そこで、訓練された制御システムが実際の実験に適用可能であるかどうかJT-60Uの実験データを用いて検証した。実験データから、ある P_{NB} の波形とその時の ΔT_i の時間発展が得られるが、制御のシミュレーションでは実験時とは異なる P_{NB} 波形となった時の ΔT_i の応答を予測する必要がある。そこで、JT-60U実験のイオン温度、電子温度分布の時間発展を再現する熱輸送モデル（実験データに基づく熱輸送モデル）を構築し、そのモデルに対する ΔT_i 制御のシミュレーションを行った。その結果、図2に示す通り、学習に使ったモデルとは異なる、実験データに基づく輸送モデルに対しても適切な ΔT_i 制御が行えることがわかった。また、示している2つの結果は、トロイダル磁場を変更した影響でITBの強さが大きく異なる2つの放電を模したモデルに対する制御結果であり、実験で現れる幅広い応答特性に適応して制御が行える可能性があることを示す結果である。

4.2.4 イオン温度勾配制御への応用 -まとめ-

応答特性の予測が困難な内部輸送障壁を伴うトカマクプラズマの制御への強化学習の応用を紹介した。統合輸送シミュレーションを用い、加熱パワーに対するイオン温度勾

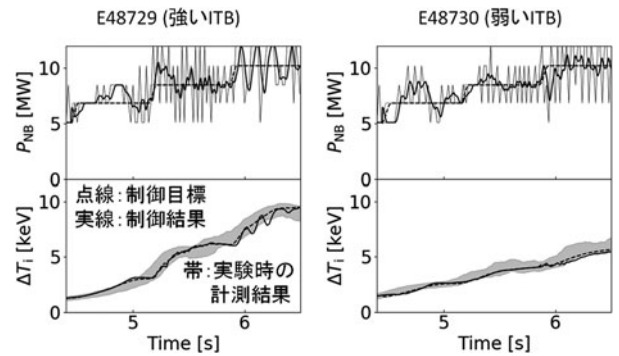


図2 JT-60Uの実験データに基づく輸送モデルに対して、強化学習を用いて訓練した ΔT_i 制御システムの動作を検証した結果。点線は実験と同じ P_{NB} 入射波形での計算結果を示す。この ΔT_i が実験時の計測結果を再現するようにそれぞれの放電に対して輸送モデルが調整されている。この ΔT_i を制御目標として制御を行った結果が実線で示されており、ITBの強さの異なる2放電を模したシミュレーションにおいて、十分よい制御を行えることが示されている。

配の多様な応答特性に対して強化学習を行うことによって制御システムを構築した。このシステムは実際に実験で現れる内部輸送障壁の応答特性を模したシミュレーションに対しても適切な制御を実現できるものであり、今後、様々な条件に適応した制御を可能とするシステムへ展開できると期待される。原型炉ではより広範なパラメータ範囲におけるロバストな制御システム、すなわち適応的な制御が必須である。強化学習を用いた制御システムの最も大きな特長はこの適応的な制御を実現できる、という点にある。

4.3 スパースモデリングによるディスラプション予知

4.3.1 機械学習を用いたディスラプション研究

ディスラプション現象は、トカマク方式の核融合プラズマに特有の突発的な崩壊現象である。ディスラプションの回避・緩和・予知はトカマク方式の核融合炉を実現するために克服しなくてはならない最重要課題の1つであるが、その物理的背景は未だ完全には解明されていない[13, 14]。

中でもディスラプションの発生を予知することをめざす研究が、JET, DIII-Dなどの実験装置で行われている[15-17]。これらの研究ではニューラルネットワークやサポートベクトルマシン(SVM)といった機械学習モデルに放電中のプラズマパラメータの値を入力し、放電がディスラプションを起こすか否かを判定するという問題設定がよくされる。このような研究では、モデルに入力するプラズマパラメータは各々の研究者の物理考察に基づいて選択されることが多い。一方、ディスラプション現象に関連するパラメータを偏りなく、適切かつ合理的に選択することができれば、予知の性能はさらに向上できることが期待される。本セクションでは、機械学習を用いたディスラプション予知研究に全状態探索を用いたスパースモデリングを導入し、入力パラメータの選択を実施した例を紹介する。

4.3.2 全状態探索を用いたスパースモデリング

分類問題における変数選択では、単に個々の変数の分布

だけから検討するのではなく、変数同士の組合せによる効果を十分に考慮する必要がある。そこで、実現可能なすべての組合せについてモデルの評価と訓練を行って最適な変数の組合せを求める、全状態探索 (Exhaustive search, ES) の手法が考案された[18]。しかし ES には、パラメータの数 N が増えると変数の組合せの数が指数関数的に増えるという問題がある。

そこで、この問題を解決するために K -スパース全状態探索 (ES- K) という拡張を用いる。ES- K では、変数の最適な組合せは K 個のパラメータからなる、すなわち K -スパースであると仮定して、考えられる ${}_N C_K$ 通りのすべての組合せを網羅的に調べ、最適な組合せを求める。

最適なパラメータの組合せを求めるために、それぞれの組合せで得られた予知器モデルの未知のデータに対する分類性能である汎化性能を比較することが必要である。ここで用いられるのが交差検証法である。交差検証法では、利用可能なデータを訓練データと評価データに分割し、訓練データを用いて訓練したモデルの評価データに対する分類性能を汎化性能とする。ここではこのプロセスを10回実施して結果を平均する10分割交差検証法を使用して、訓練データと評価データの分割によるデータの偏りの影響を除外している。

4.3.3 ディスラプション予知器モデルの作成

1) データセットの作成

ここでは JT-60U で行われた高ベータプラズマ実験中のディスラプションを対象とした。この実験キャンペーンは、プラズマ回転を利用して抵抗性壁モードの成長を抑制し、高いベータ値のプラズマの生成をめざしたものである。ディスラプション発生前における規格化ベータ値 β_N は、キンク不安定性が生じる目安である壁なし限界ベータ $\beta_{n}^{no-wall} \sim 3I_i$ を超えるものが多い[19]。

この実験キャンペーン中の放電データを用いて機械学習モデルを訓練・評価するためのデータセットを作成した。データセットには、ディスラプションに関する10種類の巨視的なパラメータ ($I_p, \beta_N, \beta_p, I_i, q_{95}, \delta, \kappa, |B_r^n|, f_{GW} = \bar{n}_e/n_{GW}, f_{rad} = P_{rad}/P_{input}$, それぞれ、プラズマ電流, 規格化ベータ, ポロイダルベータ, 内部インダクタンス, 安全係数, 三角度, 楕円度, 磁場揺動強度, グリーンワルド密度割合, 放射パワー割合) とそのうち7種類のパラメータの時間微分値, 及び空間分布を持つ4個のパラメータとその小半径方向微分値 ($V_i, dV_i/d\rho, T_i, dT_i/d\rho, \rho/a, s$, それぞれ, トロイダル回転速度, その径方向微分, イオン温度, その径方向微分, 規格化小半径, 磁気シア) からなる。

各放電は予めディスラプション放電・非ディスラプション放電に分類した。このうちディスラプション放電とは、プラズマ電流の減衰が始まった時刻を基準時刻として、ある閾値以下の時間で電流が消滅している放電のみを指すものとした。

2) サポートベクターマシンを用いた予知器モデル

SVM[7]は教師あり機械学習モデルの一種であり、多次元の入力データ x_i を予め設定したラベル $y_i \in \{-1, 1\}$ に

従って分類する分類器として用いることができる。本研究では、分類境界面の方程式が $w \cdot x + b = 0$ と記述できると仮定する線形 SVM を使用した。線形 SVM の訓練, すなわち分類境界の方程式中の重み w 及びバイアス b の決定は、両クラスからの距離が最大になるように決定される。この方法はマージン最大化 (Margin Maximization) と呼ばれる。

3) モデルの訓練と評価

モデルの訓練においては、訓練データに含まれるすべてのデータを個別のデータ点として取り扱っている。一方で、モデルの評価においては各放電を時系列データとして取り扱うことで、ディスラプション予知器の実時間での運用を模擬し、ディスラプション発生前の各時刻における予知性能を計算することを可能にしている。

ディスラプション予知器には、「ディスラプションの発生を早く正確に予知する」「ディスラプションが発生しない場合に誤って警報を出さない」の2つの異なる性能が求められる。そこで、モデルの評価基準には以下の2種類を用いた。

予測成功率 (Prediction Success Rate, PSR)

$$= \frac{\text{正しくディスラプションと判定された放電の数}}{\text{ディスラプション放電の総数}}$$

誤警報率 (False Alarm Rate, FAR)

$$= \frac{\text{誤ってディスラプションと判定された放電の数}}{\text{非ディスラプション放電の総数}}$$

ES- K の結果は、ディスラプション発生前の 200 ms 前から 30 ms 前までの 5 ms 毎の累積 PSR および累積 FAR を用いて比較した。 $K=7$ の場合の結果を示した図 3 (a) のように、縦軸に PSR, 横軸に FAR をとった平面に ES- K で求めた組合せの数をプロットした 2次元ヒストグラムを 2次元状態密度図 (2D-DoS, Density of Status) と呼ぶ。組合せ同士の比較には、理想的な予知機の性能である PSR = 100%, FAR = 0% を表す 2 D-DoS 上の点から各組合せによる性能までの距離を用いた。この距離が小さいほど予知機の性能が理想に近く、良い性能であると言える。また、図 3 (b) には、各組合せに含まれるパラメータを濃淡で示した。各列は組合せに対応し、左から理想的な性能に近い順に並んでいる。このような図をインディケータダイアグラムと呼ぶ。

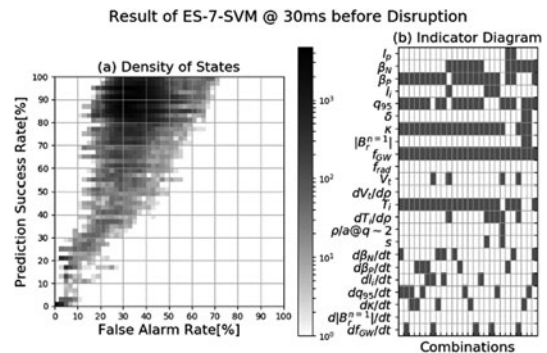


図 3 ディスラプションの発生の 30 ms 前での ES-7-SVM の結果を示す (a) 2D-DoS と (b) 理想的な性能に近い組合せに含まれるパラメータを示すインディケータダイアグラム。

4.3.4 全状態探索の結果

図4に、ES-K-SVMを実施した結果を示した。各Kにおいて最も理想的な性能に近い性能を示した組合せについて、その時の理想的な性能との2D-DoS上での距離を図4(a)に、各組合せに含まれるパラメータを図4(b)に、それぞれ示した。K=23の結果は、データセット中のすべてのパラメータを用いた場合の結果であり、適切にパラメータを選択することでディスラプション予知器の性能を向上できることが示された。最も理想的な性能に近くなったのはK=7のとき、すなわち7個のパラメータからなる組合せを用いたときであるが、その前後では距離が横ばいであり、組合せに含まれるパラメータも共通していることが読み取れる。

ここで、図3に示したK=7の場合の結果を見ると、上位の結果には β_p , q_{95} , κ , f_{GW} , T_i が共通に含まれている。SVMで得られた境界面の方程式中の重み付けを見ると、これらのパラメータの重みは他の共通でないパラメータの重みよりも大きく、中心的な役割を果たすパラメータであると考えられる。

これらのパラメータについて、上位6組の組合せ中での重みを平均し、平均化された分類境界面の方程式を次のように得た。

$$1.32\beta_p - 0.288q_{95} + 1.57\kappa + 1.58f_{GW} - 0.889T_i - 1.61 = 0$$

この式は、ディスラプションの起こりやすい領域とそうでない領域との境界面であり、抽出されたパラメータに基づくディスラプションの発生しやすいパラメータ領域の表現と解釈することができる。

4.3.5 スパースモデリングによるディスラプション予知 -まとめ-

機械学習を用いたディスラプション予知研究にスパースモデリングを導入し、機械学習モデルに入力するプラズマパラメータを選択した研究を紹介した。入力パラメータを適切に選択することでディスラプション予知の性能を向上できることが示されたほか、抽出されたパラメータに基づいて、ディスラプションの発生しやすいパラメータ領域を表現することができた。

スパースモデリングの特長は、ここで取り上げたディスラプション予知のように要素還元的なアプローチではモデル化が困難な課題に対して、高次元データから情報を抽出することによってモデルや仮説を提供できることである。

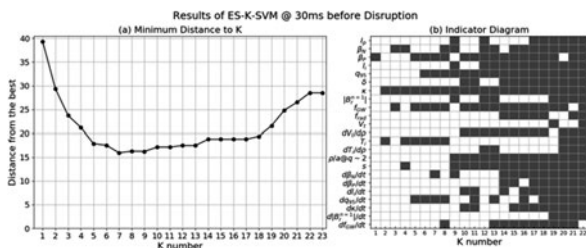


図4 ES-K-SVMの各Kで最も良い性能を示した組合せの(a)理想的な性能からの最小距離と(b)それぞれの組合せに対応するインディケータダイアグラム。

今後異なる研究対象におけるデータ駆動型のアプローチでも応用されることが期待される。

4.4 議論とまとめ

トカマクプラズマ運転に限らないが、複雑なシステムの挙動を予測し、制御するために、モデルとなる微分方程式を解くのではなく、機械学習によるデータ駆動型アプローチが注目されている。ここで、機械学習と言っても魔法の杖ではなく、多種多様な方法論があり、それぞれに特徴と対象に対する適不適がある。さらに、学習であるからには、既存のデータに基づく、あくまで帰納的帰結であるから、学習(経験)していないものへの汎化誤差への丹念な対応が、特に原型炉への研究開発の加速と学術研究としての拡がりのためには欠くことができない。そこでは標本選択バイアスへの注意と並んで、物理的な洞察や専門的な知見を盛り込むことが鍵となる。

例えば、ジャイロ運動論による輸送モデリングをニューラルネットワークによって効率化する研究で、温度臨界勾配が電子系とイオン系で同じ閾値を取るという物理からの要求を取り込むことによって精度を格段に向上させた例がある[20]。また、4.3で述べたディスラプション予知では、分離境界面を物理パラメータの線形方程式で表現したが、物理の自己相似性を鑑みれば、指数で表現する方が仮説の構築に、より役立つことができるかもしれない。

本章では機械学習という方法論にあっても源流と思想が異なり、対照的とも言えるニューラルネットワークによる深層強化学習と全状態探索によるスパースモデリングの例を紹介した。それぞれ、より外乱に対して強い適応的制御への展開、黑白の判定だけでなく灰色度合いの重要なパラメータによる表現を通じた仮説へのヒントという大きな広がりを期待させるものである。

謝辞

4.2で紹介した研究はJSPS 科研費 JP19K14697の助成を受けたものである。深層強化学習についてオープンソースの深層学習フレームワークであるChainer[21]と、その強化学習ライブラリChainerRL[22]を利用した。また、有益な助言、議論をいただいた量子科学技術研究開発機構鈴木隆博氏にも感謝申し上げます。

4.3で紹介した研究は量子科学技術研究開発機構のトカマク炉心プラズマ共同研究により実施され、またJSPS 科研費 JP17H03508の助成を受けたものである。さらに研究の遂行にあたっていただいた東京大学大学院岡田真人教授および五十嵐康彦助教からのご指導に感謝申し上げます。

参考文献

- [1] 特集「物理学とAI」人口知能 33 (2018年7月) pp.319.
- [2] シリーズ「人口知能と物理学」日本物理学会誌 74 No.1 (2019) から掲載開始.
- [3] D. Silver *et al.*, Nature 529, 484 (2016).
- [4] V. Mnih *et al.*, Nature 518, 529 (2015).

- [5] T. Haarnojo *et al.*, arXiv preprint arXiv:1812.05905 (2018).
- [6] 科学研究費助成事業 新学術領域「スパースモデリングと高次元データ駆動科学創成への支援と広報」(2013-2017) 研究代表者：岡田真人 研究成果報告書, <https://kaken.nii.ac.jp/ja/file/KAKENHI-ORGANIZER-25120001/25120001seika.pdf>
- [7] C. Cortes and V. Vapnik, *Machine learning* **20**, 273 (1995).
- [8] M. Yoshida *et al.*, *Fusion Eng. Des.* **84**, 2206 (2009).
- [9] R.S. Sutton and A.G. Barto, *Reinforcement Learning: an Introduction* (MIT Press, Cambridge, 2018).
- [10] Z. Wang *et al.*, arXiv preprint arXiv:1611.01224 (2016).
- [11] T. Wakatsuki *et al.*, *Nucl. Fusion* **59**, 066022 (2019).
- [12] N. Hayashi and JT-60 Team, *Phys. Plasmas* **17**, 056112 (2010).
- [13] T. Hender *et al.*, *Nucl. Fusion* **47**, S128 (2007).
- [14] A.H. Boozer, *Phys. Plasmas* **19**, 058101 (2012).
- [15] A. Murari *et al.*, *Nucl. Fusion* **58**, 056002 (2018).
- [16] C. Rea *et al.*, *Plasma Phys. Cont. Fusion* **60**, 084004 (2018).
- [17] J.K. Harbeck *et al.*, *Nature* **568**, 526(2019).
- [18] Y. Igarashi *et al.*, *J. Phys. Soc. Jpn* **87**, 044802 (2018).
- [19] G. Matsunaga *et al.*, *Nucl. Fusion* **50**, 084003 (2010).
- [20] K.L. van de Plassche *et al.*, in proceedings of 45th EPS Conference on Plasma Physics (Prague, 2018) P2.1086.
- [21] Chainer.
<https://chainer.org>
- [22] ChainerRL Library.
<https://github.com/chainer/chainerrl>