

利用者名：コンピュータサイエンス学部 コンピュータサイエンス学科 教授 生野 壮一郎



Title:	Jacobian-AIME: A Local and Functional Approach to Interpreting Nonlinear Maps for Physics-Informed Neural Networks
論文名 (和訳)	Jacobian-AIME：物理情報ニューラルネットワークにおける非線形写像の局所的・関数構造に基づく解釈手法
Authors:	矢野 行祐(東京工科大学 大学院生)、中西 崇文(東京工科大学 教授)、生野 壮一郎(東京工科大学 教授)
Journal:	IEEE Access, Vol. 14, pp. 54726–54746, DOI: 10.1109/ACCESS.2026.3681926
掲載年月	2026 年 4 月

研究概要：本研究では、偏微分方程式に従う物理場を予測する Physics-Informed Neural Networks (PINNs) を、物理法則と整合した形で説明するための新手法 Jacobian-AIME を提案しました。従来の LIME や SHAP では入力座標を独立に摂動するため、予測点の物理的近傍から外れた説明になりやすい。また、従来の AIME ではモデル全体を一つの逆写像で平均化するため、渦や境界層のような局所構造を捉えにくいという課題がありました。Jacobian-AIME は、各評価点においてモデルのヤコビアンから局所逆演算子を構成し、Local Feature Contribution Vector (LFCV) として、どの入力方向が予測にどの程度寄与したかをベクトル場で表すことにより、局所的で安定な説明を可能にしました。

研究背景：近年、科学・工学分野では、PINNs などの深層学習モデルが、材料開発、流体解析、構造設計、気候シミュレーション、デジタルツインなどに活用されつつあります。しかし、予測精度が高くても、なぜその予測に至ったのかを説明できなければ、安全性や信頼性が求められる現場で利用は困難です。特に、PINNs の入力である空間座標は、単に独立した特徴量ではなく、偏微分方程式や境界条件を通じて互いに結び付く連続的な物理場を構成します。そのため、一般的な説明可能 AI 手法で前提とされる特徴量の独立性や大域的平均化では、物理現象に固有の局所的な因果関係を十分に説明できませんでした。

研究成果：ヤコビアン の擬似逆行列に基づく位置依存の逆演算子を定式化し、Tikhonov 正則化により数値的不安定性を抑制しました。また、LFCV が出力のずれを再現するために必要な最小ノルムの入力摂動として解釈できることを示し、多出力場への拡張と Global Feature Importance (GFI) による大域的な特徴重要度評価を行いました。2 次元 Poisson 方程式では、中心近傍に大きな LFCV が現れ、解の対称性を反映した放射状の局所感度構造を可視化しました。(図 1)。さらに $Re=1000$ の蓋駆動キャビティ流れでは、上壁・側壁の境界層、下部隅の二次渦、主渦の圧力コアなど、流体現象に対応する感度構造を可視化しました(図 2)。Friedman 検定および Wilcoxon 符号付き順位検定により、Jacobian-AIME が LIME/SHAP とは統計的に異なる説明分布を与えることが確認されました。

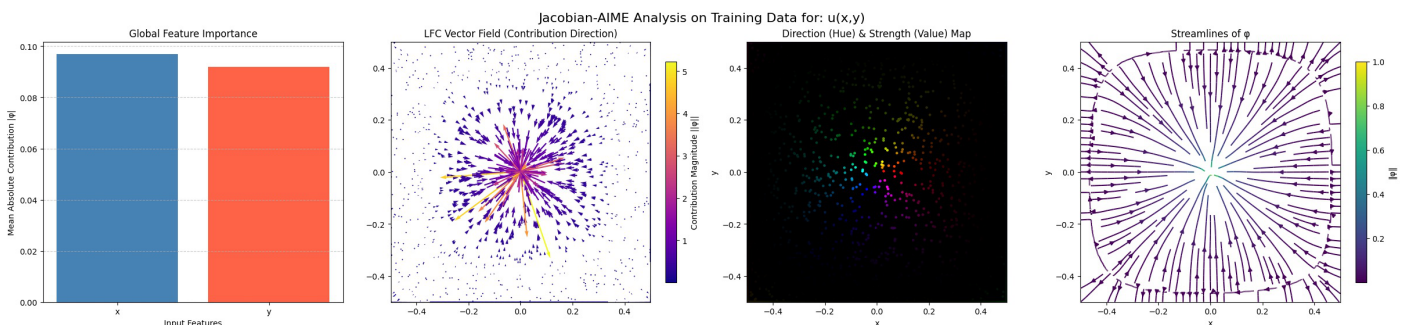


図 1 2 次元 Poisson 方程式に対する Jacobian-AIME 解析。GFI、LFCV、方向・強度マップ、流線により、解の対称性と局所感度を可視化した。

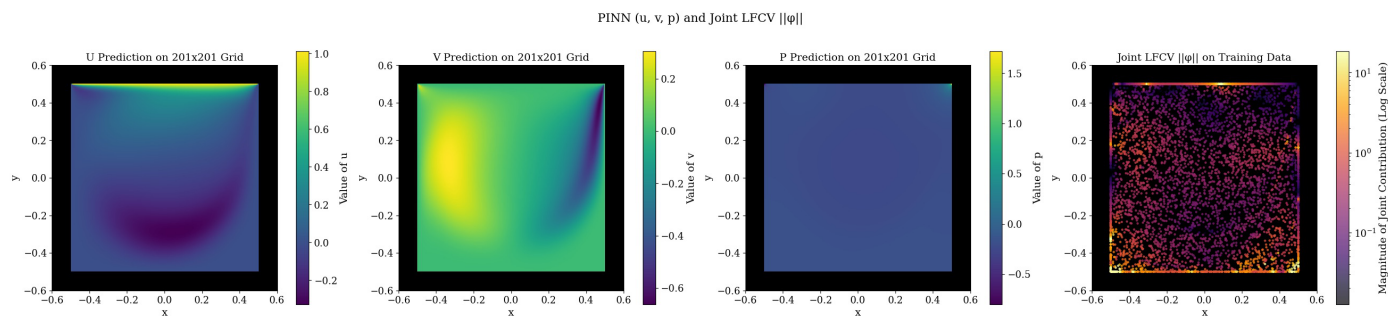


図2 蓋駆動キャビティ流れに対する PINN 予測 (u, v, p) と joint LFCV。境界層や渦に対応する高感度領域を確認できる。

社会への影響：本成果により、PINNs や科学機械学習モデルの予測根拠を、物理法則から外れにくい形で確認できます。これにより、シミュレーション結果の検証、設計判断の透明化、異常な予測結果の発見、AI モデルの説明責任の向上が期待されます。材料、流体、構造、医療・創薬、気候、デジタルツインなど、信頼性の説明が重要な分野で AI を安心して活用するための基盤技術になると考えられます。

専門用語：

Physics-Informed Neural Networks (PINNs)：偏微分方程式や境界条件を損失関数に組み込み、物理法則を満たすように学習するニューラルネットワーク。

Explainable AI (XAI)：ブラックボックス化しやすい AI モデルについて、予測に影響した要因や判断根拠を人が理解できる形で示す技術。

ヤコビアン：入力をわずかに変化させたときに出力がどのように変化するかを表す微分係数の行列。局所的な感度を示す。

擬似逆行列：通常の逆行列が存在しない場合でも、最小二乗の意味で逆向き関係を求めるための行列。Jacobian-AIME では局所逆演算子に用いられる。

Local Feature Contribution Vector (LFCV)：各点で、出力のずれを説明するために必要な入力方向と大きさを表すベクトル。寄与の強さだけでなく向きも示す。

Global Feature Importance (GFI)：各点の LFCV を空間的に平均し、入力特徴量が全体としてどの程度重要かを示す指標。