

# 論文解説：可逆関数推定の難しさ - 生成モデルを念頭に

奥野彰文<sup>\*1,2,3</sup>, 今泉允聡<sup>4,2,3</sup>

<sup>1</sup> 統計数理研究所 数理・推論研究系 <sup>2</sup> 統計数理研究所 統計的機械学習研究センター

<sup>3</sup> 理化学研究所 AIP センター <sup>4</sup> 東京大学 総合文化研究科

## 要旨

本稿は Electronic Journal of Statistics に採択された我々の原著論文: Okuno and Imaizumi (2024) に関する解説です。解説の平易さを優先するため、より厳密な記述については原著論文をご参照ください。

**キーワード:** 可逆関数推定, 一対一対応, 生成モデル, 制約の強さ, 推定の難しさ。

## 1 研究背景

非空の集合  $\mathcal{X}, \mathcal{Y} \subset \mathbb{R}^d$  について, ある関数  $f: \mathcal{X} \rightarrow \mathcal{Y}$  が可逆であるとは,

- ▷ どんな  $y \in \mathcal{Y}$  に対しても,  $f(x) = y$  となる  $x \in \mathcal{X}$  が一意に存在する

ことを言います。このような  $x$  を  $y$  の逆元と呼び, 特に  $x = f^{-1}(y)$  と書きます。関数は 1 つの  $x$  につき 1 つの出力  $y = f(x)$  を定めるので, 可逆関数とはつまり,

- ▷  $f$  を使って  $\mathcal{X}$  の各点を  $\mathcal{Y}$  の点に移し,
- ▷  $f^{-1}$  を使って  $\mathcal{Y}$  の各点を  $\mathcal{X}$  の点に移す,

という意味で  $\mathcal{X}, \mathcal{Y}$  の要素間の 1 対 1 対応を表現しているとも解釈できます。可逆関数は実は昨今話題の生成モデルと密接な関りがあります。

### 可逆性はなぜ重要か? -生成モデルとの関連-

何か特定の単語を入力するとそれに対応した画像を自動で生成して出力してくれるなどといった, 生成モデルと呼ばれる機械学習技術が昨今注目を集めています。2022 年頃にも Midjourney<sup>a</sup>や DALL-E<sup>b</sup>, Stable Diffusion<sup>c</sup>など大きな話題を呼びましたから, 試しに利用したことのある方も多いかと思います。

生成モデル研究の進展は日々目覚ましく, Okuno and Imaizumi (2024) を執筆した 2021 年当時と本稿を執筆

している 2023 年末ではまた主要技術の動向が異なりつつありますが, 実はこれら生成モデルの一部は, (きわめてシンプルに抽象化すれば) 可逆関数を推定する技術とも思うことができます。可逆性が保証された特殊な構造のニューラルネットを特に Normalizing Flow と呼び, 近年盛んに研究されています。

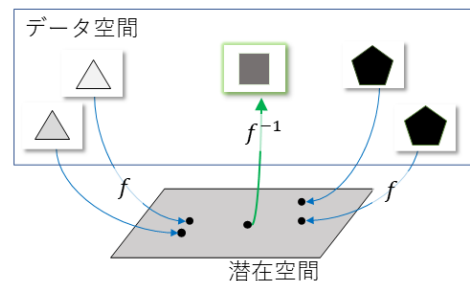


図 1: 生成モデルは一対一対応 (可逆関数) を学習。

生成モデルはまずデータ空間の元 (訓練データ) と潜在空間の元の 1 対 1 対応, つまり可逆関数  $f$  を学習して利用します。図 1 を見てください。ここでは学習した可逆関数の活用法の例として, 三角形・五角形データに対応する潜在空間データの元を, 潜在空間内で補完して逆関数  $f^{-1}$  によりデータ空間に戻すこと, つまり

$$f^{-1} \left( \frac{f(\text{三角形}) + f(\text{五角形})}{2} \right) \quad (1)$$

を考えています。(1) は三角形と五角形に対応する潜在的な“意味”の平均を取り, その平均と対応するデータを表すので, (1) により何かしらの意味で 2 つの特徴が混ざった新しいデータ (この場合, おそらく四角形の画像) が生成できると期待できます。この例に限らず, 実際に 2 つの顔写真を混ぜ合わせて, 現実には存在しない新たな顔の写真を生成したり, 2 つの絵柄を混ぜ合わせて新たな絵柄を生成したりと, 同様の手続きによって様々なデータを生成することができます。データの混ぜ合わせに本質的に寄与しているのはやはりデータ空間と潜在空間の間の対応関係を表す可逆関数  $f$  ですから, 生成モデルの学習とは, きわめてシンプルに抽象化すれば可逆関数  $f$  の推定であると言い換えることができるのです。

\* 責任著者, okuno@ism.ac.jp

<sup>a</sup> <https://www.midjourney.com/home/>

<sup>b</sup> <https://openai.com/dall-e-3>

<sup>c</sup> <https://ja.stability.ai/stable-diffusion>

## 2 本研究の問いと貢献

### 可逆関数の推定はどれくらい難しいのか？

我々の興味は生成モデルの学習，つまり可逆関数をデータから推定することがどのくらい難しいのか，数学的に評価することにあります。

ある関数  $f_*$  を推定するとき，例えば  $f_*$  が線形である ( $f_*(x) = A_*x + b_*$ ，ただし  $A_*, b_*$  はパラメータ行列/ベクトル) というのは強い制約です。  $f_*$  に線形性が課されていれば，例えば  $f_*$  が連続であることしか仮定していない場合に比べて，データから容易に関数を推定できそうです。では  $f_*$  が線形でない (つまり非線形の) 場合，例えば  $f_*$  が可逆であるならば関数の推定は簡単になるのでしょうか？可逆性は強い制約なのか，それとも弱い制約なのでしょうか？この問題について考えるため，技術的な制約から，以降は

$$\mathcal{X} = \mathcal{Y} = [-1, 1]^d$$

としましょう。

### 前段: $d = 1$ 次元ではどうなるか

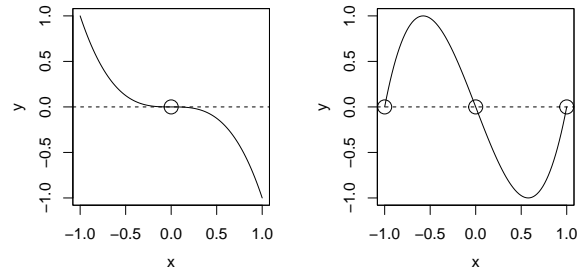
簡単のため，まず手始めに  $f_*$  に連続性を仮定し，一番簡単な設定:  $d = 1$  を考えましょう。少し考えてみると，実は，連続な 1 次元可逆関数は単調関数<sup>d</sup>に限られることがわかります。例えば可逆関数  $f(x) = x + 1$  は単調増加関数ですし，ほかにも単調減少関数  $f(x) = -x^3$  は可逆関数です。一方で，例えば  $f(x) = 0$  となる点が  $x = -1, 0, 1$  の 3 つもあり非可逆な  $f(x) = c \cdot \{x^3 - x\}$  (ただし  $c := 3\sqrt{3}/2$  は  $f(x)$  が  $[-1, 1]$  全域に値をとるための定数) は単調関数ではありません (図 2)。つまり， $d = 1$  次元の場合において可逆関数の推定とは単調関数の推定であると言い換えることができます。

単調関数推定の身近で有名な例に薬剤投与の量と効果 (反応) の関係性の議論があります。一般に薬剤を投与すればするほどその効果は強いと考えられますから，投与量を共変量  $x$  としたとき， ( $x$  があまりにも大きくなければ) その効果量  $y$  は  $x$  の増加に応じた単調増加関数となりそうですね。こうした応用から，単調関数の推定は統計学の歴史の中で盛んに議論されてきました。その理論的性質もまた非常によく研究されています。

紙面の都合で詳細は省略しますが，結論だけ述べると， $d = 1$  次元の場合，関数の単調性はリプシッツ性<sup>e</sup>と同等か，少なくともそれより弱い制約だということが示され

<sup>d</sup>  $x$  が増加するとき，  $f(x)$  が必ず増加する関数を単調増加関数，  $f(x)$  が必ず減少する関数を単調減少関数とよび，これら 2 つを合わせて単調関数と呼びます。

<sup>e</sup> 関数  $f$  がリプシッツ性を持つとは，ある  $L \geq 0$  が存在して，任



(a)  $f(x) = -x^3$  (b)  $f(x) = c \cdot \{x^3 - x\}$

図 2: (a) は  $f(x) = y$  となる  $x$  が必ず一意に定まるから可逆関数。(b) は例えば  $f(x) = 0$  となる  $x$  が  $x = -1, 0, 1$  の 3 つもあり，可逆関数ではない。

ています。したがって，可逆性もそれと同じ強さの制約だとわかります。

### 本研究の問い: $d = 2$ 次元ではどうなるか

ではこれらの議論を一般の  $d \geq 2$  次元に拡張できるのかというと，数学的には大きな困難が伴います。たとえば  $d \geq 2$  の場合，可逆関数は単調関数と限りません。次元が増えて回転の自由度が生じるので，可逆関数を数学的に特徴付けることが ( $d = 1$  次元の単純な場合に比べて) 難しいのです。詳細は割愛しますが，例えばある  $d = 2$  次元可逆関数について，関数を記述するために重要となる，要素ごとのレベル集合

$$L_{f_j}(y_j) = \{\mathbf{x} \mid f_j(\mathbf{x}) = y_j\} \quad (2)$$

は以下のような複雑な形状になります。

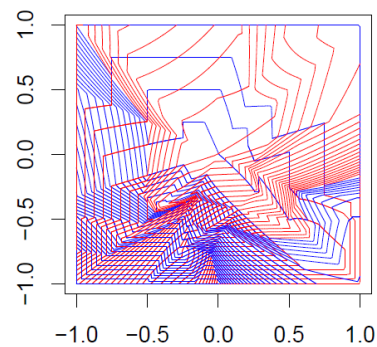


図 3: 青:  $L_{f_1}(y_1)$ , 赤:  $L_{f_2}(y_2)$

さらに厳しい問題もあります。可逆性というのは非常に小さな空間内で破綻しうるので，例えばある推定量  $\hat{f}_n$  が可逆関数  $f_*$  に収束するからと言って，(どれだけ大きな  $n$  をとっても)  $\hat{f}_n$  は可逆と限りません。それどころか，可逆な恒等関数に一樣収束するが (すべての  $n \in \mathbb{N}$

意の  $\mathbf{x}, \mathbf{x}'$  について  $\|f(\mathbf{x}) - f(\mathbf{x}')\|_2 \leq L\|\mathbf{x} - \mathbf{x}'\|_2$  を満たすことを言います。

について) どのような  $\mathbf{y}$  をとっても逆元  $\mathbf{x}$  が一意に定まらないような関数を構成することすらできます。

一例として、 $d = 1$  の場合の図 4(左) を見てください。真の可逆関数  $f_*$  をノンパラメトリック回帰で推定すると、関数自体は可逆関数  $f_*$  に収束するものの、微小領域で振動が起これると微小領域で単調性が保たれない場合があります、この図では  $\hat{f}_n$  は明らかに一意な逆元を持ちません<sup>f</sup>。通常のノンパラメトリック回帰手法は推定量が真の関数に収束しますが、この収束は推定量の可逆性を導かないのです。したがって、可逆性を担保するためには、( $d = 1$  次元の場合でいうと) 単調性などをうまく使わないと既存の理論がうまく転用できません。この  $d = 1$  次元の理論が、次元を上げて回転の自由度が増した  $d = 2$  次元に拡張できるのかは非自明でした。

### 本研究の貢献

本研究では対象を  $d = 2$  の場合に絞り、可逆関数をデータから推定する難しさを評価しました。特に、数学的なツールとして要素ごとのレベル集合を用いた逆関数の表現<sup>g</sup>を提案し、データから推定した角度を利用して関数の出力をうまく回転させることで、可逆関数をうまく特徴付けました。この結果を用いて

- 情報理論的な最適性理論と組み合わせることで、 $d = 2$  次元においても、関数の可逆性はリプシッツ性と同等か、少なくともそれより弱い制約でしかないことを証明しました。
- 最適性を持つ可逆なノンパラメトリック推定量を具体的に構成しました。

#### 最適な可逆推定量 ( $d = 1$ 次元の場合)

提案した推定量は以下の 2 段階で定義され、 $d = 1$  次元の場合は簡単に説明できます。

- (1)  $k$  近傍推定量など、好きな一致推定量で関数を推定して、 $[-1, 1]$  を一定間隔に区切った  $x_i = -1 + i/m$  と  $\hat{y}_i = \hat{f}_n^{(1)}(x_i)$  を求めます ( $i = 0, 1, 2, \dots, 2m$ )。
- (2) 次に隣接する 2 点  $(x_i, \hat{y}_i), (x_{i+1}, \hat{y}_{i+1})$  を線形に補完することで、 $\hat{f}_n^{(2)}(x)$  を区間  $[x_i, x_{i+1})$  ごとに定めれば関数  $\hat{f}_n^{(2)}$  は区分線形になりますから、先ほどのような振動を消去することが

<sup>f</sup> 本研究では、ここで利用した必ずしも可逆でない推定量  $\hat{f}_n$  を第一ステップ推定量  $\hat{f}_n^{(1)}$  とし、この推定量を可逆に修正する形で可逆な推定量  $\hat{f}_n^{(2)}$  を構成します。

<sup>g</sup>  $\mathbf{f}(\mathbf{x}) = (f_1(\mathbf{x}), f_2(\mathbf{x}))$  と  $\mathbf{y} = (y_1, y_2)$  について、(2) を用いて  $\mathbf{f}^{-1}(\mathbf{y}) = L_{f_1}(y_1) \cap L_{f_2}(y_2)$  という表現を提案しました。

でき、 $\hat{f}_n^{(2)}$  は単調となります (図 4(右))。

$d = 2$  次元で生じる回転の自由度の解消など詳細は割愛しますが、我々はこのアイデアを  $d = 2$  次元に拡張しました (図 5)。

人工データでの推定結果は図 6 を見てください<sup>h</sup>。これらの成果により、例えば実用的に用いられている可逆なニューラルネット (Normalizing Flow) が理論的に最適であるのか判断するなど、ある種のベンチマークとして利用できることが期待できます<sup>i</sup>。

### 3 今後の課題

#### $d \geq 3$ 次元ではどうなるか

ここまでは  $d = 1, 2$  の場合を考えましたが、これらの議論が  $d \geq 3$  の場合にも適用できるかという問題が残っています。最後に、 $d \geq 3$  への拡張について私見を交えて解説したいと思います。

まず、Okuno and Imaizumi (2024) では可逆かつリプシッツ連続な関数を考え、その推定はリプシッツ連続性しか課さない場合の推定と同程度に難しいことを示すことで、可逆性がリプシッツ性と同じかそれより弱い制約でしかないことを示しました。一般にリプシッツ性は微分可能性より弱い制約なので、ノンパラメトリック回帰の理論解析でよく使われます<sup>j</sup>。一方で、今回考えたリプシッツ性は、(既存理論における関数収束の議論を超えて) 可逆性を論じるには弱すぎる制約であることが徐々にわかってきました。この問題について少し紹介しておきます。

前節と図 4 で紹介した推定量は  $d = 1$  次元のものですが、 $d = 2$  次元で同様の推定量を定義するためには、図 5 のように  $[-1, 1]^2$  を一定間隔に区切った  $m \times m$  のグリッドを利用します。各区画  $I$  の 4 頂点と、その像を結ぶ四角形をうまく対応付けて補間することで推定量を構成しているのですが、どれだけ推定の粒度を細かくしても、頂点の対応が捻じれてしまうような病的な可逆関数が存在します (Okuno and Imaizumi, 2024, Appendix D.5)。同様の問題が Daneri and Pratelli (2014) などでも報告されています。例えばいたるところ微分不可能な

<sup>h</sup> 実験コードは <https://github.com/oknakfm/NPIR>

<sup>i</sup> 詳細は割愛しますが、Neural autoregressive flow や Sum-of-squares flow など構造の異なる様々な Normalizing flow についての万能近似能力が近年盛んに研究されています。近年の大きな進展としては、石川さんや手嶋さんらのグループがより実用的で理論解析の難しい Coupling-based flow についての万能近似能力を証明されており (Teshima et al., 2020)、数学的な仮定がさらに緩和されています (Ishikawa et al., 2023)。

<sup>j</sup> 本研究では著者の 2 人が数理統計の理論をバックグラウンドに持つので、議論の中で自然にリプシッツ制約を課しました。

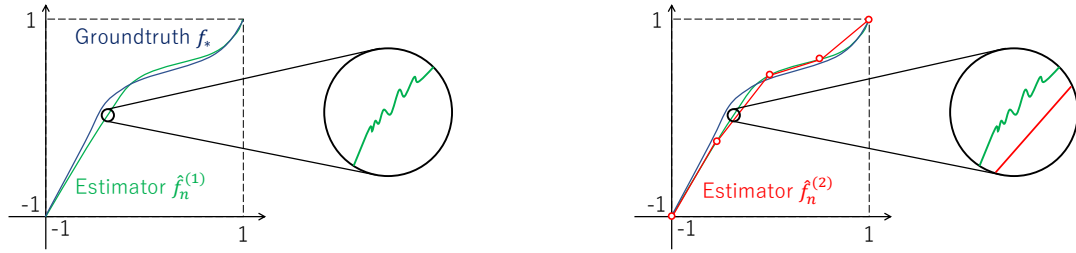


図 4: 提案した 2 段階推定を  $d = 1$  次元の場合に表したもの (左: 第 1 段階, 右: 第 2 段階).

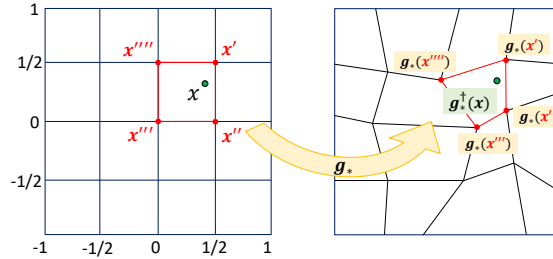


図 5: 提案した推定量の  $d = 2$  次元への拡張.  $d = 1$  次元のときは  $x$  を含む区間  $[x', x'']$  の 2 端点  $x', x''$  の像を直線で補間して単調性を担保したが,  $d = 2$  次元のときは  $x$  を含む四角形の 4 頂点  $x', x'', x''', x''''$  の像からなる四角形を補間する. (なお  $g_*$  は  $f_*$  から回転の自由度を消去したもの.) 図は Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>) の下で利用可能な Okuno and Imaizumi (2024) Fig.5 から転載されました.

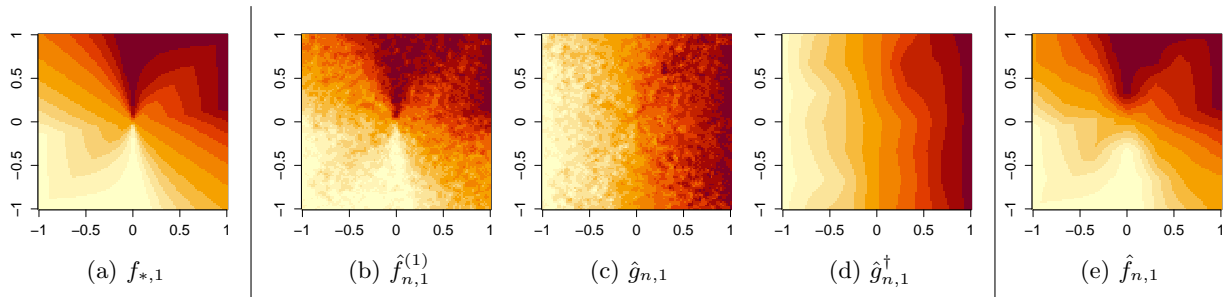


図 6: 真の可逆関数  $f_*(\mathbf{x}) = (f_{*1}(\mathbf{x}), f_{*2}(\mathbf{x}))$  に従うノイズ入りのデータを生成して  $f_{*1}$  を推定した. (a) 真の関数  $f_{*,1}$ , (b) 1 段階目の推定量 ( $k$  近傍法)  $\hat{f}_{n,1}^{(1)}$ , (c) さらに回転の自由度を打ち消したもの, (d) さらに四角形を用いた補間 (図 5) により推定量を滑らかにしたもの, (e) さらに回転を戻したもの. (e) が提案した 2 段階目の推定量であり, 通常のノンパラ推定量 (b) より滑らかで, 可逆である. 図 (a)–(e) は Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>) の下で利用可能な Okuno and Imaizumi (2024) Fig.9 (f)–(j) から転載されました.

関数のような, 数学的には考えられるが実際の応用では考え難い, このような病的な関数のせいで理論解析が非常に複雑になってしまっていて, これを  $d \geq 3$  次元以上に持ち越すことは現実的には難しそうです. 一般の次元を解析するためには, 関数に何らかの仮定を追加することが現実的だろうと考えています.

## 謝辞

草稿について東京医科大学の原田和治博士にコメントをいただきました. 厚く御礼申し上げます.

## 参考文献

- Daneri, S. and Pratelli, A. (2014). Smooth approximation of bi-Lipschitz orientation-preserving homeomorphisms. *Annales de l'I.H.P. Analyse non linéaire*, 31(3):567–589.
- Ishikawa, I., Teshima, T., Tojo, K., Oono, K., Ikeda, M., and Sugiyama, M. (2023). Universal approximation property of invertible neural networks. *Journal of Machine Learning Research*, 24(287):1–68.
- Okuno, A. and Imaizumi, M. (2024). Minimax analysis for inverse risk in nonparametric planer invertible regression. *Electronic Journal of Statistics*, 18(1):355–394.
- Teshima, T., Ishikawa, I., Tojo, K., Oono, K., Ikeda, M., and Sugiyama, M. (2020). Coupling-based invertible neural networks are universal diffeomorphism approximators. In *Advances in Neural Information Processing Systems*, volume 33, pages 3362–3373. Curran Associates, Inc.