

「思い込み」の認知過程の定式化に向けて：画像回転課題と事前制約の関係の分析

Toward a Formalization of the Cognitive Process of "Thinking-into": Analysis of the Relationship between Image Rotation Task and Prior Constraint

亀井 暁孝[†], 日高 昇平[†]
Akimichi Kamei, Shohei Hidaka

[†] 北陸先端科学技術大学院大学

Japan Advanced Institute of Science and Technology

概要

People tend to choose one or a few interpretations even when there are innumerable interpretations that can explain the phenomenon. In this research, we call the tendency of cognitive processing of "thinking-into" and advanced research for the purpose of mathematical formulation. Specifically, we proposed an image rotation task as a toy model, and conducted a numerical experiment on the relationship between prior-constraint and learning efficiency.

キーワード：思い込み, 情報圧縮, クラス算, モデル制約

1. はじめに

人は、現象を説明可能な解釈が無数に存在するときでも一つあるいは少数の解釈を選ぶ傾向にある。恣意的な解釈がある状況でも人が特定の規則性を見出す傾向を図1に模式図で表した。画像(a)は"Coffer illusion"という一種の錯視を起こす曖昧画像の一つで、この画像には(c)二つの解釈(円と四角)が存在する。人はこの画像に対し、例えば、円が図であるという解釈(c)を一つに定める。しかし、円と四角のどちらを図あるいは地と解釈するかには、いずれにも優劣や正当性があるわけではない。それでも多くの人、典型的に初めに四角が映ると解釈し、円が見える可能性を排除している。さらに言えば、"Coffer illusion"は既に無数の解釈から2択へと解釈が絞られているともいえる。こうした論理的には複数存在する解釈の候補から少数を無意識に選択し信じる人の認知的な情報処理を、本研究では「思い込み」と呼ぶ。

思い込みは無数の潜在的な可能性を一瞬で絞り込む効率的な情報処理とも言える一方で、その思い込みが成り立たない状況下では不利にも働く。本研究では、「思い込み」の数理的な定式化を目指し研究を進めた。

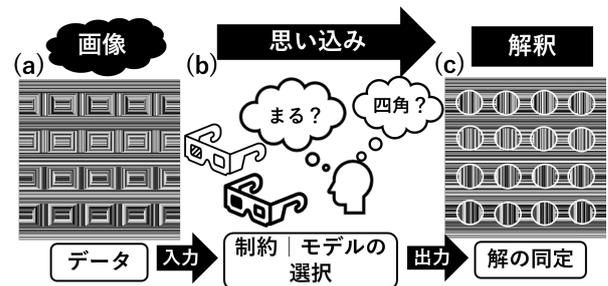


図1 (a) 四角と円の2種類の解釈のある"Coffer illusion"[2]の視覚データが与えられた場合、(b) そのデータのどの部分を図・地と解釈するかにより、(c) 知覚される像が変わる。(a),(c)の画像は© 2006 Anthony Norciaの画像およびそれを加工したもの。

2. 「思い込み」とクラス算

関連する先行研究に、論理学者のクリプキが提示した哲学的思考実験である「クラス算」がある[1]。クラス算とは、特定の範囲の数の「クラス演算 \oplus 」の結果(e.g., $1 \oplus 2 = 3$)は標準的な和(加算)と同等に見えるが、その範囲を超えた数のクラス算は和算とは全く異なる結果(e.g., $58 \oplus 2 = 5 \neq 58 + 2$)を導く演算を指す。我々の「プラス+」演算の解釈は有限事例の観察に基づくため、過去に観察した事例を超えた範囲で、我々がプラスと思い込んでいる記号が実は「クラス」である可能性を排除できない。こうして「クラス算」は「+」記号の解釈が一意に定まらないという一種の規則のパラドックスを導く。それにも関わらず、我々は典型的にはプラス記号はいわゆる「加法演算」を指すと解釈する。つまり、クラス算は演算規則の解釈を例に帰納推論だけでは成立しない「思い込み」的な推論方法の存在を指摘していると言える。これを言い換えれば、観測された現象を説明可能とする解釈が複数存在するとき、(観測事例の範囲を超えた規則性を対

象とするため) 多数の事例学習を必要とせず, 特定解釈の選択を可能とする推論方法は何かという問いになる. 本研究では, クラス算を「思い込み」の概念的な鑄型と捉え, この推論方法の数理的な定式化を目指す.

3. トイモデルとしての画像回転課題

前節で述べた哲学的思考実験のクラス算には, 計算機実装可能な定式化が存在しない. これを踏まえ, 本研究では本質的にクラス算の特徴を捉えた課題として画像回転課題 [3] を考える. 本課題は, 同じ回転関係にある画像対 (A,B),(C,D),... を複数観測し, ある画像 X に対する未知画像 Y は何であるかを問う課題である. 図 2 は画像回転課題の事例で, 画像対がもし 90 度回転の対応にあったならば, 人は少数の画像対を見て未知画像 Y はきっと 90 度回転しているに違いないと推測する. しかし, デジタル画像では同じ画素値を持つドットは見た目では区別がつかない. そのため, 仮にドットが 90 度回転の対応でなくとも結果的に「90 度回転」したように見える対応付けは無数に存在する. しかし, 人は典型的には 90 度回転という特定の規則を推定する. 従って, 画像回転課題は規則の選択に優劣がつかない状況から, 多数の事例学習を要せずに解釈を絞り込むという意味でクラス算と本質的に同じ性質を持つ. 本研究では, 予備実験として, 回転規則の推定に必要十分未満のサンプル数において, モデルに事前制約を与えることで実際に潜在規則の学習が可能となるか数値実験を行った.

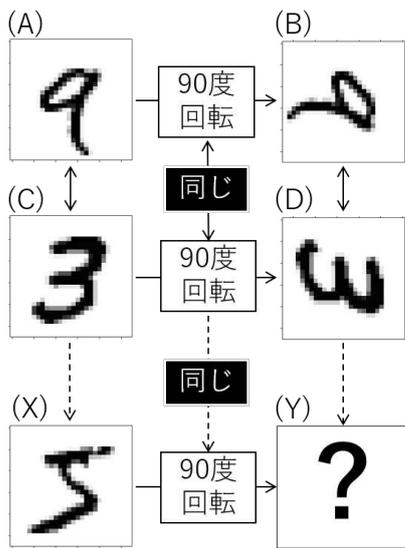


図 2 画像回転課題: MNIST の手書き文字データセットを用いた 90 度回転における事例

4. 数値実験による分析

計算モデルの変数が満たす制約は, データの持つ潜在的な意味やその解釈を反映しているため, 思い込み推論とはモデルの制約を適切に選択することと同一視することができる. 従って, 予備実験として人手で様々な制約を与え, 制約が満たす性質と規則の学習効率にどのような関係があるか分析を行った.

画像回転課題は以下のような制約付き最適化問題として定式化される. 2 つの正方行列を $X, Y \in \mathbb{R}^{n \times n}$ とし, その i 行 j 列の要素を $X_{i,j}$ と書くとき, どの $i, j \in \{1, 2, \dots, n\}$ に対しても $Y_{j, n-i+1} = X_{i,j}$ であるとき, Y を X を時計回りに 90 度回転させた行列と呼ぶ. また, ベクトル化演算子 $\text{vec} : \mathbb{R}^{m \times k} \rightarrow \mathbb{R}^{mk \times 1}$ を, 任意の行列 $X \in \mathbb{R}^{m \times k}$ に対して

$$\text{vec}(X) = (X_{1,1}, \dots, X_{m,1}, \dots, X_{1,k}, \dots, X_{m,k})^\top$$

と定義する. このとき, ある関数 $f : \mathbb{R}^{mk \times mk} \rightarrow \mathbb{R}$ を事前制約とした場合, 以下のような最適化問題となる.

$$\hat{W} = \arg \min_{W \in \mathbb{R}^{n^2 \times n^2}} f(W) \tag{1}$$

$$\text{subject to } \text{vec}(Y) = W \text{vec}(X) \tag{2}$$

一例として, 行列の l_2 ノルム (フロベニウスノルム) の二乗の場合 $f(W) = \|W\|_2^2 := \text{tr}(W^\top W)$ ($\text{tr}(\cdot)$ は行列の跡, $(\cdot)^\top$ は転置を表す) になる. 多くの文脈では, “制約” が (1) のように等号で与えられ, “目的関数” が (2) のように最小化される. しかし, 本課題では, データが少数の不良設定問題を扱うため, (2) の “誤差” ($\text{vec}(Y) - W \text{vec}(X)$) は常に 0 であるため, この関係が逆転することに注意が必要である.

具体的に行った数値実験では, 課題の計算を単純化するため $n = 4 \times 4$ 行列サイズのデジタル画像をランダムに生成しベクトル化したもの $\text{vec}(X_i)$ (図 4) と 90 度回転の関係にある画像ベクトル $\text{vec}(Y_i) = P \text{vec}(X_i)$ を用意し, それらの k 対の画像の集合 $\{(\text{vec}X_i, \text{vec}Y_i)\}_{i=1, \dots, k}$ を 1 セットの学習データとした. これらデータの画像対間の潜在的な変換は図 3 に示す置換行列 $P \in \mathbb{R}^{n^2 \times n^2}$ に固定し, P を最適化 (1)-(2) で推定すべき行列 \hat{W} の “正解” と定義した. 置換行列 P は, クロネッカー積 \otimes と j 番目の要素のみ 1 でそれ以外 0 のベクトル

$$e_{n,j} := (0, 0, \dots, 1, \dots, 0)^\top \in \mathbb{R}^n$$

を用いて以下のように定義される.

$$P = \sum_{i=1}^n \sum_{j=1}^n e_{n^2, nj-i+1} \otimes e_{n^2, (i-1)n+j}^T \quad (3)$$

推定する変数(変換行列 W)の次元は、画像サイズから $256(=n^2 = 16^2)$ であり、1 事例の学習につき 16 個の制約式が得られる。つまり、16 事例の画像対を学習すれば、 W を推定するのに必要十分な 256 個の制約式を得ることができ(1)、どんな制約関数 f でも \hat{W} は求めるべき置換行列 P に一致する。従って、何らかの事前制約関数 f に対して、16 事例未満の事例数で置換行列 P を推定することが可能となるか確認する必要がある。なお、数値実験では推定に用いる 16 事例分のデータセットをランダムに 10 セット生成して得た置換行列 P と推定行列 \hat{W} の二乗予測誤差 $\|P - \hat{W}\|_2^2$ を算出し、それらの平均と標準誤差の範囲をプロットして確認した。制約関数 f として、 f_1, \dots, f_8 用いて数値実験を行った。

$$f_1(W) = \|W\|_1 \quad (4)$$

$$f_2(W) = \|W^4 - I\|_2^2 \quad (5)$$

$$f_3(W) = \|W^8 - I\|_2^2 \quad (6)$$

$$f_4(W) = \|W^T W - I\|_2^2 \quad (7)$$

$$f_5(W) = \|W\|_2^2 \quad (8)$$

$$f_6(W) = \|W^5 - I\|_2^2 \quad (9)$$

$$f_7(W) = \{\det(W)\}^2 \quad (10)$$

$$f_8(W) = \|W^T W - W W^T\|_2^2 \quad (11)$$

これらの制約関数を選んだ理由は、求めるべき置換行列 P が持つ性質を、制約が満たす場合と満たさない場合において、回転規則の学習効率に違いがあるか分析するためである。例えば、置換行列 P が持つ性質の一つに疎な行列であることが挙げられる。制約関数 f_1 は一般にスパース制約と呼ばれ、疎な行列を効率的に求めることができる。つまり、制約関数 f_1 は P が持つ性質を満たす制約であるといえる。数値実験では、 P の性質を満たす制約として、 f_1 : スパース制約、 f_2 : 位数 4 の巡回群制約、 f_3 : 位数 8 の巡回群制約、 f_4 : 直交行列制約を選択した。また、 P の性質を満たさない制約として、 f_5 : L2 ノルム制約、 f_6 : 位数 5 の巡回群制約、 f_7 : 非正則行列制約、 f_8 : 対称行列制約を選択し数値実験を行った。

5. 数値実験の結果と考察

各制約関数に対する置換行列 P と推定行列 \hat{W} の二乗予測誤差の結果を図 5 から図 12 に示す。実際に必

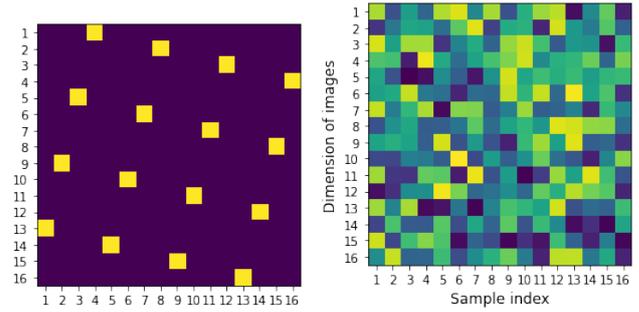


図 3 人の認識に対応する置換行列 P : 画像サイズは 16×16 で変数の次元は 256.

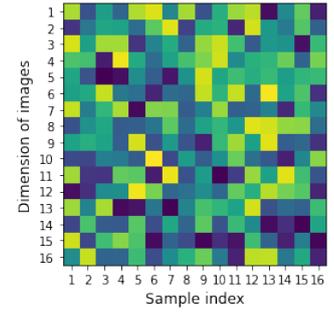


図 4 ランダムに生成した回転前の画像ベクトル: 横軸がサンプルのインデックスであり、16 サンプル分に相当する。

要十分未満のサンプル数において回転規則を学習可能とする制約が存在することが確認された。具体的に、制約関数 f が f_1 : スパース制約(図 5)、 f_2 : 位数 4 の巡回群制約(図 6)、 f_3 : 位数 8 の巡回群制約(図 7)の場合であった。これは、制約関数 f が推定すべき規則を表す行列 W の性質と合致し、 $f(W)$ が局所的に最小になる集合に正解 P が含まれると考えられる場合に少数事例での学習が可能であることを示唆する。加えて、少数事例で学習できない制約の関数(f_4 : 直交行列制約(図 8)、 f_5 : L2 ノルム制約(図 9)、 f_6 : 位数 5 の巡回群制約(図 10)、 f_7 : 非正則行列制約(図 11)、 f_8 : 対称行列制約(図 12))の場合、学習規則の予測二乗誤差が線形的に減少する傾向を持つ点が示唆された。従って、この予測誤差の分析により、必要十分少数未満の数の事例で学習を可能とする制約を、少数事例から判別するための条件が得られると期待される。

6. まとめと今後の展望

本研究では、論理的には無数に存在する解釈の候補から少数を無意識に選択し信じる人の認知的な情報処理を「思い込み」と定義し、哲学的思考実験のクワス算との関係を整理した。また、具体的な分析を行うためのトイモデルとして画像回転課題を提案した。

さらに、予備実験として、実際にモデルに事前制約を与えることで、推定に必要な十分でない少数事例の学習によって、人の認識に対応する解釈を一意に定めることが可能であることを確認した。この結果から、適切な制約関数が必要十分少数未満の数で推定可能となる条件に関する示唆が得られた。

本研究の提案する「思い込み」の定式化は、日高・高橋ら[4]の提案する“未知領域”の考え方と本質的な類似性を持つ。すなわち、得られた少数のデータ標本

への適合性だけでなく、未知の規則性の空間まで拡張した整合性により、尤もらしい仮説・規則を導く必要のある問題である。

今後は、予備実験の結果を分析するとともに、代数学の群論などを用いた情報圧縮の数理的解釈が可能であるか分析を進める予定である。

文献

- [1] 黒崎宏. ウィトゲンシュタインのパラドックス: 規則・私的言語・他人の心. 産業図書, 1983.
- [2] Norcia, A. (2006). The Best Illusion of the Year Contest Neural Correlate Society.
- [3] 河井翔太郎 (2020). 人らしく視覚的タスクを行うための学習モデルにおける制約に関する研究 北陸先端科学技術大学院大学修士論文研究 (未公刊)
- [4] 日高昇平 & 高橋康介 (2019). 未知領域を含むオブジェクト同定による窓問題知覚の説明 日本認知科学会第 36 回大会論文集, O2-1, pp. 16-18.

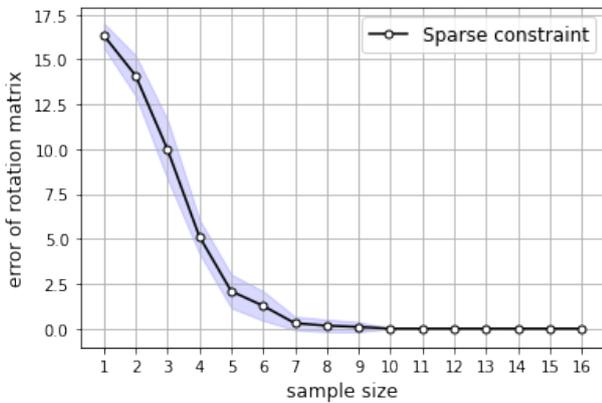


図 5 スパース制約における回転誤差とサンプルサイズの関係:16 サンプル未満で学習可能

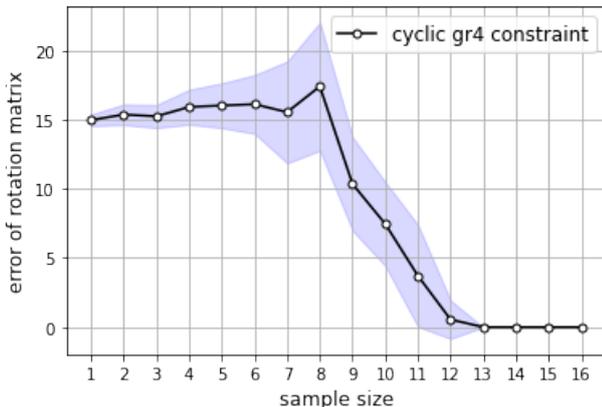


図 6 位数 4 の巡回群制約における回転誤差とサンプルサイズの関係: 16 サンプル未満で学習可能

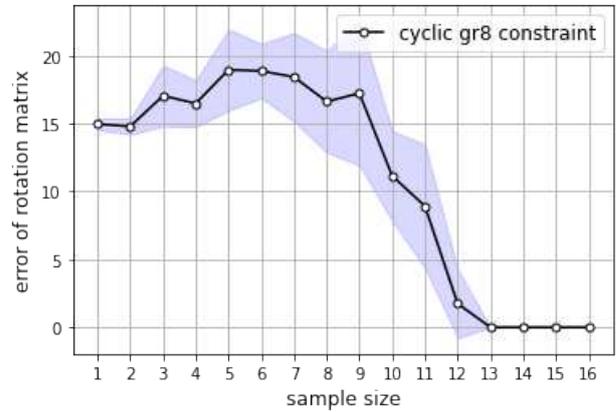


図 7 位数 8 の巡回群制約における回転誤差とサンプルサイズの関係: 16 サンプル未満で学習可能

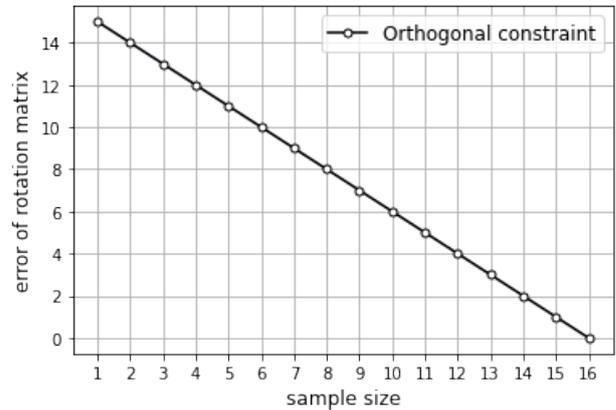


図 8 直交行列制約における回転誤差とサンプルサイズの関係

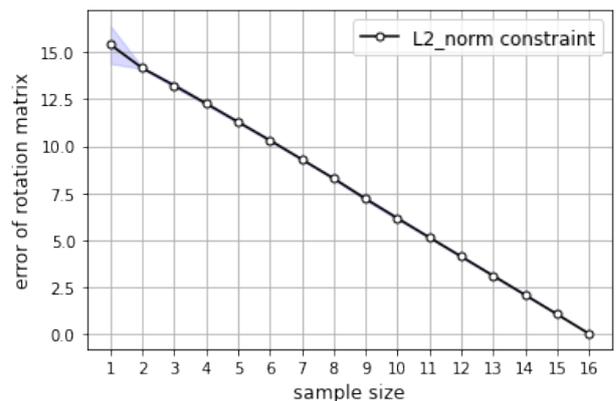


図 9 L2 ノルム制約における回転誤差とサンプルサイズの関係

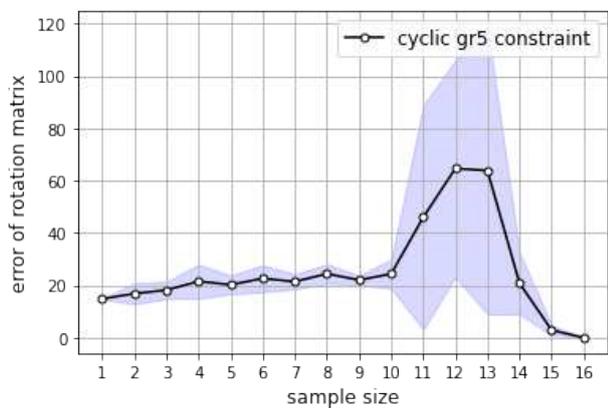


図 10 位数 5 の巡回群制約における回転誤差とサンプルサイズの関係

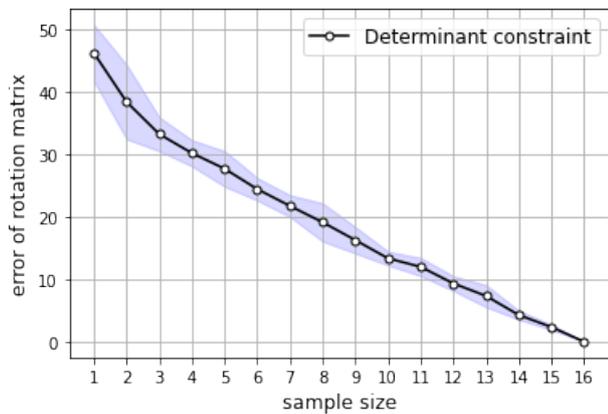


図 11 非正則行列制約における回転誤差とサンプルサイズの関係

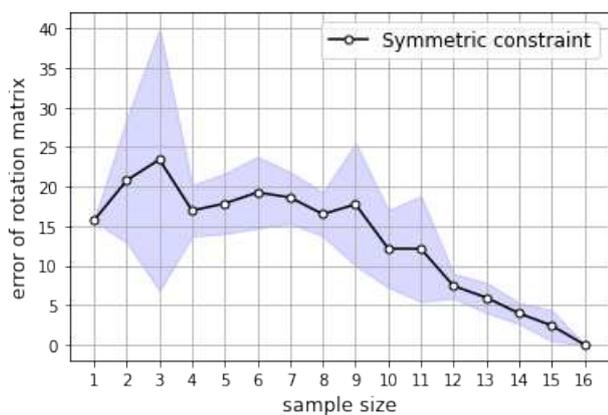


図 12 対称行列制約における回転誤差とサンプルサイズの関係