

個人の記憶パラメータに関する パラメータ推定手法の実証的な比較検討

An Empirical Comparison of Two Estimation Method for Individual Memory Parameters

新堀 耕平[†], 森田 純哉[†]
Kohei Shimbori, Junya Morita

[†] 静岡大学

Shizuoka University

shimbori.kohei.21@shizuoka.ac.jp

概要

個々の人間には固有の記憶の傾向があり、日常生活において様々なエラーを引き起こす。本研究では、エラーを引き起こす特定の個人属性や感情状態を表現するモデルパラメータを、データから推定する手法を検討する。モデルパラメータの推定手法には複数のものが存在する。本論文では、ACT-Rのパラメータを推定する手法として、勾配法に基づくものと最尤法に基づくものを比較し、両者の差異を比較検討する。

キーワード: ACT-R, 認知モデル (cognitive model)

1. 背景

個々の人間には、固有の記憶の傾向がある。そして、その傾向はエラーのパターンとして表出される。また、記憶の傾向は、個人の属性や感情状態など、様々な内的要因によって形作られる (Schacter, 2002)。よって、個人が有する記憶の傾向を、客観的に表現できるのであれば、その個人に合わせた教育や医療のアプローチが可能になる。しかしながら、これらの要因は相互に影響し、時間的にも変動する。よって、特定の記憶の傾向と内的要因を対応づけることは容易ではない。本研究では、記憶エラーをシミュレーションする認知モデルを認知アーキテクチャ ACT-R (Anderson, 2007) を用いて構築し、個人属性・感情状態に適合するパラメータを検討する。

ACT-R は、適用される認知課題の広さに利点があるとされ、個人パラメータを推定するための多くの研究がこれまでに行われてきた。しかしながら、現在の認知アーキテクチャのコミュニティにおいて、パラメータ推定のアプローチは標準化されていない (Kangasrääsiö et al., 2019)。本研究は、ACT-R によるパラメータ推定に関する 2 つの手法を実証的に比較し、複雑な認知プロセスの背後に存在する内部状態を探る方法について知見を得ることを目的とする。

2. データ

記憶の個人差を扱うために、オンライン実験として実施された数列再生課題で得られたデータを扱った。本実験は、新堀ら (2024) など、著者らの先行研究でも示されている。クラウドソーシングにより、募集された 50 人が参加した。実験参加者に対して、10 桁のランダムな数列が 2 秒間提示され、参加者はそれを可能な限り思い出して回答することを求められた。全 30 試行の課題終了後、参加者は個人属性、課題実施中の感情評価に関するアンケートに回答した。具体的には、年齢、性別、現在の体調、最終学歴、暗記力の自己評価、および日本語版 PANAS (川人ら, 2012; 佐藤ら, 2001) の指標を課題実施時の感情評価として収集した。

3. モデル

本研究では、著者ら (新堀ら, 2024) がこれまでに構築したパラメータ推定を、既存のパラメータ推定手法である Houpt ら (2022) によるアプローチと比較する。

3.1 新堀ら (2024): 正解と回答の距離ヒストグラムを指標とする勾配法

新堀ら (2024) は前節で示した実験によって得られた記憶エラーを再現するため、ACT-R Tutorial Unit 5 に含まれる Grouped モデルを一部変更して使用した。このモデルは「1234567890」という単純な数列の再生をシミュレーションする。数列を (123)(456)(7890) のようにグループ化してエンコードしており、個々の数字が含まれるグループの位置と、その数字のグループ内での位置を手がかりに、各要素の想起を試みる。

要素を想起する際、モデルは記憶として持っている数字全てに対して、活性値の計算を行う。活性値の定義は以下の式のように表される。

$$A_i = \sum_l PM_{li} + \epsilon \quad (1)$$

l は検索条件の数、 M_{li} は類似度、 P は類似度の重み、

ϵ はノイズを表す。 P を制御することによって、検索対象と同じ属性を持つ(類似する)要素が想起される確率が変わる。この重みは ACT-R の内部パラメータの一つである mp (mismatch penalty) により制御される。

また、全要素の活性値が閾値を超えないことによって、想起の失敗が発生する。この閾値は ACT-R のパラメータ rt (retrieval threshold) により設定される。主にこれら二つのパラメータを直接調整することにより、実験で得られたデータに対するフィッティングを行う。フィッティングは、提示された数列と回答された数列の編集距離のヒストグラムの一致に基づく。一致度の評価指標としては、ヒストグラムインターセクション(以降 hi とする)を用いる。hi はヒストグラムの各ビンでの最小値の総和として求められる。

3.2 Houpt ら (2022): MCMC による最尤法

Houpt ら (2022) は、最尤法に基づく ACT-R のパラメータ推定のツールである ACTRModels.jl (Houpt et al., 2022) を発表した。このツールのチュートリアルには、ACT-R Tutorial において用意されているのに対応する Grouped モデルが含まれている。このモデルは、オリジナルの ACT-R を、Julia 言語に移植することで構築されている。基本的にオリジナルのモデルと同様の構造を持っているが、パラメータ推定のアプローチが異なる。モデルは数字を宣言的記憶 M に、属性 $Q = \{group, position, retrieved, number\}$ を持つチャンク c として保持している。ここで、group は number が含まれるグループの位置、position は number のグループ内での位置、retrieved は number が想起されたかどうかを管理するフラグである。

オリジナルのモデルと同様、モデルは数列の各要素に対して活性値を計算することで、数列の想起を行う。要素 c_m の活性値 a_m は以下の式で表される。

$$a_m = blc + \rho_m + \epsilon_m \quad (2)$$

ここで、blc は活性値のベースレベル、 ρ_m は類似度によるペナルティ、 ϵ は平均 0、標準偏差 σ の正規分布に従うノイズの値である。

このモデルでは具体的に、数列内の各要素の想起確率、尤度関数がそれぞれ以下の式 2, 3 のように定義されている。想起対象 r_i に対する要素 $y_i = c_r(\text{number})$ の想起確率は式 2 のようなソフトマックス関数で与えられる。

$$\Pr(Y_i = y_i \mid \delta, \tau; \mathbf{r}_i) = \frac{e^{\frac{\mu_r}{\sigma}}}{\sum_{c_k} e^{\frac{\mu_k}{\sigma}} + e^{\frac{\mu_{m'}}{\sigma}}} \quad (3)$$

表 1: 2 モデルによる各パラメータ推定値の比較

| id | mp | | rt | |
|------|--------|---------|--------|---------|
| | 新堀ら | Houpt ら | 新堀ら | Houpt ら |
| 7 | 0.964 | 0.739 | -0.428 | -0.218 |
| 8 | 0.940 | 0.673 | -0.524 | -0.169 |
| 9 | 0.940 | 0.646 | -0.333 | -0.110 |
| 11 | 0.964 | 0.635 | -0.667 | -0.161 |
| 16 | -0.161 | 0.680 | -0.333 | -0.054 |
| 21 | 0.916 | 1.100 | -0.333 | 0.118 |
| 22 | 0.964 | 0.803 | -0.428 | -0.193 |
| 27 | 0.916 | 1.278 | -0.333 | 0.157 |
| 29 | 0.940 | 0.570 | -0.285 | -0.117 |
| 30 | 0.964 | 0.776 | -0.476 | -0.101 |
| 31 | 0.940 | 0.720 | -0.524 | -0.144 |
| 36 | 1.036 | 0.674 | -0.333 | -0.178 |
| 40 | 0.940 | 0.589 | -0.524 | -0.157 |
| 43 | 1.036 | 0.695 | -0.715 | -0.210 |
| 47 | 0.893 | 0.737 | -0.619 | -0.188 |
| 平均 | 0.951 | 0.754 | -0.457 | -0.115 |
| 標準偏差 | 0.010 | 0.048 | 0.034 | 0.028 |

$$\mathcal{L}(\delta, \tau; \mathbf{Y}, \mathbf{r}_i) = \prod_{i=1}^N \Pr(Y_i = y_i \mid \delta, \tau; \mathbf{r}_i) \quad (4)$$

ここで、 μ_m は予想されるチャンク c_m の活性化値、 σ はノイズの量を制御する。

この想起確率を実験データに近似するために、mp, rt の事前分布を仮定し、MCMC サンプリングによりデータ生成、実験データと比較することで尤度関数の近似を行い、事後分布を求める。

4. 結果

4.1 パラメータ推定結果の比較

両手法とも、単一の参加者に対するパラメータ推定には数時間を要した。そのため、50 名の参加者に対する推定結果を得る時間的コストは大きいものとなる。よって、今回の報告では、ランダムに選択した 15 人の実験データのみを対象として、モデルによるパラメータ推定、および 2 モデルによる推定値を比較した。

表 1 に、各モデルによる個人ごとのパラメータ推定値を示す。id は便宜的に振った実験参加者の ID 番号、mp, rt 列の両手法によって推定された、個人の mp パラメータおよび rt パラメータを示している。平均値を見ると、両手法では推定値に差異が生じていることがわかる。mp に関しては新堀らの手法によって大きな値(大きいペナルティ)が見積もられ、rt に関しては Houpt らの手法によって大きい値(高い閾値)が見積もられている。また、mp に関しては、新堀らの手法においては、参加者間の標準偏差が小さくなっていることも見て取れる。

表 2: 異なる手法間のパラメータ推定値の相関

| | mp | | rt | |
|----|--------|--------|--------|---------|
| | 新堀ら | Houptら | 新堀ら | Houptら |
| mp | 1 | -0.308 | -0.235 | -0.479 |
| | Houptら | 1 | 0.292 | 0.826 * |
| rt | 新堀ら | | 1 | 0.524 * |
| | Houptら | | | 1 |

*は無相関検定の結果 $p < .05$

各手法の推定値間でピアソンの積率相関係数を算出したところ (表 2), mp パラメータ同士に手法間の相関は見られなかったが, rt パラメータ同士には有意な正の相関が見られた. また, Houpt らの手法による mp パラメータと rt パラメータの間に強い相関が観察された.

4.2 確率モデルによる推定値と個人属性・感情状態との相関

得られた個人パラメータと, アンケートで得られた個人属性・感情評価に関する項目, および課題成績 (正解数列と回答数列の編集距離, 報告された数列の桁数) との相関係数を算出した. なお, ここでの相関係数は, アンケート項目が順序尺度であることから, スピアマンの指標を利用した. 結果を表 3 に示す.

推定されたパラメータは, Houpt らの手法による mp を除いて, 編集距離によって測定される課題成績と相関した. 新堀らの手法によって推定された mp が高くなることで, 課題成績が向上し (編集距離が縮小), 両手法によって推定された rt が高くなると課題成績が低下した (編集距離が増大).

また, 報告された数列の桁数との相関は, Houpt らの手法によって見積もられた rt のみにおいて負の相関が得られた. この相関は, 検索される活性値の閾値が高くなると報告される桁数が減少するという結果であり, パラメータの想定に即している.

一方, アンケート項目との相関をみると, 今回のサンプルにおいては, Houpt らにより推定された mp と複数の感情評価項目の間に幾つかの有意な相関が見られた. また, Houpt らの手法で推定された rt と, 「決心した」という感情評価項目の間に-0.536 の相関が見られた. 表末尾の編集距離, 桁数との相関にも見られる通り, rt が増加するほど, 記憶の想起が困難になる. このことから, 「決心した」と感じた人ほど rt の高いモデルとの適合が高いと解釈でき, 回答を諦めやすい傾向にあったと推察する.

表 3: 個人モデルとアンケート, 課題成績との相関

| | mp | | rt | |
|---------|---------|---------|---------|---------|
| | 新堀ら | Houptら | 新堀ら | Houptら |
| 年齢 | -0.329 | 0.330 | 0.254 | 0.240 |
| 性別 | 0.469 | 0.045 | 0.209 | -0.454 |
| 体調 | -0.022 | 0.354 | 0.358 | 0.132 |
| 最終学歴 | -0.289 | -0.202 | -0.150 | 0.151 |
| 暗記力自己評価 | 0.176 | -0.109 | -0.226 | -0.397 |
| 強気な | 0.182 | -0.479 | -0.106 | -0.277 |
| やる気がわいた | -0.071 | 0.734* | -0.065 | -0.105 |
| 活気のある | 0.147 | -0.653* | -0.195 | -0.211 |
| 熱狂した | -0.180 | 0.314 | 0.011 | -0.067 |
| 興味のある | 0.244 | -0.726* | -0.344 | -0.366 |
| 興奮した | 0.110 | -0.577* | -0.348 | -0.248 |
| 誇らしい | 0.229 | 0.357 | -0.100 | -0.137 |
| 機敏な | -0.018 | -0.004 | -0.601* | -0.190 |
| 決心した | 0.291 | -0.213 | -0.150 | -0.536* |
| 注意深い | 0.127 | -0.138 | 0.008 | -0.134 |
| ポジティブ | 0.159 | -0.388 | -0.390 | -0.368 |
| 恐れた | 0.149 | -0.655* | -0.268 | -0.189 |
| おびえた | 0.253 | 0.421 | 0.089 | -0.106 |
| うろたえた | 0.131 | 0.427 | -0.494 | -0.013 |
| 恥ずかしい | 0.044 | 0.077 | -0.545* | 0.046 |
| うしろめたい | 0.245 | -0.458 | -0.014 | -0.124 |
| びりびりした | -0.078 | -0.474 | -0.295 | 0.033 |
| 苦悩した | 0.133 | 0.271 | 0.136 | 0.057 |
| イライラした | -0.025 | -0.572* | 0.012 | 0.032 |
| 神経質な | 0.191 | -0.718* | -0.188 | -0.323 |
| 敵意を持った | 0.164 | 0.010 | 0.161 | -0.284 |
| ネガティブ | 0.291 | -0.625* | -0.220 | -0.199 |
| 編集距離 | -0.653* | 0.079 | 0.713* | 0.579* |
| 桁数 | 0.489 | -0.291 | -0.395 | -0.933* |

*は無相関検定の結果 $p < .05$

5. まとめ

今回の予備的な試行において, 手法間の共通性および差異が観察された. 手法間の共通性は, rt の推定値において見られた. 両手法で推定された rt は有意な相関を示し, ともに正解数列と報告数列の間の距離と正の相関を示した. 一方で, 両手法には主に mp の推定において差が見られた. 新堀らの手法によって推定された mp は小さい個人差を示し, Houpt らの手法による mp は rt との高い相関を示した. それぞれの手法には, 個人差の表現, 独立性の観点で課題が示されるものとなったと見ることもできる. 一方, それぞれの手法によって推定されたパラメータは異なる観点から捉えられた個人差とみることもできる. 今後, パラメータ推定に利用するサンプル数の増加, あるいはより要因を明確に操作した実験のデータを対象とすることで, パラメータ推定手法の比較を実施していく必要がある. それにより, 個人パラメータを推定する 2 つのアプローチの実証的に整理し, 両者の組み合わせに

関する実践的な指針を示すことを目指す。

文献

- Anderson, J. R. (2007). *How can the human mind occur in the physical universe?:* Oxford University Press.
- Houpt, F. C. R., Joseph W., & Gunzelmann, G. (2022). “Developing analytic likelihood functions for ACT-R.”
- Kangasrääsio, A., Jokinen, J. P. P., Oulasvirta, A., Howes, A., & Kaski, S. (2019). “Parameter Inference for Computational Cognitive Models with Approximate Bayesian Computation,” *Cognitive Science*, 43(6), e12738, DOI: <https://doi.org/10.1111/cogs.12738>.
- Schacter, D. L. (2002). *The seven sins of memory: How the mind forgets and remembers:* HMH.
- 佐藤徳・安田朝子 (2001). 日本語版 PANAS の作成, 性格心理学研究, 9(2), 138–139.
- 川人潤子・大塚泰正・甲斐田幸佐・中田光紀 (2012). 日本語版 The Positive and Negative Affect Schedule (PANAS) 20 項目の信頼性と妥当性の検討< 論文, 広島大学心理学研究, 11, 225–240.
- 新堀耕平・西川純平・長島一真・森田純哉 (2024). 記憶に関する推定パラメータと個人特性・内部状態との関係, Technical Report 13, 静岡大学情報学部, 静岡大学創造科学技術大学院, 静岡大学創造科学技術大学院, 静岡大学大学院情報学領域.