解説

運動主体感に着目したリハビリへのモデルベースドアプローチ

Model-based Rehabilitation from Sense of Agency

矢野 史朗^{*1} 近藤 敏之^{*1} 前田 貴記^{*2} ^{*1}東京農工大学 ^{*2}慶應義塾大学

Shiro Yano^{*1}, Toshiyuki Kondo^{*1} and Takaki Maeda^{*2} ^{*1}Tokyo University of Agriculture and Technology ^{*2}Keio University

1. はじめに

運動主体感 (Sense of Agency) とは、自己が営為の作用主体 (agent) であるという感覚である。身体や環境に物理的变化を引き起こしたのは自分自身であるという主観的感覚、と定義される [1]。

運動主体感は、様々な分野で注目されている。Human Computer Interaction (HCI) の分野や [2]~[4] Brain Computer Interface (BCI) の分野 [5] ではユーザに適切な操作感を持たせる必要から、定量的に運動主体感を計測し、設計に反映する研究が行われている。精神医学の分野では、統合失調症患者が過剰あるいは過小に運動主体感を持つことが知られており [6]、運動主体感の成立過程を研究することで、諸症状の理解が深まるのではないかと期待されている。また発達心理学の分野では、乳児の成長過程における運動主体感の発生時期やメカニズムに注目が集まっている [7]。運動主体感は長らく神経科学者や行動主義心理学者の研究対象とはならなかったが、近年はこれらの分野からも研究が多く報告されている [8]。そのほか、法律学の分野においても、責任の定量化に運動主体感の定量化を応用する研究が試みられている [5]。

運動主体感の定量的計測方法や理論的研究は現在でも徐々に提案されつつある状況にある。本稿ではこれらのうち、理論的研究および周辺の議論を紹介する。実験的計測手法について興味がある方は次の文献が参考になるだろう [5]。また、神経科学的な知見の紹介も最小限に留めた。参考文献として文献 [9] [10] を挙げておく。後半では、運動リハビリテーションや精神医学的な応用の可能性について、筆者らの研究事例を紹介しつつ考え方を紹介する。

2. 運動主体感とその成立過程のモデル

2.1 運動主体感とは

運動主体感は、身体や環境に物理的变化を引き起こしたのは自分自身であるという主観的感覚と定義される。

運動主体感の成立過程には、主に 2 種類の考え方があられる [11]。一つは、ヒトは自身の運動に先立って運動結果を予測しており、実際の結果と予測が一致した際に運動主体感を得るという考え方である。もう一つは、運動後にその運動を行ったか否かについて改めて振り返って判断するという考え方である。

前者はコンパレータモデル (comparator model) とよばれ、後者は見かけの心的因果 (apparent mental causation) モデルとよばれる [12]。これらは主体感について推論するタイミングによる分類である。後者は運動主体感とは、運動遂行後に経験を想起、連想して行う推論によって生じる錯覚であると強調しており、このため retrospective/postdictive な process とされる [5] [11]。これに対し、前者は運動計画時および運動遂行中に観測される情報を用いて推論するとされており、このため prospective/predictive な process とされる [11] [13]。

主体感の成立には predictive process と postdictive process の片方ないし両方が必要であるとする cue integration モデルも提案されている [11]。運動前後といったタイミングにかかわらず、記憶や経験など内部に保有している情報 (internal cue) と、外部から逐一到達する情報 (external cue) を統合して判断するという考え方である [14]。

以下ではコンパレータモデルと cue integration モデルの詳細をそれぞれ紹介する。

2.2 成立過程のモデル：コンパレータモデル

2.2.1 コンパレータモデルとは

コンパレータモデルは当初、ヒトの運動制御のモデルとして提案され、その後、運動主体感のモデルとしても用いられるようになった [15] [16]。

コンパレータモデルにおいて運動主体感とは、事前に予測していた運動の結果が、実際の結果と一致することで発生すると考える (図 1)。予測と結果の不一致度を予測誤差

原稿受付 2017 年 6 月 2 日

キーワード: Sense of Agency, Bayesian Statistics, Rehabilitation

^{*1}〒184-8588 小金井市中町 2-24-16

^{*2}〒160-8582 新宿区信濃町 35

^{*1}Koganei-shi, Tokyo

^{*2}Shinjuku-ku, Tokyo

図1 コンパレータモデル [11] [17] を改変

(comparator error, prediction error) と呼び、誤差がゼロのときに運動主体感が発生するとする (zero comparator error) [1] [5] [17]. 図1で順モデル (forward model) とは、運動の結果、観測されるであろう感覚情報を予測するモデルである。運動制御信号 (motor command) は身体運動を生み出すために用いられるとともに、脳内で複製され順モデルの引数としても用いられる。この複製信号を遠心性コピー (efference copy) とよぶ。ヒトは予測誤差を最小化するように運動制御や順モデルを更新するとされている [5] [17].

2.2.2 コンパレータモデルの問題点と周辺分野

コンパレータモデルでは予測誤差がゼロの時、運動主体感が発生すると考える (zero comparator error) [1]. しかし実験的には、予測誤差が微小であれば運動主体感に大きな影響を与えないことが分かっており [6] [18], 本モデルの主張である予測誤差と運動主体感の関係は実験と合致しないと指摘されることがある [11].

筆者らは、この問題点がコンパレータモデルにおける予測誤差の計量方法の不適切さに由来すると考えて研究してきた [19]. まずはヒトの運動学習分野における予測誤差の研究について紹介し、その次に筆者らの研究を2.2.4項で述べる。

予測誤差最小化により順モデルが学習されるという考え方は古くから研究されている [20]. この考え方では、予測信号も観測信号も (確率) 変数と捉え、その自乗誤差またはそれに類する関数を損失関数と考えることが特徴である [21] ~ [23]. 一方, Wolpert はベイズ推定の考え方を採用し、順モデルを尤度関数と事前分布に分割した [24]. 彼はベイズの定理による順モデルの更新を紹介し、尤度関数が予測誤差に相当するとしている。尤度関数の具体例としては、時間知覚モデル [25] で正規分布が用いられている。また予測誤差とは異なる誤差の話になるが、感覚モダリティ選択 [26] モデルでもやはり正規分布が用いられている。

コンパレータモデルの2種類の見方を紹介した。筆者らはコンパレータモデルをベイズ推定として再定式化し、実験データと整合性のある結果を得た [19]. 研究を紹介する前に、ベイズ推定と勾配法との関係を紹介したい。

図2 Exponentiated gradient による事前分布の更新. 太線が目的関数 $L(\theta)$, 細線が事前分布, 破線が事後分布

2.2.3 勾配法としてのベイズ推定

認知科学や神経科学において、ヒトの認知学習や運動学習をベイズ推定になぞらえながら理解しようという研究が2000年初頭から行われてきた [27] [28]. これを Bayesian coding hypothesis (Bayesian brain) とよぶ。Bayesian coding hypothesis は『知識は脳内で条件付き確率分布 (事前分布, 事後分布) によって表現されている。ベイズの定理に従って事前分布に観測値の情報が加わり、更新される』という仮説である [27] [29].

こういった分野では、ベイズ推定の説明はベイズの定理 (条件付き確率分布の定義) から始めるのが主流のようである [29] [30]. 本稿ではこのような導入は他書に譲り、文献 [31] [32] に沿って最適化問題との関連からベイズ推定を解説する。

最適化問題 (1) を考える。 $L: \mathbf{R}^n \rightarrow \mathbf{R}$ は目的関数。

$$\min_{\theta=(\theta_1, \dots, \theta_n) \in \mathbf{R}^n} L(\theta) \quad (1)$$

また、類似の問題として、最適化問題 (2) を考える

$$\begin{aligned} \min_{\{p(\theta_i)\} \in \mathbf{R}_+^n} \sum_{i=1}^n p(\theta_i) L(\theta_i) \\ \text{s.t. } \sum p(\theta_i) = 1 \end{aligned} \quad (2)$$

L が1階微分可能とすると、問題 (1) の局所解が

$$\theta^{k+1} = \theta^k - \alpha \nabla L(\theta^k) \quad (3)$$

の勾配法により得られるのは周知のとおりである。一方、問題 (2) の解は Exponentiated Gradient として知られる式 (4) の反復で得られることが知られている (付録 A, 図2)。

$$p^{k+1}(\theta_j) = \frac{\exp\left(\frac{-L(\theta_j)}{\beta}\right) p^k(\theta_j)}{\sum_{i=1}^n \exp\left(\frac{-L(\theta_i)}{\beta}\right) p^k(\theta_i)} \quad (4)$$

この手法では、まず各 θ_i に対してその重み $p^k(\theta_i)$ が与えられており、またその評価値 $L(\theta_i)$ が対応するとする。更新の前後では、評価値 $L(\theta_i)$ が小さいほど θ_i に応じた重

み $p^k(\theta_i)$ を強め、逆に評価値が大きいほど重みを弱めていく (図 2). この手続きにより最適解が得られることが知られている (付録 A).

さて最適化問題 (2) の特殊な場合を考える. 統計モデル $P(x|\theta_i)$ から見た観測値 $x := \{x_1, \dots, x_m\}$ の負の対数尤度を $L(\theta_i)$ とする.

$$L(\theta_i) = - \sum_{l=1}^m \log P(x_l|\theta_i) \quad (5)$$

この L を式 (4) に代入すると次式が得られる. 特に $\beta = 1$ のとき, これはベイズの定理とよばれる [33].

$$p^{k+1}(\theta_j) = \frac{\prod_{l=1}^m P(x_l|\theta_j)^{\frac{1}{\beta}} p^k(\theta_j)}{\sum_{i=1}^n \prod_{l=1}^m P(x_l|\theta_i)^{\frac{1}{\beta}} p^k(\theta_i)} \quad (6)$$

$1/\beta$ は逆温度とよばれ, 各更新ステップでの事前分布と事後分布の変化を KL 距離で測ったとき, どの程度の変化を許容するか調整する役割を持つ (付録 A). そのため, 学習率ともよばれる.

対数尤度関数は周知のとおり, 統計モデルと真の分布 (観測値の生成分布) の KL 距離をモンテカルロ近似したものである. すなわち式 (5) によりこの KL 距離が, 可能な範囲で徐々に最小化されるように更新される. Bayesian coding hypothesis は, ヒトが真の分布と統計モデルの距離をベイズの定理によって最小化するという仮説と言える.

以上に示したようにベイズの定理は勾配法の一つである. このためヒトによる学習手続きのいくつか, ベイズの定理や近似計算でモデル化できることは不自然ではない. ただベイズの定理はあくまで解法の一つに過ぎない. 例えばより高次の認知的推論などではまた異なる解法が取られているなど柔軟に考える必要があるだろう.

2.2.4 成立過程のモデル: 予測分布の尤度

さて運動主体感に関する筆者らの提案モデルを紹介する. ステップ k の事後分布を用いて

$$P^k(x) = \sum_{i=1}^n P(x|\theta_i) p^k(\theta_i) \quad (7)$$

は予測分布とよばれる. 筆者らの提案モデルは, 新たに観測値 X を得たとき, この予測分布から見た尤度 $P^k(X)$ を運動主体感と定義するものである [19].

従来の見方では予測誤差がゼロであるとき運動主体感が成立しているとする. そして非ゼロの場合, 運動主体感が成立しないとされるが, それが実験事実に反することは前述した通りである. 本モデルでは予測分布のピーク付近に対応する観測値に対しては自己主体感が高く, 分布の裾野に対応する観測値に対しては自己主体感が低くなる. また統計モデルを正規分布に限定していないことも特徴の一つ

である. 統計モデル, 真の分布が特異な場合, 従来のコンパレータモデルや正規分布に限定したモデルでは説明できない挙動を示すことが予想される [33]. 筆者らは事前分布を用いない手法として確率的勾配降下法を用いた更新則を比較対象とし, ある実験課題において本モデルが実験データをよく説明することを確認した [19].

2.3 成立過程のモデル: Bayesian cue integration

2.3.1 モデルの紹介

Moore, Fletcher らによる Bayesian cue integration モデルを紹介する [34].

このモデルでは事前分布, 事後分布が運動主体感に関する確率分布となる. モデルが提案された論文では, 運動主体感を A , 運動感覚情報を SM , 固有受容感覚を PR と表記し, 次のように事前分布が更新されるとしている [34].

$$P(A|SM, PR) = \frac{P(SM, PR|A)P(A)}{P(SM, PR)} \quad (8)$$

同論文には上の更新式に加え, SM と PM 以外にも参考のできる情報があれば同様の要領で用いるという旨が記載されている. 一般化して観測情報を $\mathbf{D}$ とするならば, 本モデルは次のようになる.

$$P(A|\mathbf{D}) = \frac{P(\mathbf{D}|A)P(A)}{P(\mathbf{D})} \quad (9)$$

本モデルの問題点の一つは, A がどのような根元事象から構成されるか明記されていないことである. 論文では, “Accurate attribution of agency” のモデルである旨や, 自己, 他者を明示的に回答させる課題 “Agency Judgement” [35] と整合的である旨が述べられているため, 主体感の有無という 2 値上の確率分布であると考えられる.

そのほかの問題点として, 2.2.3 項で説明したように統計モデル $P(\mathbf{D}|A)$ は目的関数に相当するのであるが, この統計モデルが定まらなければ更新則を用いることができない. つまり, cue (記憶や経験) と主体感の関係を統計モデルとして構築, 記憶することが必要になる. これらがコスト面で現実的なのか検討が必要である.

実は予測分布の尤度モデルによっても cue integration を実現できる. Cue を事前分布や統計モデルに反映することでこれが可能になる. つまり Moore-Fletcher のモデルは cue integration の一意な表現ではないということも問題点の一つだろう.

2.3.2 Cue integration と予測分布の尤度の関係

筆者らのモデルの問題点は, 主体感の有無を表すような離散的な変数がないことであろう. Bayesian cue integration モデルはこの変数を含んでいる.

これを踏まえると, Bayesian cue integration モデルは, 主体感に関する言語的で高次認知的な推論を表現している

と言えよう。筆者らのモデルはより低次の推論を表現していると考えられる。

3. 運動主体感に着目した、リハビリへのモデルベース・アプローチ

3.1 運動主体感と介入

運動介助等において、行為者の運動主体感を損なわないことは重要だろう。ここでは筆者のモデルから導き出せる仮説の中で、主体感の向上に役立つ可能性のある考え方をまとめる。これらは検証の未着手法な仮説である。例示した方法も確立した方法ではないので注意されたい。

運動主体感は、運動後の環境を予測できるようになるにつれて高くなる。図2で言えば、事前分布が最適解に近づくにつれて運動主体感が高くなる。つまり最適解に近づかせる介入が有効であろうというのが、基本的考え方である。

ベイズ推定には介入できる部分が四つある。統計モデル、事前分布、逆温度、真の分布である。以下それぞれ説明する。

3.2 統計モデルへの介入

最適な統計モデルを得るには、早期から統計モデルの候補となる集合を有しておくことが重要である。もし統計モデルの候補集合に妥当なモデルが含まれなければ、学習をしても運動主体感は低いままとなる。

体験型訓練や座学によって、本人の有する統計モデルを強化する手法が考えられる。後々本人の学習により精緻化されると考えられるので、この時点では把握を促すだけでも有効と考えられる。

3.3 事前分布への介入

運動に先行する刺激や文脈により、運動主体感に影響が出ることが知られている（プライミング効果）[5]。事前分布への介入はこれに近いと思われる。通常は式(6)によって事前分布は徐々に最適解に収束する（付録A）が、この事前分布に介入し、最適解に誘導する手法である。

2種類の介入方法が考えられる。事前分布に直接働きかける方法と、間接的に働きかける方法である。例えば前者は、各統計モデルの信頼度を直接伝える方法である。後者は、簡単な例題を先に解かせ、複雑な課題にとって望ましい事前分布へ誘導する方法が該当する。

また事前分布がガウス分布族に制約されるとすると、事前分布の標準偏差 σ が学習率となる（付録A.3）。つまり幅の狭い事前分布では学習が遅くなり[36]、言い換えれば、自分の経験に固執し過ぎると学習が遅くなる。事前分布を広げるような介入方法も有効と思われる。

3.4 逆温度への介入

逆温度は、事前分布を事後分布に更新する際、事前分布と観測事実のどちらを重視するかを調整する。実際、式(6)で $\beta \rightarrow \infty$ とすれば、事後分布と事前分布は等しくなる。例えば身体に大きな物理変化があり、新しい観測事実を重

視する必要がある際、逆温度が重要となる。

逆温度への介入方法は検討が済んでいないが、一つ紹介する。付録(A.2)から、目的関数を $L \rightarrow \infty$ のように大きくできれば、相対的に β の役割が小さくなると分かる。このことから、予測が成功した時により喜ぶようにする方法によって、学習を速められると言える。

3.5 真の分布への介入

基本的な考え方としては、行為者が予測しやすいように観測値を操作するということになる。近年の研究で誤差が大きい場合に、予測誤差を低く見積もる傾向があると報告されている[22]。例えば身体に大きな物理変化があり予測誤差が大きくなってしまった場合、予測しやすいサブゴールを与えることで学習の役に立つ可能性がある。

介入方法としては、目標到達運動課題などを用いて行為者の目標を指定し、その目標が達成できたかのように情報提示するという方法がある[37]。

4. 終わりに

本稿では、運動主体感の定義、そしてその成立過程に関する仮説を紹介した。また筆者らの仮説を含め、ヒトや動物を研究対象とした分野でベイズ推定を扱う研究が増えていることを踏まえ、ベイズの定理と勾配法の関係について紹介した。最後に、筆者らの仮説から導かれる、運動主体感を向上させ得る介入方法について大別して紹介した。本文中で述べているが、これらの介入方法は現時点では検証の未着手法な仮説である。

本稿の話題と関連して二つ問題意識を述べたい。本文で紹介したBayesian Coding Hypothesisは、動物が尤度最大化の問題をベイズ推定で解いているという仮説であった。本文中でも述べたように、ベイズの定理は勾配法の一つという解釈が可能であり、動物の学習の一部をベイズの定理でモデル化できることは不自然ではない。ただ、これは尤度最大化の一解法に過ぎず、この解法のみで考えるのは不自然に感じられる。筆者らは運動主体感の成立過程をベイズ推定でモデル化した。このあたりが本モデルを批判的に検討する切り口の一つであると考えている。

また現在、ヒトや動物を研究対象とした分野では、認知発達や運動学習の原理を予測の尤度最大化に求める仮説が受け入れられている[38]。しかしShadmehrらが言うように、運動学習では尤度以外の規範も重要となる[23]。また確率最適制御や強化学習の分野においても、期待報酬の最大化は依然として重要で有用な問題設定である[39]~[41]。これらの問題意識が、今後、本稿の読者の議論に繋がれば幸いである。

謝 辞 本研究は、JSPS 科研費新学術領域研究「脳内身体表現の変容機構の理解と制御」（課題番号 26120002, 26120005）および JSPS 科研費若手研究 B（課題番号

17K12737) の助成を受けた。

参考文献

- [1] P. Haggard and V. Chambon: "Sense of agency," *Current Biology*, vol.22, no.10, pp.R390–R392, 2012.
- [2] R. Minohara, et al.: "Strength of intentional effort enhances the sense of agency," *Frontiers in Psychology*, vol.7, p.1165, 2016.
- [3] M. Cornelio, et al.: "Agency in mid-air interfaces," *ACM CHI Conference on Human Factors in Computing Systems*, pp.2436–2439, 2017.
- [4] D. Coyle, et al.: "I did that! Measuring users' experience of agency in their own actions," *Proc. of the SIGCHI Conference on Human Factors in Computing Systems*, ACM, pp.2025–2034, 2012.
- [5] P. Haggard: "Sense of agency in the human brain," *Nature Reviews Neuroscience*, vol.18, no.4, pp.196–207, 2017.
- [6] T. Maeda, et al.: "Reduced sense of agency in chronic schizophrenia with predominant negative symptoms," *Psychiatry research*, vol.209, no.3, pp.386–392, 2013.
- [7] J.S. Kelso: "On the Self-Organizing Origins of Agency," *Trends in cognitive sciences*, vol.20, no.7, pp.490–499, 2016.
- [8] P. Haggard: "Human volition: Towards a neuroscience of will," *Nature Reviews Neuroscience*, vol.9, no.12, pp.934–946, 2008.
- [9] 村田：脳の中にある身体。pp.79–108. 東京大学出版会, 2009.
- [10] A. Murata, W. Wen and H. Asama: "The body and objects represented in the ventral stream of the parieto-premotor network," *Neuroscience research*, vol.104, pp.4–15, 2016.
- [11] N. Wolpe, and J.B. Rowe: "Beyond the "urge to move": objective measures for the study of agency in the post-Libet era," *Frontiers in human neuroscience*, vol.8, p.450, 2014.
- [12] D.M. Wegner and T. Wheatley: "Apparent mental causation: sources of the experience of will," *Am. Psychol.*, vol.54, no.7, pp.480–492, 1999.
- [13] P. Haggard: "Conscious intention and motor cognition," *Trends in cognitive sciences*, vol.9, no.6, pp.290–295, 2005.
- [14] J.W. Moore: "What is the sense of agency and why does it matter?," *Frontiers in Psychology*, vol.7, p.1272, 2016.
- [15] D.M. Wolpert, Z. Ghahramani and M.I. Jordan: "An internal model for sensorimotor integration." *Science*, vol.269, no.5232, p.1880, 1995.
- [16] S.J. Blakemore, C.D. Frith and D.M. Wolpert: "The cerebellum is involved in predicting the sensory consequences of action," *Neuroreport*, vol.12, no.9, pp.1879–1884, 2001.
- [17] M. Synofzik, G. Vosgerau and A. Newen: "Beyond the comparator model: a multifactorial two-step account of agency," *Consciousness and cognition*, vol.17, no.1, pp.219–239, 2008.
- [18] P. Fourneret and M. Jeannerod: "Limited conscious monitoring of motor performance in normal subjects," *Neuropsychologia*, vol.36, no.11, pp.1133–1140, 1998.
- [19] S. Yano, et al.: "Learning process and Sense of Agency: Bayesian learning or not," *IEEE Micro-NanoMechatronics and Human Science* 2016, pp.1–4, 2016.
- [20] M.I. Jordan and D.E. Rumelhart: "Forward models: Supervised learning with a distal teacher," *Cognitive science*, vol.16, no.3, pp.307–354, 1992.
- [21] S. Cheng and P.N. Sabes: "Modeling sensorimotor learning with linear dynamical systems," *Neural computation*, vol.18, no.4, pp.760–793, 2006.
- [22] M.K. Marko, et al.: "Sensitivity to prediction error in reach adaptation," *J. neurophysiology*, vol.108, no.6, pp.1752–1763, 2012.
- [23] R. Shadmehr, M.A. Smith and J.W. Krakauer: "Error correction, sensory prediction, and adaptation in motor control," *Annual review of neuroscience*, vol.33, pp.89–108, 2010.
- [24] D.M. Wolpert and Z. Ghahramani: "Computational principles of movement neuroscience," *nature neuroscience*, vol.3, pp.1212–1217, 2000.
- [25] M. Jazayeri and M.N. Shadlen: "Temporal context calibrates interval timing," *Nature neuroscience*, vol.13, no.8, pp.1020–1026, 2010.
- [26] K. Wei and K. Körding: "Relevance of error: what drives motor adaptation?," *Journal of neurophysiology*, vol.101, no.2, pp.655–664, 2009.
- [27] D.C. Knill and A. Pouget: "The Bayesian brain: the role of uncertainty in neural coding and computation," *Trends in neurosciences*, vol.27, no.12, pp.712–719, 2004.
- [28] D.M. Wolpert and M.S. Landy: "Motor control is decision-making," *Current opinion in neurobiology*, vol.22, no.6, pp.996–1003, 2012.
- [29] Pouget, Alexandre, et al.: "Probabilistic brains: knowns and unknowns," *Nature neuroscience*, vol.16, no.9, pp.1170–1178, 2013.
- [30] K.P. Kording, J.B. Tenenbaum: "Causal inference in sensorimotor integration," *Advances in Neural Information Processing Systems*, pp.737–744, 2007.
- [31] S. Shalev-Shwartz: "Online learning and online convex optimization," *Foundations and Trends® in Machine Learning*, vol.4, no.2, pp.107–194, 2012.
- [32] M.K. Warmuth: "A Bayes rule for density matrices," *NIPS*, pp.1457–1464, 2005.
- [33] 渡邊：「ベイズ統計の理論と方法」, ベイズ統計の理論と方法, 2012.
- [34] J.W. Moore and P.C. Fletcher: "Sense of agency in health and disease: a review of cue integration approaches," *Consciousness and cognition*, vol.21, no.1, pp.59–68, 2012.
- [35] M. Synofzik, et al.: "Misattributions of agency in schizophrenia are based on imprecise predictions about the sensory consequences of one's actions," *Brain*, vol.133, no.1, pp.262–271, 2010.
- [36] S. Yano, T. Maeda and T. Kondo: "Slow dynamics perspectives on the embodied-brain systems science," *Neuroscience research*, vol.104, pp.52–55, 2016.
- [37] W. Wen, et al.: "Goal-directed movement enhances body representation updating," *Frontiers in Human Neuroscience*, vol.10, 2016.
- [38] K. Friston, J. Kilner and L. Harrison: "A free energy principle for the brain," *Journal of Physiology-Paris*, vol.100, no.1, pp.70–87, 2006.
- [39] 松原：「確率最適制御の最近の動向: 確率推論による解法」, システム/制御/情報, vol.59, no.10, pp.369–374, 2015.
- [40] M.P. Deisenroth, G. Neumann and J. Peters: "A survey on policy search for robotics," *Foundations and Trends® in Robotics*, vol.2, issue 1-2, pp.1–142, 2013.
- [41] M. Miyashita, R. Hirotani, S. Yano and T. Kondo: "Experiment of Reinforcement Learning with Extremum Seeking," *IEEE Sixth ICT International Student Project Conference*, in press.
- [42] S. Arora, H. Elad and K. Satyen: "The Multiplicative Weights Update Method: a Meta-Algorithm and Applications," *Theory of Computing*, vol.8, no.1, pp.121–164, 2012.
- [43] Erven, Tim V., et al.: "Adaptive hedge," *Advances in Neural Information Processing Systems*, pp.1656–1664, 2011.
- [44] A. Wibisono, A.C. Wilson and M.I. Jordan: "A variational perspective on accelerated methods in optimization," *Proc. of the National Academy of Sciences*, pp.7351–7358, 2016.
- [45] 宮下, 廣谷, 矢野, 近藤: "ブラックボックス最適化としての PI2 アルゴリズム導出", システム・情報部門学術講演会, GS03-3, 2016.
- [46] E. Theodorou, B. Jonas and S. Schaal: "Reinforcement learning of motor skills in high dimensions: A path integral approach," *Robotics and Automation (ICRA)*, 2010 IEEE International Conference on. IEEE, pp.2397–2403, 2010.

付録 A. Exponentiated gradient

付録 A.1 Exponentiated gradient 導出

最適化問題 (A.1) を考える.

$$P^{k+1}(\theta) = \arg \min_{P(\theta) \in \mathbf{R}_+^n} J_k - \lambda \sum_{i=1}^n (p(\theta_i) - 1) \quad (\text{A.1})$$

ただし $P(\theta) = \{p(\theta_1), \dots, p(\theta_n)\}$ とし, J_k は次とする.

$$\begin{aligned} J_k &= \sum_{i=1}^n p(\theta_i) L(\theta_i) + \beta \text{D}_{\text{KL}}[P(\theta), P^k(\theta)] \\ &= \sum_{i=1}^n p(\theta_i) \left(L(\theta_i) + \beta \log \frac{p(\theta_i)}{p^k(\theta_i)} \right) \end{aligned} \quad (\text{A.2})$$

λ は等式制約 $\sum_i P(\theta_i) = 1$ を満たすための Lagrange 未定乗数である. さて式を整理すると (A.1) の目的関数は (A.3) となる. そのため解は (A.4) と求まる.

$$\text{D}_{\text{KL}} \left[P(\theta), P^k(\theta) \exp \left(\frac{\lambda - L(\theta)}{\beta} \right) \right] \quad (\text{A.3})$$

$$p^{k+1}(\theta_j) = p^k(\theta) \exp \left(\frac{\lambda - L(\theta_j)}{\beta} \right) \quad (\text{A.4})$$

Lagrange 未定乗数は, $\sum_{i=1}^n p^{k+1}(\theta_i) = 1$ によって定まるので, 最終的に解 (A.5) を得る.

$$p^{k+1}(\theta_j) = \frac{\exp \left(\frac{-L(\theta_j)}{\beta} \right) p^k(\theta_j)}{\sum_{i=1}^n \exp \left(\frac{-L(\theta_i)}{\beta} \right) p^k(\theta_i)} \quad (\text{A.5})$$

(A.5) のように, $L(\theta_i)$ の関数を $P^k(\theta_i)$ に乗ずることで更新する手法を Multiplicative weights algorithm (MW) とよぶ [42]. MW の中で (A.5) は, Exponentiated gradient あるいは Hedge 等とよばれる. 本手法は関数 $L(\theta)$ の微分可能性を必要としない点や, 関数 $L(\theta)$ が期待値の形をしている場合, オンライン, バッチ学習等への変形が容易な点が特徴的である. また上記以外の導出方法も知られている [31]. 手法 (A.5) は様々な派生が存在しており, 現在も基礎研究 [32] [43] [44] や応用研究 [40] [45] が進んでいる.

付録 A.2 漸近性

(A.5) を再帰的に用いると次が成り立つ.

$$\begin{aligned} p^k(\theta_j) &= \frac{\exp \left(\frac{-kL(\theta_j)}{\beta} \right) p^0(\theta_j)}{\sum_i \exp \left(\frac{-kL(\theta_i)}{\beta} \right) p^0(\theta_i)} \\ &= \frac{p^0(\theta_j)}{\sum_i \exp \left(-kL'_{ij} \right) p^0(\theta_i)} \end{aligned} \quad (\text{A.6})$$

ただし (A.6) の L'_{ij} は

$$L'_{ij} = \frac{L(\theta_i) - L(\theta_j)}{\beta}. \quad (\text{A.7})$$

(A.6) の分母を Z_{kj} とし, $A := \arg \min L(\theta)$ とすると, $\theta_j \notin A$ の場合は次が成り立つ.

$$\lim_{k \rightarrow \infty} Z_{kj} \rightarrow +\infty \quad (\text{A.8})$$

一方で $\theta_j \in A$ の場合は有限にとどまる. 以上から (A.6) は集合 A を台とする確率分布に収束すると分かる. 収束後の台を広げる目的で, エントロピー正則化等が用いられる.

付録 A.3 近似解法

事前分布は脳内でどのような表現をしているだろうか. (A.5) の近似手法の内, 事前分布の代わりに確率変数を用いる手法を紹介する [41] [46].

事前分布が正規分布 $\mathcal{N}(\mu^k, \sigma^2)$ とする. すると (A.5) から

$$\begin{aligned} \mu^{k+1} &= \sum_{i=1}^n \theta_i p^{k+1}(\theta_i) \\ &= \frac{\sum_{i=1}^n \theta_i \exp \left(\frac{-L(\theta_i)}{\beta} \right) p^k(\theta_i)}{\sum_{i=1}^n \exp \left(\frac{-L(\theta_i)}{\beta} \right) p^k(\theta_i)} \end{aligned} \quad (\text{A.9})$$

とできる. 事前分布からのサンプルは次のように表記できる:

$$\hat{\theta}_j^k = \mu^k + \sigma \varepsilon_j \text{ where, } \varepsilon_j \sim \mathcal{N}(0, 1) \quad (\text{A.10})$$

これを用いて (A.9) をモンテカルロ近似すると

$$\mu^{k+1} = \mu^k + \sigma \frac{\sum_j \varepsilon_j \exp \left(\frac{-L(\hat{\theta}_j^k)}{\beta} \right)}{\sum_j \exp \left(\frac{-L(\hat{\theta}_j^k)}{\beta} \right)} \quad (\text{A.11})$$

となり, 平均値のみ更新する近似手法が導かれる.

矢野史朗 (Shiro Yano)

2012 年東京大学大学院工学系研究科精密機械工学専攻博士課程修了. 立命館大学専門研究員を経て, 2015 年より現職. 非線形現象のモデル化, 計算論的精神医学, 強化学習に関する研究に従事. 博士 (工学).

近藤敏之 (Toshiyuki Kondo)

1999 年名古屋大学大学院工学研究科計算理工学専攻博士後期課程修了. 東京工業大学助手, 東京農工大学講師, 准教授を経て, 2014 年より現職. 運動学習, BCI リハビリシステムに関する研究に従事. 博士 (工学).

(日本ロボット学会正会員)

前田貴記 (Takaki Maeda)

1995 年慶應義塾大学医学部卒業. 横浜市立市民病院, 桜ヶ丘記念病院を経て, 2010 年慶應義塾大学大学院医学研究科博士課程単位取得退学. 慶應義塾大学医学部精神神経科助教を経て, 2012 年より現職 (同専任講師). 統合失調症の精神病理学, 病態生理学研究に従事.

精神科専門医. 精神保健指定医. 博士 (医学).