解 説

運動学習の統一理論モデル

—運動学習における誤差の予測の重要性—

瀧 山 健

東京農工大学 工学部 電気電子工学科*

A Unified Model of Motor Learning—Importance of Prospective Error—

Ken Takiyama

Department of Electrical and Electronic Engineering, Tokyo University of Agriculture and Technology*

概要

腕の到達運動と外乱を用いた実験パラダイムにより、身体運動学習に関わる脳内メカニズムの理解が進められてきた。特に、実験と理論の融合的アプローチが広く進められ、様々な理論的枠組が提案されるに至っている。しかしながら、様々な理論的枠組が場当たりの提唱され、詰まるところ身体運動学習に関わる脳内メカニズムにおける最も重要な要素は何か、という問題は未解決なままである。著者は近年、“我々は運動する前に暗黙の内に運動誤差を予測している”と想定する誤差の予測モデルを提案した。この誤差の予測モデルは既存の理論的枠組では再現することができなかったランダム学習を再現するための数学的要請に基づいたモデルであり、誤差の予測モデルのみが予測できる現象を行動実験に基づき実証することができ、別々の既存モデルで再現されてきた様々な現象を同一パラメータで再現することが可能であった。興味深いことに、著者は意思決定の知見を一切考慮せずに誤差の予測モデルを提案したものの、この誤差の予測モデルは意思決定の分野でしばしば用いられる価値関数を運動学習の分野に応用したモデルとして解釈できることを本稿にて示す。また、誤差の予測モデルを意思決定へと応用することで、意思決定に関する様々な現象を統一的に解釈できる可能性に対しても言及する。

1. 意思決定研究と身体運動研究との関連性について

野球では変化球を投げるか直球を投げるか意思決定し、目標の位置にボールを投げられなかったときには、目標の運動と実際の運動の間の誤差を小さくするように運動学習する。このように、日常生活における運動は意思決定→運動学習→意思決定…の繰り返しである(図1)。近年、身体運動学習・身体運動制御研究に意思決定研究の要素を取り入れた融合的研究が一世を風靡している¹⁻³⁾。例えば、モデルベース、モデルフリー学習⁴⁾という概念が運動学習でも頻繁に議論され始めている³⁾。事実、意思決定と運動学習は非常に密な関係にある。意思決定は報酬最大化⁵⁾、運動学習は評価関数最小化でモデル化され⁶⁾、どちらも最適化の枠組みに基づきモデル化されてきた。運動学習と意思決定には同時期に同様のモデルが提案されることも

あり(学習には複数の時定数に関わる複数時定数モデル^{7,8)}、カルマンフィルタモデル^{9,10)})、更に腕の到達運動に基づいた意思決定課題では、意思決定と運動制御のどちらにも運動野が関わること¹¹⁾など、共通した脳領域が関わることも報告されている。以上から身体運動と意思決定は密な関係にあり、一方の知見を他方に流用することで新たな知見が得られることが期待できる。

本稿では、著者が近年提案した新たな運動学習モデル¹²⁾を紹介する。著者は既存の数値モデルでは説明不可能であった運動学習に関連する現象を再現する条件を数学的に探り、そのモデルを提案するに至った。興味深いことに、数学的に導かれたそのモデルは、意思決定の分野でしばしば用いられる価値関数の概念を運動学習に適用したモデルであった。著者のモデルと従来の意思決定のモデルとで大きく異なる点は、著者のモデルにおいては価値関数を学習信号や行動選択の基準

* 〒184-8588 東京都小金井市中町 2-24-16

図1 意思決定と運動学習の関係性。日常生活において意思決定と運動学習は密接な関係にある。

として用いるのではなく、神経活動を決める入力として用いる点である。本稿では第一に、著者のモデルを説明するために必要な予備知識を詳述する。より理解を深めるためには様々な疾患の患者を対象とした運動学習実験の知見を紹介すべきではあるがスペースの関係上それは著者の以前の原稿（瀧山 健，計算論的神経科学から眺めたりハビリテーション—現状と展望—，2013，日本神経回路学会誌）に委ねることとする。第二に、著者が提案する新たな運動学習モデルを詳述する。最後に、著者の提案する新たな運動学習モデルを意思決定へ適用することを試みる。

2. 身体運動の複雑性

常日頃、我々は苦もなく身体を動かすことができる。実はこれはとても不思議な事である。例えば、重力を無視できる肩と肘の水平面運動を考えてみよう。この肩と肘が従う運動方程式は、肩関節の角度を q_1 ，肘関節の角度を q_2 ， $\mathbf{q} = (q_1, q_2)$ ， $\dot{y}$ を y の時間微分 (n 個ドットがあれば n 階微分) とすると、

$$\mathcal{M}(\mathbf{q})\ddot{\mathbf{q}} + \mathbf{C}(\dot{\mathbf{q}}) = \boldsymbol{\tau} \quad (1)$$

$$\mathcal{M}(\mathbf{q}) = \begin{pmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{pmatrix} \quad (2)$$

$$m_{11} = m_1 l_{c1}^2 + I_1 + m_2(l_1^2 + l_{c2}^2 + 2l_1 l_{c2} \cos q_2) + I_2$$

$$m_{12} = m_{21} = m_2 l_1 l_{c2} \cos q_2 + m_2 l_{c2}^2 + I_2$$

$$m_{22} = m_2 l_{c2}^2 + I_2$$

$$\mathbf{C} = \begin{pmatrix} -h\dot{q}_2 & -h(\dot{q}_1 + \dot{q}_2) \\ h\dot{q}_1 & 0 \end{pmatrix} \begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \end{pmatrix} = \boldsymbol{\Gamma} \begin{pmatrix} \dot{q}_1 \\ \dot{q}_2 \end{pmatrix} \quad (3)$$

と記述される。ただし、 $h = m_2 l_1 l_{c2} \sin q_2$ ， $\boldsymbol{\tau} = (\tau_1, \tau_2)$ は各関節に外部から与えるトルクである。各々のパラメータの意味は図 2a を参照していただきたい。肘と肩の単純な運動を想定しても、目標の運動を達成するためには (1)～(3) 式に適した運動指令を生成しなくてはならない。我々の身体運動は複雑な計算が必須であることは明白であろう。しかしながら、日常生活

において自然と複雑な身体を動かしているこの状況こそ、身体運動研究の醍醐味であり、未だに矛盾のない説明がなされていない部分でもある。

複雑な制御対象（身体）を動かすために最も単純な方策の一つが、PID フィードバック制御器である。目標の運動 d と実際の運動 x との差分に比例した運動指令 $u = K(d - x)$ を生成する枠組みが最も単純な PID フィードバック制御器である。近年では脳が最適なフィードバック制御器のような役割を果たしていることが示唆され¹³⁾，支持を得ている¹⁴⁾。しかしながら、素早い運動に着目する限り、フィードバック制御器だけでは不十分であることは明白である。我々の感覚系は時間遅延を伴う。例えば機械的に肩や肘にトルクを発生させて腕を強制的に動かしたとしよう。このとき、“腕が動いた”という情報が脳活動に反映されるまで、おおよそ 50 msec 程度の時間遅延が存在する¹⁵⁾。すなわち動き始めにおいては、フィードバック制御指令 $u = K(d - x)$ を計算するために x の情報を使うことができない。たった 50 msec じゃないかと思われるかもしれないが、ゴルフスイング等の非常に素早い運動では 50 msec もあれば運動の大部分は実行し終えているであろう。実際にシミュレーションをしてみると 50 msec の時間遅延がある PID フィードバック制御器では非線形なダイナミクスに従う肩と肘を動かすことは難しい（図 2b, c）。したがって、少なくとも素早い運動の出だしの部分においては、脳はフィードフォワード制御器のような役割を担っているのではないかと考えられている。フィードフォワード制御器とは、 x の情報を用いずに何かしらの関数 $f(\cdot)$ を用いて $u = f(d)$ として運動指令を決定する制御器である。この $f(\cdot)$ はもちろん我々の身体のダイナミクスをある程度反映させたものでないと効果的な運動指令を生成できないため、 $f(\cdot)$ は身体のモデルを反映させた関数であると考えられる（図 2d）。この身体のモデルをしばしば内部モデルと呼ぶ。以下では運動学習とは内部モデルないしはフィードフォワード制御器の学習である、と捉えて議論を進めていく。

図2 2関節腕のシミュレーション。(a):肩関節の中心を原点,肩関節角度 q_1 ,上腕の慣性モーメント,質量,長さ,肩関節から上腕の重心までの距離を l_1, m_1, l_1, l_{c1} とする。同様にして肘関節角度,前腕のパラメータを定義する。(b):肩関節トルク,肘関節トルクをフィードバック制御器により発生させたときの運動の軌道。赤い実線が肘関節と手先の理想的な初期位置と終点位置を表しており,運動時間500 msec,手先距離10 cmの腕の到達運動をシミュレートした。黒い点線は20 msec毎の運動の軌跡を表している。実際の肩関節角度,肘関節角度は時間遅れなしでモニタリングできている状況を想定している。(c):実際の肩関節角度,肘関節角度は50 msecの時間遅れをもってモニタリングできている状況を想定している。(d):目標の関節角度軌道が入力情報となる基底関数にフィードバック誤差学習を用いて腕の内部モデルを学習させた後,フィードフォワード制御により同様の運動をシミュレートした結果。

3. 身体運動学習の実験系

幼いころ練習を重ねて自転車に乗ることができるようになったこと,新たに購入したキーボードではミスタイピングしてしまうが次第に使い慣れていくこと,ギターを始めてすぐにはFのコードを弾くことさえ難しいが次第に上達していくこと,次第に水泳が上達していくこと,麻痺してしまった腕もリハビリテーションを積み次第に動かせるように回復していくこと,これらは全て身体運動学習の例である。多くの方が上記の内少なくとも一つは経験があるように,身体運動学習は日常生活において身近に経験することができる脳機能である。「脳」機能である,と述べた理由は,サル

を対象とした実験において運動学習前後で運動野(運動に関連する脳領域)の脳活動が変化すること¹⁶⁾,ラットのスライス実験では運動学習前後で運動野内での可塑性の特性が変化すること¹⁷⁾,ヒトを対象とした実験では特に小脳疾患(spinocerebellar ataxia)患者は運動学習能力が低下している状況があること¹⁸⁾などが報告されていることから,運動学習は主に脳が関わる学習過程であると推測できるためである。中枢神経系の特定の領域に焦点を当てて議論を行うことは本稿の主目的ではないため,皮質,小脳,大脳基底核など皮質下領域を含めて広い意味で「脳」という言葉を使わせていただく。

我々の身体運動は,脳を制御器,身体(保持している

図3 運動学習実験系。(a)：身体運動は脳を制御器，身体を制御対象と見なしたスキームとして捉えることができる。(b)：被験者には椅子に座ってもらい身体をある程度固定する。モニター下に配置しているロボットアームを操作することで(「手」と表記された黒い点線)モニターに表示されているカーソル(緑色の丸)をターゲット(黒い丸)に向かって動かす。視覚回転外乱を印加したときには実際のロボットアームの軌道とカーソルの軌道の間にずれが生じる。(c)：各トライアルの開始時に，被験者はカーソルをモニターに表示される初期位置(黄色い丸)の上に2秒間ほど留めておかなくてはいけない。(d)：視覚回転外乱を印加した後はカーソルの動きが予測していた動きと異なるため，予測誤差が生じる。(e)：横軸が視覚回転外乱を印加し始めてからのトライアル数，縦軸が予測誤差。被験者は徐々にカーソルをターゲットに向かって真っ直ぐ動かすように学習していく。点線が実データ，実線が状態空間モデルないしは運動プリミティブモデルでフィッティングした結果。(f)：典型的な腕の到達運動の速度プロファイル，横軸が時間，縦軸が速度，速度が最大値をとる時刻をピーク時刻と呼び，ピーク時刻における運動方向を運動指令または視覚回転外乱課題を達成するための運動記憶と見なす。(g)：運動プリミティブの枠組み。(h)：力場学習の汎化関数(視覚回転外乱学習の汎化関数もおおよそ同様のもの)。 $\theta = 0$ の運動方向で外乱を学習した後に，様々なターゲット方向にてテストトライアルを行い運動学習効果がどのように発揮できるかを検証した結果。横軸がテストトライアルでの運動方向，縦軸が発揮できた運動学習効果を表している。点線が実データ($N = 6$, mean $\pm$ s.e.m.)，実線が運動プリミティブモデルによりフィッティングした結果。

道具も含む)を制御対象とした模式図で捉えることができる(図 3a)。はじめに、運動を実行する前の段階(運動準備段階・運動計画段階)において目標となる運動 d が脳内で表現される。それに応じて脳活動 A が定まり、脳内相互作用や脳-脊髄相互作用の影響 W を受けて最終的に運動指令 x を身体に送り運動が実行される。予想とは異なる結果が得られたとき(例えば目標となる運動を達成できなかったとき)、視覚情報や体性感覚情報などの感覚情報を通じて予測誤差 e を把握することができる(本稿では予測誤差のことを運動誤差、単に誤差と呼ぶこともある。予測誤差と運動誤差は本来異なるものであるが、ここでは単純のためどちらも同じ意味であることを想定する)。例えばボールを投げた後に思い描いた場所にボールが到達しなかったときには予測誤差を視覚的に確認することができ、体操競技において自らが思い描いた動作ができなかったときには予測誤差を体性感覚を通じて確認することができる。素早く正確な運動に着目したとき、この予測誤差を小さくするためにフィードフォワード制御器が学習されていく過程を運動学習と見なすことができる。

運動学習過程を検証するためにはどのような実験系を用いるべきであろうか。我々は日常生活において予測誤差を経験することは稀である。そこで、しばしば予測誤差を人工的に付与することで、運動学習過程を観察する。例えば下記の実験系はその一例である(図 3b~d)。被験者は椅子に座り、マニピュランダムを握る(しばしば右利きの被験者を募集し右手で握ってもらう)。上体を動かさないようにベルトで固定する。また、手首を動かさないように固定する。そしてモニターを見つつマニピュランダムを動かしてもらう。マニピュランダムはコンピュータマウスとほとんど同じ役割を果たすものと思っていた方がいい構わない。すなわち、被験者にはマニピュランダムを動かしてモニターに映っているカーソルを動かしてもらう。我々が日常生活で使用しているような形でモニターを設置することもあるが、モニター画面が地面と水平になるように置くことが多い。そのモニターの下でマニピュランダムを動かしてもらうため、被験者は実際のマニピュランダムの動きを視覚的に確認することはできない。また、多くの場合では仮想的な床を設定するなどして、マニピュランダムは上下方向には動かず、身体から見て左右方向、前後方向のみに動くように設定する。すなわち、3次元の動きではなく、2次元の動きに限定して実験を行う。通常の場合ならば、マニピュランダムを前方向に動かせばカーソルは前方向に動き、右方向に動かせばカーソルも右方向に動く。まずはモニター上に黄色いカー

ソルとして提示されている初期位置におおよそ2秒ほどカーソルを合わせてもらい(図 3c)、そして10 cmの距離のところにターゲット(黒色のカーソル)を提示し、そのターゲットに向かっておおよそ500 msec以内にカーソルを動かしてもらうことが被験者に課せられる課題である。そしてしばしばマニピュランダムが自動的に初期位置まで戻るように設定されている。この一連の動作を1トライアル(もしくは1試行)と定義し、以下話を進めていく。課題中、可能な限りターゲットに向かって真っ直ぐな軌道を描くようにカーソルを動かすことを被験者は要求される。そんなの簡単じゃないか、腕を伸ばしてマニピュランダムを動かせばいいだけだろう、と多くの方は感じるかもしれない。実際に経験すると意外とこの課題は難しい。我々の運動系はこんな簡単な運動を生成する際にもばらつきを消すことはできないのだ、ということを痛感する。この状況に十分慣れてもらった後に、突然マニピュランダムの動きとは異なるカーソルの動きを提示する。すなわち、外乱を印加することで予測誤差を人工的に生成する。しばしば用いられる課題は視覚回転外乱である。視覚回転外乱では、マニピュランダムの位置に回転行列を掛けた値をカーソルの位置としてモニターに表示する(図 3b, d)。これにより、ターゲットに向かってカーソルを動かそうと思っていたものの、突如としてそれが達成できなくなる。図 3a と照らし合わせると、制御対象が突如変化して思い描いたカーソルの運動と実際のカーソルの運動が異なり、予測誤差が生じることに相当する。そしてトライアルを重ねる毎に被験者が予測誤差を減少させるように運動学習する様子を観察することができる(図 3e)。この運動学習曲線は、 t トライアル目の運動方向を x_t (ターゲットの角度とカーソルの運動方向のずれを x_t と定義する)、 t トライアル目の予測誤差を e_t とすると、状態空間モデル

$$x_{t+1} = Ax_t + Be_t \quad (4)$$

で表現できることが知られている(図 3e, 正確な統計モデルとしての状態空間モデルでは状態ノイズ、更には観測モデルが定義されるべきであり、そのような正確な統計モデルに基づいた議論もあるが¹⁹⁾、本分野の通例に倣いここでは状態ノイズ、観測ノイズ共に0、観測行列は単位行列とするような状態空間モデルを用いる)。上記の視覚回転外乱では多くの場合 t トライアル目に印加する回転外乱の角度を p_t (被験者にとっては未知の値であるが、実験者にとっては自身が設定する値でありデータ解析において既知の値)、実際に描いたカーソルの軌道と初期位置とターゲットを結ぶ直線と

のずれを e_t とし、運動指令は $x_t = p_t - e_t$ として計算することができる。すなわち、状態空間モデルは外乱 p を補正するように徐々に運動記憶を獲得する様子をモデル化している。運動方向の調節とはすなわち運動指令の調節と見なすことができるため、以下では運動方向 = 運動指令とみなして議論を進めていく。また、運動指令と誤差の値はどちらも学習の指標として用いることが可能であり、本稿ではどちらの値も用いる。運動指令の調節とはすなわち外乱 p を補正するような運動記憶の形成に外ならない。そのため、 x_t を運動記憶と呼ぶこともできるであろう。本来、運動方向や誤差の値はトライアル中においても時々刻々と変化する値ではあるが、先行研究に倣い単純化²⁰⁾、手先の運動速度がピーク値をとった時刻での値に着目して話を進めよう。これはかなり大胆な単純化である。すなわち、本来は運動方向、誤差は運動実行中の時刻の情報を含んでいるはずであるものの、それらは全て速度がピーク値をとる時刻（図 3f 参照、本稿ではピーク時刻と呼ぶ）での値で置き換えてしまおうというものである。著者の経験上、軌道全体を考えてもピーク時刻での値のみを考えても、得られる結果は定性的に変わらない（これらは違うという議論も存在する²¹⁾）。(4) 式における A は忘却率と呼ばれ ($0 \leq A \leq 1$)、試行の間に忘れてしまう運動記憶の量を定量化するパラメータであり、 B は学習率と呼ばれ ($B \geq 0$)、各試行において獲得する運動記憶の量を定量化するパラメータである。状態空間モデルや後述するその発展モデル（複数時定数モデル）は腕の到達運動のみならず、歩行運動²²⁾や眼球運動²³⁾など広範な運動学習パラダイムに適用可能であることが報告されている。

以上のように、制御対象を人工的に変化させることで予測誤差を付与し、運動学習過程を観測する。著者が提案する新たな運動学習モデルの妥当性を視覚回転外乱実験を用いて示しているため、ここでは視覚回転外乱の例を説明した。その他には今まで経験したことがない力を人工的に付与する力場学習²⁴⁾という実験系も広く用いられている。

4. 身体運動学習の数理モデル

この一連の運動学習過程を扱う数理的枠組みの一つが運動プリミティブ^{6, 12, 20, 25)}である（図 3g）。“運動プリミティブ”という言葉は脊髄のとある領域を刺激した際に誘発される動きを表すこともあるが²⁶⁾、ここでは主に運動プリミティブは脊髄の神経活動ではなく脳内での神経活動 A により表現されているであろうと想定されている。

運動プリミティブの枠組みでは、フィードフォワード制御器 $f(d)$ を d に関する非線形関数である脳活動 $A(d)$ の重み付け和 $\sum_{i=1}^N W_i A_i(d)$ で表すことを考える。なぜならば、我々の身体運動学習は下記の 2 つの特徴を示し、運動プリミティブはこれらの特徴を同時に説明することができるためである。その特徴とは、

1. 我々の身体は非線形なダイナミクスを示すため ((1)~(3) 式)、生成されるべき運動指令は関節角度に依じた非線形な関数であること、

2. 運動学習における学習曲線は線形の漸化式で表すことができること（図 3e, (4) 式）、である。すなわち、運動制御の非線形性、運動学習の線形性を同時に担保できる枠組みの一つが運動プリミティブである。運動制御の非線形性に関しては本稿においては深く触れることはないものの、非線形な関数 $A_i(d)$ の線形和 $\sum_{i=1}^N W_i A_i(d)$ は d に関する任意の関数を任意の精度で近似できることが数学的に証明されている²⁷⁾。したがって、運動プリミティブの枠組みは重み付け和のパラメータ W をうまく調節することで非線形な身体ダイナミクスに対して適切な運動指令を生成することが可能である。

本分野において初めてこのような枠組みを導入した研究⁶⁾に倣い、議論を進めていく。当該研究では、運動プリミティブの活動を定める入力は、関節角度ではなく、初期位置からターゲットまでを直線で結ぶ手先の理想軌道であると想定している。目標となる関節角度軌道の方がより良い入力である可能性は全く否定できないが、ここでは単純のため、手先の初期位置からターゲットまでを直線で結ぶ理想軌道が入力であると想定して話を進めよう。更に、前述の状態空間モデルのように、ピーク時刻での理想軌道、運動方向、誤差の値を代表値として議論を進める。

ピーク時刻での t トライアル目の理想軌道はターゲットの角度 θ と等価であると考えられるため、運動計画段階、運動準備段階において運動プリミティブの活動はベクトル $\mathbf{A}(\theta) = (A_1(\theta), \dots, A_N(\theta))$ として決定される (N は運動プリミティブの数)。そして運動実行において、運動指令 $x_t(\theta) = \sum_{i=1}^N W_i A_i(\theta)$ が決定する。外乱 p (例えば視覚回転外乱ならば $p = 30^\circ$ など) により予測誤差 $e = p - x$ が生じたとき、2乗誤差 $E = \frac{1}{2}e^2$ を最小化するように、最急降下法 $W_{i,t+1} = W_{i,t} - \eta \frac{\partial E}{\partial W_{i,t}}$ により結合荷重を学習する。ただし、 η は学習率である。最急降下法を用いる妥当性は示されていないものの、通例に倣いここでは最急降下法を想定する。多くの場合、 W に関する忘却率 λ も導入し、

$$W_{i,t+1} = \lambda W_{i,t} + \eta e_t A_i(\theta) \quad (5)$$

として学習方程式を立てる．運動学習中には2乗誤差以外にも最適化している値はあると期待されるが，その周りの議論は拙著²⁸⁾を参照していただきたい．式(5)に $A_i(\theta')$ を掛けて(θ' は次の試行において提示されるであろうターゲットの角度を表す)，全ての i に対して足すと，

$$x_{t+1}(\theta') = \lambda x_t(\theta') + \eta e_t \sum_{i=1}^N A_i(\theta) A_i(\theta') \quad (6)$$

という運動指令に関する状態空間モデルを得ることができる． t トライアル目のターゲットの角度は θ であるため $x_t(\theta')$ は実際に観測されていない運動指令であるが，運動プリミティブの枠組み上ではその運動指令を(6)式に基づき推定することができる．注目していただきたい点は，(5)式は神経活動を模擬した運動プリミティブに関するパラメータ，すなわち行動実験のみでは観測できないミクロなパラメータであることに対し，(6)式は行動実験で観測できるマクロな値である運動指令に関する学習方程式である．運動プリミティブの枠組みはミクロな世界とマクロな世界をつないでくれる枠組みであることを意味しており，この枠組みを通じて，観測した行動からミクロな現象に対する予測が立てられるのではないかという(淡い)期待を抱かずにはいられない．

5. 汎化関数

さて，(6)式の意味を更に突き詰めていこう．今，実験中にターゲットは1箇所しか現れないと仮定してみよう(そのような実験も多々ある)．すると， $\theta' = \theta$ となり，(6)式は

$$x_{t+1}(\theta) = \lambda x_t(\theta) + \eta e_t \sum_{i=1}^N A_i^2(\theta) \quad (7)$$

と書くことができる．これは忘却率 λ ，学習率 $\eta \sum_{i=1}^N A_i^2(\theta)$ の状態空間モデルである($\sum_{i=1}^N A_i^2(\theta)$ は N 個の要素の足し算を含むため，適切なスケールが必要である)．一方， $\theta' \neq \theta$ のとき，(6)式が意味することは，ターゲット θ において運動学習した学習効果は，忘却率 λ ，学習率 $\sum_{i=1}^N A_i(\theta) A_i(\theta')$ の状態空間モデルを通じてターゲット θ' への運動指令に反映されるということである(忘却率が θ, θ' に依存しないことに違和感を感じる読者の方々は拙著²⁸⁾を参照していただきたい)．すなわち， $\sum_{i=1}^N A_i(\theta) A_i(\theta')$ が大きければ大きいほど， θ への運動で学習した学習効果

は θ' の運動へ転移し，小さければ小さいほど転移しない．この $\sum_{i=1}^N A_i(\theta) A_i(\theta')$ は汎化関数と呼ばれている．実際のところ，データを計測すると図3hのような汎化関数を計測することができる．この汎化関数を説明できる条件の一つが，プリミティブの活動を周期ガウス関数

$$A_i(\theta) = \exp\left(-\frac{1}{2\sigma^2} \|\theta - \varphi_i\|^2\right) \quad (8)$$

を用いてモデル化することであり，確かにデータにフィッティングすることができる(図3h, wrapped Gaussian関数でも von-Mises関数でも良いだろう)．ただし， $\varphi_i \in (0, 2\pi]$ は i 番目のプリミティブの最適方位であり， $\theta = \varphi_i$ となるターゲットにおいてこのプリミティブは最も高い活動度を示し， $\theta = \varphi_i + \pi$ となるターゲットにおいてこのプリミティブは最も低い活動度を示す．そして $\|\theta - \varphi_i\| = \|\theta - \varphi_i + 2\pi\|$ であり， $-\pi \leq \theta - \varphi_i < \pi$ においては $\|\theta - \varphi_i\| = \theta - \varphi_i$ とする．また，しばしば最適方位は一様分布していることを想定する(本来一様分布ではないことは既に知られているが²⁹⁾，あくまで単純化するためである)．上記の条件の下で周期ガウス関数をガウス関数とみなして議論をすると(σ がある程度小さければ正しい想定となる)， $N \rightarrow \infty$ においてサンプル平均は分布平均に置き換えることから，汎化関数は

$$\begin{aligned} \sum_{i=1}^N A_i(\theta) A_i(\theta') &= \\ \frac{N}{2\pi} \int_0^{2\pi} d\varphi \exp\left(-\frac{1}{2\sigma^2} \left(\varphi - \frac{\theta + \theta'}{2}\right)^2 - \frac{1}{4\sigma^2} (\theta - \theta')^2\right) \end{aligned} \quad (9)$$

より，

$$\sum_{i=1}^N A_i(\theta) A_i(\theta') = \frac{N}{2} \sqrt{\frac{\sigma^2}{\pi}} \exp\left(-\frac{1}{4\sigma^2} (\theta - \theta')^2\right) \quad (10)$$

を得る．この汎化関数が意味することは， θ への運動で学んだ学習効果はターゲット θ への運動において最も発揮され， θ から大きく離れたターゲットへの運動では学習効果が発揮しづらくなるというものである．これは実際の行動データを説明しうることは図3hに示した通りである．

6. 運動学習実験から得られている様々な知見とそのモデル

運動学習の数理モデルは運動プリミティブだけでは

なく、様々提案されている。幾つかの数理モデルを紹介する。

6.1 ベイズモデル

手が見えない状況において、自らが動かした手の位置を推定させる実験を課したとき、それらのデータはカルマンフィルタにより説明できることが報告されている³⁰⁾。すなわち、体性感覚や視覚などの感覚系を通じて観測された感覚情報を利用して我々は手先の位置を確認することができるが、感覚系をはじめとした、神経情報処理過程にはノイズが含まれる（それは同じ刺激を加えてもニューロンの活動は毎回異なることなどに由来するであろう）。ノイズを含む情報のみから手先の位置を推定しては、不確実性の高い推定値しか得ることができない。おそらく我々は手先の動きが従うダイナミクスのモデルを予め事前知識として獲得しており、そのモデルに基づいた予測情報と、体性感覚や視覚などの感覚系を通じて観測された感覚情報を統合して、手先の位置の推定を行っているであろう、と示唆した結果である。この研究の概念を拡張し、ベイズ推定の枠組みが運動学習に適用できることを示唆した研究が2004年に報告された³¹⁾。それを紹介しよう。

ベイズ統計に詳しい方には当たり前のことではあるが、ベイズの枠組みは全ての現象を説明できるだけの説明能力を有しており³²⁾（とても複雑な階層ベイズを考えれば明らかであろう），“行動実験の結果をベイズ推定の枠組みで説明できる”ことは自明である。しかしながら、ある実験結果に基づいたベイズモデルが他の実験結果を説明できるか否かを検証することは脳がどのような事前知識を有しているのか検証するためには有効なアプローチであり、本当に不確実性という情報を利用しているのか、脳を模擬したニューラルネットワークモデルにはどのような事前知識が埋め込まれているのか、など興味深い仮説を検証するためには有効な考え方である。

閑話休題、運動の修正量は不確実性に影響されることを示唆する研究成果が報告された³¹⁾。用いた実験系は前述の腕の到達運動課題である（動きの検出にはモーションキャプチャーを用いている）。しかしながらワークスペースは20 cmと前述した腕の到達運動実験を比べると広く、腕の軌道が10 cmのところ到達した時のみ手先の位置を視覚情報としてモニター上にフィードバックする。例えば、実際の腕の位置よりも運動方向とは垂直な方向に1 cmずらすことで、中間地点から終点位置に到達するまでに得られた予測誤差に応じて、どの程度運動を修正したのかを計測することができる。更に、視覚的にフィードバックするカーソルに

ブラーをかけ（正確な実験系は当該研究³¹⁾を参照していただきたい）、視覚情報を曖昧にする。被験者は自らの手先の位置 x_p （不確実性 σ_p^2 ）と視覚的に与えられたカーソル位置 x_v （不確実性 σ_v^2 ）に基づき、正確なカーソル位置ないしは正確なカーソル位置と手先位置のずれ $\hat{x}$ の推定を行い、適切に運動を修正するであろうと想定されている。つまり、体性感覚情報としての手先位置 x_p と視覚情報としてのカーソル位置 x_v は共にガウス分布 $\mathcal{N}(x_p|\hat{x}, \sigma_p^2)$, $\mathcal{N}(x_v|\hat{x}, \sigma_v^2)$ で表現されていると想定すると ($\mathcal{N}(y|\mu, \sigma^2)$ は平均 μ 、分散 σ^2 の y に関するガウス分布)、 $\hat{x}$ の事後分布は

$$\mathcal{N}(\hat{x}|x_v, x_p, \sigma_v^2, \sigma_p^2) \propto \mathcal{N}(x_v|\hat{x}, \sigma_v^2) \mathcal{N}(x_p|\hat{x}, \sigma_p^2) \quad (11)$$

となり ($\propto$ は比例するという意味であり、これは規格化定数を見捨れているため比例記号を用いている)、この事後分布を最大化する $\hat{x}$ の推定量は

$$\hat{x} = \frac{\sigma_v^2}{\sigma_v^2 + \sigma_p^2} x_p + \frac{\sigma_p^2}{\sigma_v^2 + \sigma_p^2} x_v \quad (12)$$

として得ることができる。つまり、我々が視覚情報と体性感覚情報をベイズ最適に統合しているならば、 σ_v^2 が大きくなればなるほどに $\hat{x}$ は x_p に近い値を示し、視覚的に提示されたカーソル位置の情報は無視するようになり、この実験系では運動を修正しなくなるであろうと予想される。

事実、視覚情報が不確実であればあるほど、運動の修正量が小さくなることが報告された（図11aに実データを掲載している。実データは <http://crcns.org/datasets/movements/dream/> からダウンロード可能）。前述の状態空間モデルや運動プリミティブモデルでは、例えば視覚情報が曖昧なときには誤差情報が $e + \xi$ というようにノイズで濁されたものと想定してみても、このノイズは運動の修正量には平均的には影響を与えないため（2乗予測誤差を計算すれば明らかではあるが、残留誤差を大きくするだけである）、不確実性の影響を説明することはできない。既存の枠組みでは説明することができない不確実性の影響が、実際に実験的に示唆されたのである。

また、(12) 式に基づくと、体性感覚情報の不確実性 σ_p^2 を上げることで、 $\hat{x}$ は x_v に近い値となり、視覚情報に依存した運動の修正が行われるように思える。実際、視覚回転外乱の外乱値を気付かれない程度にトライアル毎に変動させることで体性感覚情報の不確実性を向上させ、より視覚情報に依存した運動の修正が見られることが報告されている⁹⁾。

6.2 複数時定数モデル

2006年に提案され、今でも多くの運動学習研究者に影響を与え続けているモデルが複数時定数モデルである⁷⁾。この複数時定数モデルは、“運動記憶を獲得する速度は速いが、忘却する速度もまた速い”とする速い過程と“運動記憶を獲得する速度は遅いが、忘却する速度もまた遅い”とする遅い過程を含む。

この複数時定数モデルは状態空間モデル((4)式)を発展させたものである。従来の状態空間モデルでは、いくつかの眼球運動(サカド)に関連する実験結果を再現できない(腕の到達運動に着目しておいていきなり眼球運動の話をするのはいささか唐突ではあるが、実際にこのモデルはそのように導入されているため致し方あるまい)。再現できない現象とは、同じ外乱を2回学習する実験において、1回目の学習速度よりも2回目の学習速度の方が速い、というセービングスという現象である³³⁾。通常、課題を始める前の運動指令は $x_0 = 0$ を想定する。通常の状態空間モデル((4)式)では、 $x_t = 0$ となるトライアルが存在する場合、それは課題を体験する前と全く同じ状態に戻ってしまったと解釈される。一方、我々は一度経験したことがある課題ならば、その課題を学習した運動記憶は一定時間保持されているであろうと期待できる。例えばテニスを初めて経験した時と2回目に経験した時とでは、上達度合が異なることを体感できるであろう。この記憶の保持を従来の状態空間モデルでは表現することができない。

記憶の保持という概念を導入してセービングスを再現しようと試みたモデルが複数時定数モデルである。 t トライアル目の運動指令 x_t は速い過程により獲得された運動指令 $x_{f,t}$ と遅い過程により獲得された運動指令 $x_{s,t}$ の足し算

$$x_t = x_{f,t} + x_{s,t} \quad (13)$$

であると想定する。各々の過程は

$$x_{f,t+1} = A_f x_{f,t} + B_f e_t \quad (14)$$

$$x_{s,t+1} = A_s x_{s,t} + B_s e_t \quad (15)$$

に従いトライアル毎に変化すると想定される。ただし、 $0 < A_f < A_s < 1$ かつ $0 < B_s < B_f$ である。すなわち、速い過程は素早く誤差を減らすように学習が進むものの($B_f > B_s$)、運動記憶はトライアル毎に素早く忘れていってしまう($A_f < A_s$)。一方、遅い過程はゆっくり誤差を減らすものの、運動記憶を忘れていく速度は遅い。このようにモデル化をすると、図4aのようにセービングスを再現することができる。重要な

点は、 $p > 0$ 外乱ありトライアル $\rightarrow p < 0$ 外乱ありトライアル $\rightarrow p > 0$ 外乱ありトライアルと変化する際、 $p < 0$ 外乱ありトライアルにおいて遅い過程に $p > 0$ の外乱を補正する運動記憶が保持されていることである(図4a 水色の線)。これにより、再び $p > 0$ 外乱ありトライアルを経験した時に、 $p > 0$ を補正する運動記憶が蘇り素早い運動学習を達成することができる。

更に、複数時定数モデルならではの予測が行動実験により実証された(モデルは眼球運動の実験結果に触発されているが、行った実験は腕の到達運動の実験である)。その予測とは自発的回復である。自発的回復とは、一度忘れたように見える記憶が再び蘇ってくる現象である。自発的回復自体は古典的条件付けの分野などで古く知られているものの、運動記憶においては実証されていなかった。複数時定数モデルを用いて、 $p > 0$ の外乱ありトライアル $\rightarrow p < 0$ の外乱ありトライアル $\rightarrow$ エラークランプトライアルという実験をシミュレーションしてみよう(図4b)。エラークランプトライアルとは、細かい議論はさておき(詳しく知りたい方は文献³⁴⁾を参考にいただきたい)、強制的に $e_t = 0$ として誤差を0に固定するトライアルであると考えていただきたい。複数時定数モデルのシミュレーションからは、エラークランプトライアルが始まった直後は $x_t < 0$ となるものの、エラークランプトライアルを続けていくとあたかも $p > 0$ の外乱ありトライアルの記憶が蘇ったかのように、 $x_t > 0$ と変化していく(図8bに実データが掲載されているため、参照されたい)。複数時定数モデルは自発的回復を初めて予測できるモデルであり、自発的回復の実証により複数時定数モデルは市民権を獲得するに至った。

7. 誤差の予測モデル

以上、状態空間モデル、運動プリミティブモデル、ベイズモデル、複数時定数モデル、と様々なモデルを紹介した。しかしながら、果たしてどのモデルが最も我々の身体運動学習の脳機能を表現できているモデルなのだろうか。運動プリミティブモデルは身体制御の非線形性と運動学習の線形性を両立させる枠組みであるが、ベイズモデルのように不確実性を考慮することはできず、複数時定数モデルのようにセービングスや自発的回復を再現することはできない。同様にして特定のモデルは特定の現象しか再現できず、どのモデルが最も我々の身体運動学習の脳機能を表現できているモデルなのかは全く定かでない。更に、これらのモデルでは再現できない現象が報告されている。それが、構造学習またはランダム学習である³⁵⁾。ランダム学習では、

図4 複数時定数モデル. (a): セービングスのシミュレーション結果. 横軸がトライアル数 t , 縦軸が運動指令 x_t を表す. 黒い点線が外乱 p_t を表している. 通常の状態空間モデル (緑実線) ではセービングスを再現することができないものの, 複数時定数モデル (赤実線) ならば再現することができる. (b): 自発的回復のシミュレーション結果. 複数時定数モデルはエラークランプトライアル試行において, 直前に学んでいた $p = -1$ の運動記憶ではなく, それ以前に長時間学んでいた $p = 1$ の運動記憶が蘇ることを予測する. 実際の実験結果は図 8b に掲載しているため, 参照していただきたい.

外乱がランダムに変化するもののその平均が 0 であることを想定する. つまり, 外乱 p が確率分布 $P(p)$ からサンプリングされることを想定し, K 回シミュレーションを行ったときの t トライアル目の外乱の平均値を $E[p_t] = \frac{1}{K} \sum_{k=1}^K p_{t,k}$ とする. $E[p_t] = 0$ の外乱を経験したときに通常の状態空間モデルでは,

$$E[x_{t+1}] = AE[x_t] + B(E[p_t] - E[x_t]) = (A - B)E[x_t] \quad (16)$$

かつ $x_0 = 0$ より, ランダムな外乱からは何も学ばないという予測が生じる. 同様に上記の運動プリミティブ, ベイズモデル, 複数時定数モデルでもランダムな外乱からは何も学ばないという予測が生じる. しかしながら, $E[p] = 0$ かつ $E[p^2] \neq 0$ となる外乱を経験することで, その後印加される外乱への適応が速くなることが報告された³⁵⁾. これがランダム学習という現象である (詳細な議論は先行研究³⁵⁾を参照していただきたい).

7.1 数学的考察

今回紹介する我々の研究¹²⁾では, 運動プリミティブの枠組みにおいて今まで再現されていなかったランダム学習を再現するための数学的条件を探り, 新たなモデルを提案した. 運動プリミティブの学習則 ((5) 式) から, 十分トライアル数を経た後の平衡状態 ($W_{i,t+1} = W_{i,t} = W_i$ となった状態) を計算すると,

$$\mathbf{W} = \eta E[p\mathbf{A}^T] \left((1 - \lambda)\mathbf{I} + \eta E[\mathbf{A}\mathbf{A}^T] \right)^{-1} \quad (17)$$

を得る. ただし, $\mathbf{W} = (W_1, \dots, W_N)$ であり, $\mathbf{A}$ を決定する入力情報は未知であることを想定しているため, $\mathbf{A}(\theta)$ ではなく $\mathbf{A}$ と記している. また, ここでは, K 回のシミュレーションにおいて $\mathbf{W}$ は毎回同じ結合荷重に収束するという仮定, もしくは $\mathbf{W}$ のトライアル変化量は $\mathbf{A}$ のトライアル変化量よりも十分速いもしくは遅いという仮定をおき, $E[\mathbf{W}\mathbf{A}\mathbf{A}^T] = \mathbf{W}E[\mathbf{A}\mathbf{A}^T]$ を想定した. この想定自体は本質的ではなく本来仮定しなくても良いものではあるが, 見通しが立てやすいため本研究ではこのように想定した. (17) 式は,

$$E[p\mathbf{A}^T] \neq 0 \quad (18)$$

がランダム学習を再現するための十分条件であることを意味している. 既存の運動プリミティブモデルは運動プリミティブの活動は目標の運動方向により決定されること ($\mathbf{A} = \mathbf{A}(\theta)$) を想定しており, θ_t と外乱 p_t の値は無相関であるため $E[p\mathbf{A}^T] = 0$ となりランダム学習を再現することはできなかった.

(18) 式が達成される条件を探ってみよう. それはすなわち, $\mathbf{A}$ を決定する入力は何か, という問題を考えるに等しい. (18) 式が達成されるためには, $\mathbf{A}_t$ と p_t が関連していれば良いため, 最も単純な考え方は $\mathbf{A} = \mathbf{A}(p)$ として外乱の大きさにより運動プリミティブの活動が決定するモデルである. 通常, 被験者に対して外乱の

大きさは観測できない隠れ変数であり、脳活動を決定する入力として隠れ変数を用いることは明らかにおかしい。また、このモデルは後述するように (図 9) モデルとしての再現能力は低いため、妥当なモデルでないことが容易に想像できる。では、どのような情報が運動プリミティブの活動を決定しているのであろうか。外乱 p の情報を含む何か $\mathbf{A}$ を決定する入力情報であれば (18) 式は達成されるであろう。

外乱 p の情報を含む何か、と言われて思い浮かぶものは予測誤差 $e = p - x$ である。外乱 p は隠れ変数であるが、予測誤差は被験者が観測できる値である。事実、誤差の情報により決定する神経活動も報告されている³⁶⁾。そのため、誤差が $\mathbf{A}$ を決定するモデルはそこまで悪くないと期待できる。しかしながら、そのモデルは大きな問題を含んでいる。 $\mathbf{A}$ は運動開始前に決定されるべき値である一方、誤差 e は運動が終了した後に観測できる値であり、誤差 e が $\mathbf{A}$ を決定することは不可能である。そこで我々は、「将来に渡り観測されるであろうと予測される誤差 (Prospective error) P_t 」が運動プリミティブの活動を決定するというモデルを提案した。それはすなわち、誤差の予測値を $\hat{e}_t$ として

$$P_t = \sum_{r=0}^{\infty} \gamma^r \hat{e}_{t+r} \quad (19)$$

と定義される。 P_t は

$$\hat{e}_t = P_t - \gamma P_{t+1} \quad (20)$$

を満たすため、 P_t の予測誤差 $e_t - \hat{e}_t$ を用いて P_t が

$$P_{t+1} = P_t + \alpha(e_t - \hat{e}_t) \quad (21)$$

に従い計算できる。そしてこの P_t が運動プリミティブの活動を決定することで、ランダム学習が再現できると期待できる。

以下では最も単純な場合 $\gamma = 0$ を想定して

$$\hat{e}_{t+1} = \hat{e}_t + \alpha(e_t - \hat{e}_t) \quad (22)$$

$$A_{i,t} = \exp\left(-\frac{1}{2\sigma^2}\|\hat{e}_t - \varphi_i\|^2\right) \quad (23)$$

$$W_{i,t+1} = \lambda W_{i,t} + \eta e_t A_i(\hat{e}_t) \quad (24)$$

として誤差の予測 $\hat{e}_t$ のアップデート、誤差の予測に応じた運動プリミティブの活動度、結合荷重が学習されていくモデルを考える (図 5a)。 φ_i は誤差の予測の最適方位であり、例えば誤差が生じないと予測されたとき ($\hat{e}_t = 0$) に最も活動する運動プリミティブもいれば、誤差が正の値を取るであろうと予測されたとき ($\hat{e}_t > 0$) に最も活動する運動プリミティブもいる、と

いう想定である。このように想定すると (18) 式を達成することができ、ランダム学習を再現することができ (図 5b, c)。厳密に言うならば、 $I = \beta p + \gamma x$ かつ $\mathbf{A} = \mathbf{A}(\hat{I})$ ならば (18) 式は達成できるが、ここでは $\beta = 1$, $\gamma = -1$ に限定して話を進める。なぜならば、小脳に誤差の予測を表現しているニューロンがいるであろう、ということが報告されているため^{37,38)}、今回は誤差の予測に焦点を絞り話を進める (我々の仮説を覆す実験が提唱されるならばそれは大歓迎である)。また、おそらく $\gamma \neq 0$ の方がデータフィッティングは良くなることが期待できるが、ここではパラメータ数が可能な限り少なくなるようにモデルを考えている。

7.2 強化学習と運動学習の対応関係

運動制御・運動学習における 2 乗誤差 $\frac{1}{2}e^2$ は、強化学習においてはマイナスをかけて報酬関数と捉えることができる³⁹⁾。つまり、 t トライアル目において得られる報酬関数 r_t を

$$r_t = -\frac{1}{2}e_t^2, \quad (25)$$

その推定量を $\hat{r}_t$ として、将来得られるであろう報酬の総量である価値関数

$$V_t = \sum_{r=0}^{\infty} \gamma^r \hat{r}_{t+r} \quad (26)$$

を最大化することで、2 乗誤差を最小化して目標の運動が達成できる制御則が得られる。この価値関数は

$$\gamma V_t - V_{t+1} = \hat{r}_t \quad (27)$$

を満たすため、報酬予測誤差、または Temporal difference error $\delta_t = r_t - \hat{r}_t = r_t - \gamma V_t + V_{t+1}$ を用いて

$$V_{t+1} = V_t + \eta \delta_t = V_t + \eta(r_t - \hat{r}_t) \quad (28)$$

として価値関数を推定することができる。

運動学習では符号まで考えた誤差すなわちベクトル誤差を考え、強化学習では誤差の符号は考えずスカラー値として誤差を考える。この違いはあるものの、著者が提案する誤差の予測 ((21) 式) と強化学習における価値関数 ((28) 式) とは同じ形のアップデート則でモデル化できるため、密接な関係にあることは明白である。異なる点は、著者が提案する誤差の予測モデルでは、 P_t は強化学習のように行動選択に用いるのではなく、運動プリミティブの活動を決定する入力情報として用いる点である。

図5 誤差の予測モデルとそのシミュレーションと実験的実証。(a): 誤差の予測モデル。次の運動における誤差の予測という新たな要素が加わっている。(b): ランダム学習のシミュレーション。横軸がトライアル数、黒い細線が外乱 p_t 、黒い太線が運動指令 x_t を表す。黄色いエリアでは外乱がランダムに変化し、青い変化では定数値の外乱、そして白いエリアでは外乱の値は0になっている。(c): 外乱が一定値をとっているトライアル(青いエリア)におけるランダム学習の結果。黄色いエリアにおいて p_t が完全にランダムに変化することを想定したグループがグループ1、2回毎に変化することを想定したグループがグループ2、3回毎に変化することを想定したグループがグループ3である。(d): 実験設定。従来の設定と同様。ターゲットは $\theta = 0$ の一箇所のみ提示される。(e, f): 用いた実験系。図5(b)のランダム学習とはほぼ同様の実験設定であるものの、仮想的なトンネルを用いてランダム学習中(図5(e)の黄色とグレーのエリア)の腕の動きはターゲットに向かって動くように固定されている。その一方で外乱の値はトライアル毎に変化するため、カーソルの動きもまたトライアル毎に変化し、予測誤差が生じる。(g): 横軸は外乱 p_t 、縦軸は次のトライアルにおける運動指令の修正量に比例した値 $C(x_{t+1} - x_t)$ を表している。ここで計測できる値は運動方向ではなく仮想的なトンネルに加えられた力の値であるため、比例定数を掛け算することで運動方向から力へと変換している。(h): 誤差の予測モデルが予測する実験結果。一定値をとる外乱への学習速度は、グループ2の被験者の方がグループ1の被験者よりも速くなることが予想される。この実験結果は誤差の予測モデルのみが予測できるものである。(i): 実際の実験結果 ($N = 24$, mean $\pm$ s.e.m.)。 (j): 得られた学習データにブートストラップを適用して取りうる全ての学習速度の値を計算した結果。横軸が学習速度の値、縦軸がその学習速度がサンプリングされた頻度を表している。各色の実線は各ヒストグラムの平均値を表している(中央値でも最頻値でも結果は本質的に変わらないことは確認済みである)。

7.3 実験的検証

上記のように数学的考察により導かれた誤差の予測モデルであるが、誤差の予測モデルの妥当性を実証しなくてはならない。なぜならば、モデルパラメータ (α) を増やしたのだから再現能力が高まることは当たり前ではないか、と言われればその通りであるためである。「誤差の予測」という要素を増やすことは妥当なことである、と実証しなくてはならない。そこで我々は、“誤差の予測モデルのみが予測できる現象を実験で示すことができるならば、それは誤差の予測モデルの妥当性を支持していると言えるであろう”と考えた。

このモデルの提案段階において、誤差の予測モデルのみが予測できる現象とは、“達成不可能な運動であれ、少なくとも2回連続で同じ運動を練習することで運動学習効果は高まる”というものである。なぜならば、1回目の運動において誤差が観測できた後に、2回目の運動開始時には誤差がある程度予測できるようになるためである。例えばランダムに変化する視覚回転外乱を考えると、 $p = 30^\circ$ の後に再び $p = 30^\circ$ となると、 30° の誤差を経験した次のトライアルではおよそ 30° の誤差が観測されるであろうと予測でき、予測と大方一致する 30° の誤差を経験することになる。そして次いで $p = -30^\circ$ の後に再び $p = -30^\circ$ となると、およそ -30° の誤差を経験した次のトライアルではおよそ -30° の誤差が観測されるであろうと予測でき、予測と大方一致する -30° の誤差を経験することになる。このように、外乱が切り替わるトライアルは全くのランダムなので予測も運動の修正もできないものの、外乱が切り替って2回目のトライアルにおいてはおよそ正しい値に誤差を予測することができ、誤差を修正することができるであろう。これを繰り返すことにより、この誤差の予測モデルにおいてランダム学習が促進されていくであろうと期待できる。

この誤差の予測モデルの予測を行動実験に基づき検証した。実験のプロトコルは、まず十分マニピュレーションに慣れてもらい、その後ベースライントライアル→トレーニングトライアル→ウォッシュアウトトライアル→テストトライアル、と進める。テストトライアルは 30° の視覚回転外乱である(図5d~f)。トレーニングトライアルでは以下の様な奇抜な状況を体験してもらっている。ターゲットは正面の一箇所 ($\theta = 0$) しか提示されない。外乱は $p = -45, -30, -15, 0, 15, 30, 45$ のうち一つが毎試行ランダムにサンプリングされる。毎試行完全にランダムなグループをグループ1、2試行は同じ外乱がサンプリングされるグループをグループ2とした。更に、仮想的にトンネルを作成し、被験者

の腕の動きはターゲットに向かって真っ直ぐ動くように固定される。すなわち、腕の動きは $\theta = 0$ に向かうように固定されるので、カーソルの動き = 外乱の角度、となる状況を作成する。カーソルがターゲットにまっすぐ向かって動き、課題が達成できる確率(外乱の大きさが0になる確率)は $1/7$ である。このような工夫は先行研究で提案されてきたモデルの予測と分離するために行っている。動かした腕の方向が運動プリミティブの活動に影響すると想定したモデルや⁴⁰⁾、課題が達成できることが運動学習に影響を与えると想定するモデル⁴¹⁾では、グループ1とグループ2の間に差はないであろうと想定される状況を作り上げている。このような一見変わった環境においても、仮想的なトンネルにかけた力を計測することで、被験者はカーソルが右にずれたら次のトライアルでは左に修正しようとしていたことがわかる(図5g)。このデータを再現するように誤差の予測モデルのパラメータフィッティングを行った結果、 $R^2 = 0.9950$ の決定係数にてフィッティングできることがわかった。被験者は $1/7$ の確率でしか課題が達成できない状況においても運動の修正を行っており、図5gが示すように外乱が大きければ大きいほど修正量は単調増加するわけではなかった。これは後述するように既に報告されている結果でもあるが、誤差の予測モデルはこの結果“も”説明できるということを強調しておきたい。

このような状況下において、誤差の予測が運動学習に関わっているならば、ランダム学習が生じているグループ2の被験者の方がグループ1の被験者よりもテストトライアルでの学習が速くなっていることが期待できる(図5h)。事実、確かに予測が正しいことが実証できた(図5i)。また、ブートストラップにより起こりうる全ての学習曲線をランダムにリサンプリングし、指数関数フィッティングにより学習速度を計算した結果、やはりグループ2の被験者の方が学習速度が速いことがわかる(図5j)。更に、同じ外乱を3回連続で経験するグループ3 ($N = 12$) の方がグループ2よりも学習速度が統計的に有意に速いことも示され、これもまた誤差の予測モデルの予測通りの結果である(ここではグループ3のシミュレーション結果は示していない)。ここでの学習速度とは、誤差のトライアル変化に対して指数関数

$$e_t = a \exp(-bt) + c \quad (29)$$

のフィッティングを行い、 b を学習速度と呼んでいる(b が大きければ誤差は素早く減少し、小さければゆっくり減少することを意味する)。以上から、誤差の予測

モデルのみが予測できる実験結果が確かに行動実験に基づき実証された。

図 5i に示したデータの内、グループ 1 の学習曲線のデータを最も説明しうる誤差の予測モデルのパラメータを推定した結果、 $R^2 = 0.8638$ の決定係数にてフィッティングが可能であった。そしてグループ 1 の実験結果のみから推定したパラメータを用いてグループ 2 の学習曲線を予測した結果、 $R^2 = 0.7967$ にて実験結果の予測が可能であった。これらの比較的高い決定係数は、誤差の予測モデルが著者らの行った実験のデータを説明、予測できていることを意味している。

7.4 シミュレーション

行動実験からその妥当性を実証した誤差の予測モデルであるが、本当に妥当性が高いならば、先行研究で報告されてきた行動実験データも説明できることが期待できる。そこで、バイズモデル、複数時定数モデルのどちらか一方で再現されてきたいくつかの現象の説明を試みた。以下のシミュレーションにおけるパラメータは全て上記の行動実験データ（グループ 1 のデータのみ）にフィッティングしたものである。

7.4.1 不確実性の影響

“バイズモデル”のセクションで前述したように、不確実性が運動学習に影響を与えるという結果が報告されている。これはすなわち、外乱学習において、視覚情報の不確実性を σ_o^2 、自らの運動のばらつきを σ_s^2 としたとき、学習率 B が

$$B = \frac{\sigma_s^2}{\sigma_o^2 + \sigma_s^2} \quad (30)$$

として決定することに等しい。すなわち、視覚情報が不確実であればあるほど誤差の情報は運動の修正に反映されないことを意味する ($\sigma_o \rightarrow \infty$ において $B \rightarrow 0$)。

誤差の予測モデルでは、ブラーがかかったことにより、誤差の観測にノイズが加わることを想定する。すなわち、誤差の項を $e_t + \xi_t$ として置き換える。 ξ_t は平均 0、分散 σ_o^2 のガウス分布から毎試行サンプリングされることを想定する。このようにして誤差の予測モデルをシミュレーションすると、図 6a, b のように不確実であればあるほど学習速度が遅くなることが説明できる（図 6c は先行研究の結果⁴²⁾）。また不確実性の影響は解析的に示すことが可能である。運動プリミティブの活動にガウス関数を想定して、 $e_t - \hat{e}_t$ が小さい状況、すなわち誤差がある程度予測できている状況を想定すると、誤差の予測モデルは漸化式

$$x_{t+1} = \lambda \left(1 - \frac{\alpha^2}{\sigma^2} (e_t - \hat{e}_t)^2 \right) x_t + \sqrt{\pi \sigma^2} \eta \left(\frac{\alpha^2}{4\sigma^2} (e_t - \hat{e}_t)^2 \right) e_t \quad (31)$$

と一致する。これは忘却率 $A = \lambda \left(1 - \frac{\alpha^2}{\sigma^2} (e_t - \hat{e}_t)^2 \right)$ 、学習率 $\sqrt{\pi \sigma^2} \eta \left(\frac{\alpha^2}{4\sigma^2} (e_t - \hat{e}_t)^2 \right)$ の状態空間モデルと見なすことができるため、誤差が予測できる時 ($\hat{e}_t \simeq e_t$) は忘却速度が遅く、学習速度が速いことがわかる。このようにして、誤差の予測モデルは不確実性という概念を必要とせず、不確実性により学習速度が変化することを説明できる。

7.4.2 セービングス

“複数時定数モデル”のセクションで紹介したセービングスであるが、実は誤差の予測モデルでも、複数時定数という概念を想定していなくてもかかわらず、セービングスを再現することができる。複数時定数モデルは上記のランダム学習、不確実性の影響を説明することはできず、我々の行動実験の結果を予測することもできない。したがって、この時点では我々の誤差の予測モデルは複数時定数モデル、バイズモデルを含み、更に予測能力、再現能力が高いモデルであると言える。

複数の時定数を想定していなくてもかかわらず、なぜ我々の誤差の予測モデルはセービングスを説明できるのであろうか。“複数時定数モデル”のセクションと同様に、 $p > 0$ の外乱ありトライアル $\rightarrow p < 0$ の外乱ありトライアル $\rightarrow p > 0$ の外乱ありトライアルという実験をシミュレーションしてみよう（図 7a）。図 7b, c は誤差の予測モデルにおける運動プリミティブの活動のトライアル変化、誤差の予測値のトライアル変化、そして結合荷重 W_i のトライアル変化を表している。図 7b, c が示すように、1 回目の $p > 0$ の外乱ありトライアルでは誤差 e_t とその予測値 $\hat{e}_t$ が正の値をとり、 $\varphi_i > 0$ となる運動プリミティブが活動することによりこれらのプリミティブに $p > 0$ を補正するような結合荷重の値が (24) 式に従い学習されていく。 $p < 0$ の外乱ありトライアルに切り替わったときには誤差も誤差の予測値も負の値をとり、 $\varphi_i > 0$ の運動プリミティブは活動しない。結合荷重の学習則 ((24) 式) が意味することは、運動プリミティブが活動しなければ結合荷重の学習は生じないということである。すなわち、 $p < 0$ のトライアルにおいては $\varphi_i > 0$ の運動プリミティブに埋め込まれた記憶は保持され続ける。そして再び $p > 0$ に切り替わった時には $\varphi_i > 0$ となる運動プリミティブが再度活動し、 $p > 0$ を補正するような運動記憶が蘇り学習が速くなる。以上が誤差の予測モ

図6 誤差の予測モデルにおける不確実性の影響. (a): 誤差に加えるノイズの標準偏差を様々に変化させ、シミュレーションを行った結果. (b): 図6aの学習曲線に状態空間モデルをフィッティングして算出した学習率. (c): 同様の状況において報告されている実データ⁴²⁾.

図7 誤差の予測モデルによるセービングスのシミュレーション. (a): 誤差の予測モデルを用いてセービングスのシミュレーションを行った結果. 以下、青いエリアは $p > 0$ のトライアル、赤いエリアは $p < 0$ のトライアルを表す. 黒い実線は外乱 p_t ないしは運動指令 x_t , 赤い実線は同じ外乱を2度目に学習した際の学習曲線を表す. (b): 横軸はトライアル、縦軸は各々の運動プリミティブのもつ φ_i の値ないしは誤差の予測値 $\hat{e}_t$ を表す. 赤い実線は誤差の予測値のトライアル変化、黒と白の色の変化は運動プリミティブの活動のトライアル変化を表す. 白色は高い運動プリミティブの活動度、黒色は低い運動プリミティブの活動度を表す. (c): 結合荷重のトライアル変化. 横軸はトライアル、縦軸は各々の運動プリミティブのもつ φ_i の値を表す. 水色は $p > 0$ を学習した結合荷重の値 ($W_i > 0$), 紫色は $p < 0$ を学習した結合荷重の値 ($W_i < 0$) を示す.

デルがセービングスを再現できる理由である.

7.4.3 自発的回復

複数時定数モデルの証拠であると思われていた自発的回復であるが、我々の誤差の予測モデルでも再現できる (図8). $p > 0$ の外乱ありトライアルを長く続けることで誤差も誤差の予測もほぼ0となり、 φ_i が0に近い値をもつ運動プリミティブに $p > 0$ を補正するような結合荷重が学習されていく. 自発的回復を検証する実験においては $p < 0$ の外乱ありトライアルは短く設定され、誤差は0にならない程度でエラークランプトライアルに切り替える. エラークランプトライアルではもちろん誤差は0になり、予測される誤差も0になるため、 φ_i が0に近い値をもつ運動プリミティブに埋め込まれた $p > 0$ を補正する運動記憶が蘇る. これ

が誤差の予測モデルから説明する自発的回復である.

また、誤差の予測の代わりに外乱の予測値が運動プリミティブの活動を決定するモデルをシミュレーションした結果、自発的回復は再現できないことがわかる (図9). そのため、自発的回復をシミュレートすることはモデルの妥当性を検証するための有効な手段であると言える.

7.4.4 順行性干渉

$p > 0$ の外乱ありトライアル $\rightarrow$ $p < 0$ の外乱ありトライアルと変化させると、 $p < 0$ の外乱ありトライアルの学習速度は $p > 0$ の外乱ありトライアルでの学習速度よりも遅くなる. $p > 0$ を補正する運動の記憶と $p < 0$ を補正する記憶とが互いに干渉しあうために生じる現象であると考えられており、この現象を順行性

図8 誤差の予測モデルによる自発的回復のシミュレーション. (a, c, d):セービングスのシミュレーションと同様. 灰色のエリアは強制的に $e_t = 0$ と設定するエラーランプトライアルを意味する. (b):自発的回復の実データ⁷⁾.

図9 外乱の予測モデルによる自発的回復のシミュレーション. (a):外乱の予測モデルではエラーランプ中にどのような外乱の値が推定されるか未知であるため, $p = 0$ とみなす場合と $p = -30^\circ$ と見なす場合の2種類をシミュレーションした. (b, c):エラーランプ中に $p = 0$ と設定した場合の外乱の予測値, 運動プリミティブの活動度, 結合荷重のトライアル変化. (d, e):エラーランプ中に $p = -30^\circ$ と設定した場合の外乱の予測値, 運動プリミティブの活動度, 結合荷重のトライアル変化.

図10 誤差の予測モデルによる順行性干渉のシミュレーション. (a, b, c): セービングス, 自発的回復のシミュレーションと同様. (d): $p > 0$, $p < 0$ を学習している際の学習曲線の比較. (e): 図10aの青いエリアのトライアルの長さとは $p < 0$ の学習速度との関係性. (f): 順行性干渉の実データ⁴³⁾.

干渉という. この順行性干渉は複数時定数モデルで説明されてきたが⁴³⁾, 誤差の予測モデルでもまた説明することができる (図10). 解釈は上記の自発的回復とはほぼ同様である.

7.5 データフィッティング

行動実験データの予測能力, 様々な現象を再現できる再現能力から妥当性を検証した誤差の予測モデルであるが, もし誤差の予測モデルの妥当性が高いならば, 先行研究で計測された行動実験データへもフィッティングできることが期待できる.

7.5.1 不確実性に関する実験結果

“ベイズモデル”のセクションで前述した行動実験データ³¹⁾へのフィッティングを行った (図11a). ただし, 前述したように20 cmの腕の到達運動であるため我々の行動実験とはセッティングが異なる. そのため残念ながらモデルパラメータはこの実験データに適したものになるように推定している. この実験でもまたブラーを用いているため, 前述のようにブラーは誤差にガウスノイズを加え, ブラーの大きさはガウスノイズの不確実性により表現している. 図11aに示すように, 誤差の予測モデルは先行研究の結果を高い精度で説明することができる. $R^2 > 0.93$ の決定係数にて4つの条件の実験データにフィッティング可能であることがわかった. これらはモデルシミュレーションや(31)式と同様に, ブラーが大きければ大きいほど誤差

が予測しづらく学習速度が遅くなる, という解釈にて説明することができる.

7.5.2 誤差の大きさと修正の大きさに関する実験結果

ある先行研究において, 平均0となる一様分布から外乱の値 p_t をランダムにサンプリングして次のトライアルにおける運動指令の修正量 $x_{t+1} - x_t$ を計測した結果 (視覚フィードバックされるカーソルの位置が実際のカーソルから $-8, -4, -2, -1, 0, 1, 2, 4, 8$ cm ずれる状況を作る), 外乱が大きい時ほど次のトライアルでの修正量が下がるということが報告された⁴⁴⁾. すなわち, 変数 k, l を用いて学習率 B が

$$B = \frac{k}{k + le_t^2} \quad (32)$$

とモデル化されることが報告された (詳細は先行研究を参照していただきたい⁴⁴⁾). 先行研究の解釈では, 誤差が大きくなればなるほど, 自分の運動が生み出した誤差ではないと気づくため, 上記のようなベイズモデルで説明できると主張されていた. これは我々の実験結果 (図5g) でも観測されていることであり, もちろん誤差の予測モデルにてフィッティングすることが可能である (図11b, $R^2 = 0.8974$). 我々の誤差の予測モデルの解釈は, “ランダムな外乱” という言葉が肝であり, この先行研究での実験系ならば外乱は平均して

図 11 誤差の予測モデルの他の実験データへのフィッティング結果。(a)：運動学習に対する不確実性の影響を検証した実験データ³¹⁾へ誤差の予測モデルをフィッティングした結果。被験者は 1 cm の外乱の値に十分慣れた後に、視覚フィードバック情報にブラーをかけ、更に 1 cm とは異なる外乱を与える。そのため、横軸は外乱 p_t 、縦軸は運動の修正量の値 $x_{t+1} - x_t$ を表しているが、横軸の基準は 1 cm になっている。 σ_0 がブラーなし、 σ_M が弱程度のブラー、 σ_L が強程度のブラー、 σ_∞ は視覚情報なし、という実験設定を表している。(b)：外乱の大きさと修正量の関係性を検証した実験データ⁴⁴⁾に誤差の予測モデルをフィッティングした結果。横軸は外乱 p_t 、縦軸は運動の修正量の値 $x_{t+1} - x_t$ を表している。

0 であるため、 -1 cm, 1 cm の外乱で生じる誤差は予測しやすい値となり修正率が高く、最も大きい -8 cm, 8 cm の外乱で生じる誤差は最も予測しづらいものであり修正率が低いというものである ((31) 式も参照されたい)。もちろんこの実験に特化したバイズモデルはセービングスを始めとした種々の実験結果を説明することはできないが、我々の誤差の予測モデルは我々の行動実験結果を予測可能であったことに加え様々な実験結果を統一的に説明することができる。どちらのモデルの妥当性が高いかは言うまでもない。

8. 誤差の予測モデルの意思決定への応用

以上から、様々な角度から誤差の予測モデルの妥当性を支持する結果が得られた。“はじめに”のセクションで述べたように、運動学習と意思決定では類似したモデルが多々提案されている。それならば、誤差の予測モデルを意思決定に応用することで、意思決定に関する様々な実験結果を統一的に再現できるのではないかと期待できる。ここではある刺激 s に対する報酬 r を予測する課題、すなわち古典的条件付けを想定して話を進める。複数刺激存在し、softmax 関数などで予測した報酬に基づき行動選択するモデルへと拡張すれば意思決定のモデルになるため、まずは古典的条件付けを想定して話を進める。

運動学習のモデルと同様に、やはり意思決定においてもまたセービングスや不確実性による学習速度の調

節が再現できる (図 12)。これらは複数時定数モデルやバイズモデルで別々に再現されてきた現象であり、誤差の予測モデルならば統一的に再現することができる。このようにして、誤差の予測モデルを意思決定に応用することで、様々な現象を統一的に眺めることが可能になることが期待できる。

9. 議 論

誤差の予測モデルは今まで再現されていなかったランダム学習を再現するための条件を満たすモデルの一つであることを数学的に示し ((17), (18) 式)、誤差の予測モデルのみが予測できる現象を実験的に実証した (図 5)。さらに実験結果にフィッティングしたパラメータを用いて運動学習に関わる種々の実験結果を再現できること (図 6～10)、他の実験で計測されたデータにフィッティングできることを示した (図 11)。更に、意思決定と同様の枠組みに誤差の予測モデルを適用した結果、様々な現象を統一的に眺めることができた (図 12)。

意思決定や古典的条件付けの古典的なモデルである Pearce-Hall モデルをヒューリスティックに運動学習に適用したモデルも提案されており⁴⁵⁾、それらもまた我々の誤差の予測モデルと同様に様々な現象を統一的に再現できることが報告されている。それらは学習率 B が報酬や誤差の履歴に依存すると想定したものであり、例えば自発的回復が再現できない、忘却率 A の調節は誤

図12 誤差の予測モデルを意思決定に応用した結果. 運動学習と同様にセービングス (a, b), 不確実性の影響 (c, d) をシミュレーション. 点線はそのトライアルで与えられた報酬量を表している.

差の履歴ではされない, など我々の誤差の予測モデルと異なる点も多々存在する. 今後は類似モデルとの分離を始めとして誤差の予測モデルを更に発展させ, その後に意思決定に応用することで運動学習と意思決定の双方の理解を深めていく必要があると考えている.

謝辞

本稿を書く機会を与えていただいた神経回路学会誌編集委員の坂本一寛先生 (東北大学) に感謝致します. また, 本研究において大変お世話になった野崎大地教授 (東京大学 教育学研究科), 平島雅也研究員 (CiNet), 野崎研究室の皆様方 (特に元秘書の阿部香苗さん, 進矢泰子さん, 原稿をチェックしていただいた博士課程在籍中の林拓志さん), そして常に私の研究生生活を支えてくれている妻 優子と息子 拓人にこの場を借りて感謝の意を表させていただきます.

参考文献

- 1) Wolpert, D.M., Landy, M.S. (2012): Motor control is decision-making, *Current Opinion in Neurobiology*, Vol.22(6), pp.996–1003.
- 2) Resulaj, A., Kiani, R., Wolpert, D.M., Shadlen, M.N. (2009): Changes of mind in decision-making, *Nature*, Vol.461(7261), pp.263–266.
- 3) Huang, V.S., Haith, A., Mazzoni, P., Krakauer, J.W. (2011): Rethinking motor learning and savings in adaptation paradigms: Model-free memory for successful actions combines with internal models, *Neuron*, Vol.70(4), pp.787–801.
- 4) Daw, N.D., Niv, Y., Dayan, P. (2005): Uncertainty-based competition between prefrontal and dorsolateral striatal systems for behavioral control, *Nature Neuroscience*, Vol.8(12), pp.1704–1711.
- 5) Loewenstein, Y., Seung, H.S. (2006): Operant matching is a generic outcome of synaptic plasticity based on the covariance between reward and neural activity, *Proceedings of the National Academy of Sciences*, Vol.103(41), pp.15224–15229.
- 6) Thoroughman, K.A., Shadmehr, R. (2000): Learning of action through adaptive combination of motor primitives, *Nature*, Vol.407(6805), pp.742–747.
- 7) Smith, M.A., Ghazizadeh, A., Shadmehr, R. (2006): Interacting adaptive processes with different timescales underlie short-term motor learning, *PLoS Biology*, Vol.4(6), pp.e179 1035–1043.
- 8) Fusi, S., Asaad, W.F., Miller, E.K., Wang, X.-J. (2007): A neural circuit model of flexible sensorimotor mapping: Learning and forgetting on multiple timescales, *Neuron*, Vol.54(2), pp.319–333.

- 9) Burge, J., Ernst, M.O., Banks, M.S. (2008): The statistical determinants of adaptation rate in human reaching, *Journal of Vision*, Vol.8(4), 20, pp.1–19.
- 10) Behrens, T.E.J., Woolrich, M.W., Walton, M.E., Rushworth, M.F.S. (2007): Learning the value of information in an uncertain world, *Nature Neuroscience*, Vol.10(9), pp.1214–1221.
- 11) Cisek, P. (2006): Integrated neural processes for defining potential actions and deciding between them: a computational model, *The Journal of Neuroscience*, Vol.26(38), pp.9761–9770.
- 12) Takiyama, K., Hirashima, M., Nozaki, D. (2015): Prospective errors determine motor learning, *Nature Communications*, Vol.6(5925), pp.1–12.
- 13) Todorov, E., Jordan, M.I. (2002): Optimal feedback control as a theory of motor coordination, *Nature Neuroscience*, Vol.5(11), pp.1226–1235.
- 14) Scott, S.H. (2004): Optimal feedback control and the neural basis of volitional motor control, *Nature Reviews Neuroscience*, Vol.5(7), pp.532–546.
- 15) Pruszynski, J.A., Kurtzer, I., Nashed, J.Y., Omrani, M., Brouwer, B., Scott, S.H. (2013): Primary motor cortex underlies multi-joint integration for fast feedback control, *Nature*, Vol.478(7369), pp.387–390.
- 16) Li, C.-S.R., Padoa-Schioppa, C., Bizzi, E. (2001): Neuronal correlates of motor performance and motor learning in the primary motor cortex of monkeys adapting to an external force field, *Neuron*, Vol.30(2), pp.593–607.
- 17) Rioult-Pedotti, M.-S., Friedman, D., Donoghue, J.P. (2000): Learning-induced LTP in neocortex, *Science*, Vol.290(5491), pp.533–536.
- 18) Criscimagna-Hemminger, S.E., Bastian, A.J., Shadmehr, R. (2010): Size of error affects cerebellar contributions to motor learning, *Journal of Neurophysiology*, Vol.103(4), pp.2275–2284.
- 19) Cheng, S., Sabes, P.N. (2006): Modeling sensorimotor learning with linear dynamical systems, *Neural Computation*, Vol.18(4), pp.760–793.
- 20) Donchin, O., Francis, J.T., Shadmehr, R. (2003): Quantifying generalization from trial-by-trial behavior of adaptive systems that learn with basis functions: theory and experiments in human motor control, *The Journal of Neuroscience*, Vol.23(27), pp.9032–9045.
- 21) Scheidt, R.A., Ghez, C. (2007): Separate adaptive mechanisms for controlling trajectory and final position in reaching, *Journal of Neurophysiology*, Vol.98(6), pp.3600–3613.
- 22) Emken, J.L., Benitez, R., Sideris, A., Bobrow, J.E., Reinkensmeyer, D.J. (2007): Motor adaptation as a greedy optimization of error and effort, *Journal of Neurophysiology*, Vol.97(6), pp.3997–4006.
- 23) Ethier, V., Zee, D.S., Shadmehr, R. (2008): Spontaneous recovery of motor memory during saccade adaptation, *Journal of Neurophysiology*, Vol.99(5), pp.2577–2583.
- 24) Shadmehr, R., Mussa-Ivaldi, F.A. (1994): Adaptive representation of dynamics during learning of a motor task, *The Journal of Neuroscience*, Vol.14(5), pp.3208–3224.
- 25) Yokoi, A., Hirashima, M., Nozaki, D. (2011): Gain field encoding of the kinematics of both arms in the internal model enables flexible bimanual action, *The Journal of Neuroscience*, Vol.31(47), pp.17058–17068.
- 26) Mussa-Ivaldi, F.A., Bizzi, E. (2000): Motor learning through the combination of primitives, *Philosophical transactions of the Royal Society of London, Series B, Biological Sciences*, Vol.355(1404), pp.1755–1769.
- 27) Pouget, A., Sejnowski, T.J. (1997): Spatial transformations in the parietal cortex using basis functions, *Journal of Cognitive Neuroscience*, Vol.9(2), pp.222–237.
- 28) Takiyama, K. (2015): Context-dependent memory decay is evidence of effort minimization in motor learning: a computational study, *Frontiers in Computational Neuroscience*, Vol.9(4), pp.1–10.
- 29) Scott, S.H., Gribble, P.L., Graham, K.M., Cabel, D.W. (2001): Dissociation between hand motion and population vectors from neural activity in motor cortex, *Nature*, Vol.413(6852), pp.161–165.
- 30) Wolpert, D.M., Ghahramani, Z., Jordan, M. (1995): An internal model for sensorimotor integration, *Science*, Vol.269, pp.1880–1882.
- 31) Körding, K.P., Wolpert, D.M. (2004): Bayesian integration in sensorimotor learning, *Nature*, Vol.427(6971), pp.244–247.
- 32) Jones, M., Love, B.C. (2011): Bayesian fundamentalism or enlightenment? On the explanatory status and theoretical contributions of Bayesian models of cognition, *Behavioral and Brain Sciences*, Vol.34(4), pp.169–188.
- 33) Kojima, Y., Iwamoto, Y., Yoshida, K. (2004): Memory of learning facilitates saccadic adaptation in the monkey, *The Journal of Neuro-*

- science, Vol.24(34), pp.7531–7539.
- 34) Scheidt, R.A., Reinkensmeyer, D.J., Conditt, M.A., Rymer, W.Z., Mussa-Ivaldi, F.A. (2000): Persistence of motor adaptation during constrained, multi-joint, arm movements, *Journal of Neurophysiology*, Vol.84(2), pp.853–862.
 - 35) Braun, D.A., Aertsen, A., Wolpert, D.M., Mehring, C. (2009): Motor task variation induces structural learning, *Current Biology*, Vol.19(4), pp.352–357.
 - 36) Kitazawa, S., Kimura, T., Yin, P.B. (1998): Cerebellar complex spikes encode both destinations and errors in arm movements, *Nature*, Vol.392(6675), pp.494–497.
 - 37) Ferrera, V.P., Barborica, A. (2010): Internally generated error signals in monkey frontal eye field during an inferred motion task, *The Journal of Neuroscience*, Vol.30(35), pp.11612–11623.
 - 38) Popa, L.S., Hewitt, A.L., Ebner, T.J. (2012): Predictive and feedback performance errors are signaled in the simple spike discharge of individual Purkinje cells, *The Journal of Neuroscience*, Vol.32(44), pp.15345–15358.
 - 39) Doya, K., Samejima, K., Katagiri, K.-i., Kawato, M. (2002): Multiple model-based reinforcement learning, *Neural Computation*, Vol.14(6), pp.1347–1369.
 - 40) Castro, L.N.G., Monsen, C.B., Smith, M.A. (2011): The binding of learning to action in motor adaptation, *PLoS Computational Biology*, Vol.7(6), pp.e1002052 1–14.
 - 41) Pekny, S.E., Criscimagna-Hemminger, S.E., Shadmehr, R. (2011): Protection and expression of human motor memories, *The Journal of Neuroscience*, Vol.31(39), pp.13829–13839.
 - 42) Wei, K., Körding, K.P. (2010): Uncertainty of feedback and state estimation determines the speed of motor adaptation, *Frontiers in Computational Neuroscience*, Vol.4(11), pp.1–9.
 - 43) Sing, G.C., Smith, M.A. (2010): Reduction in learning rates associated with anterograde interference results from interactions between different timescales in motor adaptation, *PLoS Computational Biology*, Vol.6(8), pp.e1000893 1–14.
 - 44) Wei, K., Körding, K.P. (2008): Relevance of error: what drives motor adaptation?, *Journal of Neurophysiology*, Vol.101(2), pp.655–664.
 - 45) Herzfeld, D.J., Vaswani, P.A., Marko, M.K., Shadmehr, R. (2014): A memory of errors in sensorimotor learning, *Science*, Vol.345(6202), pp.1349–1353.