

遅れ確率制御の仮説と実験

(株)ソニーコンピュータサイエンス研究所
大平 徹¹

1 はじめに

我々がある状況や物体をコントロールしようとするときにはさまざまな要素がくみあわされて行われていると考えられます。特に人間の反応時間は100ミリ秒のオーダーであり、工学的なシステムと比べてきわめて「遅い」システムであるという事実に鑑みれば、フィードバックだけではなく、予測などもかなり活用されていると考えられます。しかし、これらのような工学的な観点からの視点では捉えきれないような制御を、生体は行っていないかというのが、ここでの議論の背景にある疑問かつ課題です。例えば、最近も自転車に乗るロボットの開発がありました。人間と同様の遅い反応時間でも、自転車や一輪車に乗るロボットを作ることは、従来の制御理論やシステムでは難しいのではないかと予想します。

このような疑問に関連して、前回は報告しましたが、人間の指先で倒立棒を制御する実験では興味深い結果が得られています [1, 2]。この実験では実際に棒の動きを記録しているのですが、明らかにその動きの特徴的な時間スケールは、人間の反応時間の限界である100ミリ秒より短い場合が頻繁に観測されました。また被験者は訓練により、この倒立棒の制御をよりうまく行うことができるようになることも明らかになりました。反応時間よりも短い時間での変化が存在し、しかもノイズにより必ずしも明解な予測がしがたいこのような状況での制御やその上達においては上記にあげた2つの制御パラダイムとは違う、別の制御パラダイムが使われている可能性が示唆されていると考えます。

このような状況での新しい制御パラダイムの一つの仮説として「遅れ確率制御」(Delayed Stochastic Control)[3, 4]を考えています。すなわち人間は倒立棒制御のような訓練において、フィードバックやフィードフォワード制御では自身の反応時間の限界のため十分な制御ができないということを認識して、意識的もしくは無意識的に指先の動きに自身の反応遅れの値に対して、適度な揺れ(ノイズ)を取り入れるのではないかとという仮説です。

前回はこの仮説の理論的背景として斥力をもつ遅れランダムウォークについて述べましたが、今回はこの仮説を実験的に検証できないかという試みをいくつか行いましたのでこれについて紹介したいと思います。

2 理論的側面：斥力をもつ遅れランダムウォーク

まず、前回は紹介した斥力をもつ遅れランダムウォークを簡単に触れておきます。ノイズと遅れを含むシステムの記述として遅れを含むランダムウォークを提案してその性質を調べてきました [5, 6, 7]。このモデルは次のステップをとる確率がウォーカーの現在位置ではなく、ある一定時間前の位置によって決まるというランダムウォークです。このモデルの一つのケースとして、原点からのバイアスが反発の方向に働くようなモデルを考えます。バイアスのかかり方でいろいろモデルを考えることができますが、ここでは原点を境にしてバイアスが、非連続的にスイッチするようなモデルを考えます。物理的には原点を頂点としてさかさV字の丘の上にウォーカーがいるような場合に相当します。具体的には以下のように遷移確率が定義される、1次元整数値をとる離散時間、離散空間のランダムウォークです。

$$P(X(t+1) = X(t) + 1 | X(t-\tau) > 0) = p, \quad (1)$$

$$P(X(t+1) = X(t) + 1 | X(t-\tau) = 0) = \frac{1}{2}, \quad (2)$$

$$P(X(t+1) = X(t) + 1 | X(t-\tau) < 0) = 1 - p. \quad (3)$$

ここで、 $X(t)$ はウォーカーの時刻ステップ t における位置、 p はバイアス、 τ は遅れを表します。ここで $p > 0.5$ とすることで、原点から斥力が働くようにできます。

物理的に考えるとウォーカーは坂をどんどん落ちていくようになり、原点からどんどん離れていくばかりの状況と分かります。従って、定常確率分布は存在しませんし、統計的にも平均は対称性を用いて $\langle X(t) \rangle = 0$ となりますが、分散(もしくは位置の二乗平均)はどちらも有限値には収束しません。特に遅れがあるときにはその挙動の理解をより困難とします。このモデルで興味深い特徴は、初期条件の取り方にもよりますが、 p と τ を調節してやることにより、原点から拡散していくスピードを遅らせることができるようになります。別の表現をすれば p と τ の値がうまくかみ合うと、不安定な原点の周りでウォーカーをより長く存在させることができるということです。この特徴を表現する統計量として、平均初通過時間 (average first passage time) を使います。これは

¹ E-mail: ohira@csl.sony.co.jp

ウォーカーがある地点から別の地点に初めて到達するまでにかかる平均時間です。ここでのわれわれのモデルにおいては、原点を出発したウォーカーが初めて目的地 $\pm X^*$ に到達するまでの時間と定義します。別の見方をすればウォーカーが $\pm X^*$ の範囲の中にとどまっていた平均時間ともいえます。

この平均初通過時間 $\langle L \rangle$ はこのモデルで遅れがない場合においては、導出はやや複雑なので省略しますが以下のように求められます。

$$\langle L \rangle = 2 \left(\frac{q}{q-p} \right) \left(\frac{1 - \left(\frac{q}{p}\right)^{X^*}}{1 - \frac{q}{p}} \right) + \frac{X^*}{p-q}, \quad (p \neq 0.5). \quad (4)$$

ここでは、 $q \equiv 1-p$ としています。上記は引力と斥力の両方について成り立ちます。また対称単純ランダムウォーク $p = q = 0.5$ の場合にはより簡単に

$$\langle L \rangle = (X^*)^2 \quad (5)$$

で与えられます。

遅れがあるときには残念ながら、解析的な記述は困難です。しかし、数値計算によって調べると「最適安定」が p と τ の値を調整してやることにより現象として現れます。

初期状態としては $t \in (-\tau, 0)$ の時間についてはウォーカーは対称単純ランダムウォーク $p = q = 0.5$ を行い、 $t = 0$ でのウォーカーの位置を原点として空間軸を設定します。この初期状態で p と τ として、目的地 X^* の値の様々な組み合わせでそれぞれ 1000000 サンプルの斥力遅れランダムウォークの数値計算を行い、平均初通過時間 $\langle L \rangle$ を計算しました。

平均初通過時間は p と X^* が固定された状況では τ が増加していくと始めは増加しますが、やがてピークを打って減少に転じます。物理的にはこれはピークのときに目的地 $\pm X^*$ に到達するまでの平均時間ももっとも長くなるということです。別の言い方をすればピークのときにウォーカーは、原点の周りでもっとも安定しているということです。

この現象についてさらに理解を進めるために、近似的にこれを捉える数式表現を求めます。数値的に以下の形で、ピークの周辺の挙動を表すことができます。

$$\langle L(\tau) \rangle = (1 + \alpha \tau_n e^{\beta \tau_n}) \langle L(\tau = 0) \rangle. \quad (6)$$

ここで、 α と β は定数、 τ_n は「規格化」された遅れで

$$\tau_n \equiv \tau \frac{p-q}{X^*} \quad (7)$$

で定義します。この規格化は目的地までの距離 X^* をウォーカーの「速度」 $p-q$ で割った特徴的な時間で、遅れの時間 τ を割ったものですが、これにより τ_n は無次元となります。

ここで図 1 にこの近似式と数値計算の結果を示しました。違う値の (p, X^*) によって得られた数値計算によるカーブと、上記の式で α と β を適切に選んだカーブがほぼ重なることが分かります。また、ピークにおける平均初通過時間は、遅れがないときに比較してほぼ 1.7 倍となっています。

この現象は、遅れとノイズの「共鳴」によって最適な安定が得られたという解釈もできます。メカニズムは違いますが、このような「遅れ確率共鳴」については別のモデルで数理解析的な考察をしてきました [8, 9]。ここでの現象も同じような観点から解析が可能かもしれませんが、まだ課題として残されています。

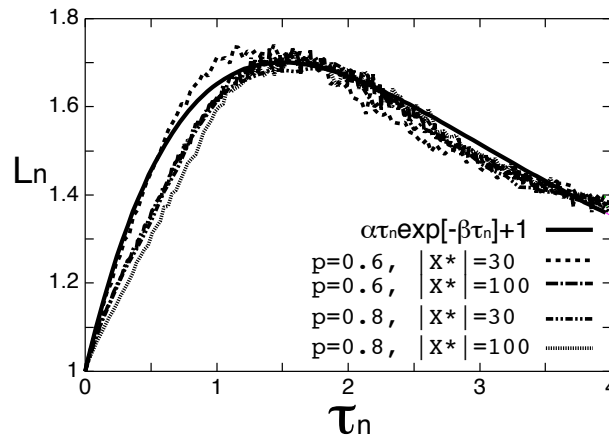


図 1: 規格化された τ_n の関数としての $\langle L_n \rangle \equiv \frac{\langle L(\tau) \rangle}{\langle L(\tau=0) \rangle}$ 。パラメータの組 (p, X^*) の値は $(0.6, 30)$ 、 $(0.6, 100)$ 、 $(0.8, 30)$ 、 $(0.8, 100)$ である。理論式の調整された定数は $\alpha = 1.27$ 、 $\beta = 0.67$ である。

3 実験的試み：揺らしながらの倒立棒制御

実際上記のようにノイズと遅れの適度な組合せが、より安定した制御を可能にするのかを考えてみました [4]。倒立棒制御をする際に意図的にノイズを加えてみて、これが実際によりより制御につながるのかを試みました。ノイズを加える方法はいくつかこころみました。倒立棒制御をしながら足をゆすってもらったり、以下の図 2 に示したように、棒のバランスを取っている手と反対の手に物を持って振ってもらったりをしました。個人差は大きいのですが、約 6 割程度の人で、自覚的に倒立棒制御が容易になり、倒れるまでの時間の計測にも現れるという結果を得ました (図 3)。より精密な実験は現在行っているところですが、ここでの結果は個人の反応時間に応じての適切なノイズがより安定した制御につながるという仮説を支持する方向にあります。



図 2: 物を振りながら倒立棒制御をする被験者の例。

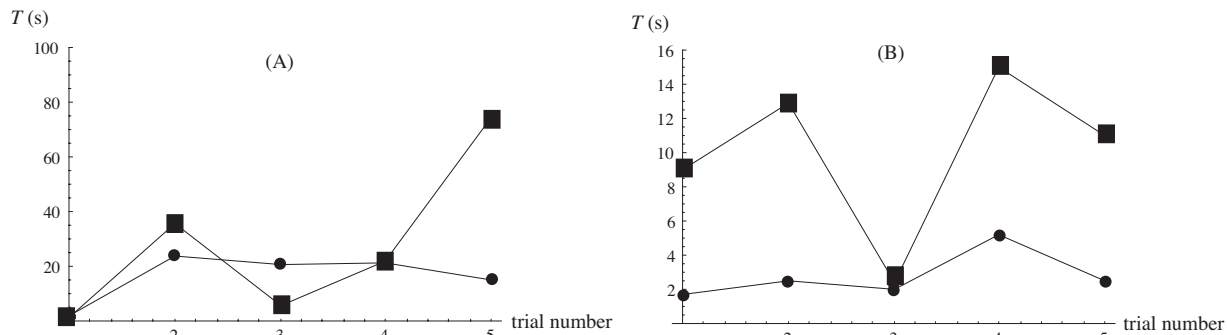


図 3: (A) 体を揺らした場合 (四角) と揺らさない場合 (丸) での倒立棒制御の時間 T (秒) の試行回数 (trial number) ごとの比較。この被験者は事前の練習を行うことなく 5 回試行した。5 回目には制御時間に明らかな違いが現れた。(B) 数時間の練習を行った別の被験者における制御時間の比較。この被験者も揺らした場合の方が長い時間の制御を行えることが明確であった。

4 まとめ

ここで議論しているような遅れ確率制御はまだ仮説の域をでません。しかし、この実験の結果は生体の制御が工学的な制御理論からでは捉えきれない可能性を示しているといえます。視点としても、伝統的な制御理論では漸近安定性が重視されるのに対して、生体では振動や揺らぎの不安定な側面を内包しながら制御が行われている可能性があります。また、制御部分のシステムを切り出して考察していく方法も、生体制御ではあてはまらずに、システム全体のバランスが重要である可能性があります。このような視点を工学的にも活用していけるのかどうかは今後の課題です。

参考文献

- [1] J. L. Cabrera and J. G. Milton: On-Off Intermittency in a Human Balancing Task, *Physical Review Letters*, 89 158702 (2002).
- [2] J. L. Cabrera, R. Bormann, C. Eurich, T. Ohira and J. Milton: State-dependent Noise and Human Balance Control, *Fluctuations Noise Letters*, 4, L107 (2004).
- [3] T. Ohira and T. Hosaka: Repulsive Delayed Random Walk *Artificial Life and Robotics*, Vol.9 (2004, in press).
- [4] T. Ohira and T. Hosaka: Control in Delayed Stochastic Systems, In the proceedings of DETC05, 2005 ASME Design Engineering Technical Conference, (Long Beach, U.S.A., September 2005).
- [5] T. Ohira and J. G. Milton: Delayed Random Walks, *Physical Review E*, 52 3277 (1995).
- [6] T. Ohira: Oscillatory Correlation of Delayed Random Walks, *Physical Review E*, 55 R1255 (1997).
- [7] T. Ohira and T. Yamane: Delayed Stochastic Systems, *Physical Review E*, 61 1247 (2000).
- [8] T. Ohira and Y. Sato: Resonance with Noise and Delay, *Physical Review Letters*, 82 2811 (1999).
- [9] L.S. Tsimring and A. Pikovsky: Noise-induced Dynamics in Bistable Systems with Delay, *Physical Review Letters*, 87 250602 (2001).