

Hopfield 模型における誤りとノイズの効果

Noise and error in the Hopfield model

曾根 彰吾 (Shogo Sone)

静岡大学・工学部システム工学科

Dept. of Eng., Shizuoka Univ.

久保 勲生 (Isao Kubo)

大阪大学・大学院情報科学研究科

Grad. School of Info., Osaka Univ.

南 和彦 (Kazuhiko Minami)

名古屋大学・大学院多元数理科学研究科

Grad. School of Math., Nagoya Univ.

1. Introduction

生物系を含む現実の種々のシステムは、外部からの雑音の影響の中にあり、またその内部の動作にもある程度の不正確さがある。つまり、ノイズによってシステムが変更をうけたり、内部の規則的な更新において一定の確率で誤りが起きたりする中で動作して行く。このノイズや誤りはシステムにどんな影響を及ぼすのであろうか。その影響は、はたしてシステムにとってマイナスなのであろうか。

例えば生物のシステムは、時々間違えながら動作しているわけであるが、しかし結局は正しい結果に行き着く、あるいは間違いを前提にした最適の動作の仕方を自律的に発見していくのではないか、という推測ができる。

また例えば、生物の進化のプロセスは系統によってかなり異なることが知られているが、生物の集団は大きなノイズの中で変化を続け、このときむしろノイズが本質的な役割を持っているのではないか、つまり集団の進化のシステムは、ノイズの影響をうけて、それに依存して決定されるのではないか、ということが推測できる。[1]

ここでは神経系の数理モデルとして導入された Hopfield 模型を考え、Hopfield 模型における誤りとノイズの影響について、ミクロな立場から考察する。

モデリングの出発点になるこの Hopfield 模型は、脳の神経回路を模して作られた数理モデルであるが、脳の特定の部位をモデル化したものというよりも、ニューロンとシナプスの情報伝達の機構についての、理想化された数理モデルである。ただしその抽象化の結果、脳に限らず種々の問題に対して適用できる普遍性を持つ模型になっている。この標準的な模型に、ノイズと誤りの効果を個別に加えて調べることにする。結果としては興味深いことに、物事を考えたり思い出したりする際の我々の経験に、非常に似た振る舞いが得られ、それらはノイズと誤りがあるときに、はじめて現れるものである。

第 2 章で Hopfield 模型とその更新の規則、第 3 章でノイズと誤りの効果を考慮した変形について説明し、第 4 章ではその計算の結果を述べ、第 5 章でそれについて考察する。

2. Hopfield 模型

神経細胞のように ON と OFF の状態をとるサイト（ノード）の間を、エッジがつないでネットワークを形成している状況を考える。最も古典的な Hopfield 模型では、ネットワークとしては、すべてのサイトの間にエッジがある完全グラフを考える。

エッジに重みを付随させておき、初期状態から出発して適当な規則の下でサイトの状態を更新していく。更新をくり返した後に特定のパターンが再現される。エッジに付与された重みの値を決めると、それに応じて再現されるパターンが決まる。

これは記憶を思い出すプロセスの、比較的単純なモデル化であると考えることができる。この模型においてパターンを再現するプロセスは、例えば視覚から入った図形の情報を初期状態として、その初期状態に近い文字を連想する神経系のプロセスに相当すると考えることができる。図形から何を連想するのかは、それを見た人間のそれまでの学習の経験によって結果が異なる。この学習された蓄積がエッジの重みの値に相当する。

時刻 t における i 番目のシナプスの状態を $x_i(t)$ とする。シナプスが興奮状態にあるときを $x_i(t) = 1$ 、そうでないときを $x_i(t) = 0$ とする。サイト i を選び、その状態 $x_i(t)$ を周囲のシナプスの状態とニューロンの重み

を参照しながら更新する. 更新の規則は

$$x_i(t+1) = \begin{cases} 1, & u_i(t) > 0 \\ x_i(t), & u_i(t) = 0 \\ 0, & u_i(t) < 0 \end{cases} \quad (1)$$

ただし $u_i(t)$ は

$$u_i(t) = \sum_{l=1}^n w_{il}x_l(t) - \theta_i(t). \quad (2)$$

周囲のシナプスの状態から重み w_{ij} で情報を受け取り $u_i(t)$ を決め, その値と閾値 $\theta_i(t)$ との大小関係に応じて自分自身の次の状態を決定する. ここでは $\theta_i(t) = 0$ の場合を考えることにする. 古典的な Hopfield 模型では, それぞれのシナプスをランダムに選んで順に更新する. これを非同期的な状態変化とよぶ.

各ニューロンに付随するシナプス間の結合の重み w_{ij} は次の性質を持つ. まず自分自身の現在の状態は参照しない, つまり $w_{ii} = 0$ である. またシナプス間の結合に方向性はなく $w_{ij} = w_{ji}$ であるとする.

ここでエネルギー関数

$$E = -\frac{1}{2} \sum_{i,j=1}^n w_{ij}x_i(t)x_j(t) - \sum_{i=1}^n \theta_i(t)x_i(t) \quad (3)$$

を導入する. (1) にしたがって状態を更新すると, このエネルギー関数 E は同じ値に留まるか, または減少することを示そう. x_k が関係する項は, $w_{ij} = w_{ji}$ を使って計算すると

$$\begin{aligned} E_k &= -\frac{1}{2} \sum_{i=1}^n w_{ik}x_i(t)x_k(t) - \frac{1}{2} \sum_{j=1}^n w_{kj}x_k(t)x_j(t) - \theta_k(t)x_k(t) \\ &= -\left[\sum_{l=1}^n w_{kl}x_l(t) - \theta_k(t)\right]x_k(t) \\ &= -u_k(t)x_k(t) \end{aligned} \quad (4)$$

(1) にしたがって状態が更新されるとき $E_k \leq 0$ である. 実際, $u_k(t) = 0$ であり状態を変えても E が変化しないときは $x_k(t)$ を変えない. また $u_k(t) \neq 0$ であり状態変化によって E の値が増加するときにも $x_k(t)$ を変えず, $u_k(t) \neq 0$ であり状態変化によって E の値が減少するときには $x_k(t)$ の値を変化させ, E を減少させる. これを繰り返すことで状態 $(x_1(t), x_2(t), \dots, x_N(t))$

は E が極小（局所的な最小）を取る状態か、 E が変化しない状態（これは広義の極値にあると言える）でその変化をやめる。このとき極小をとる状態が、記憶されたパターンであると考える。

そこで指定されたパターンが極小値に対応するような重み w_{ij} はどうすれば得られるかということが重要になる。記憶するパターンは複数あり得るので、 $(x_1^{(s)}, x_2^{(s)}, \dots, x_N^{(s)})$ ($s = 1, 2, \dots, \alpha$) なる α 個のパターンを記憶するとして、これらの状態で E が極小になるような w_{ij} は意外と簡単に

$$w_{ij} = \sum_{s=1}^{\alpha} x_i^{(s)} x_j^{(s)} \quad (5)$$

で与えられる。これをの Heb 則とよぶ。実際

$$\begin{aligned} E &= -\frac{1}{2} \sum_{i,j=1}^n w_{ij} x_i(t) x_j(t) = -\frac{1}{2} \sum_{i,j=1}^n \sum_{s=1}^{\alpha} x_i^s x_j^s x_i(t) x_j(t) \\ &= -\frac{1}{2} \sum_{s=1}^{\alpha} \left(\sum_{i=1}^n x_i^s x_i(t) \right)^2 \end{aligned} \quad (6)$$

ただし記憶するパターンの間には相関がなく、 α 個の状態は互いに直交するとしてある。問題にしているシステムにおいてパターンが正常に再現されるような α の最大値を調べることは重要であり、これはシステムの記憶容量を求めることに相当する。

3. 方法

以上が古典的な Hopfield 模型である。これに新しい構造を加えて、変形 Hopfield 模型を作る。その規則を決めること自体がモデリングになる。今回は更新の誤りを考えるモデルと、ノイズの効果としての w_{ij} のばらつきの極端なケースとして、結合を一部切るモデルを採用し、合計 4 つの方法を考えた。システムはすべて図 1、図 2 示したように 3 つのサイトで構成されており、記憶させるパターンは $X^c = (1, 1, 1)$ の 1 つのみとした。

図 1 は古典的な Hopfield 模型と同じであり、完全グラフの上で状態を更新する。また、図 2 はノイズによりリンクが切れたグラフを考えたものである。

(I) ノイズがない場合

ノイズがない場合、古典的な Hopfield 模型と同じであり、ネットワークとしては図 1 を考える。全てのサイトは互いに繋がっているため、すべてのリンクの重みは $w_{il} = 1$ となり式 (2) より更新の規則が決まる。これに

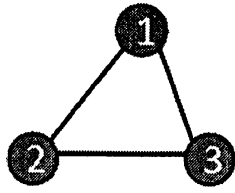


図 1: 完全グラフ

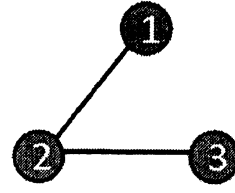


図 2: ノイズによりリンクが切れているネットワーク

より, $u_1 = x_2 + x_3$, $u_2 = x_1 + x_3$, $u_3 = x_1 + x_2$ と $u_i(t)$ が定まり, この規則をもとに式 (1) を用いて状態を更新する.

(II) 確率 $1/k$ で間違える場合

この場合ネットワークとして図 1 を用いること, また $u_i(t)$ を求めるところまでは (I) と同じである. しかし, 式 (1) を用いて更新する際, 確率 $1/k$ で本来取るはずの値と異なった値で更新される. たとえば, $u_i(t) > 0$ のときは $x_i(t+1) = 1$ となるはずが, 確率 $1/k$ で 0 になってしまうという具合である.

(III) ノイズによりリンクが切れた場合

この場合, ノイズによりネットワークの一部が切れてしまったとして図 2 を考える. これにより, $w_{12} = w_{23} = 1$, $w_{31} = 0$ となる. よって, 式 (2) より $u_1 = x_2$, $u_2 = x_1 + x_3$, $u_3 = x_2$ となる. この規則をもとに式 (1) を用いて状態を更新する.

(IV) ノイズによりリンクが切れていて, かつ確率 $1/k$ で間違える場合

この場合は (II) と (III) を合わせたものである. ネットワークの一部が切れているため, w_{il} および $u_i(t)$ は (III) のものを用い, 式 (1) を用いて更新する際, 確率 $1/k$ で本来取るはずの値と異なった値で更新される.

4. 結果

(I) の場合, 初期値 $(0, 0, 0)$ から始めると, どれだけ更新を行っても X^c

に行き着くことはない. 実際,

x_1	x_2	x_3	
0	0	0	← 1 を更新
0	0	0	← 2 を更新
0	0	0	← 3 を更新
0	0	0	⋮

しかし, 初期値 $(1, 0, 0)$ から始めると, 3step または 2step で X^c を再現することが出来た:

x_1	x_2	x_3		x_1	x_2	x_3		x_1	x_2	x_3	
1	0	0	← 1	1	0	0	← 2	1	0	0	← 3
1	0	0	← 2	1	1	0	← 3	1	0	1	← 1
1	1	0	← 3	1	1	1		1	0	1	← 2
1	1	1						1	1	1	

(II) の場合, (I) の場合と違い初期値 $(0, 0, 0)$ から始めても, 確率 $1/k$ で間違いが起こることにより最終的に X^c へに行き着くことが出来る.

x_1	x_2	x_3	
0	0	0	
			↓
0	0	0	← 間違える
1	0	0	← 思い出し始める
			↓
1	1	1	← 思い出す

ただし, $(0, 0, 0)$ 以外のときに更新が行われる際には, 間違えることによって $0 \rightarrow 1$ のはずが $0 \rightarrow 0$ となり前進しない場合が出てくる. また, $1 \rightarrow 1$ のはずが $1 \rightarrow 0$ となり 1-step 後退してしまう場合も出てくる. つまり, パターンを思い出す速度は $(0, 0, 0)$ 付近では少し遅くなり, $(1, 1, 1)$ 付近ではさらに遅くなってしまう. そして, X^c へに行き着いた後にも間違いは起き続けるために状態は振動する.

(III) の場合, (I) と同じく初期値 $(0, 0, 0)$ から始めると思い出すことが出来ず, 初期値 $(1, 0, 0)$ から始めると思い出すことが出来る. ただ (I) と違う点は, x_2 の状態が 1 になるまでは他の x_i は変化しないということである. つまり, x_2 が反転してはじめて X^c を再現するプロセスが始まるの

である：

x_1	x_2	x_3		x_1	x_2	x_3		x_1	x_2	x_3	
1	0	0	$\leftarrow 1$	1	0	0	$\leftarrow 2$	1	0	0	$\leftarrow 3$
1	0	0	$\leftarrow 2$	1	1	0	$\leftarrow 3$	1	0	0	$\leftarrow 1$
1	1	0	$\leftarrow 3$	1	1	1		1	0	0	$\leftarrow 2$
1	1	1						1	1	0	$\leftarrow 3$
								1	1	1	

(IV) の場合, (II) と (III) の特性を併せ持っているため, 初期値 $(0, 0, 0)$ からでも思い出すことが出来る. またその際, x_2 が 1 となるように間違いが起こる方が x_1, x_3 が間違えるよりも思い出す速度が速くなる.

5. 考察

(I) では, 状態の更新が規則通りであるとき, 大部分の始状態からは記憶が再現されるものの, 初期値 $(0, 0, 0)$ からは更新が起こらず記憶が戻らない. しかし (II) で誤りが確率 $1/k$ で起きると, それをきっかけにして更新がはじまり, X^c が再現される. この誤りは最初はこのように有効であるがパターンの再現の速度を遅くし, 目標に近づくにつれそのマイナスは大きくなる. このことはまるで, 迷っている状態では雑音の中で考えるのがいいが, 細部をつめて完全に思い出そうとする時には静寂が好ましい, という我々の経験を再現しているかのようである.

この更新の誤りは, Hopfield 模型における熱揺らぎの効果に相当する. 更新においてサイトの状態を変化させることによってエネルギー関数 E の値が E_1 から E_2 ($< E_1$) に変化する場合を考えよう. このときそれぞれの状態が実現する確率が Boltzmann 因子に比例するなら, 状態が変化する確率と変化しない確率はそれぞれ

$$\frac{e^{-\beta E_2}}{e^{-\beta E_1} + e^{-\beta E_2}} \quad \frac{e^{-\beta E_1}}{e^{-\beta E_1} + e^{-\beta E_2}} \quad (7)$$

ただし T を温度として $\beta = 1/T$ である. つまり状態が更新されて E が減少すべきときにも, 誤って更新されない確率が 0 ではない. ここで $T \rightarrow 0$ とすると $E_1 > E_2$ であるので両者はそれぞれ 1 と 0 になり, このとき状態は完全に規則通りに更新される. これが誤りのない場合に相当する. このとき更新の起きない始状態は, 1-step の更新では E が変化しないような E の値の平坦な領域に含まれる. 誤りがあるおかげでランダムに移動し, こ

の領域から抜け出して E の極小へと移動することができる。このことは Hopfield 模型に関して、古くから調べられてきたことであった。完全グラフでないネットワーク上の解析については例えば [2] の結果がある。

(Ⅲ) の場合にはネットワークの一部を切ったために、ハブになるサイト (ノード、あるいは神経細胞) が現れた。いまの場合サイト x_2 がハブになっており、更新についての重要な役割を担っている。各サイトをランダムに更新していくと、 x_2 が更新されるまではパターンの再現のプロセスが先に進まない。しかしハブである x_2 が更新されるとそれをきっかけにプロセスが進行しはじめる。これはちょうど、はじめはなかなか思い出せないでいるのが何かのきっかけで一気に思い出すという、我々の経験を連想させる。

代表的なネットワーク、例えば scale-free, small world, random graph の上での Hopfield 模型については [2]-[7] で調べられている。特にハブの影響については [5] で扱われているが、これはネットワークのサイズが大きい場合の、例えば記憶容量とハブとの関係を数値的に調べたもので、上記のようにプロセスに対する影響を更新の規則自体から出発して解析したものではない。

[5] の結果を、ハブの関するいまの議論から解析することができれば、ネットワークにおけるハブの役割をより明確に知ることができるのではない。さらにネットワークの種々の特性とそのネットワーク上での Hopfield 模型の性質との関係を明らかにすることは面白い。また発火していないリンクをノイズにより切断されやすくして Hopfield 模型に優先的選択を導入するなど、複雑ネットワーク上のニューラルネットワークにはいろいろな可能性があるように思われる。

参考文献

- [1] 日本数理生物学会ニュースレター 第 61 号 (2010) 研究会報告「生物現象に対するモデリングの数理」
- [2] J. Torres, J. Marro P. Garrido, J. Cortes, F. Ramos and M. Munoz: Biophysical Chemistry 115 (2005) 285-288.
- [3] D. Stauffer, A. Aharony, L. Costa and J. Adler: European Physical Journal B32 (2003) 395-399.

- [4] J. Torres, M. Munoz, J. Marro and P. Garrido: Neurocomputing 58-60 (2004) 229-234.
- [5] E. Rodrigues, M. Barbosa and L. Costa: cond-mat/0507677.
- [6] P. Zhang and Y. Chen: Physica A 387(2008) 1009-1015.
- [7] T. Autani: Kyoto University, Master Thesis 2010