

混沌系工学特論 配布資料 #2

担当：井上 純一 (情報科学研究科棟 8-13)

URL : http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/

平成 16 年 11 月 1 日

目次

1.9 補足：クロストークの数値的評価について	18
1.9.1 サンプリングとヒストグラム	18
1.9.2 ガウス分布からのずれの検出：高次キュムラント	18
1.10 非対称に結合した回路網の動作特性	20
1.10.1 非対称 Hebb 則とリミットサイクル解の出現	20
1.10.2 計算機シミュレーションで動作を確かめる	22

1.9 補足：クロストークの数値的評価について

1.9.1 サンプリングとヒストグラム

前小節で記憶容量の評価をする際に、各ニューロンにおけるクロストーク：

$$C_i^\nu \equiv -\frac{\xi_i^\nu}{N} \sum_j \sum_{\mu \neq \nu} \xi_i^\mu \xi_j^\mu \xi_j^\nu \quad (40)$$

の統計的性質を調べたわけであるが、これは独立なランダムパターン ξ^μ の和であるから、中心極限定理よりこの確率変数は平均がゼロ、分散が $\sigma^2 = p/N$ のガウス分布に従うはずであるとして話を進めた¹。しかし、実際に様々 ξ^μ を変えて C_i^ν をサンプリングし、ヒストグラムを作った上で、それから C_i^ν の従う分布の大まかな形を予想することもできる。その例を図 12 に載せる。この図は $p = 50$, $N = 500$ (このとき中心極限定理から予想される分散は $\sigma^2 = p/N = 0.1$ である) と選び 1000 回のサンプリングからヒストグラムを作っている。もちろん、中心極限定理は和の要素数が十分に大きな場合の話であるから、数値的に得られたヒストグラムには統計誤差が残るが、中心極限定理からの結果とまあまあ良く合っている。また、このサンプリングから平均や分散を求めることもできる。

1.9.2 ガウス分布からのずれの検出：高次キュムラント

上の例は答えがわかっている場合の確認を示したに過ぎないが、場合によっては数値的評価が有効になることもある。(40) 式では回路網の状態が想起パターンに一致した場合、すなわち、 $S \rightarrow \xi^\nu$ のときのクロ

¹ 中心極限定理に関する証明は例えば少々資料は古いが http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/ENS/STATPHYS/stat99-2.pdf の **問題 3**、及びその解答 http://chaosweb.complex.eng.hokudai.ac.jp/~j_inoue/ENS/STATPHYS/stat99-2ans.pdf を見てください。

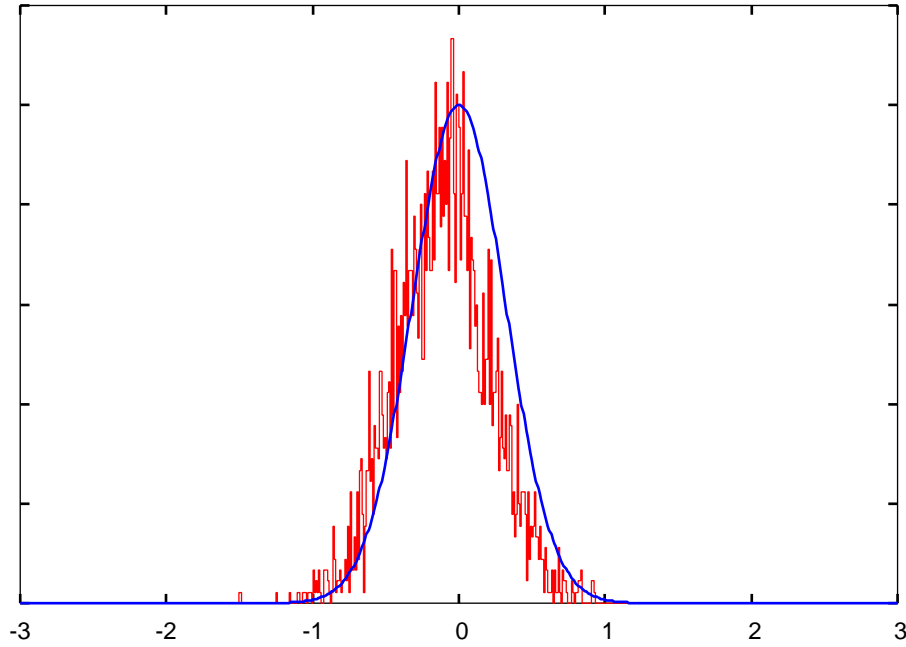


図 12: C_i^μ の ξ^μ を様々変えたサンプリングによるヒストグラムと中心極限定理から予測される答え : $P(C_i^\nu) = (1/\sqrt{2\pi\sigma^2}) \exp[-(C_i^\nu)^2/2\sigma^2]$. ここでは $p = 50, N = 500$ ($\sigma^2 = p/N = 0.1$) であり 1000 回のサンプリングからヒストグラムを作った.

ストークを求め、この想起パターン ξ^ν が安定である条件を求めたわけであるが、実際には回路網の状態 S はダイナミクス :

$$S_i(t+1) = \text{sgn} \left(\sum_{j=1}^N w_{ij} S_j(t) \right) \quad (41)$$

に従って時間的に変化しているわけであり、後の節で詳しく見るが、回路網のダイナミクスを調べる場合には各ステップで時間依存するクロストーク :

$$C_i(t) \equiv -\frac{S_i(t)}{N} \sum_j \sum_{\mu \neq \nu} \xi_i^\mu \xi_j^\nu S_j(t) \quad (42)$$

がどのような分布に従うか、を知る必要が出てくる。この (42) 式の場合には和の中に現れる各項はもはや独立ではないので中心極限定理に従うと断言することができない。場合によっては $C_i(t)$ は依然としてガウス分布かも知れないし、ガウス分布からずれているかもしれない。このようなとき、以下に説明する高次のキュムラントを数値的に調べることによってガウス分布からのずれを検出することができる。

まずは、注目する分布関数 $p(x)$ のフーリエ変換で分布 $p(x)$ の特性関数 :

$$\phi_x(k) \equiv \langle e^{ikx} \rangle = \int_{-\infty}^{\infty} dx e^{ikx} p(x) = \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle x^n \rangle \quad (43)$$

を定義する。ここで最後の変形では e^{ikx} を展開し、項別に積分したものをモーメント :

$$\langle x^n \rangle \equiv \int_{-\infty}^{\infty} dx x^n p(x) \quad (44)$$

で書き換えてある。さて、(43) 式は $\log(1+x)$ の展開を用いると次のように書き換えることができることに注意しよう。

$$\begin{aligned}
 \phi_x(k) &= 1 + \sum_{n=1}^{\infty} \frac{(ik)^n}{n!} \langle x^n \rangle \\
 &= \exp \left[\log \left\{ 1 + \sum_{n=0}^{\infty} \frac{(ik)^n}{n!} \langle x^n \rangle \right\} \right] \\
 &= \exp \left[\log \left\{ 1 + (ik) \langle x \rangle + \frac{(ik)^2}{2} \langle x^2 \rangle + \frac{(ik)^3}{6} \langle x^3 \rangle + \dots \right\} \right] \\
 &= \exp \left[1 + (ik) \langle x \rangle + \frac{(ik)^2}{2} \langle x^2 \rangle + \frac{(ik)^3}{6} \langle x^3 \rangle + \dots \right. \\
 &\quad \left. - \frac{1}{2} \left\{ (ik) \langle x \rangle + \frac{(ik)^2}{2} \langle x^2 \rangle + \frac{(ik)^3}{6} \langle x^3 \rangle + \dots \right\}^2 \right. \\
 &\quad \left. + \frac{1}{3} \left\{ (ik) \langle x \rangle + \frac{(ik)^2}{2} \langle x^2 \rangle + \frac{(ik)^3}{6} \langle x^3 \rangle + \dots \right\}^3 - \dots \right] \equiv \exp \left[\sum_{n=1}^{\infty} \frac{(ik)^n}{n!} C_n(x) \right] \quad (45)
 \end{aligned}$$

最後の等号の両辺で n の次数の同じもの同士を等しいをおけば

$$C_1(x) = \langle x \rangle \quad (46)$$

$$C_2(x) = \langle x^2 \rangle - \langle x \rangle^2 \quad (47)$$

$$\begin{aligned}
 C_3(x) &= \langle x^3 \rangle - 3\langle x \rangle \langle x^2 \rangle + 2\langle x \rangle^3 \\
 &\dots \dots \dots \quad (48)
 \end{aligned}$$

が得られる。この $C_n(x)$ を確率変数 x の n 次のキュムラントと呼ぶ。これからわかるように $C_1(x)$ は x の平均、 $C_2(x)$ は分散に等しく、 $p(x)$ がガウス分布のときには $C_3(x)$ 以降、高次のキュムラントはゼロとなる (各自確かめてみる。また、ここでの考え方に慣れるために図 12 に示したように、問 2 で作成したプログラムを用いて C_i^y をサンプリングし、実際に 3 次以降のキュムラント $C_3(x) \dots$ がゼロになるか否かを確かめてみると良いと思う)。従って、回路網のダイナミクスの過程で高次のキュムラントをプロットし、それがゼロか否かを確認することにより、確率変数がガウスか否かを確認することができる。

1.10 非対称に結合した回路網の動作特性

1.10.1 非対称 Hebb 則とリミットサイクル解の出現

ここまででパーセプトロンと呼ばれる神経細胞のモデルを多数つなげたとき、我々が日常的に「連想記憶」と呼んでいる現象と同じような動作が実現できることを学んだ。この際、どのように素子どうしをつなげれば良いのか、が本質的には重要であり、うまく連想記憶を実現するためには任意の神経素子 i と j とを p 個のパターン ξ^μ ($\mu = 1, \dots, p$) から作られる Hebb 則：

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^\mu \xi_j^\mu \quad (49)$$

として選べば良かった。ここで、この式 (49) を次のように一般化しておこう。

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^p \sum_{\nu=1}^p \xi_i^\mu A_{\mu\nu} \xi_j^\nu \quad (50)$$

すると、我々がここまで調べてきたのは $A_{\mu\nu} = \delta_{\mu,\nu}$ の場合に相当する。添え字 i と j を交換しても変わらないわけであるから、この場合を対称結合と呼ぼう。

ところで現実の脳の回路網では、あるニューロンから他のニューロンへの結合が対称であるとは限らない。むしろ、「脳神経系の多様性」という観点からは非対称であると考えた方が自然であるような気がする。従って、今まで我々が調べてきた結合の対称性を破り、非対称な結合を用いた場合の回路網の動作特性を調べておくことは意味のあることであろう。

そこで、ここでは行列 $A_{\mu\nu}$ を $A_{\mu\nu} = \delta_{\mu+1,\nu}$ と選ぶ。すなわち

$$w_{ij} = \frac{1}{N} \sum_{\mu=1}^p \xi_i^{\mu+1} \xi_j^\mu \quad (51)$$

とする。もはやこの場合には添え字 i, j の交換に対して不変ではないことに注意しておく。また、添え字 μ に関しては周期的な境界条件を課しておくことにする。つまり、 $\mu = p + 1$ は $\mu = 1$ である。

具体的に $p = 3$ の場合を書き下してみると

$$w_{ij} = \frac{1}{N} (\xi_i^2 \xi_j^1 + \xi_i^3 \xi_j^2 + \xi_i^1 \xi_j^3) \quad (52)$$

となる。

このときの回路網の動作を簡単に考えてみよう。そこで今、回路網の状態が 1 番目のパターン $S = \xi^1$ であるとしてみよう。すると、 i 番目の神経素子の内部ポテンシャル h_i は

$$\begin{aligned} h_i &= \sum_{j=1}^N w_{ij} \xi_j^1 = \frac{1}{N} \sum_{j=1}^N (\xi_i^2 \xi_j^1 + \xi_i^3 \xi_j^2 + \xi_i^1 \xi_j^3) \xi_j^1 \\ &= \xi_i^2 + \frac{1}{N} \sum_{ij} (\xi_i^3 \xi_j^2 + \xi_i^1 \xi_j^3) \xi_j^1 \end{aligned} \quad (53)$$

となるが、それぞれのパターンの各成分が無相関であるならば上式 (53) の右辺第 2 項は $\mathcal{O}(1/N \cdot \sqrt{N}) = \mathcal{O}(1/\sqrt{N})$ のオーダーとなり、素子数 N が十分大きければ無視できる²。従って、 $h_i = \xi_i^2$ となり、全ての神経素子に対して

$$\xi_i^2 = \text{sgn}(h_i) = \text{sgn}(\xi_i^2) \quad (54)$$

が成り立つ。従って、回路網の状態は $S = \xi^2$ となるので 2 番目のパターンが安定となる。次に、 $S = \xi^2$ としたときの内部ポテンシャルは

$$\begin{aligned} h_i &= \sum_{j=1}^N w_{ij} \xi_j^2 = \frac{1}{N} \sum_{j=1}^N (\xi_i^2 \xi_j^1 + \xi_i^3 \xi_j^2 + \xi_i^1 \xi_j^3) \xi_j^2 \\ &= \xi_i^3 + \frac{1}{N} \sum_{ij} (\xi_i^2 \xi_j^1 + \xi_i^1 \xi_j^3) \xi_j^2 = \xi_i^3 + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \end{aligned} \quad (55)$$

となるので、明らかに回路網の状態は $S = \xi^3$ となり、今度は 3 番目のパターンが安定となる。そこで最後に $S = \xi^3$ としたときの内部ポテンシャルを書き出してみると $h_i = \xi_i^1$ となるので、回路網の状態は $S = \xi^1$ となり、1 番目のパターンが安定となる。以上の考察から直ちに、このような非対称結合 (51) を採用することにより、回路網は

$$\xi^1 \rightarrow \xi^2 \rightarrow \xi^3 \rightarrow \xi^1 \rightarrow \dots \quad (56)$$

のような周期的な運動 (リミットサイクル) を生成する。

² ξ^1, ξ^2, ξ^3 のそれぞれが無相関であるのならば、 $\xi_i^3 \xi_j^2 \xi_j^1$ 等はランダムに ± 1 をとる変数となる。原点を中心とした歩幅 1 のランダム・ウォークの t ステップ後の減点からのずれは \sqrt{t} であったことを思い出すと、 $\sum_{j=1}^N \xi_i^3 \xi_j^2 \xi_j^1$ のオーダー (stochastic order) は \sqrt{N} となる。

1.10.2 計算機シミュレーションで動作を確かめる

ここでも例によって簡単なシミュレーションで回路網の動作を確かめておこう。

まずは $p = 3$ の場合. 回路網に埋め込むパターンをわかりやすいように図 13 のような 3 つの 2 値画像に選ぶ. また, 回路網のある時刻での状態をモニタリングするために, 次の 3 種類の重なりを各ステップで観



図 13: 回路網に埋め込む 3 つのパターン. 左から, ξ^1, ξ^2, ξ^3 . 画像サイズは 80×50 .

察することにしよう.

$$m_1(t) = \frac{1}{N} \xi^1 \cdot S(t) = \frac{1}{N} \sum_{j=1}^N \xi_j^1 S_j(t) \quad (57)$$

$$m_2(t) = \frac{1}{N} \xi^2 \cdot S(t) = \frac{1}{N} \sum_{j=1}^N \xi_j^2 S_j(t) \quad (58)$$

$$m_3(t) = \frac{1}{N} \xi^3 \cdot S(t) = \frac{1}{N} \sum_{j=1}^N \xi_j^3 S_j(t) \quad (59)$$

つまり, 先ほどの考察が正しければ, 回路網の状態 S は全ての時刻で ξ^1, ξ^2, ξ^3 どれかに一致しているはずなので, 各時刻の各パターンからの近さを重なりで測ろうというわけだ. 結果を図 14 に載せる. この図からきれいなリミットサイクルが確認できる. 図 15 に最初の 6 ステップの回路網の状態を載せる. この図で $t = 1$ のときの画像が ξ^1 から若干ずれているのは, 初期条件を ξ^1 に 20% 程度のノイズをかけた初期状態かた出発したためである. しかし, 十分に時間が立ち, 「リミットサイクル・アトラクタ」に軌道が落ち込んだ後にはきれいに $\xi^1 \rightarrow \xi^2 \rightarrow \xi^3 \rightarrow \xi^1 \rightarrow \dots$ という画像が現れる.

次に $p = 4$ の場合を考える. 埋め込むパターンの画像としては図 16 の 4 枚の画像を選ぶ. この図 16 の 4 番目のパターン ξ^4 は全ての要素が 1 であるようなパターンであることに注意しておこう. また, $p = 3$ の場合に観察した重なり m_1, m_2, m_3 の他に 4 番目のパターンと回路網の近さを測る量として

$$m_4(t) = \frac{1}{N} \xi^4 \cdot S(t) = \frac{1}{N} \sum_{j=1}^N \xi_j^4 S_j(t) \quad (60)$$

を導入し, 合計 4 つの重なりを各ステップで観察する. 結果を図 17 に載せる. この図 17 から明らかに $p = 3$ の場合と異なる点はパターン ξ^1, ξ^4 と回路網の重なりが時間が経過した後であっても 1 にはならない点である (図 18 も合わせて参照).

これは奇妙に思えるかもしれないが, 図 16 からわかるように, これら 2 つのパターン ξ^1, ξ^4 はお互い画像が黒っぽくかなり強い相関を持っている. 従って, (53) 式の右辺第 2 項のように我々が $\mathcal{O}(1/\sqrt{N})$ の寄与として無視した項がもはや無視できなくなり, $\mathcal{O}(1)$ の量になってしまっていることが原因であると考えられる.

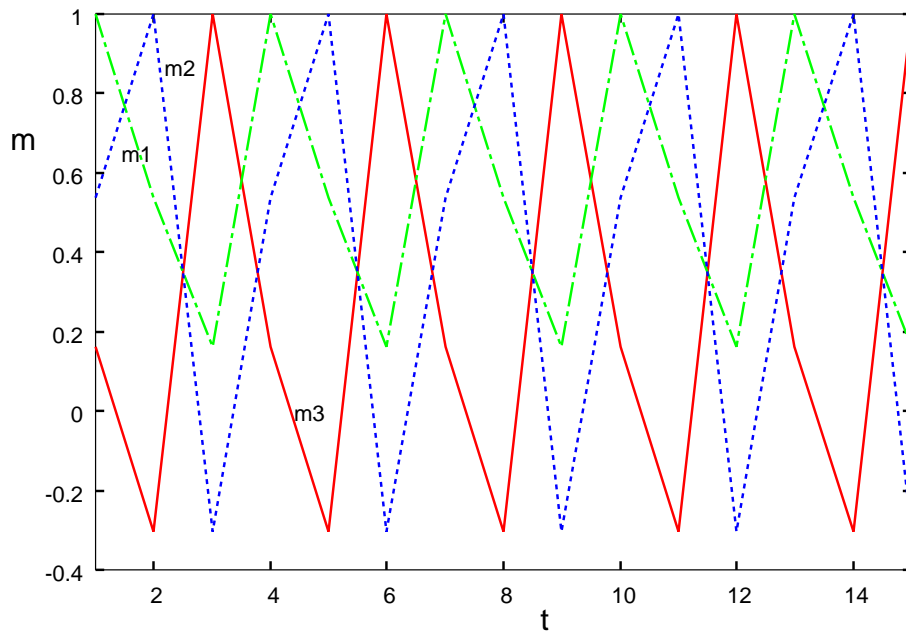


図 14: $p = 3$ の場合の 3 つの重なり m_1, m_2, m_3 の時間変化.

具体的には、回路網のある時点での状態が ξ^4 であるとき、素子 i の内部ポテンシャルは

$$h_i = \xi_i^1 + \frac{1}{N} \sum_j \xi_i^2 (\xi_j^1 \xi_j^4) + \mathcal{O}\left(\frac{1}{\sqrt{N}}\right) \tag{61}$$

となり、上式右辺の第 2 項は ξ^1, ξ^4 の相関が強くなるとオーダー $\mathcal{O}(1)$ の量であり³、第 1 項と同じ寄与を及ぼし、これがパターン ξ^1 の想起を阻害しているわけである。

問 3 : リミットサイクルとしてのパターン分離性がこの節の最後に述べたようにパターン間の相関によって阻害されるとするのであれば、直交パターンを 4 つ持ってきてそれらを用いて回路網を構成すれば全て 4 つのパターンがきれいに分離された形でリミットサイクルが実現できることが予想される。そこでこれを計算機シミュレーションにより確かめよ。直交パターンとしては前回説明したように各成分にランダムに ± 1 を割り振るのが手っ取り早い³が、他の方法で直交パターンを作る場合にはそれを用いても良い。具体的には m_1, m_2, m_3, m_4 の時間発展を図示せよ。

³ 極端な話、 $\xi^1 = \xi^4$ であるとすれば、(61) 式の右辺第 2 項は $(1/N) \sum_j \xi_i^2 (\xi_j^1 \xi_j^4) = \xi_i^2$ となり、これが第 1 項のシグナルに対するノイズとして寄与する。



図 15: 回路網の状態の変遷. 左上から右下に $t = 1, \dots, 6$ の回路網の状態.



図 16: 回路網に埋め込む 4 つのパターン. 左から, $\xi^1, \xi^2, \xi^3, \xi^4$. 画像サイズは 80×50 .

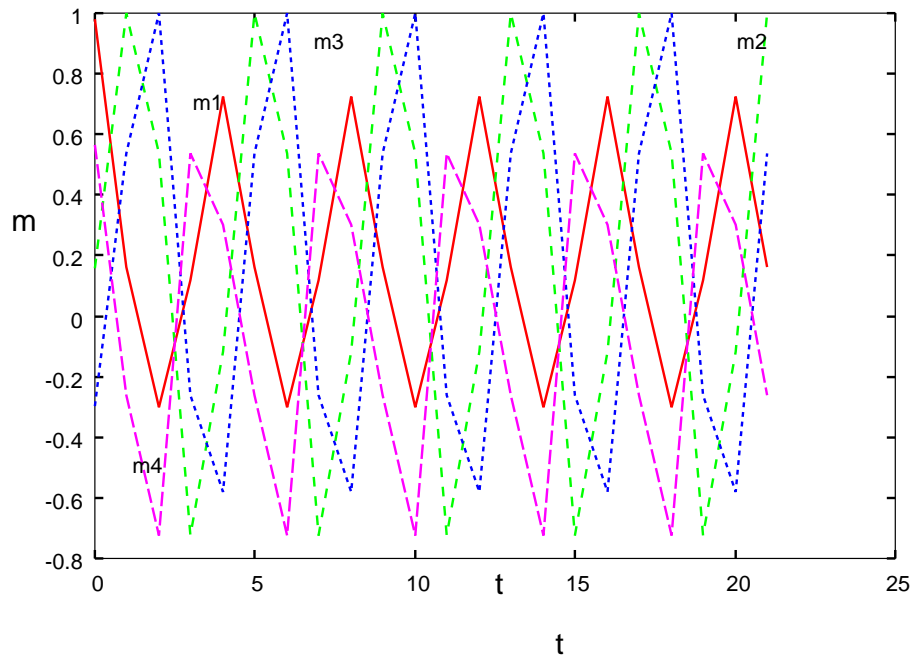


図 17: $p = 4$ の場合の 4 つの重なり m_1, m_2, m_3, m_4 の時間変化.



図 18: $p = 4$ の場合の回路網の状態の変遷. 左上から右下に $t = 1, \dots, 6$ の回路網の状態. $t = 1, 4, 5$ のとき, ξ^1, ξ^4 に対応するパターンがうまく分離されていない.