神経回路網特論
2011 年 5 月 6 日

課題 1: 自己組織化モデル

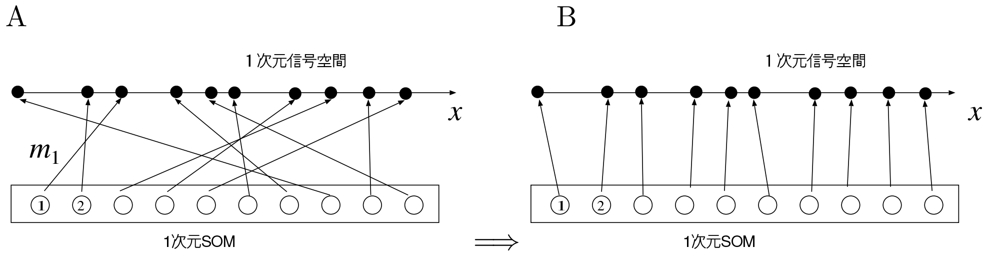
提出締切 5 月 20 日（金）正午

1. 信号 $\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}}$（T は転置）の各要素 x_i は区間 [0,1) の一様分布にしたがう．このような 2 次元の信号 $\boldsymbol{x}$ を 100 個生成し，平面（横軸 x_1，縦軸 x_2）上にプロットせよ．

2. 信号 $\boldsymbol{x}$ の 各要素 x_i が平均 $\boldsymbol{\mu} = (\mu_1, \mu_2)$, 分散 σ^2 の正規分布にしたがう場合を考えよう．ここで，$\sigma = 1.0$，μ_1, μ_2 は区間 [-5,5) の一様分布にしたがう乱数とする．このような信号 $\boldsymbol{x}$ を 100 個生成し，平面上にプロットせよ．

3. 前問において，確率的に中心位置 $\boldsymbol{\mu}$ を得ることを 3 回繰り返し，得られた $\boldsymbol{\mu}$ を $\boldsymbol{\mu}^1, \boldsymbol{\mu}^2, \boldsymbol{\mu}^3$ とする．σ の値を $0.3.0, 8, 2.0$ とした 3 種類のデータ群をそれぞれ 100 個生成し，平面上に重ねてプロットせよ（○，×，＋など異なる記号を使い表現すること）．

4. [**1 次元の信号空間に 1 次元 SOM を適用**]



素子数 $n = 10$ の 1 次元神経場モデル（自己組織化マップ，SOM）を考えよう．各素子は 1 次元の信号 x を受け取り，i 番目の素子は入力空間に対して m_i の重みをもつ（m_i は，一般的には，参照ベクトルとよばれている）．ここで $m_i, i = 1, \cdots, n$ の初期値は，区間 [0,1) の一様分布にしたがう乱数としておく（図 A）．SOM は，入力 x に対し，以下のルールにしたがい，一つの素子 c だけが興奮する（これを勝者とよぶ）．

$$c = \underset{i}{\operatorname{argmin}} |m_i - x| \tag{1}$$

21 通りの入力 $x = 0.0, 0.05, 0.10, 0.15, 0.20, \cdots, 1.0$ に対する勝者（素子の番号 c）を求めよ．

5. 区間 [0,1) の一様分布にしたがうランダムな入力信号 x を生成し，これに対する勝者を計算せよ．これを 1000 回くりかえし，各素子が勝者になった回数を求めよ．

6. モデルに入力信号 x を 1 回与えるごとに，以下の式にしたがい係数 $m_i, i = 1, \cdots, n$ の値を更新せよ（これを学習という）.

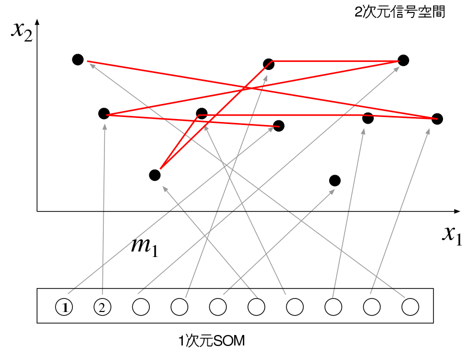
$$m_i \quad := \quad m_i + \alpha \exp\Big\{-\frac{|c-i|^2}{2\sigma^2}\Big\}(x - m_i) \tag{2}$$

ここで $\alpha > 0$ は学習の強さを表す係数，$\sigma > 0$ は，勝者素子が，どのくらい遠く離れた素子まで影響を与えるかを定める定数である（exp{} の部分は近傍関数とよばれている）．パラメータの値は，たとえば $\alpha = 0.2, \sigma = 0.8$ にしておく．これで 2000 回の学習をおこなえ．ここで 100 回の学習ごとに，前問 4,5 で求めた統計量を計算しなさい（このときは学習しない）.

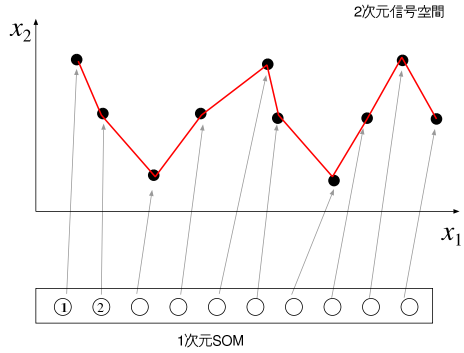
7. 図 B は学習後の様子を描いた例である．初期状態から，学習が進む様子を，この図 A,B を参考に作成せよ（本質的な部分が表現できればよい）．α, σ の値を適当な値に変えて実験せよ.

8. [**2 次元の信号空間に 1 次元 SOM を適用**]

C



D



素子数 $n = 10$ の 1 次元 SOM の各素子が 2 次元の信号 $\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}}$ を受け取る場合を考えよう．i 番目の素子は入力空間に対して $\boldsymbol{m}_i = (m_{i1}, m_{i2})^{\mathrm{T}}$ の重みをもつとする．ここで $m_{i1}, m_{i2}, i = 1, \cdots, n$ の初期値は，区間 [0,1) の一様分布にしたがう乱数とする（図 C）．入力信号 $\boldsymbol{x}$ が $[0,1) \times [0,1)$ 平面上に一様に分布する場合を考えよう．前問の場合と同様に，SOM は，入力 $\boldsymbol{x}$ に対し，一つの素子 c だけが興奮する.

$$c \quad = \quad \underset{i}{\operatorname{argmin}} ||\boldsymbol{m}_i - \boldsymbol{x}|| \tag{3}$$

入力信号 $\boldsymbol{x}$ を 1 回与えるごとに，以下の式にしたがい係数 $m_i, i = 1, \cdots, n$ の値を更新せよ.

$$m_i \quad := \quad m_i + \alpha \exp\Big\{-\frac{|c-i|^2}{2\sigma^2}\Big\}(\boldsymbol{x} - \boldsymbol{m}_i) \tag{4}$$

図Dは学習後の様子を描いた例である．初期状態から学習が進む様子を，この図C,Dを参考に作成せよ（入力信号$\boldsymbol{x}$の平面に，各素子の$\boldsymbol{m}_i$をプロットし，素子配列上で近傍関係にある素子どうしの$\boldsymbol{m}_i$を線で結べばよい）．パラメータの値は，たとえば$\alpha = 0.2, \sigma = 0.8$にしておく．1000回の学習をおこなえ．素子数を$n = 20, 50$など何通りかの値に変え，適宜，各パラメータの値，学習回数を調整して計算機実験をおこない，その結果を考察せよ．

9. [2 **次元の信号空間に** 2 **次元** SOM **を適用**]

素子数$n = 10 \times 10$の2次元SOMを考えよう．各素子は2次元の信号$\boldsymbol{x} = (x_1, x_2)^{\mathrm{T}}$を受け取り，$i$番目の素子は入力空間に対して$\boldsymbol{m}_i = (m_{i1}, m_{i2})^{\mathrm{T}}$の重みをもつとする．ここで$m_{i1}, m_{i2}, i = 1, \cdots, n$の初期値は，区間[0,1)の一様分布にしたがう乱数とする．入力信号$\boldsymbol{x}$が$[0,1) \times [0,1)$平面上に一様に分布する場合を考えよう．SOMは，入力$\boldsymbol{x}$に対し，式(3)にしたがい，一つの素子cだけが興奮する．モデルに入力信号$\boldsymbol{x}$を1回与えるごとに，以下の式にしたがい係数$\boldsymbol{m}_i, i = 1, \cdots, n$の値を更新せよ．

$$\boldsymbol{m}_i \;\; := \;\; \boldsymbol{m}_i + \alpha \exp\Big\{ -\frac{||\boldsymbol{r}_c - \boldsymbol{r}_i||^2}{2\sigma^2} \Big\} (\boldsymbol{x} - \boldsymbol{m}_i) \qquad (5)$$

ここで$\boldsymbol{r}_i$は，i番目の素子の配列上（神経場）での位置である．配列上の位置とは，$n = 10 \times 10$の場合，$i = 1, \cdots, 100$であり，$\boldsymbol{r}_1 = (0, 0)$，$\boldsymbol{r}_2 = (1, 0), \boldsymbol{r}_3 = (2, 0), \cdots, \boldsymbol{r}_{11} = (0, 1), \boldsymbol{r}_{12} = (1, 1), \cdots, \cdots, \boldsymbol{r}_{100} = (9, 9)$などと考えればよい．$\boldsymbol{r}_i = (x, y)$と書くと，$x, y = 0, \cdots, 9$　として$i = x + 1 + 10y$と表現できる．パラメータの値は，たとえば$\alpha = 0.2, \sigma = 0.8$にしておく．20,000回の学習をおこなえ．

10. 前問において，入力信号$\boldsymbol{x}$の確率分布を適当に設計し，生成したデータをモデルに与え，どう学習が進むか，いろいろ試して考察せよ．講義中に例は示すが，面白い振る舞いをする例を独自に探せ．入力信号の次元は3次元以上にしてもよい．

注意：　課題が手に負えない場合，はやめにA-334かA-423まで，相談にくること．

補足：　「脳はどういう風に動いているのか」，このモデルを通して感じ取れる可能性のある性質をいくつか挙げておく．

(a) トポグラフィックマップの形成：でたらめな結合が，整然とした結合に変化していく．

(b) 入力空間において，入力頻度の高い領域ほど多くの細胞が割り当てられる．

→ 入力信号 $\boldsymbol{x}$ の確率分布を一様分布から偏ったものにしてみる．

(c) 一部の細胞（素子，ニューロン）が壊れても，他の細胞が代わりに働くようになる．

→ 学習の途中で一部の細胞を取り除いてみる．

(d) 高次元の情報を低次元でうまく表現できる．

→ 課題 8. がこれに対応している．

細かい点（学習回数など）は自分で適当に決めて進めればよい．その他，課題に曖昧な点があった場合も適当に解釈していい．その場合，どこをどう解釈したか，そう解釈した理由等をレポートに記述すること．レポートの最後には，感想，質問などを記述してほしい．今後のため,「理解できなかった」という記述よりは，具体的にどこの部分から分からなくなったか記述してもらえれば助かる．

レポートの LaTeX を使った簡単な書き方は

`http://www.cs.miyazaki-u.ac.jp/~date/lectures/latex/latexreport.html`

を参照．