複製・競合を考慮した学習アルゴリズムの適用

2120014　小澤翔太　　情報基盤工学講座　指導教員 奥原 浩之

要約

動径基底関数ネットワーク（**Radial Basis Function Network：RBFN**）は，任意の非線形関数に対し，基底関数を足し合わせることで関数近似を効果的に行うことができるニューラルネットワークのうちの一つであり，様々な問題に適用されている．しかし，**RBFN**には冗長なニューロンの存在や，基底関数が不足していることによって関数近似ができない場合があるといった問題点がある．

そこで本研究では，このような問題に対して，ニューロン間の競合や，基底関数の複製を考慮した**RBFN**を提案する．そして，新たなアルゴリズムを機械学習へ適用し，さらなる効率化・改善を目的とする．

キーワード：シナプス結合荷重，シナプス可塑性方程式，競合動径基底関数ネットワーク，複製，学習の高速化

1 はじめに

関数近似問題やパターン識別に適したニューラルネットワークの一つに**RBFN**がある（図**1**参照）．

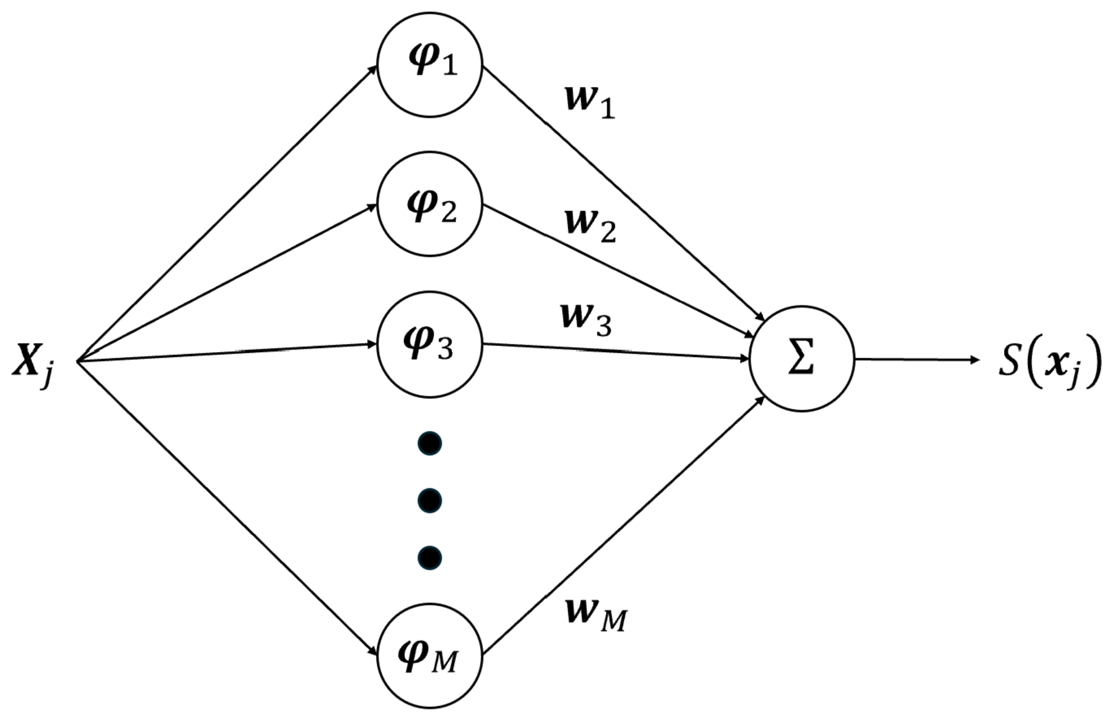


図1 **RBFN**の構造

RBFNは階層型ニューラルネットワークに比較してニューロンごとの局所的な学習が可能であるなどの優れた点を持つ．しかし，**RBFN**では未知の非線形関数を近似するため，あらかじめ必要なニューロン数が不明であり冗長なニューロンを必要とする．一般に，ニューロンの増加は学習の遅延化や過学習の問題を生じることが知られている．

これらの問題を解決するために，適者生存型学習則に基づいたシナプス可塑性方程式を適用した，競合動径基底関数ネットワーク（**Competitive RBFN：CRBFN**）が提案されている[1]．**CRBFN**は，シナプス結合荷重間に競合を生じさせ，学習に必要なニューロンのみが自然に生き残るようになっており，冗長なニューロンの削減を図ることができる．しかし，**CRBFN**では基底関数を追加する機能は無く，基底関数の数が足りない場合は，関数近似自体が不可能となる．そこで，**CRBFN**に基底関数を複製して追加する機能を加えた，複製・競合動径基底関数ネットワーク(**Reproductive CRBFN：RC-RBFN**)が提案されている[2]．

そこで本研究では，**RC-RBFN**が効率よく動径基底関数を削除あるいは追加するニューラルネットワークであることを示したのち，機械学習のアルゴリズムに適用して従来より学習が効率的に行われていることを示す．

2 複製・競合を考慮した動径基底関数ネットワーク

2.1 競合動径基底関数ネットワーク

ニューラルネットワークは大きく分けて，素子であるニューロン，それらを結合するシナプス，そして動作規則により構成される．なかでも，記憶にもっとも関係した情報処理は，シナプスにおいて行われているとされる．記憶には種々のものが考えられるが，本論文では短期記憶と長期記憶に着目し，短期記憶はニューロンの発火頻度，長期記憶は細胞膜の特性の変化により生じるものとする．シナプスの可塑性を記述する方程式は，これらの要因を含んだものとなっていなければならない．また，実際の生体では，シナプス結合の性質が興奮性，抑制性であるのかは送り出すニューロンにより決まる(**Dale**則)．さらに，微小な領域では成長や活動に必要な神経成長因子（**Nerve Growth Factor：NGF**）は競合によってシナプスに摂取される．

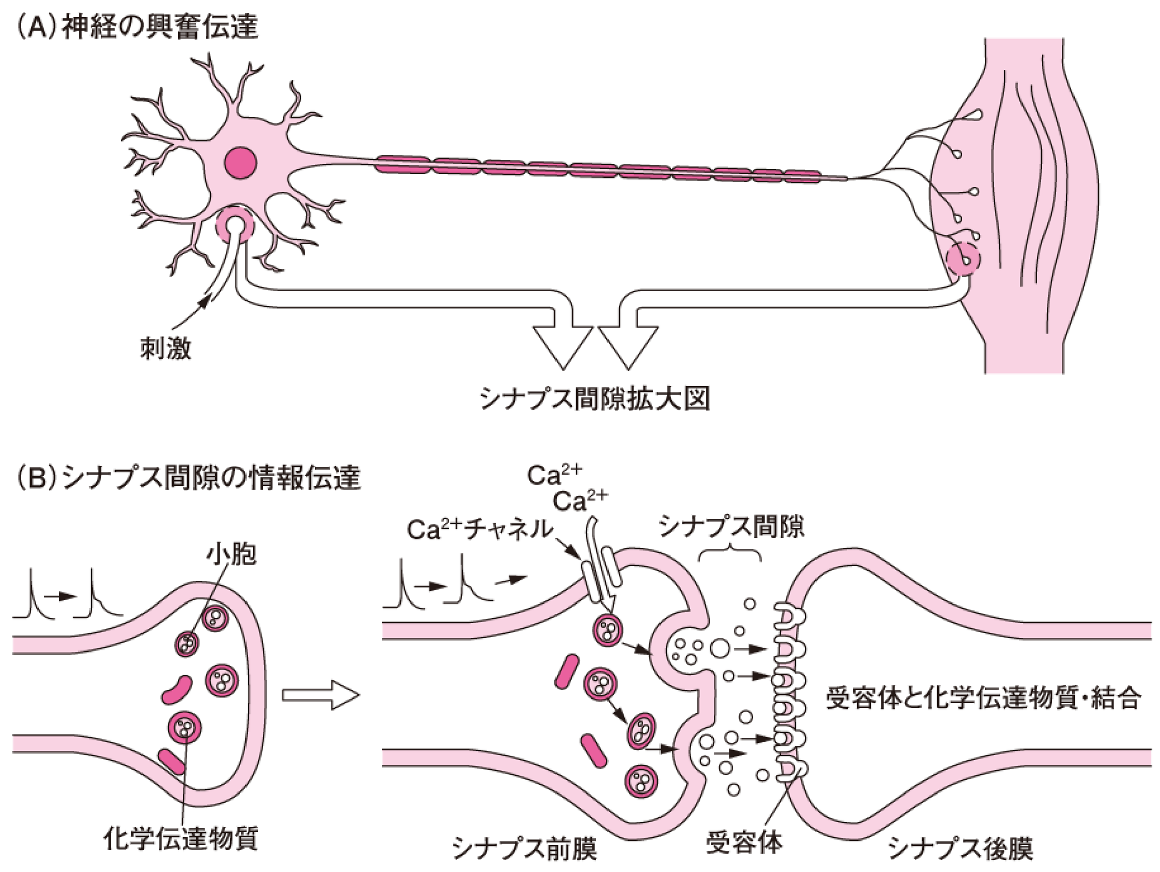


図2 微小領域における**NGF**摂取の様子

これらの事実もシナプス可塑性のモデル化において重要な要因であると考えられる．そこで，発火頻度や膜の特性変化を生じる物質の時間変化と，生理学的拘束条件である**Dale**則や微小な領域での競合を考慮したシナプス間感度の大きさの時間変化は以下の方程式に従うものとする（図**2**参照）．

Dale則を考慮したシナプス可塑性方程式

$$\frac{dw_{ik}^{j}}{dt} = (\alpha_{ik}^{j} - \sum_{h} \gamma_{ik}^{jh} w_{ik}^{h}) w_{ik}^{j}$$

$$\alpha_{ik}^{j} = \int_{x \in B_{ik}} \eta_{ik}(x)\, \xi_{ik}^{j}(x) dx \qquad \gamma_{ik}^{jh} = \int_{x \in B_{ik}} \xi_{ik}^{j}(x)\, \xi_{ik}^{h}(x) dx$$

$$\xi_{ik}^{j}(x) = \mu_i \exp\left\{-\frac{1}{2}\left(\frac{x - m^j}{\sigma^j}\right)^2\right\}$$

α_{ik}^{j}：内的自然増加率　　γ_{ik}^{jh}：第j,hニューロン間の競合の効果

ξ_{ik}^{j}：第jニューロンの微小領域B_{jk}におけるシナプス前終末発火頻度　　η_{ik}：ξ_{ik}^{j}が作用する第iニューロンの細胞膜におけるシナプス後発火頻度

図3 シナプス可塑性方程式

このような，競合を考慮したシナプス可塑性方程式に従う**RBFN**が**CRBFN**である．

2.2 ターミナルアトラクタと基底関数の複製

先ほど述べた**CRBFN**では，競合に負けたシナプス結合荷重は平衡状態で**0**になるが，平衡解への漸近は指数関数的に行われるので原理的には有限時間で平衡状態へ到達することはできない．

そこで，あるシステムが最終的に安定した「学習結果」や「最適解」に収束する，というターミナルアトラクタ（**Terminal Attractor**）の概念を適用して，与えられた時刻t^*で平衡解へ収束するように修正されたシナプス可塑性方程式を導出する．まず，動的システムの安定性を解析するために使われる数学的なツールである**Lyapunov**関数とその時間変化，そして与えられた時刻t^*で平衡解へ収束するように修正されたシナプス可塑性方程式を以下のように定義する(図**3**参照)．

Lyapunov関数とその時間変化

$$V(\boldsymbol{w}_{ik}) = \frac{1}{2}\int_{x \in B_{ik}} \{\eta_{ik}(x) - s_{ik}(x)\}^2 dx\,, \quad \frac{dV(\boldsymbol{w}_{ik})}{dt} = -\frac{V(\boldsymbol{w}_{ik}^{0})^{R} V(\boldsymbol{w}_{ik})^{\frac{1}{r}}}{Rt^*}$$

$$\frac{dw_{ik}^{j}}{dt} = \frac{\left(\alpha_{ik}^{j} - \sum_{h=1}^{N_{ik}} \gamma_{ik}^{jh} w_{ik}^{h}\right) w_{ik}^{j}}{\sum_{j=1}^{N_{ik}} w_{ik}^{j}\left(\alpha_{ik}^{j} - \sum_{h=1}^{N_{ik}} \gamma_{ik}^{jh} w_{ik}^{h}\right)^2} \times \frac{V(\boldsymbol{w}_{ik}^{0})^{R} V(\boldsymbol{w}_{ik})^{\frac{1}{r}}}{Rt^*}$$

$\boldsymbol{w}_{ik} \equiv \left[w_{ik}^{1}, w_{ik}^{2}, \dots, w_{ik}^{N_{ik}}\right]^{\mathrm{T}} \in R^{N_{ik}}$　　$\boldsymbol{w}_{ik}^{0}$：$\boldsymbol{w}_{ik}$の初期値

r：任意の奇数　　$R : R = \frac{r-1}{r}$　　$s_{ik}(x) = \sum_{j=1}^{N_{ik}} w_{ik}^{j} \xi_{ik}^{j}(x)$

図4 **Lyapunov**関数とその時間変化

ここまでで，シナプス結合荷重間の競合を考慮したシナプス可塑性方程式を利用することにより学習の効率化を図る手法が提案されてきた．ところが，**CRBFN**は冗長な動径基底関数を削除する能力をもつものの，必要とされる動径基底関数を追加する能力は備えていない．ニューラルネットワークに関数近似を行うために必要な数のニューロンが存在しない場合は，関数近似を行うこと自体が不可能となる．

そこで，新しい動径基底関数を追加する能力を備えたニューラルネットワークとして**RC-RBFN**が提案されている．この**RC-RBFN**はシナプス可塑性方程式に関する考察から得られるものであり，必要な動径基底関数を効率的に追加することができる．簡単のため以降はとあるiとkについてのみ考えるものとし，添え字のi, kは省略する．

ここで，動径基底関数の複製アルゴリズムに論じるにあたって自由エネルギーを以下のように導出する（図**4**参照）．自由エネルギーはシステムのエネルギーとエントロピー（無秩序さ）のバランスを表す量であり，機械学習においては自由エネルギーの最小化を通じてパラメータの推定を行う．

自由エネルギーの導出

累積2乗誤差関数：$E(m) = \frac{1}{2}\sum_x E(x,m) = \frac{1}{2}\sum_x \{\eta(x) - s(x)\}^2$,

分配関数：$Z_\beta(m) = -\sum_x exp\{-\beta E(x,m)\}$

とすると条件付き確率密度関数$p_\beta(x|m)$は確率の正規化と，2乗誤差関数$E(x,m)$の条件付き期待値

$$\langle E(m) \rangle_\beta = \sum_x p_\beta(x|m) E(x,m)$$

が一定となるという二つの制約のもとで，エントロピー

$$S_\beta(m) = -\frac{1}{\beta}\sum_x p_\beta(x|m) \log p_\beta(x|m)$$

を最大化する確率密度関数として導出できる．このとき，自由エネルギーは

$$F_\beta(m) = -\frac{1}{\beta}\log Z_\beta(m)$$

で定義される．

図5 自由エネルギーの導出

このような自由エネルギーは，データのクラスタリングのための手法であるメルティング[4]においても同様に定義されている．メルティングとは，$m = x$かつβが∞である初期状態から，徐々にβを0へ近づけていきながら，パラメータmを自由エネルギー$F_\beta(m)$の最急降下方向に更新していくものである．その結果，パラメータmは徐々に同じ値をとりはじめ，最終的に一つの値に収束する．

3 提案手法

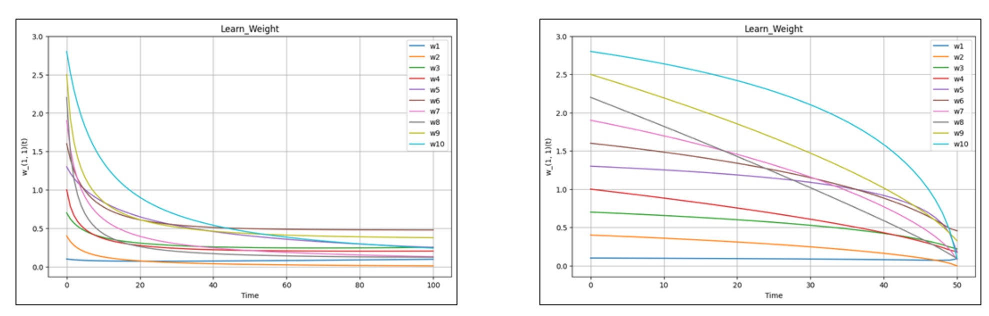
本研究で提案するシステムの概要について説明する．本研究では機械学習の学習アルゴリズムとして，

4 数値実験並びに考察

ターミナルアトラクタの比較（仮）数値実験では，問題数を**10**，解答者を**20**人としてデモデータを用意し，**k**平均法によりクラスタリングを行った．用意したデータは，正答率，解答時間，自信度でありそれぞれの値を標準化し，５つにクラスタリングした．数値実験の結果を図**5**に示す．

それぞれのクラスタに対して考察した．まず０に対しては，自信度と正答率がともに高く，解答時間は平均的であった．よって，問題をよく理解できている人と考察できる．次に１は，正答率と自信度が低く，解答時間が長いため，自分自身が不得意と認識していると考えられる．２は，自信度が高いが，正答率が低くなっているため，客観的指標と主観的指標のずれがみられ，誤った認識をしていると考えられる．３は，自信度が高い割には正答率が低めであり，解答にも時間がかかっていることから，比較的理解ができているが，理解できていないところがあるように考えられる．４は，正答率に対して自信度が低いことから，偶然問題が当たった可能性があり，理解度が正答率に対して，伴っていないように考えられる[7]．

TA適用前後の結果



	w_1	w_2	w_3	w_4	w_5	w_6	w_7	w_8	w_9	w_{10}
初期値	0.1	0.4	0.7	1.0	1.3	1.6	1.9	2.2	2.5	2.8
適用前	0.097	0.011	0.252	0.204	0.251	0.478	0.131	0.123	0.377	0.242
適用後	0.078	0.008	0.220	0.180	0.266	0.446	0.118	0.102	0.343	0.207

➡ 学習回数を少なくしながら，同様の結果を得ることができた

図6 **TA**適用前・適用後のシナプス結合荷重

5 おわりに

本研究では，**FlexSim**を用いて正誤データや解答時間のデータ，自信度データを取得し，解答者の特徴や理解度を可視化した．今後の課題として，実際の処方せんの問題の解答データで数値実験を行い解析することと，臨場感と接遇マナーの追加を行うことを考えている．

参考文献

[1] 奥原 浩之, 尾崎 俊治, “適者生存型学習則を適用した競合動径基底関数ネットワーク”, 電子情報通信学会論文誌, pp. 3191-3199, 1997

[2] 奥原 浩之, 佐々木 浩二, 尾崎 俊治, “環境の変化に適応できる複製・競合動径基底関数ネットワーク”, 電子情報通信学会論文誌, pp. 941-951, 1999

[3] 奥原 浩之, 尾崎 俊治, “Dale 則を考慮したシナプス可塑性方程式の解析”, システム制御情報学会論文誌, pp. 718-720, 1995

[4] Wong, Y.-. fai . Clustering Data by Melting. Neural Computation 1993, 5 (1), 89–104. https://doi.org/10.1162/neco.1993.5.1.89.