

はじめに
やったこと
文献調査 (1)
文献調査 (2)
文献調査 (3)
まとめ

進捗報告

小澤 翔太

富山県立大学 情報システム工学科

2024 年 8 月 9 日

中間発表の仮テーマにしようと考えているもの

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の教師無し学習または強化学習への適用

本研究の目的

教師あり学習のアルゴリズムである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を教師無し学習に組み込むことで、従来のアルゴリズムと比較した場合の処理速度の向上を図る。

文献調査

- ・ Dale 則を考慮したシナプス可塑性方程式の解析
- ・ 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
- ・ 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

はじめに

やったこと

文献調査 (1)

文献調査 (2)

文献調査 (3)

まとめ

背景

生体では、ニューロン同士のシナプス結合の性質は信号を送り出すニューロンによって決まっているとされる。ニューロンの微小領域では、活動に必要な物質は競合が行われシナプスに摂取される。これらの原理を明らかにすることは、シナプス可塑性をモデル化するにあたって重要なことであると考えられる。

目的

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 則や微小領域での競合を考慮したシナプス可塑性方程式を解析することで、シナプスでは分岐を利用した情報処理が可能であることを示す。

Dale 則を考慮したシナプス可塑性方程式の解析

Dale 則：ニューロンは単一な性質であり，シナプス間感度が興奮性であるか抑制性であるかは送り出すニューロンによって決まる．

$$\frac{d\mu_j w_{ij}^k}{dt} = (G_i^k - \frac{1}{\alpha_i^k} \sum_{h \neq j} \mu_h w_{ih}^k + \overline{\varepsilon \eta_i^k \xi_{ij}^k}) \mu_j w_{ij}^k - \frac{\beta_i^k}{\alpha_i^k} (w_{ij}^k)^2 + f_{ij}^k \quad (1)$$

μ_j ：第 j ニューロンが興奮性か抑制性であるかの識別子

w_{ij}^k ：シナプス間感度

G_i^k ：微小領域への神経成長因子 (NGF) の供給速度

$\alpha_i^k, \beta_i^k, \varepsilon$ ：正の定数

η_i^k ：信号を受け取る側のニューロンの発火頻度

ξ_{ij}^k ：信号を送る側のニューロンの発火頻度

f_{ij}^k ：NGF と環境因子に依存する揺らぎ

シナプス可塑性方程式の探索機能における競合と分岐

$$\dot{x}_i = (\lambda_i - \alpha x_i - \sum_{j \neq i} x_j) x_i \quad (i = 1, \dots, N) \quad (2)$$

$x_i \geq 0$, $\alpha > 1$, $\lambda_1 > \lambda_2 > \dots > \lambda_N > 0$ であり, この系は以下のような分岐特性があり, それをシミュレーションした結果を示す.

• $\frac{\lambda_N}{\lambda_1} > \alpha > 0$ の場合

$x_i = \frac{\lambda_i}{\alpha}$ ($i = i_0 : i_0 \in 1, \dots, N$), $x_i = 0$ ($i \neq i_0$)

$\lambda = [1.0, 2.0, 3.0, 4.0, 5.0]$ とすると, $5 > \alpha > 0$ となり, $\alpha = 2$ とした.

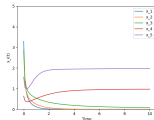


図 1

• $\frac{\lambda_n}{\lambda_1} > \alpha > \frac{\lambda_{n+1}}{\lambda_1}$ の場合

$$x_i = \frac{\lambda_i}{\alpha} \quad (i = i_0 : i_0 \in 1, \dots, n), \quad x_i = 0 \quad (i \neq i_0)$$

$\lambda = [1.0, 2.0, 3.0, 4.0, 5.0], n = 3$ とすると, $4 > \alpha > 3$ となり,
 $\alpha = 3.5$ とした.

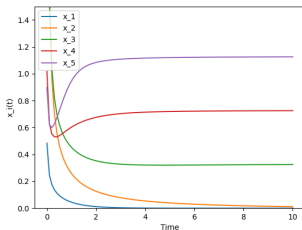


図 2

• $1 + \Lambda_1 > \alpha > \frac{\lambda_2}{\lambda_1}$ の場合

$$x_i = \frac{\lambda_i}{\alpha} \quad (i = 1), \quad x_i = 0 \quad (i = 2, \dots, N)$$

$\lambda = [4.0, 3.0, 1.0, 2.0, 5.0]$ とすると, $4 > \alpha > \frac{1}{4}$ となり, $\alpha = 1$ とした.

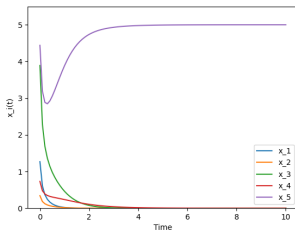


図 3

ただし, $\Lambda_n = \sum_{j=1}^n (\frac{\lambda_j}{\lambda_{n+1}} - 1)$ であり, $\Lambda_0 = 0$ である.

・ $1 + \Lambda_n > \alpha > 1 + \Lambda_{n-1}$ の場合

$$x_i = \frac{(\alpha+n-1)\lambda_i - \sum_{j=1}^n \lambda_j}{(\alpha-1)(\alpha+n-1)} \quad (i = 1, \dots, n), \quad x_i = 0 \quad (i = n+1, \dots, N)$$

$\lambda = [5.0, 4.0, 3.0, 2.0, 1.0]$, $n = 3$ とすると, $7 > \alpha > 2.5$ となり,
 $\alpha = 3$ とした.

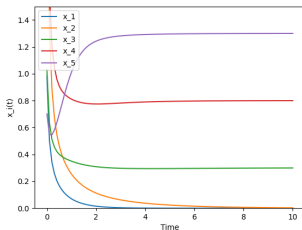


図 4

・ $\alpha > 1 - \Lambda_{N-1}$ の場合

$$x_i = \frac{(\alpha + N - 1)\lambda_i - \sum_{j=1}^n \lambda_j}{(\alpha - 1)(\alpha + N - 1)} \quad (i = 1, \dots, N)$$

$\lambda = [5.0, 4.0, 3.0, 2.0, 1.0]$ とすると, $\alpha > 10$ となり, $\alpha = 14$ とした.

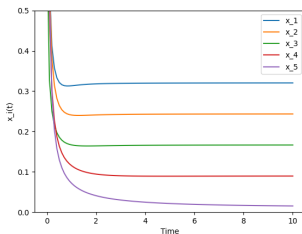


図 5

背景

動径基底関数ネットワーク (RBFN) はニューロンごとに局所的な学習が可能であるという特長がある。しかし、未知の非線形関数を近似しようとした場合、必要なニューロン数が不明なために冗長なニューロンが必要となる。これを解決する方法として、望ましい出力と相関が高い入力を伝達しているシナプス結合荷重が生き残り、生き残ったシナプス結合荷重により伝達される入力に変化を受ける入力ニューロンの入出力特性のみを調節することが挙げられる。

目的

シナプス結合荷重間に入出力の相関に応じた競合が生じるようなシナプス可塑性方程式を導出し、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 (CRBFN) を提案する。

RBFN

非線形関数を $\eta(X)$, 動径基底関数を $\xi^i(X)$ とすると d 次元のベクトル $X_j \in R^d (j = 1, 2, \dots, N)$ が入力されたとき, 第 i 入力ニューロンは入力ベクトル X_j に対して

$$\xi^i(X_j) = \exp\left(-\frac{1}{2}(X_j - m_i)^T \sum_i^{-1} (X_j - m_i)\right) \quad (3)$$

を出力する. 出力された $\xi^i(X_j)$ はシナプス結合荷重の重みを付加されたのち, 出力ニューロンへ伝達されて足し合わされて

$$s(X_j) = \sum_i^N \xi^i(X_j) \quad (4)$$

が出力される.

具体的な関数近似方法

RBFN では以下の 2 乗誤差評価関数を減少させることにより、関数近似を行う。

$$E(W) = \frac{1}{2} \sum_j^M (\eta(X_j) - s(X_j))^2 \quad (5)$$

ここに Delta ルール (誤差を最小化するようにニューラルネットワークの重みを調整するアルゴリズム) を適用することにより、パラメータを取得していく。

Dale 則を考慮したシナプス可塑性方程式

$$\frac{dw_{ik}^j}{dt} = (\alpha_{ik}^j + G_{ik} - \sum_{h=1}^{N_{ik}} \gamma_{ik}^{jh} \mu^h w_{ik}^h) w_{ik}^j + f_{ik}^j \quad (6)$$

μ^j : 第 j ニューロンが興奮性か抑制性であるかの識別子

w_{ik}^j : シナプス間感度

G_{ik} : 微小領域への神経成長因子 (NGF) の供給速度

α_{ik}^j : 正の定数

f_{ik}^j : NGF と環境因子に依存する揺らぎ

γ_{ik}^{jh} : 競争係数 (第 j ニューロンと第 h ニューロンが同時に NGF を消費することによる競合効果)

望ましい時間 t^* で平衡解へ収束する修正版シナプス可塑性方程式

$$\frac{dw_{ik}^j}{dt} = \frac{(\alpha_{ik}^j - \sum_{h=1}^{N_{ih}} \gamma_{ik}^{jh} \mu^h w_{ik}^h) w_{ik}^j}{\sum_{j=1}^{N_{ik}} \mu^j w_{ik}^j (\alpha_{ik}^j - \sum_{h=1}^{N_{ih}} \gamma_{ik}^{jh} \mu^h w_{ik}^h)^2} \times \frac{V(w_{ik}^0)^R V(w_{ik})^{\frac{1}{r}}}{R t^*} \quad (7)$$

μ^j : 第 j ニューロンが興奮性か抑制性であるかの識別子

w_{ik}^j : シナプス間感度

G_{ik}^j : 微小領域への神経成長因子 (NGF) の供給速度

α_{ik}^j : 正の定数

f_{ik}^j : NGF と環境因子に依存する揺らぎ

γ_{ik}^{jh} : 競争係数 (第 j ニューロンと第 h ニューロンが同時に NGF を消費することによる競合効果)

関数近似への適用

簡単な 1 次元の非線形関数 $\eta(x)$ を例にして, 提案された CRBFN が RBFN と比較して, 高速に関数近似が可能なことを示す. $\eta(x)$ を以下のように定義する.

$$\eta(x) = 3N(-1.5, 1) + 2N(1, 0.5) \quad (8)$$

サンプルは $x \in [-5, 5]$ の間から 0.1 ずつ 101 個取り, 教師信号として $(x_j, \eta(x_j)) j = 1, 2, \dots, 101$ を与える.

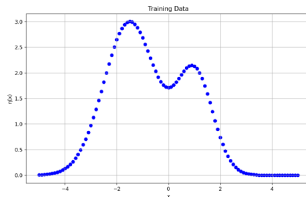


図 6: 教師信号の分布図

RBFN との比較

学習前のシナプス結合荷重と学習が収束したときのシナプス結合荷重を図 7, 学習が収束するまでの学習回数を図 8 に示す. 学習収束判定は $E(x) \leq 0.0001$ とした.

	w^1	w^2	w^3	w^4	w^5	w^6
初期値	0.431	0.898	0.324	0.927	0.389	0.161
RBFN	0.290	1.880	2.065	1.054	0.109	0.109
CRBFN	0.000	0.000	2.990	1.986	0.000	0.000

図 7: 学習前・学習後のシナプス結合荷重

	全体ループ	局所ループ
RBFN	219	10723224
CRBFN	11	232710

図 8: 学習が収束するまでの学習回数

背景

競合動径基底関数ネットワーク (CRBFN) は競合によって冗長なニューロンを消滅させることが可能であり、その結果として学習の高速化と過学習の回避が行われる。しかし、CRBFN には新しい動径基底関数を追加する機能がないため教師信号が変化するような環境の変化には対応できないという課題が残っている。

目的

シナプス可塑性方程式に関する考察から得られた、新しい動径基底関数を追加する手法を CRBFN に組み込んだニューラルネットワークとして、複製・競合動径基底関数ネットワーク (RC-RBFN) を提案する。

また、提案した RC-RBFN の性能をシミュレーションによって確認する。

CRBFN における Dale 則を考慮したシナプス可塑性方程式

$$\Delta w_j = \varepsilon(\alpha_j(\phi) - \sum_{k=1}^M \gamma_{jk}(\phi) \mu_k w_k) w_j \quad (9)$$

μ^k : 第 k ニューロンが興奮性か抑制性であるかの識別子

w_j : 第 j ニューロンのシナプス結合荷重 (興奮性なら正, 抑制性なら負となる).

$\alpha_j(\phi)$: 内的自然増加率

$\gamma_{jk}(\phi)$: 競争係数 (第 j ニューロンと第 k ニューロンが同時に NGF を消費することによる競合効果)

シナプス結合荷重 w_j の変数変換

$$y_j^2 = \mu_j w_j \quad (10)$$

y_j の定義域は任意の実数であり、このとき (9) 式は以下のようになる.

$$\frac{dy_j}{dt} = \left(\frac{\alpha_j(\phi)}{2} - \sum_{k=1}^M \frac{\gamma_{jk}(\phi)}{2} w_k^2 \right) y_j \quad (11)$$

ただし, ε は省略した.

このとき, ポテンシャルを求めることが可能であり, ポテンシャル $V(w, \phi)$ は

$$V(w, \phi) = - \sum_{j=1}^M \left(\frac{\alpha_j(\phi)}{4} \mu_j w_j - \sum_{k \neq j}^M \frac{\gamma_{jk}(\phi)}{4} \mu_j w_j \mu_k w_k - \frac{\gamma_{jj}(\phi)}{8} w_j^2 \right) \quad (12)$$

と求められる.

確率密度関数の導出

その結果、累積 2 乗誤差関数 $E(w, \phi)$ は以下のようにになる。

$$E(w, \phi) = 4V(w, \phi) + \frac{1}{2} \sum_{i=1}^N \eta^2(x_i) \quad (13)$$

以上より $E(w, \phi)$ の最小化は、ポテンシャル $V(w, \phi)$ の最小化と等価であることが分かる。

ここで、(11) 式の y_j はポテンシャル $V(w, \phi)$ の最急降下方向に更新されて極小解に収束し、その解から逃げられなくなる。そこで、極小解から脱出するための手法として更新則を以下のように、ノイズを考慮した離散近似した見本過程で与えることが考えられる。

$$y_j(t + \Delta t) = y_j - \frac{\partial V(y, \phi)}{\partial y_j} \Delta t + \sqrt{Q \Delta t} n_j(t) \quad (14)$$

$n_j(t)$: 平均 0, 分散 1 の正規分布に従う独立な確率変数
 Q : 任意の正の定数

確率密度関数の導出

このとき、学習終了時に CRBFN の平均ベクトル、共分散行列とシナプス結合荷重が満たす同時確率密度 $p_{\beta}(w, \phi)$ は

$$p_{\beta}(w, \phi) = Z_{\beta}^{-1} \exp(-\beta V(w, \phi)) \quad (15)$$

ここで、 $\beta = \frac{2}{Q}$ 、 Z_{β} は分配関数である。また、(15) 式はポテンシャル $V(w, \phi)$ と累積 2 乗誤差関数 $E(w, \phi)$ の関係式 (13) より

$$p_{\beta'}(w, \phi) = Z_{\beta'}^{-1} \exp(-\beta' E(w, \phi)) \quad (16)$$

と書き直すことができる。ここで、 $\beta' = (2Q)^{-1}$ であり、 $Z_{\beta'}$ は分配関数である。

確率密度関数の導出

ここで，教師信号 $\eta(x)$ を (17) 式で与える．

$$\eta(x) = 3N(-1.5, 1) + 2N(1, 0.5) \quad (17)$$

この場合，近似する非線形関数 $\eta(x)$ の複雑さに対して，必要とされている動径基底関数が十分に存在していないため，累積 2 乗誤差関数を 0 にすることが不可能である．しかし，この動径基底関数のパラメータ m が従う条件付き確率密度関数は導出可能で，図 9 のようになる．

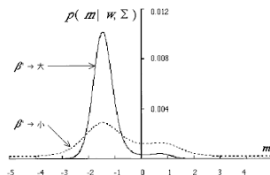


図 9: パラメータ m の条件付き
確率密度関数

複合・競合動径基底関数ネットワークの更新則

RC-RBFN のパラメータ m_j^k の更新則を

$$\Delta_{\beta'} m_j^k = \langle \Delta m_j^k \rangle_{\beta'} \quad (18)$$

で与えることとする。ここで,

$$\Delta_{\beta'} m_{j[i]}^k = -\varepsilon \frac{\partial E(x_i, m_{j[i]}^k)}{\partial m_{j[i]}^k} \quad (19)$$

である。特に $\beta' = 0$ であり, 初期の状態が $m_{j[i]} = m_j (\forall_i)$ である場合は

$$\Delta_0 m_j^k = \Delta m_j^k \quad (20)$$

であることが示される。

複製する位置の決定法

得られた結果を確認するため、以下から適当に 5 点選んだ (表 1).

$$\eta(x) = 3N(-1.5, 1) + 2N(1, 0.5) \quad (21)$$

また、 β' を 0, パラメータ Σ を徐々に小さくしていったときに $\Delta_0 m_{[i]} = 0$ となる点をプロットしたのが図 10 である.

表 1: 教師信号の値

x_i	-4.00	-3.50	-1.00	0.50	3.00
$\eta(x_i)$	0.13	1.82	2.68	1.96	0.04

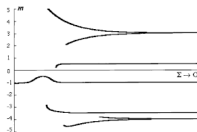


図 10: $\Delta_0 m = 0$ となるパラメータ m

今後やること

- 論文中の数式の意味をより深く理解する.
- アルゴリズムを教師無し学習や強化学習に適用できるように論文や技術資料を読み進める.