数理解析研究所講究録
1132巻 2000年 84-87

非線形大域結合写像モデルによるカオスを利用した最適化

奥原 浩之[†], 田中 稔次朗[†], 石井 博昭[††]

[†] 広島県立大学 経営学部 経営情報学科

[††] 大阪大学 大学院 工学研究科

1. はじめに

動径基底関数ネットワーク (Radial Basis Function Networks: 以後 RBFN)[1] では未知の非線形関数を近似するために，予め必要なニューロン数が不明である．このことが学習の遅延化や過学習の問題を引き起こしている．そこで，我々は先に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 [2] を導出することにより，学習の効率化をはかることができる競合動径基底関数ネットワーク (Competitive Radial Basis Function Networks: 以後 CRBFN)[3] を提案した．さらに，シナプス可塑性方程式に関する考察から，必要な動径基底関数を効率的に追加することもできる複製・競合動径基底関数ネットワーク (Reproductive and Competitive Radial Basis Function Networks: 以後 RC-RBFN)[4] を提案した．しかしながら，これらすべてのニューラルネットワークでは学習則として最急降下法を適用しているため，局所最適解に収束することもあり，大域的最適解が得られるかどうかは初期値に大きく依存する．

従来から最適化問題の大域的最適解に到達するために，確率的なノイズや決定論的なカオスを利用したニューラルネットが提案されている．カオスを発生する非線形大域的結合写像を工学的な問題解決のために応用した例としては，井上ら [5] のカオスニューロ・コンピュータや野沢 [6] による Hopfield モデルを大域結合写像として定式化したモデルが知られている．これらの研究では，カオスの特長である決定論的な推移と微小変位の指数関数的な増幅により，従来のシミレーテッド・アニーリングにとってかわるセルフ・アニーリングを利用することで優れた学習効果が得られることが示されている．

しかしながら，非線形大域的結合写像を組合せ最適化問題に適用した研究は数多いが，関数近似問題に適用した研究はほとんどない．その理由は関数近似問題と非線形大域的結合写像モデルの関係が明確にされていないからであると考えられる．そこで，本研究では関数近似問題の効率的な解法である動径基底関数ネットワークにおいて非線形大域的結合写像を導出する．また，動径基底関数ネットワークのシナプス結合荷重は組合せ最適化問題の状態変数と等価であるとみなせることも示す．さらに得られた結果を動径基底関数ネットワークに適用することにより，関数近似問題においても初期値に依存しない大域的最適解の探索を実現する．

2. 動径基底関数ネットワークによる関数近似問題の解法

RBFN は非線形関数 $\eta(\mathbf{x})$ を動径基底関数の足し合わせで近似するニューラルネットワークである．動径基底関数としては規格化されたガウス型活性関数などが用いられる．ここでは，M 個の入力ニューロンと 1 個の出力ニューロンからなる基本的な RBFN について説明するが，複数個の出力ニューロンからなる RBFN への拡張は容易である．

d 次元の第 i 入力ベクトル $\mathbf{x}_i \in R^d$, $(i = 1, 2, \cdots, N)$ は全ての入力ニューロンに入力される．第 j 入力ニューロン $(j = 1, 2, \cdots, M)$ はパラメータ ϕ_j をもつ．パラメータ ϕ_j は平均ベクトルと共分散行列の集合 $\{\mathbf{m}_j, \mathbf{\Sigma}_j\}$ であるものとする．ここで，$\mathbf{m}_j = [m_i^1, m_i^2, \cdots, m_i^d]^{\mathrm{T}}$ であり，$\mathbf{\Sigma}_j$ はその逆行列 $\mathbf{\Sigma}_j^{-1}$ の 第 kl 要素に σ_j^{kl} をもつ $d \times d$ の行列である．また，$\mathbf{\Sigma}_j$ は正定値対称行列である．第 j 入力ニューロンは入力ベクトル $\mathbf{x}_i$ に対して

$$\xi(\mathbf{x}_i, \phi_j) = \exp\{-\frac{1}{2}(\mathbf{x}_i - \mathbf{m}_j)^{\mathrm{T}}\mathbf{\Sigma}_j^{-1}(\mathbf{x}_i - \mathbf{m}_j)\} \tag{1}$$

を出力する．ここで，添字の T はベクトルの転置を示す．以後，このような出力を行う入力ニューロンのことを動径基底関数ということとする．出力値 $\xi(\mathbf{x}_i, \phi_j)$ はシナプス結合荷重 w_j, $(0 \leq w_j \leq 1)$ を通して出力ニューロンへ伝達され，出力ニューロンでこれらは足し合わされ

$$s(\mathbf{x}_i, \mathbf{w}, \phi) = \sum_{j=1}^{M} w_j \xi(\mathbf{x}_i, \phi_j) \tag{2}$$

が出力される．ここで，$\mathbf{w} = [w_1, w_2, \cdots, w_M]^{\mathrm{T}} \in R^M$ であり，ϕ で集合 $\{\phi_1, \phi_2, \cdots, \phi_M\}$ を表す．ニューラルネットワークによる関数近似は，非線形関数 $\eta(\mathbf{x})$ をネットワークの出力 $s(\mathbf{x}, \mathbf{w}, \phi)$ で表すことである．そのため，RBFN による関数近似は累積二乗誤差関数をエネルギー関数とみなし

$$E(\mathbf{w}, \phi) = \frac{1}{2}\sum_{i=1}^{N} E(\mathbf{x}_i, \mathbf{w}, \phi) \tag{3}$$

の値を減少させることにより実現される．ここで，

$$E(\mathbf{x}_i, \mathbf{w}, \phi) = \{\eta(\mathbf{x}_i) - s(\mathbf{x}_i, \mathbf{w}, \phi)\}^2 \tag{4}$$

は二乗誤差関数である．つまり，RBFN が学習により獲得しなければならないのは，第 j 動径基底関数のシナプス結合荷重 w_j，パラメータ $\mathbf{m}_j$ ならびにパラメータ $\mathbf{\Sigma}_j$ である．

一般の RBFN の学習アルゴリズムは式 (3) の累積二乗誤差関数に最急降下法を適用した

$$\frac{dw_j}{dt} = -\epsilon\frac{\partial E(\mathbf{w}, \phi)}{\partial w_j}, \quad \frac{dm_j^k}{dt} = -\epsilon\frac{\partial E(\mathbf{w}, \phi)}{\partial m_j^k}, \quad \frac{d\sigma_j^{kl}}{dt} = -\epsilon\frac{\partial E(\mathbf{w}, \phi)}{\partial \sigma_j^{kl}} \tag{5}$$

で与えられる．ただし，ϵ は適当な正の定数であり，m_j^k はパラメータ $\mathbf{m}_j$ の第 k 要素である．

3. 関数近似問題における非線形大域結合写像モデル

既に我々は CRBFN や RC-RBFN のシナプス結合荷重 w_j の学習則を導出するために，Dale 則を考慮したシナプス可塑性方程式である適者生存型学習則を提案しているが，そこでは，$\alpha_j(\phi)$ を内的自然増加率

$$\alpha_j(\phi) = \sum_{i=1}^{N}\eta(\mathbf{x}_i)\xi(\mathbf{x}_i, \phi_j) \tag{6}$$

として定義し，第 j ニューロンと第 k ニューロンとの競合の効果を表すために，$\gamma_{jk}(\phi)$ を競争係数

$$\gamma_{jk}(\phi) = \sum_{i=1}^{N}\xi(\mathbf{x}_i, \phi_j)\xi(\mathbf{x}_i, \phi_k) \tag{7}$$

として定義している．これら内的自然増加率と競争係数を用いると，関数近似問題における累積二乗誤差関数 $E(\mathbf{w}, \phi)$ は

$$\begin{aligned} E(\mathbf{w}, \phi) &= \frac{1}{2}\sum_{i=1}^{N}\left[\eta(\mathbf{x}_i)^2 - 2\eta(\mathbf{x}_i)\sum_{j=1}^{M} w_j\xi(\mathbf{x}_i, \phi_j) + \left\{\sum_{j=1}^{M} w_j\xi(\mathbf{x}_i, \phi_j)\right\}\left\{\sum_{k=1}^{M} w_k\xi(\mathbf{x}_i, \phi_k)\right\}\right] \\ &= \frac{1}{2}\sum_{j=1}^{M}\sum_{k=1}^{M}\gamma_{jk}(\phi)w_j w_k - \sum_{j=1}^{M}\alpha_j(\phi)w_j + \frac{1}{2}\sum_{i=1}^{N}\eta(\mathbf{x}_i)^2 \end{aligned} \tag{8}$$

となる．右辺第 3 項はパラメータ $\{\mathbf{w}, \phi\}$ に依存しない定数項となるため，あらためて関数近似問題におけるエネルギー関数は

$$E(\mathbf{w}, \phi) = \frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M}\gamma_{ij}(\phi)w_i w_j - \sum_{i=1}^{M}\alpha_i(\phi)w_i \tag{9}$$

で定義できることが示される．

ここで，巡回セールスマン問題に代表される組合せ最適化問題の目的関数は一般に

$$F(\mathbf{v}) = -\frac{1}{2}\sum_{i=1}^{M}\sum_{j=1}^{M}T_{ij}v_i v_j + \sum_{i=1}^{M}I_i v_i \tag{10}$$

で与えることができる．v_i は $0 \leq v_i \leq 1$, $(i = 1, 2, \cdots, M)$ を満たす．このとき，目的関数の係数を

$$\gamma_{ii}(\phi) = -T_{ii},\ \gamma_{ij}(\phi) = \gamma_{ji}(\phi) = -\frac{T_{ij} + T_{ji}}{2},\ \alpha_i(\phi) = -I_i \tag{11}$$

の関係を用いて変換するとエネルギー関数 $E(\mathbf{w}, \phi)$ と等価となることもわかる．

ところで，シナプス結合荷重の学習則は累積二乗誤差関数 $E(\mathbf{w}, \phi)$ の値を減少させるために

$$\frac{dE(\mathbf{w}, \phi)}{dt} = (\mathrm{grad}E(\mathbf{w}, \phi))^{\mathrm{T}} \frac{d\mathbf{w}}{dt} \leq 0 \tag{12}$$

となるように決定される．ここで，

$$\mathrm{grad}E(\mathbf{w}, \phi) = \left[\frac{\partial E(\mathbf{w})}{\partial w_1}, \frac{\partial E(\mathbf{w}, \phi)}{\partial w_2}, \cdots, \frac{\partial E(\mathbf{w}, \phi)}{\partial w_M}\right]^{\mathrm{T}} \tag{13}$$

である．そのため，$\mathbf{w} \in [0, 1]^M$ において半正定値行列となる $\mathbf{S}(\mathbf{w})$ を用いて

$$\frac{d\mathbf{w}}{dt} = -\mathbf{S}(\mathbf{w})\ \mathrm{grad}E(\mathbf{w}, \phi) \tag{14}$$

で与えることができる．ここで，$\mathbf{S}(\mathbf{w}) \in \Re^M \times \Re^M$ は対角要素が $a(1 - w_i^2)$, $(a \geq 0)$ で，その他の要素が 0 となる行列である．これらの結果，エネルギー関数 $E(\mathbf{w}, \phi)$ の極小点あるいは最小点へ状態が収束する RBFN の連続型シナプス結合荷重の動作規則は

$$\frac{du_i(t)}{dt} = \sum_{j=1}^{M} \gamma_{ij}(\phi) w_j(t) - \alpha_i(\phi) \tag{15}$$

$$w_i(t) = \frac{1}{2}\{1 + \tanh(a u_i(t))\} \tag{16}$$

で与えられる．

RBFN の連続型シナプス結合荷重更新則を離散型シナプス結合荷重更新則とするために一次オイラー差分をとると

$$u_i^{(n+1)} = u_i^{(n)} + \Delta t \left(\sum_{j=1}^{M} \gamma_{ij}(\phi) w_j^{(n)} - \alpha_i(\phi)\right) \tag{17}$$

となり，

$$u_i^{(n+1)} = u_i^{(0)} + \Delta t \left\{\sum_{j=1}^{M} \gamma_{ij}(\phi) \sum_{k=0}^{n} \left(w_j^{(k)} - \alpha_i(\phi)\right)\right\} \tag{18}$$

が導かれる．ここで，初期値 $u_i^{(0)}$ を $-\alpha_i(\phi)$ で与えるとすると，離散型シナプス結合荷重の動作規則は

$$u_i^{(n+1)} = \Delta t \sum_{j=1}^{N} \gamma_{ij}(\phi) \sum_{k=0}^{n} \left(w_j^{(k)} - \alpha_i(\phi)\right) - \alpha_i(\phi) \tag{19}$$

$$w_i^{(n)} = \frac{1}{2}\left\{1 + \tanh\left(\alpha u_i^{(n)}\right)\right\} \tag{20}$$

で与えられることとなる．そこで，次のような変数

$$p_i^{(n)} = \Delta t \sum_{k=0}^{n-1} \left(w_j^{(k)} - \alpha_i(\phi)\right) \tag{21}$$

を導入すると離散型シナプス結合荷重の動作規則は

$$p_i^{(n+1)} = p_i^{(n)} + \Delta t \left(w_i^{(n)} - \alpha_i(\phi)\right) \tag{22}$$

$$w_i^{(n)} = \frac{1}{2} + \left\{1 + \tanh\left(-a \left.\frac{\partial E(\mathbf{p}, \phi)}{\partial p_i}\right|_{p_i^{(n)}}\right)\right\} \tag{23}$$

で表される非線形大域結合写像となる．この非線形大域結合写像によって発生するカオスを利用することにより，評価関数が累積二乗誤差関数やエネルギー関数で与えられる最適化問題で，最急降下法に比較して初期値に依存せずに大域的最適解を求めることが可能となる．

非線形大域結合写像モデルを利用した関数近似問題の解法

Step 1 学習終了の判定値 ϵ を与え,α に大きな値，$p_i^{(n)}$ に適当な値を初期値として与える．

Step 2 式 (24) に従い $w_i^{(n)}$ を求める．

Step 3 式 (5) で m_j^k ならびに σ_j^{kl} を更新する．

Step 4 評価関数の値が ϵ より小さければ終了．

Step 5 評価関数の値が ϵ より大きな値に収束すれば α を増加する．

Step 6 評価関数の値が増加すれば α を減少する．

Step 7 Step 2 へ戻る．

4. まとめ

本研究では関数近似問題の効率的な解法である動径基底関数ネットワークにおいて非線形大域的結合写像を導出した．また，動径基底関数ネットワークのシナプス結合荷重は組合せ最適化問題の状態変数と等価であるとみなせることを示した．さらに得られた結果を動径基底関数ネットワークに適用することにより，関数近似問題においても初期値に依存しない大域的最適解の探索を実現できることを示した．

参考文献

[1] J. Park, and I. W. Sandberg, “Universal approximation using radial basis function networks,” *Neural Computation*, **3**, pp. 246-257, 1991.
[2] 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, **8**, No. 12, pp. 718-720, 1995.
[3] 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,” 信学論 (D-II), vol. J80-D-II, no. 12, pp. 3191-3199, 1997.
[4] 奥原浩之, 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,” 信学論 (D-II), vol. J82-D-II, no. 5, pp. 941-951, 1999.
[5] M. Inoue and S. Fukushima, “A Neural Network of Chaotic Ocillators,” *Prog. of Theor. Phys.*, **87** p. 771, 1992.
[6] H. Nozawa, “A Neural Network Model as a Globally Coupled Map and Application based on Chaos,” *Chaos*, **2** p. 377, 1992.