LG-002

SpikeProp の誤差曲面に対する一考察
Shape of Error Surface in SpikeProp

藤田 優† 高瀬 治彦† 北 英彦† 林 照峯†
Masaru Fujita Haruhiko Takase Hidehiko Kita Terumine Hayashi

1 はじめに

スパイキングニューラルネットワークは，ネットワーク内部の情報表現に，スパイクを用いるニューラルネットワークである．これは，少ないビット数で情報を表すことができ，耐ノイズ性にも優れているため，近年注目されている [1]．スパイキングニューラルネットワークの一種に，SpikeProp と呼ばれるモデルがある [2]．これは，スパイクのタイミングにより情報を表現し，ユニット間の結合を工夫することで学習を容易にしたモデルである．

本論文では，SpikeProp を LSI 等のハードウェアで実装することを想定し，その際の誤差曲面の解析を行った．誤差曲面とは，ネットワークが持つパラメータ（結合荷重）に対する誤差を曲面として表したものであり，学習ではこの曲面の極小点を探索する．そのため，その形状は学習の成否，学習速度に多大な影響を及ぼす．具体的には，離散化した時刻のもとネットワークを動作させることの影響について議論する．

2 SpikeProp

本論文の対象とする SpikeProp[2] について簡単に説明する．

SpikeProp ではスパイキングニューロン（以下ニューロン）を使用している．ニューロン j がシナプス前ニューロン i からのスパイクを d だけ時間遅れしたうえで受け取る．このとき，ニューロン j の内部電位 $x_j(t)$ は式 (1) で定義されるスパイク応答関数 $\varepsilon(t)$ に結合荷重 w_{ij} を乗じた分だけ変化する．この関係は式 (2) で表される．

$$\varepsilon(t) = \begin{cases} 0 & (t < 0) \\ \dfrac{t}{\tau} e^{1-t/\tau} & (t \geq 0) \end{cases} \tag{1}$$

$$x_j(t) = \sum_{i \in \Gamma_j} w_{ij} \varepsilon(t - t_i - d) \tag{2}$$

ここで Γ_j はニューロン j の前ニューロンの集合，t_i はシナプス前ニューロン i の発火時間である．そしてこの内部電位 $x_j(t)$ がしきい値 θ を超えるとニューロン j はスパイクを出力する．

SpikeProp では，図 1 に示す特殊なニューロン間の結合を持つ階層型ネットワークを使用する．この特殊な結合は，1 本の結合につき m 本の副結合により構成されている．この副結合それぞれに決まった時間遅れ d^k と結合荷重 w_{ij}^k が定められている．副結合の時間遅れ d^k には $1, 2, 3, \cdots, m$ のような一定間隔の定数を用い，すべての結合において同じ値を用いる．これによりニューロン j の内部電位 $x_j(t)$ は式 (3) となる．

$$x_j(t) = \sum_{i \in \Gamma_j} \sum_{k=1}^{m} w_{ij}^k \varepsilon(t - t_i - d^k) \tag{3}$$

SpikeProp における学習は，式 (4) で定義される誤差関数を小さくするように，勾配を用いた降下法により，結合荷重で調整することで出力スパイクのタイミングを調整する．

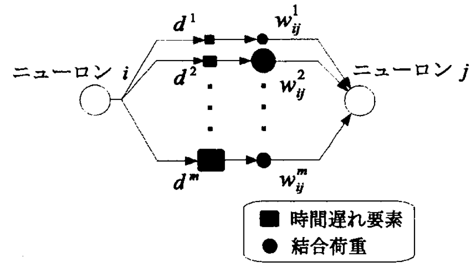

図 1 SpikeProp でのニューロン間の結合

$$E = \frac{1}{2} \sum_{i \in \Gamma_j} (t_j^a - t_j^d)^2 \tag{4}$$

ここで t_j^a は出力ニューロンの発火時刻，t_j^d は教師の発火時刻である．勾配を用いた降下法での学習は，結合荷重 w を初期値から式 (5) を用いて更新をくりかえすことを行う．

$$w \leftarrow w - \eta \frac{\partial E}{\partial w} \tag{5}$$

しかしこの学習法では学習が遅いという欠点がある．この欠点を解消するために，収束の速い非線形最適化手法がしばしば用いられる．この手法は，誤差曲面の形状の特徴をとらえることにより探索の高速化をはかっている．そのため，誤差曲面の形状が，その収束の速さに大きな影響を与える．

3 誤差曲面の調査と考察

この章では，スパイキングニューラルネットワークの誤差曲面の形状を調査し，この学習に収束の速い学習法を適用できるのかを検討する．スパイキングニューラルネットワークはスパイクのタイミングにより情報をコーディングする．そのため，各ユニットで内部電位がしきい値を超える時刻を式（3）にしたがい計算し，伝搬する必要がある．しかし，高集積化が容易なデジタル回路として実装する場合，時刻を離散化したうえで時々刻々と各ニューロンで内部電位としきい値の比較を繰り返すことでスパイクの出力タイミングを求め，スパイクを伝搬する方法が現実的である．

そこでこの章では，時刻を離散化した場合のスパイキングニューラルネットワークの誤差曲面の形状を調査する．収束の速い最適化手法は，あるベクトル (特に勾配) の方向に直線的に探索または 2 次近似を行い，誤差が最小となる点を見つけ出すことを繰り返す．そのためここでは，誤差曲面そのものではなく，勾配方向の誤差の変化を調査する．

実験は，Bhote らの時間版 XOR 問題 [2] を，式（5）を用い学習することで行った．ネットワークは，バイアスを含む 3 個の入力，5 個の中間層ニューロンと 1 個の出力層ニューロンからなる．各ニューロン間の副結合は 16 本とし，それぞれ 1 から 16 の時間遅れを持たせた．また，時刻を離散化した際の単位時間 Δt は 0.01 とした．これを学習した際の誤差の変化を図

† 三重大学大学院工学研究科

2に示す．次に，学習開始時，学習途中 (80回学習時点) と学習終了時 (426回学習時点) においてそれぞれ，誤差の勾配方向の変化を調査した．以下の図では，いずれも，横軸は勾配ベクトルの倍率 (式 (5) の右辺の η に相当) を，縦軸はその地点での誤差を表している．

図3は，発火時刻を離散化せずに求めた場合の，学習開始時点近傍の誤差の変化である．図4は，時刻を離散化した場合の同じ地点の誤差の変化を表す．また，図5，図6にはそれぞれ学習途中，学習終了地点の近傍において，時刻を離散化した場合の誤差の変化を示した．

図3と図4を比較すると，時刻の離散化により誤差曲面に細かい凹凸が発生していることが分かる．また図4から図6を比べることで，学習の進行に伴い凹凸が小さくなるが，学習終了時でも小さい凹凸が発生していることが分かる．

図4から図6のように誤差曲面に凹凸が多いと，非線形最適化 (準ニュートン法，共役勾配法など) の手法 [3] を利用して，学習を高速化しようとしても，凹凸により曲面の形状を正しく近似できないため，効果があがらない．それどころか，学習 (探索) により誤差曲面の最小値を正しくとらえることができず，探索が極小値で終了してしまう．これを解消するには，探索点を複数にすることで，各探索点での凹凸の影響を受けにくくする手法が考えられる．

4 まとめ

本論文では SpikeProp における誤差局面について調査し，時刻を離散化した場合ではある程度学習が進むまでは誤差曲面の表面に多数の凹凸が存在し，学習が進むにつれてこの状態が多少解消されていくことがわかった．以上の結果を踏まえ，時刻を離散化した場合で一般的な非線形最適化手法をどのように適応させるか考えていきたい．

参考文献

[1] Wolfgang Mass and Christopher M. Bishop: Pulsed Neural Networks, The MIT Press (2001)

[2] Sander M. Bhote, Joost N. Kok, and Han La Pourté: Error-backpropagation in temporally encoded networks of spiking neurons , Neurocomputing, Vol.48, pp.17–37 (2002)

[3] 金谷健一：これなら分かる 最適化数学－基礎原理から計算手法まで－, 共立出版 (2005)

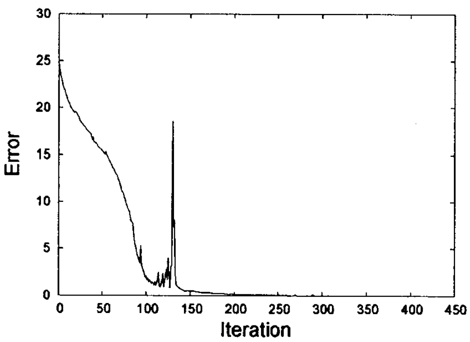

図2 学習回数と誤差の関係

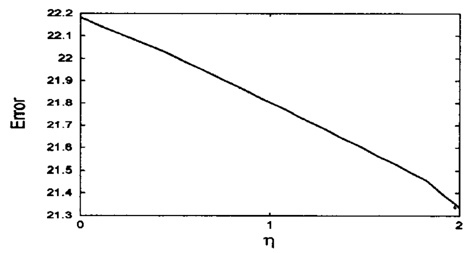

図3 時刻を離散化していない場合の学習前における誤差曲面

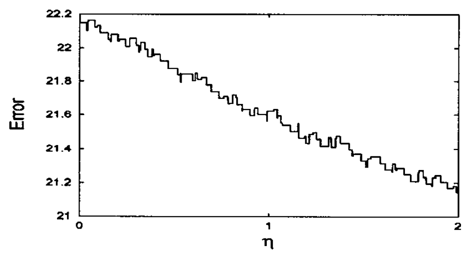

図4 時刻を離散化した場合の学習前における誤差曲面

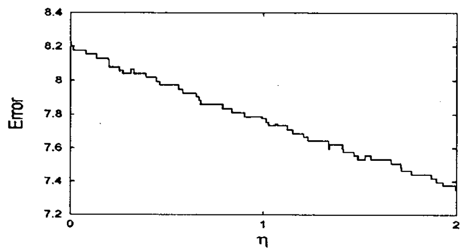

図5 時刻を離散化した場合の学習途中 (学習回数80回) における誤差曲面

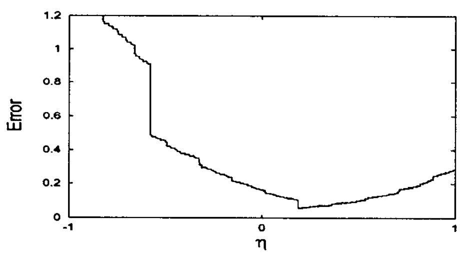

図6 時刻を離散化した場合の学習終了時付近 (学習回数200回) における誤差曲面