

遺伝アルゴリズムによる耐故障ニューラルネットワークの実現

1D-2

丹 康雄

南谷 崇

東京工業大学 工学部

1 はじめに

最近、ニューラルネットワーク(以下NN)を用いた情報処理が盛んに研究、応用されている^①。特に、Rumelhart型に代表される階層型NNは学習、汎化の能力を持つため、広く応用されるに至っている。これらNNの特長のひとつに耐故障性があることは古くから言われてきたが、従来の研究は実験によりその能力の存在を示すにとどまるものがほとんどであり、耐故障能力の数量化や耐故障性獲得方法などについての研究は少ない。

本論文では、階層型NNにおける素子間結合の切断に対する耐故障性を学習によって得る手法として、遺伝アルゴリズムを応用した学習方法を提案する。

2 階層型ネットワークの耐故障性

階層型NNのうち最も広く用いられているのがRumelhartらの提案した誤差逆伝搬(Error Back Propagation 以下BP)モデルである。これはS字状関数を出力関数とする素子を層状に配置してネットワークを構成し、各素子間の荷重を出力における誤差評価関数の値が減少するように最急降下法で変化させ、目的とする写像関係を獲得(学習)するものである。つまり、ネットワークという形で関数を与え、そのパラメータ推定を最急降下法で行なうものといえる。

素子の出力関数が飽和特性を持っていることから、飽和領域で動作している素子では入力の変動が出力にはほとんど影響を与えなくなる。この特性をうまく活かすようなネットワークを形成することにより、学習ボタンからずれたものに対しても適切な出力を与える汎化性や、ネットワークの構成要素が故障しても全体としては正しい動作を続ける耐故障性といった能力が得られるようになる。

本論文では、結合の断線に対する耐故障性を向上させることを考える。ここで、次のような結合を考える。

[定義] ある結合を切断したとき、ネットワーク出力の学習ボタンに対する最大誤差の絶対値が判定基準Cよりも大きくなれば、その結合を必須結合と呼ぶ。□

必須結合はすなわち切断不可能な結合であるから、存在しないことが望ましい。ある特定の素子に役割が集中してしまうとこれに付随する結合は必須結合となるであろうことは明らかである。従って、必須結合を作らないためには複数の素子に役割を分担させる必要がある。

通常のBP学習では余分な素子を用意してもうまく分担

Fault-tolerant multi-layer neural networks with genetic algorithm

Yasuo Tan and Takashi Nanya

Faculty of Engineering, Tokyo Institute of Technology

される保証はない。BP学習はあくまでも学習ボタンに対する誤差を減らすもので、内部の表現形態がどうなるかは一般にわからないためである。

何らかの方法を用いて強制的に重みを分散させるなどの方法も考えられるが、問題によっては学習が困難になる場合もあり、また、表現を分散化することで耐故障性が期待したように向上する保証はない。よって、より直接的に耐故障性を向上させるような手法が求められる。

3 遺伝アルゴリズムによるNNの学習

最近、NNの学習に遺伝アルゴリズム^②(以下GA)を応用する研究が報告されている^③。GAは複数の個体を含んだ集団を用意し、その中で優れた性質を持ったものが多く残るような操作を繰り返して、集団全体の平均的な評価値を山登り的に向上させるというものである。

単純なGAは、解こうとする問題を遺伝子の形で表現するコーディングと、評価、増殖・淘汰、交叉、突然変異の繰り返し過程からなる。我々は、このアルゴリズムを階層型NNの学習に以下のように適用した。

0) コーディング 個体は、それぞれ異なった性質を持つが、これらは染色体の中にあらかじめ遺伝子の形で記述されており、その情報を基に個体になったときに性質が現われるものとする。この染色体中での表現を遺伝子型(genotype)、個体になったときの性質を表現型(phenotype)と呼ぶ。性質を遺伝子の形に記述することを遺伝子へのコーディングという。我々の場合、階層型NNの重みを遺伝子と考え、これを一次元に並べたものを染色体とした。

1) 評価 各個体について、その表現型が環境に対してどれくらい望ましい性質を持っているかを示す適合度(fitness)という値を求める。我々は出力の平均二乗誤差を学習ボタンに対する適合度、ネットワーク中に存在する必須結合の割合の低さを耐故障性に対する適合度とし、これらの重み付き和を個体の最終的な適合度とした。

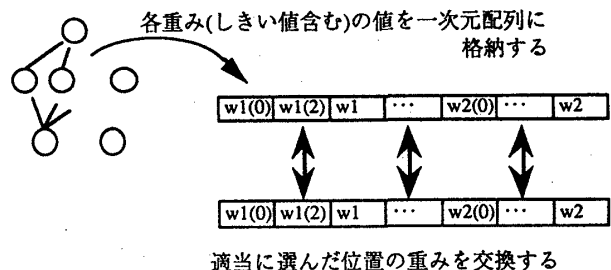


図1 ネットワークの遺伝子表現と交叉方法

2) 増殖・淘汰 適合度に応じて個体の複製を作る。適合度が高いものはより多くの複製をつくり、一方、低いものは自分の複製をつくる機会を失い、淘汰される。

3) 交叉 個体間で遺伝子レベルでの情報のやり取りを行ない、新たな個体を生みだす。最も簡単には、二つの個体の染色体をある点で文字通り交叉させ、部分的には親と同じだが、全体としては新しい個体を二つづつ作る。我々の場合は、ある確率に従って二つの個体間で遺伝子(荷重)の値を交換することにした。(図1)

4) 突然変異 ある(低い)確率で遺伝子の情報を書き換える。これによって、親の代では全く存在しなかった遺伝子が出現することになる。我々の例では、ある確率で選出した荷重(遺伝子)について、その値を元の値の±50%の範囲でランダムに増減した。

以上の1)から4)までのステップを繰り返すことにより、集団における平均適合度が向上する。この過程の中で、2)の増殖・淘汰が平均適合度の直接的な山登りに対応する。この操作のみを繰り返すと集団内でもともと存在した最も優秀な個体に全個体が揃った段階で学習が止る。山登りにおける局所解である。これに対して、GAでは3)交叉と4)突然変異によって局所解からの脱出が起こりうる。

評価において、染色体の持つ情報から作った個体が個体として成り立たないことがある。このような個体は直ちに死滅することから、その遺伝子を致死遺伝子という。交叉や突然変異を行なって容易に致死遺伝子が出来てしまうとGAは成り立たない。現在のところ一般的なコーディング手法というものは存在せず、いかにうまく問題を遺伝子の形にコーディングするかがこのアルゴリズムにおける最も重要なポイントとなっている。以上に述べた我々の方法の場合は、どのような荷重の値であってもNNとしては成り立つから、致死遺伝子は存在しないことになる。

この手法が最急降下法や焼き鈍し法などとは異なる点として、一つには集団による多点探索であることが挙げられるが、我々の目的にとって最も重要な点は、過程において評価関数は単に評価に使われるだけであるということである。すなわち、各個体の適合度を求めるために評価関数は存在するが、これは単純に評価をするためだけのものであり、これを基に学習規則やダイナミクスを導くことはない。したがって、評価関数には微分可能等、関数上の制約がなく、かなり複雑なものも用いることができる。よって、上記のように、耐故障性に対する評価が学習ボタンへの誤差に対するものと同様に与えることができ、学習ボタンを実現すべくネットワークが学習するのと同様に耐故障性が高まっていくと期待できるわけである。

4 実験

中間層素子を8個持ったネットワークにXORを学習させ、通常のBP学習と比較した。XORは中間層素子2個で学習が可能であるから、6個の余分な素子があることになる。また、

結合の数は、しきい値も含めて合計33本となっている。

BP学習の方は最も一般的な慣性項付きの学習方法で行ない、学習率は0.9、学習の終了判定は全てのボタンについて誤差の絶対値が0.1よりも小さいこととした。GAでは個体数500、突然変異確率10%とし、終了条件は集団中で最も優秀な個体がBP同様の条件と同時に必須結合を持たなくなることとした。

異なった初期値を使って8回ずつ実験を行なったところ、GA、BP 何れも全て与えたボタンの学習に成功した。しかしGAで必須結合数が0となって停止したのは1例のみで、それ以外は時間切れによる打ち切りである。この場合、何れも集団中の全ての個体が非常に類似した形になっていた。

表1に結果を示す。これは、必須結合の判定基準Cを0.1(学習終了条件と同じ評価)および0.5(出力を二値化して評価)にしたときの、各ネットワークでの必須結合数である。

表1 学習によって得られたネットワーク中の必須結合数

	実験 No.	1	2	3	4	5	6	7	8	平均
C=0.1	BP	22	22	23	24	17	22	22	24	22.0
	GA	5	2	4	4	4	3	3	0	3.1
C=0.5	BP	5	4	3	2	5	4	6	5	4.3
	GA	2	2	1	4	0	1	2	0	1.5

通常のBPでは、断線してしまうと学習終了判定基準を満たさなくなってしまう結合が全体の2/3程度を占めるのに対し、GAを用いたものでは、1/10程度となっている。また、必須結合の存在しないネットワークもGAでは得ることができた。

5 むすび

耐故障性の高い階層型NNを得る手法として遺伝アルゴリズムを用いた学習法を提案し、シミュレーションによる実験でその有効性を示した。耐故障性向上に望ましいであろう制約下で通常の学習を行なう手法と異なり、直接的に耐故障性を向上させるものとなっている。

本手法は汎化性など、耐故障性同様、従来直接的には向上させる方法がなかった能力についても応用できるものと思われる。

文献

- (1) 丹, 南谷, "階層型ニューラルネットの耐故障性に関する一考察", 第42回情処全大, 6, pp.104-105 (1991)
- (2) 河野, 丹, 南谷, "断線に対する耐故障性をもつニューラルネットワークの一構成法", 第44回情処全大 (1992)
- (3) 名古, 丹, 南谷, "内部表現の分散化によるニューラルネットワークの耐故障化の一手法", 第44回情処全大 (1992)
- (4) 市川 芳明, "ニューラルネットワークによる教師なし学習制御", 計測自動学会, SICE'90, JS 37-5 (1990)
- (5) Goldberg D. E., "Genetic Algorithms in Search Optimization and Machine Learning", Addison-Wesley (1989)
- (6) Hertz J., Krogh A. and Palmer R.G., "Introduction to the Theory of Neural Computation", Addison-Wesley (1991)