

5N-1 変異遺伝子を導入した遺伝的アルゴリズム MGGA による構造学習

○高濱 徹行

広島市立大学情報科学部[†]

阪井 節子

広島修道大学商学部[‡]

1 はじめに

一般に、遺伝子の中には正常な遺伝子とは性質の異なる変異遺伝子 (mutant gene) が存在している。変異遺伝子は、正常な遺伝子の塩基配列において置換、挿入、欠損などの変化が生じた場合に発生する。このような変異遺伝子が体細胞内の DNA に発生した場合には、異常増殖する体細胞を生じるなどの遺伝子病を生じることがある。また、変異遺伝子が生殖細胞の DNA に生じた場合には、子孫にまで伝わる遺伝病を生じることがある [1]。

遺伝的アルゴリズム (GA) は、生物の遺伝機構を模倣した最適化手法であるが、変異遺伝子についてはあまり考慮されてこなかった。GA では、コード化の評価規範として、遺伝子の並びである染色体がすべて問題空間の解候補に対応付けることができるという健全性 [2] を重視してきたため、致死遺伝子や変異遺伝子の積極的な導入は避けられてきた。

本研究では、変異遺伝子を積極的に利用した遺伝的アルゴリズムである「MGGA(GA with Mutant Gene)」を提案する。変異遺伝子が発現すると、その遺伝子に対応する形質の発現が不完全になったり、発現しなくなるなどの現象が起きる。これは、遺伝子型を表現型に変換する関数が変異遺伝子の影響を受けることを意味する。

本研究は、ニューラルネットワークにおける構造学習 [3] のような、システムを記述するパラメータの単純化に応用できると考えられる。

2 MGGA

GA では、各個体のもつ遺伝情報を伝える実体として染色体を仮定し、染色体を遺伝子の並びで表現する。遺伝子の並びを $G = g_1 g_2 \cdots g_L$ 、表現型への変換関数を h 、適応度関数を f と仮定すれば、個体の適応度は $f(h(G))$ で与えられる。

MGGA は、変異遺伝子を導入し、変異遺伝子により形質の発現が影響を受けるという状況をモデル化したアルゴリズムである。MGGA における個体は以下の情報

を持つ。

遺伝子の並び (G) $G = g_1 g_2 \cdots g_L$ とする。

状態の並び (S) 状態の並びを $S = s_1 s_2 \cdots s_L$ とすると、各状態 s_i はある遺伝子 g_i が変異遺伝子であるかどうかを示している。状態の値は、対応する遺伝子の変異遺伝子ならば 1、そうでなければ 0 となる。

ここでは、一度変異した遺伝子は正常な遺伝子に戻ることはないと仮定する。また、変異遺伝子の近くの遺伝子は変異しやすいと仮定する。

適応度 表現型への変換関数 h^m は状態に依存することになる。したがって、個体の適応度は $f(h^m(G, S))$ で与えられる。

なお、変異遺伝子は生殖細胞に発生したと仮定し、次世代に遺伝させる。

MGGA のアルゴリズムを以下に示す。

```
MGGA()  
{  
  t=0;  
  変異遺伝子のない個体集団  $P(t)$  を生成;  
  while(!収束条件を満足) {  
     $P' = P(t)$  から個体を選択;  
    for(each 個体対  $p, q$  in  $P'$ ) {  
      確率  $P_c$  で  $p$  と  $q$  を交叉;  
      確率  $P_m$  で  $p, q$  の  $G$  を突然変異;  
      for(all  $s_i$  in  $p, q$  の  $S$ )  
        if( $s_i == 0$ )  
          確率  $P_d \cdot (R_d \cdot \text{len}(g_i) + 1)$  で  $s_i = 1$ ;  
    }  
     $t = t + 1$ ;  
     $P(t) = P'$ ;  
  }  
}
```

ただし、 P_c は交叉率、 P_m は (可逆的) 突然変異率である。 P_d は不可逆的突然変異率 (以下損傷変異率と呼ぶ) であり、正常な遺伝子のみが損傷し、一度損傷すると戻ることではない。 $\text{len}(g_i)$ は g_i に隣接する変異遺伝子数を

*Structural Learning by Genetic Algorithm with Mutant Genes (MGGA)

[†]Faculty of Information Sciences, Hiroshima City Univ.

[‡]Faculty of Commercial Sciences, Hiroshima Shudo Univ.

返す関数である。 R_d は損傷変異上昇率であり、 g_i に隣接する変異遺伝子数に比例して損傷変異率を上昇させる際の比例係数となる。

3 ニューラルネットワークの構造学習

3.1 表現型への変換

回帰モデル、ニューラルネットワーク、RBF ネットワーク、ファジィ推論ルールなどは複数の実数値パラメータで記述されるモデル記述系である。本研究では、このような記述系に対する遺伝子型から表現型への変換関数 h^m として、変異遺伝子数の増加に比例して形質の発現が減少する線形の変換関数を提案する。

実数値パラメータを $p = (p_1, p_2, \dots, p_i, \dots, p_n)$ とし、遺伝子の並び G において p_i を表現する部分を G_i 、それに対応する状態の並びを S_i とすると、区間 $[L, H]$ の値を取る変数 p_i は以下のように定義できる。

$$p_i = A_i \{L + (H - L) \text{bin}(G_i) / (2^{l_i} - 1)\} \quad (1)$$

$$A_i = 1 - \text{one}(S_i) / l_i \quad (2)$$

ただし、 l_i は並びの長さであり、並びはビット列で表現する。 bin はビット列を $0 \sim 2^{l_i} - 1$ の整数に変換する関数、 one は '1' であるビットの数を返す関数である。また、通常 $L < 0, H > 0$ であり、 A_i は G_i 中で損傷していない有効な遺伝子の割合を示している。

3.2 実験

以下の関数を学習するために、2入力1出力3層の階層型ニューラルネットワークを用いた。学習データは、 $\{(x_1, x_2) | x_1, x_2 = 0.0, 0.2, \dots, 1.0\}$ の36点において、関数値に平均0、分散0.01の正規乱数を加えた値を用いた。

$$f(x_1, x_2) = x_1 x_2^2 \quad (3)$$

中間層を6ユニット、すなわち、最大パラメータ数を25 ($=3 \times 6 + 7$) とした。結合荷重および閾値は10ビットで表現し、 $[L, H] = [-10, 10]$ 、適応度は教師データとの平均2乗誤差 σ^2 とし、適応度を最小化した。MGGAの設定は、個体数50、 $SP = 2$ (Selective Pressure for Linear Ranking)、1点交叉、 $P_c = 0.8$ 、 $P_m = 0.005$ 、 $P_d = 0.002$ 、 $R_d = 1.0$ とした。実験は、世代数300まで、10回繰り返して行った。

各実験について、削除されたパラメータ、有効パラメータ数 (P_{eff})、平均2乗誤差 σ^2 およびAICを以下の表に示す。ただし、 nN は中間層ユニット (2つの結合荷重と1つの閾値)、 nw は結合荷重、 $n\theta$ は閾値がそれぞれ

n 個削除されたことを示している。なお、3層の階層型ニューラルネットワークでは、中間層と出力層のユニット間の結合荷重が削除された場合には、中間層ユニットの出力は無効になり、実質的にはそのユニットを削除できる。括弧内は、この知識を用いた場合の実質的な結果である。

実験 No	削除パラメータ	P_{eff}	2乗誤差	AIC
1	6w, 2 θ	17	0.00228	-80.8002
2	3w, 3 θ	19	0.00319	-64.7449
3	2N, 6w	13	0.00218	-90.4270
	(3N, 4w)	(12)		(-92.4270)
4	1N, 6w, 3 θ	13	0.00279	-81.5625
	(2N, 5w, 2 θ)	(12)		(-83.5625)
5	5w, 2 θ	18	0.00192	-85.0184
	(2N, 5w, 2 θ)	(12)		(-97.0184)
6	3w, 3 θ	19	0.00261	-72.0209
	(1N, 2w, 2 θ)	(18)		(-74.0209)
7	1N	22	0.00275	-64.1330
8	4w, 2 θ	19	0.00171	-87.1987
	(2N, 2w, 2 θ)	(15)		(-95.1987)
9	7w, 4 θ	14	0.00398	-66.7809
	(1N, 6w, 4 θ)	(12)		(-70.7809)
10	2w	23	0.00225	-69.3540
	(2N, 2w)	(17)		(-81.3540)
平均		17.7 (15.6)	0.00257	-76.2040 (-80.4040)

実質的な有効パラメータ数は平均15.6であり、平均9.4パラメータが削除された。また、実質的なAICの高い実験No.3,5,8から、中間層ユニットは4個程度とし、さらに結合荷重および閾値を数個削減可能であることが分かった。したがって、MGGAはニューラルネットワークの構造学習に有効である。

4 おわりに

本研究では、変異遺伝子を積極的に導入したMGGAを提案した。ニューラルネットワークの構造学習を行うことにより、MGGAがパラメータ値を最適化しながらパラメータ構造も最適化できる方法であることを示した。

MGGAは、誤差を最小化するだけで自然に構造が学習できる優れた方法である。しかし、長い世代に渡って実験を行うと、パラメータ数を削減しすぎて誤差を十分に減少できない場合があるため、最適化を終了する世代を適切に選ぶ必要がある。また、通常のGAにおける P_c, P_m に加えて P_d, R_d を導入したため、これらの値を適切に設定する必要がある。

今後は、これらの諸問題についてさらに検討するとともに、様々な分野への応用を試みる予定である。

参考文献

- [1] 島田和典編: “遺伝病,” 化学同人, 1993.
- [2] 北野宏明編: “遺伝的アルゴリズム,” 産業図書, 1993.
- [3] M.Ishikawa: “Structural learning with forgetting,” Neural Networks, vol.9, no.3, pp.509-521, 1996.