

## 遺伝的アルゴリズムを利用した項目選択

藤 森 進\*

### Item Selection Using A Genetic Algorithm

Susumu Fujimori

The items in a test or questionnaire must often be selected to fulfill, as much as possible, some contradictory requirements in the psychological measurement. For example, sometimes the number of items in the scale must be as small as possible in order to maintain the reliability of scale when measuring a characteristic. Under such a situation, item selection generally depends on trial and error in which unavailable items are deleted one by one from the items in the scale while considering the level so that the reliability does not decrease too much. This process requires time and effort; moreover, in this process the intentions of the researcher selecting the items may greatly affect the result. In this study, the items were automatically selected using a genetic algorithm. As a result, the validity of the model in which the genetic elite and the exactly opposite non-elite were saved at the same time was confirmed.

Key Words : genetic algorithm, item selection, test, questionnaire, psychological measurement

#### 1. 序

##### 1.1 テスト・質問紙における項目選択

心理的テストや質問紙（以下テスト）を作成したり利用する場合には、何らかの指標-例えば信頼性の推定値 係数など-を頼りに尺度を構成する必要がある。信頼性とテスト項目数との関係はスピアマン・ブラウンの公式（池田，1973）に知られるように、項目数

が多ければ多いほど信頼性も大きくなるといえる。しかし、類似項目の問題や、印刷の分量、テスト時間の制約などの要因により、項目数の増加には抵抗もある。このため通常は試行錯誤的に尺度を構成する項目を入れ替えて係数を再計算し、なるべく少ない項目数で出来るだけ大きな係数を達成するようにテスト作成者は苦勞している。このような項目分析、特に項目の取舍選択の過程では、指標のみならずテスト実施上の各種の要請、例

---

\*ふじもり すすむ 文教大学人間科学部人間科学科

えばテストの項目数や、項目内容の分布、テスト時間、あるいは項目がどの程度被験者の眼に既にさらされているかという率など様々な要因が考慮されているだろう（項目分析については、Guilford（1954）、芝（1972）などに詳しい）。これらの多数の条件、しかも尺度の項目数と信頼性に見られるように互いに相容れないことも多い諸条件を考慮して項目を選択するのは大変な作業である。また試行錯誤的方法では手間がかかるだけでなく、場合によっては項目選択を行う人の意図、あるいは意図せざる好みのような恣意性も問題になるかもしれない。項目選択に伴う作業量を軽減し、その過程を客観化する方法が求められている。

以上のような状況では、項目選択における諸条件を評価関数に適宜表現しておけば、大域的な最適値の近似解の探索方法の一つとして知られている遺伝的アルゴリズム（後述）を利用して、最適解に近い解を比較的容易に得ることができる。もちろん厳密な最適解が容易に得られるのであればそれに越したことはない。仮にテスト項目数を40としたとき、これらの部分集合からなる下位テストの可能性としては $2^{40}$ 通り 約1.1兆通りの組み合わせがあるのだから、評価関数を全ての組み合わせにおいて計算すれば、望みの基準による厳密解を求めることができる。しかし、この計算を一般のテスト作成場面で実行するのは容易ではないし、もう少し項目数が多くなれば実際上計算不可能となってしまうだろう。心理学においても厳密な解を求めることが困難な場合や、近似解でも実用上十分な場合には、（近似的な）解を比較的容易に得ることが出来る遺伝的アルゴリズム（次節参照）を利用するのが適当であろう。もちろん遺伝的アルゴリズムによる解（多くの場合は近似解）で最終結果とすることもできるし、人手によって近似解の修正を行うことも、あるいは、近似解を元にして少なくともその近傍で最適解を求めることもできるだろう。また、この方法によれば、評価関数の記述を通じて項目選

択における客観性が増すことも期待できる。遺伝的アルゴリズムの適用例は、北野(1997)に見られるように工学的分野が中心である。心理学に関連したものとしては、しかしその殆どは心理学者というよりは工学などの分野の研究者によるものであるが、人工知能に係わる研究（たとえば芝・小谷・赤沢(1998)など）や知覚、認知に係わる研究（たとえば横尾・萩原(1997)など）が海外および国内を含めて幾つか見られる程度であり、テスト項目の選択に遺伝的アルゴリズムを適用した例は見あたらない。この点からも項目選択に有効と考えられる遺伝的アルゴリズムの適用を試みる価値はあると考えられる。

この研究は以上のような意図に基づき、遺伝的アルゴリズムによる項目選択の可能性についてシミュレーションにより検討を加えるものである。

## 1.2 遺伝的アルゴリズム

遺伝的アルゴリズム（Genetic Algorithm; GA）とは、Holland, J.H.とミシガン大学の同僚により開発されたもので、自然システムの適応プロセスの要約及び厳密な説明と、自然システムの重要なメカニズムを保持するソフトウェアをデザインしようとする目的を持つ（Goldberg, 1989）。より具体的に述べれば、GAは、自然システムの中の特に生物の進化-その遺伝子の変化の過程を模倣することにより、関数の最大値あるいは最小値を与える（近似）解を求める探索アルゴリズムである。

GAでは、まず解決すべき課題を染色体の各遺伝子に何らかの形で反映させる。この方法は課題の内容に応じて反映されるもので定式化された方法というものは無い。各染色体ごとに適応度として関数の値を計算する事になる。

どの染色体を選抜して交配させるかは選択交配というが、このための方法としては適応度比例戦略が代表的なものとしてある。この戦略では、各染色体の適応度に応じた確率で、

他の染色体と交叉し生き残るように設定され、次の世代が構成される。適応度を評価する関数により交配時の染色体の生き残り確率が定められるが、その際に評価関数をそのまま利用するのではなくて関数値の違いを拡大縮小するスケージングとよばれる操作を行うことがある。

また別の選択交配の方法としてはエリート保存戦略がある。これは染色体集団の中で最も優秀なものをそのまま次世代に移行するものである。すなわちエリートは選択されても他の染色体と交叉されないので、選択交配の方法と考えるよりは世代モデルの枠組みの中で考えるほうが良いかもしれない。選択交配の方法としては、スキーマ貪欲法（相澤，1995）、あるいは期待値戦略、トーナメント戦略など色々なもの、およびそれらの組み合わせがあるが詳しくはGA関係の文献（例えば北野，1993，1995，1997；Mitchell，1996など）を参照されたい。

染色体が2つ選択された後に、これらを交叉させる。交叉は生物の場合と同様に2つの染色体の遺伝子を交換して次世代の染色体を生み出すものである。交叉の方法にもいろいろあるが最も単純なものは2つの染色体の同じ場所で染色体を切断し互いに入れ替える単純交叉と呼ばれる方法である。この他にも複数点交叉や一様交叉などの方法が知られている。

また解の多様性を保つために生物と同様に突然変異により染色体の遺伝子を変化させることも行われる。GAでは、この他にも、その適用分野の特色を生かした様々な工夫がなされており、進化の過程を繰り返して適応度の良い染色体 - すなわち解を探すことになる。

## 2. 方法

### 2.1 GAによる最適な項目数の探索

本研究では、以下のようにしてテスト項目の選択でGAを適用できるように模した。まず染色体に含まれる遺伝子の長さは、選択対象となる全項目の数  $n$  とする。染色体上の遺伝子において1は選択された項目をあらわし、

0は非選択項目とする。

項目選択における適応度評価関数（以下評価関数）については、本研究では信頼性（係数）のみの評価関数、信頼性+項目数（1項目あたり0.001のペナルティ）を加味した評価関数の2条件で行った。ペナルティの大きさは、信頼性が0~1の範囲にあることと、テスト項目数が10~50項目程度となることが多そうなことを配慮して決定した。

選択交配の方法は適応度比例戦略である。本研究ではスケージングの大きさは、評価関数の2乗とした。初期遺伝子の生成は、確率0.5で0あるいは1を割り当て、 $m$  個の染色体をランダム生成し初期値とした。交叉の方法は、染色体を1箇所交叉させる単純交叉single-crossoverである。各世代ごとに保存する染色体数は15とした。遺伝子の突然変異率は0.001である。世代モデルとしては離散世代モデル1種、また連続世代モデルとしてエリート保存モデル（2種）、エリート+非エリート保存モデル（4種）の計7種を試みた（表1）。表中の世代ギャップとは、世代交替ごとの入れ替わり率であり、例えばエリート1位のみ保存する場合は、集団の大きさが15なので  $(15-1)/15=0.93$  が入れ替わることになる。

表1 モデル

モデル番号	モデルの特徴	世代ギャップ
	離散世代モデル	1.00
	エリート保存モデル（1位のみ）	0.93
	エリート保存モデル（1及び2位）	0.86
	エリート+非エリート保存モデル（1位のみ）	0.86
	エリート+非エリート保存モデル（1及び2位）	0.73
	エリート+非エリート保存モデル（1位のみ、非エリートはエリートの対極）	0.93
	エリート+非エリート保存モデル（1及び2位、非エリートはエリートの対極）	0.86

$k-1$ 世代（現世代）での可能な親の対を全て作成し、交叉確率に従って交叉するかどうかを

決定する。同一世代に同一遺伝子個体の重複は許していないので重複染色体を削除し、モデル (離散世代モデル)なら適応度関数の上位  $m$  個を  $k$  世代 (次世代) とし、残余は淘汰するという過程を繰り返す。モデル  $\sim$  は現世代の染色体の一部を次世代にそのまま据え置く連続世代モデルである。モデル  $\sim$  は現世代の適応度の最も高かった染色体 (エリート) を残すものであり、モデル  $\sim$  は次点も保存するものである。これに対してモデル  $\sim$  と  $\sim$  は、エリートだけでなく最も適応度の低い染色体の1位、あるいは2位までを保存するものである。エリートの保存は理解できるとして、最も適応度の低い染色体を保存することは一見奇妙に感じられるかもしれない。この理由は、エリート保存戦略では、時によりその遺伝子が集団中に急速に広がり、解が局所解になってしまう危険性が指摘されることがあり、これを避ける意図をもって、適応度の低い染色体、すなわち現世代の中でエリートの遺伝子と最もかけ離れているであろう遺伝子を敢えて保存していることによる。

モデル  $\sim$ 、 $\sim$  もモデル  $\sim$ 、 $\sim$  と同様にエリートだけでなく非エリートも保存することにより遺伝子の源の多様性を確保しようとする意図においては同じであるが、モデル  $\sim$ 、 $\sim$  では現世代の中での非エリートであったのに対して、モデル  $\sim$ 、 $\sim$  では、エリートの対極となる染色体、すなわち染色体で0の遺伝子であったものは1の遺伝子に、あるいはその逆に1から0への変換を行ってエリートの対極になる非エリート染色体を作成し、これを次世代に置くものである。このようにモデル  $\sim$ 、 $\sim$  では、現世代からエリートは保存されるが非エリートは現世代からの選択ではなくエリートから作成される点が異なる。

## 2 .2 シミュレーションデータ

ここでは、(1)式の項目反応モデルに基づきテストデータを作成した。

$$p = 1 / \left( 1 + \exp \left( - D \sum_{k=1}^g a_{jk} - b_k \right) \right) \quad (1)$$

ここで  $D = 1.7$ 、また  $g$  はテストの次元数

である。 $a_k$ は次元  $k$  の識別力、 $b_k$ は次元  $k$  の困難度である。 $\sim_{ik}$ は被験者  $i$  の次元  $k$  の能力母数である。各次元とも  $a$  は平均0.85、SD 0.8、下限 - .8 (識別力であるにもかかわらずマイナスの値を許しているのに注意) 上限2.0の切断正規分布となり、 $b$  は平均 0、SD1.0の正規分布、 $\sim$  は標準正規分布に従う。通常新しい尺度を作成する場合には反転項目を考慮すれば、たいいてい場合は尺度全体の得点と項目得点は正の相関関係を持つが、それでも若干の項目は負の相関関係を持ってしまふことが起こる。このような現象を考慮してシミュレーションデータ作成の段階では、マイナスの識別力を許している。データ1と2は、基本的に2次元データとして、データ3は1次元データとして作成された。ただしデータ1、2は各次元の識別力を確率的に0に強制的に置き換えている。その確率はデータ1においては、各項目ごとに次元1の確率0.2、次元2の確率0.5である。データ2においては、次元1の確率0.1、次元2の確率0.9である。すなわち、データ1よりデータ2の1次元性が高い。全ての母数を確定した後に0から1の一樣乱数と(1)の比較により正誤パターンを生成している。1次元だけでなく2次元データを準備した理由は実際のテストや質問紙の作成場面では、項目作成に配慮しても、出来あがったものは1次元の尺度ばかりではなく2次元、あるいはそれ以上のもになってしまうことが多いことを考慮したものである。3次元以上のデータを作成しなかったのは、1次元データと2次元データの比較によっておおよその傾向がわかれば、3次元以上はその延長上にあることからある程度は推論が効くと考えた為である。

また、いずれも項目数は40項目、人数は500人とした。以上のデータに対して、G Aを各10回ずつ200世代まで繰り返した。

## 3 . 結果と考察

### 3 .1 評価関数が信頼性のみの場合

表2-1 ~ 表2-3は評価関数が信頼性のみの結

遺伝的アルゴリズムを利用した項目選択

表 2 - 1 係数の最大値 ( 評価関数は信頼性( 係数 ) )

モデル	データ 1				データ 2				データ 3			
	平均	SD	MAX	MIN	平均	SD	MAX	MIN	平均	SD	MAX	MIN
	0.849	0.012	0.869	0.829	0.817	0.015	0.846	0.784	0.879	0.014	0.905	0.864
	0.863	0.009	0.878	0.851	0.835	0.010	0.849	0.816	0.887	0.014	0.919	0.864
	0.874	0.006	0.884	0.865	0.843	0.008	0.856	0.830	0.894	0.011	0.917	0.874
	0.855	0.011	0.868	0.836	0.836	0.008	0.847	0.822	0.886	0.015	0.907	0.858
	0.854	0.012	0.873	0.832	0.831	0.012	0.852	0.804	0.892	0.011	0.905	0.876
	0.863	0.011	0.883	0.844	0.832	0.017	0.860	0.802	0.884	0.016	0.906	0.857
	0.867	0.018	0.885	0.841	0.842	0.008	0.857	0.828	0.895	0.014	0.921	0.877

表 2 - 2 世代交代回数 ( 評価関数は信頼性 )

モデル	データ 1				データ 2				データ 3			
	平均	SD	MAX	MIN	平均	SD	MAX	MIN	平均	SD	MAX	MIN
	74.500	71.697	199	1	47.100	60.952	198	5	18.500	19.007	62	3
	118.000	58.207	194	30	122.400	54.085	197	3	40.500	54.806	181	1
	113.800	66.025	198	4	131.700	54.756	193	18	100.800	69.584	200	1
	86.400	61.054	185	2	82.800	56.432	152	2	82.800	61.193	199	3
	84.400	57.037	200	11	112.300	63.198	187	5	100.600	62.984	199	2
	71.700	61.534	185	1	59.400	52.345	163	2	87.700	62.222	195	3
	127.900	54.300	199	36	125.700	73.613	200	4	83.900	68.335	193	8

表 2 - 3 採用項目数 ( 評価関数は信頼性 )

モデル	データ 1				データ 2				データ 3			
	平均	SD	MAX	MIN	平均	SD	MAX	MIN	平均	SD	MAX	MIN
	21.200	2.315	26	18	23.500	2.062	28	20	22.500	1.500	25	20
	24.200	2.315	28	20	25.700	2.610	29	20	22.800	2.857	28	19
	23.800	3.429	30	16	24.600	2.010	29	21	22.700	3.874	28	16
	22.800	1.249	25	21	24.400	3.292	30	21	22.100	2.385	27	18
	24.400	1.800	27	21	22.800	3.600	27	14	23.900	2.508	27	20
	25.200	2.821	29	19	23.800	2.135	28	20	23.000	3.162	28	19
	24.300	3.743	31	19	23.600	2.728	30	20	22.600	2.577	26	18

果である。まず初めに表2-1は、信頼性 ( 係数 ) を評価関数とした時に各モデルで達成された 係数の最大値の G A の10回の繰り返しによる平均、SD、MAX 及びMINである。データ1では、最も大きい平均を与えたモデル で0.874であり、最小はモデル の0.849となっている。データ1の40項目全てで計算したときの 係数は0.856であるからモデル などは、後に示す表2-3から分かるように、項目数を25% ~ 50%前後削減しているにもかかわらず、元のデータと同程度あるいはそれ

以上の 係数を達成していることが分かる。またデータ2の40項目全てで計算したときの 係数は0.827、データ3は0.871であり、どちらもデータ1と同様に G A で項目数の削減と信頼性の維持を同時に達成できていることが分かる。データ2でもモデル が最大の平均を与え、モデル が最小の平均となった。データ3では、モデル が最大を与え、モデル はわずかに劣って2位であり、モデル が最小の平均となっている。なおMAXではモデル は僅差ではあるがデータ1~3いずれ

も最良の成績を示している。モデル は、エリートの対極を次世代に置くものであり、ユニークな遺伝子を集団中に保持するという点で有効な方法である可能性を示したものと見えよう。図1~図3には、モデル、及びの全世代にわたる評価関数(係数)の変化の一例を示してある。離散世代モデルである図1とエリートを保存した図2、図3との関数の変化の違いが大きいことが分かる。図1からは60世代程度まで評価関数が低下傾向にあり、その後反転し横ばいから次第に上昇したが、初期の世代で達成した評価関数の最大値に200回という世代交替の回数の上限内で到達しなかったことが読み取れる。これに対して図2では、世代ごとの変化は一般に大きくなく、所々に生じている評価関数が急激に上昇しているギャップの部分を除けば通常は定常状態にあることが分かる。図3は図2とモデルの違いはあるものの基本的に類似した探索方法であることが図からも理解できる。図2及び図3からは、モデルやによれば、定常状態に陥った世代数の長さを調べれば、世代交替回数の一般的な上限(本研究では200回)を設けることなく、定常状態の繰り返し回数により計算を打ち切ることも出来るように思われる。今後の検討課題と見えよう。

表2-2は、信頼性(係数)を評価関数とした時に最大の係数に達した世代の順番、すなわち最良の染色体を得た世代までの世代交替回数の平均、SD、MAX及びMINである。データ1について見てみると、モデルによってバラツキがあるものの100回前後の世代交替を平均的に繰り返していることがわかる。しかし最大は200回、すなわち世代交替回数の上限、あるいはその近くに達しているものがある一方、何とMIN=1という初期解で最大の(とは言っても当該モデルのそのシミュレーションにおいての)に達して、以後の変化ではより良い染色体を生み出せなかった場合も結構あることが分かる。MAXが上限に近い例が多いことは、回数の上限そのものを引き上げればもう少し成績が改善される

余地がある程度あることを示しているものと思われる。データ2も、交替回数の平均のモデルごとの最大と最小が、データ1の場合よりも拡大したことの他は、データ1と同様の結果が見られる。しかしデータ3では、平均回数が全体的に短くなっていること、及びモデルのMAXが62と短くなっているのがデータ1、2と違っている。データは1から3の順に1因子性が強くなるので、表2-2からはデータが単純になるにつれて最大の係数に到達する回数は平均的には減少していることが分かる。

同じ程度の近似解を与える染色体であるならば、その染色体を生み出すまでの世代交替回数が短いモデルの方が良いであろうが、表2-1で見たように近似解の水準に違いがある以上モデルが良いとは考えられない。

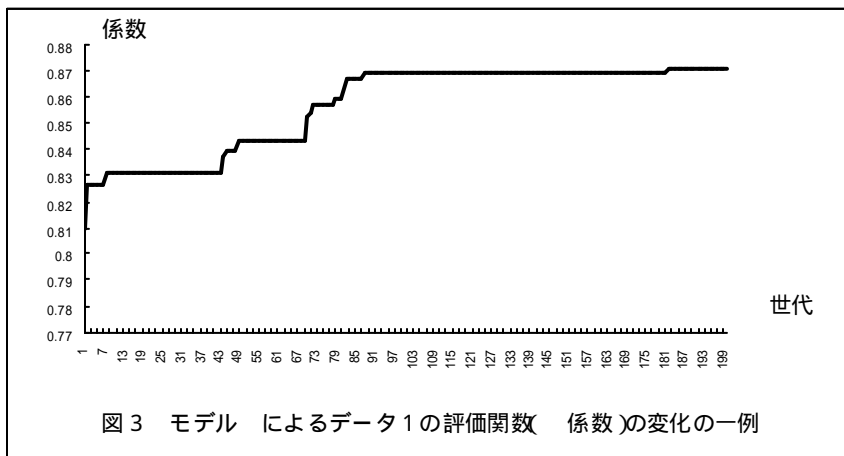
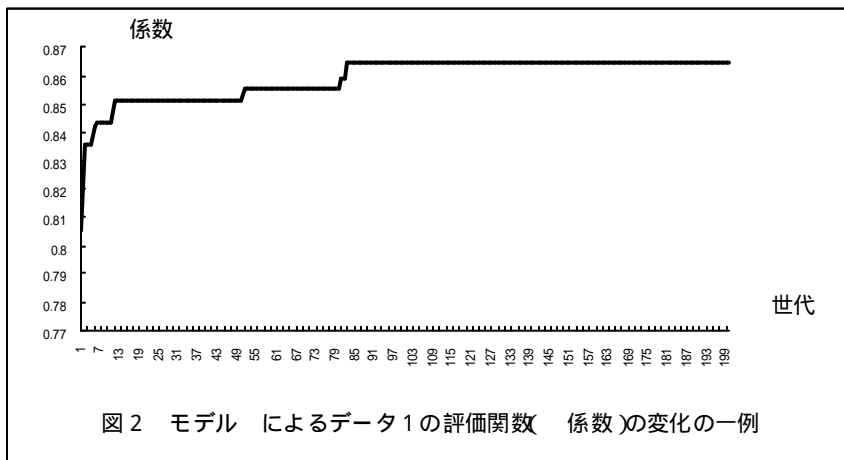
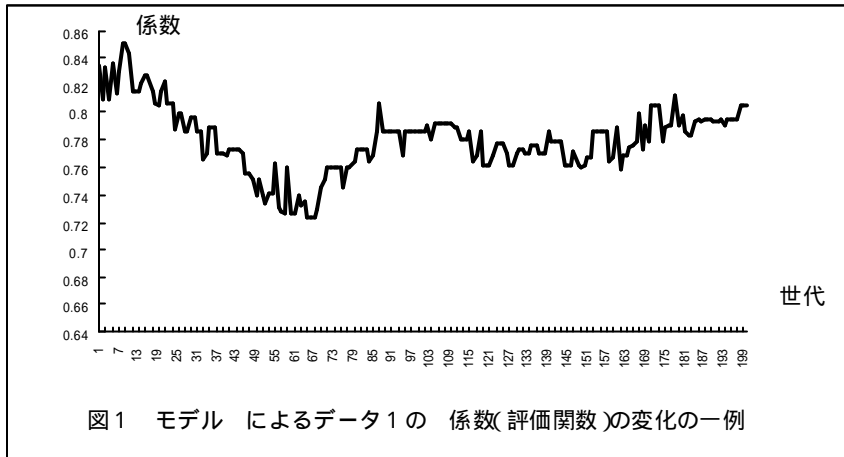
またどのモデルにも言えることだがランダムに作成された初期解を親世代に持つ世代の中に最大の評価関数を与える場合があるのは染色体の改良が必ずしも巧く行われていない可能性を示唆するものであり、今後の検討課題と見えよう。

表2-3は、信頼性(係数)を評価関数とした時に最大の係数を与えたテスト項目数である。元のデータの性質に依存するだろうが、本研究のデータ1~3では10~20項目前後、率にして25%~50%前後の項目が除かれている。評価関数として係数だけを設定したのであるから、かなり大きい割合で項目が削減されたことが分かる。

### 3.2 評価関数が信頼性+項目数の場合

表3-1~表3-4は評価関数を信頼性と項目数の増加のペナルティの和とした場合の結果である。まず初めに表3-1は、各モデルで達成された係数の結果である。もちろん係数を最大化していない以上、直接問題にすることは難しいが、テストや質問紙を分析する実務上からは、やはり重要な指標であることに変わりはない。同表の平均欄から見て、データ1~3いずれの場合も、評価関数として信頼

遺伝的アルゴリズムを利用した項目選択



性だけを設定したときと同様に、モデル と が良い成績を示していることが分かる。また表2-1と表3-1の比較から、今回の研究では評価関数の違いがあつたにもかかわらず達成された 係数の大きさそのものは、どちらの場合も大きな差が無いことが分かる。ペナルティが0.001であるのだから投入される項目数を考えれば、 係数の小数点2桁目には影響が出てきて良さそうであるが、表2-1と表3-1の違いは小数点3桁目の違いであると判断できる。これはもちろん項目増加に対するペナルティが比較的に穏やかなものであったからであろうが、項目選択においてペナルティがどのように実際に効いているのについては詳細に検討してみる必要があるだろう。

表3-2は、最大の 係数に達した世代までの世代交替回数の平均等である。データ1と2の交替回数の大小はマチマチであるがデータ3に関しては明らかに交替回数の平均が小さいと言える。またデータ1~3を通してモデル の交替回数の短さが際立っている。

表3-3は、最大の 係数を与えたテスト項目数である。項目数として10~20項目前後、率にして25%~50%前後の項目が除かれているのは表2-3と同様である。また表2-3と表3-3を比較してみると、興味深いことに評価関数を 係数だけにしたときと、項目数増に対するペナルティも加味したときでは、本来前者の項目数が後者を上回るのが自然であろうが、部分的には逆転しているケースもある。この観点からは、実施項目数の平均について逆転が生じない程度に、項目数増加に伴うペナルティの大きさをもう少し大きい値としてもよいかもしれない。この辺りは、項目数の制御をどの程度行いたいかという意図の強さにかかることであるから、実際にはシミュレーションで検討しながら個別に検討する必要があるだろう。

表3-4は、評価関数( 係数の大きさと項目数の増加のペナルティの和)の結果であるが、ここでもモデル と が相対的に良い成績を示しており、モデル が劣っていること

は今までと同様である。

これらの結果からは、3つのデータ、2種類の評価関数とも離散世代モデルのモデル は、やや成績が劣り、エリート保存のモデル 、あるいはその変形であるモデル の探索成績が相対的に良いと考えられる。

### 3.3 その他の問題とGAの可能性について

この他の問題として、たとえば各世代で残す染色体数 $m$ の決定は幾つとすべきかという問題などがある。本研究で設定した $m=15$ という値は特に理由とした根拠はない。 $m$ を多くすれば、遺伝情報の多様性も安定性も増すが計算量も急激に増加するし、少なすぎれば遺伝情報が限られしかも不安定となる。本研究で報告した $m=15$ の計算の前に、 $m=10$ で試してみたところ、やや少なすぎるような印象がしたので $m=15$ としているが、テストの項目数などとの関係においてどのような影響が生じるのかは今後の検討課題であろう。また交叉の方法などについても様々な可能性があり、テストや質問紙のデータにとって、どのようなものが適当なのかは今後の研究に待たなければならないだろう。本研究で用いたGAの手法は、非エリートの保存、特にエリートの対極を作り出して次世代に置く方法以外は、すべて基本的なものである。GAのその他の手法の採用や、より高次元のデータで検討することなども含めシミュレーションによる検討はもう少し充実する必要があるだろう。このように今後GAのテストや質問紙への適用において検討しなければならない課題は多いものの、GAによる項目選択は実務上大いに役立つそうである。なるべく少数の項目で、しかもできるだけ信頼性は低くしないように尺度を構成したい場合には、本研究で行ったように項目数の増加にペナルティを課せばよい。また、テストの項目数に希望があれば、希望するテスト項目数の前後で評価関数を大きくし、離れるに従って小さくなるように設定してGAを利用して項目選択を行えば、最適値に近い解を比較的容易に得ることができ



遺伝的アルゴリズムを利用した項目選択

表 3 - 1 係数の最大値 (評価関数は信頼性 + 項目数のペナルティ)

モデル	データ 1				データ 2				データ 3			
	平均	S D	MAX	MIN	平均	S D	MAX	MIN	平均	S D	MAX	MIN
	0.850	0.010	0.862	0.828	0.819	0.017	0.848	0.798	0.872	0.011	0.892	0.855
	0.858	0.008	0.873	0.847	0.828	0.012	0.855	0.812	0.885	0.010	0.901	0.872
	0.873	0.006	0.879	0.858	0.849	0.005	0.859	0.840	0.891	0.011	0.910	0.873
	0.860	0.010	0.876	0.845	0.836	0.008	0.845	0.823	0.882	0.011	0.900	0.867
	0.859	0.011	0.875	0.846	0.834	0.011	0.850	0.811	0.890	0.016	0.910	0.851
	0.862	0.010	0.873	0.844	0.838	0.010	0.851	0.813	0.885	0.014	0.910	0.867
	0.868	0.016	0.889	0.830	0.839	0.015	0.861	0.810	0.889	0.009	0.905	0.873

表 3 - 2 世代交替回数 (評価関数は信頼性 + 項目数のペナルティ)

モデル	データ 1				データ 2				データ 3			
	平均	S D	MAX	MIN	平均	S D	MAX	MIN	平均	S D	MAX	MIN
	50.600	61.350	183	1	55.500	78.007	196	1	4.500	3.612	10	1
	130.800	53.794	183	5	87.100	67.071	197	2	80.100	66.557	180	1
	163.100	22.056	199	120	137.300	54.681	200	38	79.800	76.971	197	1
	130.400	51.887	195	1	91.100	59.324	196	3	54.400	49.896	140	1
	77.500	56.388	193	11	114.500	56.505	170	3	80.800	66.991	179	8
	133.500	58.625	199	7	152.600	48.972	198	30	57.000	52.796	183	1
	120.400	61.743	186	2	113.700	37.884	173	57	97.500	48.380	179	7

表 3 - 3 採用項目数 (評価関数は信頼性 + 項目数のペナルティ)

モデル	データ 1				データ 2				データ 3			
	平均	S D	MAX	MIN	平均	S D	MAX	MIN	平均	S D	MAX	MIN
	19.900	2.548	24	16	23.700	3.926	29	14	23.200	2.821	28	19
	22.800	3.572	29	17	22.600	1.960	27	20	21.800	2.993	25	16
	23.600	2.458	28	20	23.100	2.508	26	18	21.300	2.830	25	15
	23.200	3.092	28	18	22.100	1.868	26	19	21.900	2.343	25	16
	24.500	2.291	28	20	22.800	2.272	27	19	21.700	2.610	28	18
	22.900	2.879	28	19	23.500	2.802	27	18	21.500	4.080	28	16
	23.300	4.076	30	18	23.100	3.590	29	17	21.700	3.318	29	18

表 3 - 4 評価関数の最大値 (評価関数は信頼性 + 項目数のペナルティ)

モデル	データ 1				データ 2				データ 3			
	平均	S D	MAX	MIN	平均	S D	MAX	MIN	平均	S D	MAX	MIN
	0.831	0.009	0.839	0.811	0.795	0.015	0.822	0.773	0.849	0.010	0.869	0.836
	0.835	0.010	0.848	0.819	0.805	0.012	0.830	0.790	0.864	0.010	0.879	0.850
	0.850	0.007	0.857	0.836	0.826	0.007	0.841	0.815	0.870	0.010	0.887	0.853
	0.837	0.011	0.849	0.817	0.814	0.008	0.824	0.803	0.860	0.011	0.876	0.844
	0.835	0.010	0.851	0.821	0.811	0.010	0.826	0.788	0.868	0.015	0.889	0.831
	0.839	0.009	0.849	0.824	0.815	0.010	0.829	0.791	0.864	0.014	0.889	0.842
	0.844	0.016	0.865	0.811	0.815	0.014	0.835	0.788	0.867	0.008	0.885	0.852

るだろう。最終的な係数に下限を設けたい場合は、もちろん当該のテストがそれを満足し得るものであったと仮定しての話であるが、項目選択の結果として得られる係数がある一定水準を下回ったら急激にペナルティが増加するような関数を評価関数に付加しておけば良いわけである。また項目反応理論（項目反応理論については芝（1990）、池田（1994）などを参照のこと）で問題になるような適用場面としては、希望するテスト情報関数に近くなるような項目群を、既存の項目プールから選び出す場合なども有力な候補となるだろう。

## 文 献

- 相澤彰子 スキーマ貪欲な遺伝的アルゴリズム（北野宏明編）「遺伝的アルゴリズム2」第1章 産業図書, 1995.
- Goldberg, D.E. Genetic algorithms in search, optimization, and machine learning. Reading, MA: Addison-Wesley, 1989.
- Guilford, J.P. Psychometric methods, 2nd ed. McGraw-Hill, 1954. 秋重義治（監訳）精神測定法 培風館, 1959.
- 池田 央 テスト 心理学研究法8 東京大学出版会, 1973.
- 池田 央 現代テスト理論 朝倉書店, 1994.
- 北野宏明編 遺伝的アルゴリズム 産業図書, 1993.
- 北野宏明編 遺伝的アルゴリズム2 産業図書, 1995.
- 北野宏明編 遺伝的アルゴリズム3 産業図書, 1997.
- Mitchell, M. An introduction to genetic algorithms. MIT Press, 1996. (伊庭齊志監訳「遺伝的アルゴリズムの方法」東京電機大学出版局, 1997.)
- 芝 直樹・小谷 学・赤沢堅造 遺伝的アルゴリズムによる階層型ニューラルネットワークの構造選択 計測自動制御学会論文集, 34, 8, 1080-1087, 1998.
- 芝 祐順 項目分析・テストの編集 肥田野直（編）テスト（心理学研究法7）東京大学出版会, 1972.
- 芝 祐順（編）項目反応理論 基礎と応用 東京大学出版会, 1990.
- 横尾裕規・萩原将文 遺伝的アルゴリズムを用いた自然画像からの複数顔領域抽出 電気学会論文誌 C 電子情報・システム部門誌, 117, 9, 1245-1252, 1997.