



HOKKAIDO UNIVERSITY

Title	「印象の推定モデル」の構築のための「判別分析」の計算的な方法に関する研究
Author(s)	大澤, 光
Degree Grantor	北海道大学
Degree Name	博士(文学)
Dissertation Number	乙第5738号
Issue Date	2001-03-23
DOI	https://doi.org/10.11501/3182591
Doc URL	https://hdl.handle.net/2115/51511
Type	doctoral thesis
File Information	000000400189.pdf



「印象の推定モデル」の構築のための
「判別分析」の計算的な方法に関する研究

大 澤 光

「印象の推定モデル」の構築のための 「判別分析」の計算的な方法に関する研究

大 澤 光

概 要

私たちは、暮らしや仕事で接するありとあらゆる対象に対して「印象」というものを“感じる”。また、「印象」というものを“持っている”。そして、私たちは、この“感じた”あるいは“持っている”「印象」によって、その対象あるいはそれに類似したものに対する態度や行動を決めている。したがって、私たちは、それを明に意識しているかどうかは別にして、毎日の暮らしや仕事のなかで、日常的に「印象」を扱っている、ということができる。

この論文では、私たちに取り組んでいる「『印象』を扱う仕事」に含まれる要素的な仕事のうちで、その対象に対する“最終的な”態度や行動を決める（つまり判断をする）ための情報である「印象」を導き出す「印象の推定」、および、これを実際の仕事に利用するときの数学的な“道具立て”（からくり）である「印象の推定モデル」の構築の方法論の1つである「判別分析」について論じている。この検討を通じて、また、「印象」を専門的に扱ってきた「心理学」の知識や経験を利用して、私たちに取り組む「『印象』を扱う仕事」の成果の“品質”を上げ、また、そのプロセスの“効率”を上げる（改善・向上する）ことが目的である。

この論文の前半（第1章）では、まず、私たちにとって「印象」をどう考えればよいかを検討し、その結果として、「印象」は「その対象が“自分にとってどうであるか”の評価の結果」と解釈するのがよいことを示す。次に、「印象」は、大別して、私たちそれぞれがその対象に対して持っている情報（つまりは知識や経験）、それをベース（手がかり）にしてその対象から新たに獲得した情報、および、その対象に対する私たちの利害と興味・関心の3つの要素をもとに作り出されて“感じる”ものであることを明らかにする。そして、それは“多元的”であり、あるものは“部分的”でありあるものは“全体的”／“総合的”であり、また、あるものは“相対的”でありあるものは“絶対的”であることなどを示す。さらに、これらの印象はいずれも、私たちが持っている情報が正しくまた確かであるという“条件つき”のものであることなどを明らかにする。

次に、私たちに取り組む「『印象』を扱う仕事」の内容と構成を検討し、これが“機能的な”意味では「印象の計測」「印象の分析」「印象の推定」などの要素的な仕事を含み、また、“プロセス的な”意味では「印象の計画」「印象の設計」「印象の実現・製作」「印象の評価」「印象の管理」などの要素的な仕事からなっていることを示す。このうち、“機能的な”意味での要素的な仕事の1つである「印象の推定」は、「『印象』を扱う仕事」が扱う対象に対する“最終的な”態度や行動を決める、つまり判断をするための情報である「印象」を導き出す重要な仕事であることを明らかにする。

さらに、「印象の推定」という仕事の内容を具体化し、その成果の品質とプロセスの効率を

上げる（改善・向上する）ためには、「推定」とは「すでに分かっていることとなにかの仮説や条件を根拠に、（仕事の目的の達成に）必要かつ十分な“まだ分からない”情報を推し量ること」と解釈するのがよいことを示し、この解釈のなかで、とくに「根拠とするデータはどのように得られたか」と「選択した方法の仮説や条件は妥当か」についての検討が重要であることを指摘する。

さて、「印象の推定」の方法のなかに“量的な”評価が含まれていたり、その方法が“複雑”であったり、あるいは、それを“自動的に”行うことが必要であるなどの場合には、数学的な「印象の推定モデル」が構築され利用される。そこで次に、（この論文の後半での検討の準備の1つとして、）実際の問題の必要に応じた「印象の推定モデル」を検討するための枠組みを検討し、併せて、「感性情報処理」を中心に、いままでに試みられた「印象の推定モデル」の構築の方法とその問題点を検討する。

この論文の後半（第2章と第3章）では、数学的な「印象の推定モデル」の構築の方法として、対象のカラー画像のデータから“計算によって得られた”特徴量（連続量）と、その対象に対して、平均的な回答者が“申告した”特定の印象の有無（離散量）を「判別分析」を利用して対応づける方法を論じる。

この「判別分析」には、すでに「解析的な方法」が開発され広く使われている。しかし、“機能的な”意味での要素的な仕事である「印象の推定」を進める上での重要な注意点の1つである「選択した方法の仮説や前提条件は妥当か」から考えると、この「解析的な方法」では、「判別分析の目的」である“もっとも判別力が高い”の意味を、計算技術的な理由などから、「相関比を最大にする」という“間接的な”解釈によっている。しかし、私たちが実際に取り扱う問題からの要請は、たとえば「誤判別数の合計を最小にする」など、より“直接的な”解釈、あるいは、実際の問題の細かな事情への配慮を必要としたものであることが多い。

この検討の結果をもとに、（判別分析の「解析的な方法」ではなく、）遺伝的アルゴリズムを利用した「計算的な方法」によって、特定の仮説や条件を前提とはせずに、実際に取り扱う問題の性格と必要に応じた「判別分析」の解（つまり判別式）を見つけ出すことができる、新しい“柔軟な”方法と計算アルゴリズムを開発した。当然のことではあるが、この新しい“柔軟な”「計算的な方法」で見つけ出される「判別分析」の解は、（仕事の必要に応じて）設定された目的をより“満たしている”という意味で、従来の「解析的な方法」で導かれる解よりも、論理的にも“判別力の高い”解である。

この新しい“柔軟な”「計算的な方法」では、判別式の形式として「1次式」や「2次式」だけでなく、「絶対値関数」などを含めて、どんな形式の関数でも“柔軟に”取り入れることができる。また、分析対象のデータのなかに、（判別得点の分布の上ではほかの対象からは離れ

てしまっている、という意味での) “特異な”つまり“異質な”ものが含まれている場合にも、これを考慮した解を見つけ出すことができる。

さらに、この新しい“柔軟な”「計算的な方法」には、データが複雑であり、“単純な”判別分析では完全には判別できない場合に、“単純な”判別分析の結果、誤判別を生じる可能性のあるすべての対象に対して、別途に“局所的な”判別分析を行うことによって、全体として“完全な”判別ができる解を見つけることができる「2段階判別方式」(含、逐次的な判別方式)や、2つの判別式を“同時に”使ってその組み合わせによって“完全な”判別ができる解を見つけることができる「二重判別方式」も含んでいる。また、1つの判別式で計算した判別得点の値の大きさで、3つ以上の集団(カテゴリー)に分類された対象を判別できる解を見つける「重判別分析」の方法をも含んでいる。

そして、この判別分析の「計算的な方法」を利用して、実際のビジネスバッグ(カバン)のカラー画像から“計算によって得られた”特徴量と、ビジネスバッグに対して実際の回答者が“申告した”特定の印象の有無を対応づける「印象の推定モデル」の構築を試みて、この「計算的な方法」が有効であることを確認した。

さて、判別分析を利用した「印象の推定モデル」は、“連続的な”説明変数と“離散的な”被説明変数を対応づけるものであった。ここで、説明変数のなかに“離散的な”ものを含むことができれば、たとえば対象の仕様や特徴(離散量)と、その対象に対する特定の印象の有無(離散量)を対応づける、数学的な「印象の推定モデル」を構築することができる。このため、代表的な“離散的な”変数と対応関係の1つである「論理変数と論理関係」に対して、この「計算的な方法」を適用し、この方法がそのままの形で適用できることを確認した。これは「数量化理論2類」に対する「計算的な方法」の開発に相当する。

さらに、判別分析を利用した「印象の推定モデル」の被説明変数のなかに“連続的な”ものを含むことができれば、たとえば対象のカラー画像を代表する特徴量(連続量)と、「SD法」などの方法によって計測したその対象に対する特定の印象の大きさ(連続量)を対応づける、数学的な「印象の推定モデル」を構築することができる。このため、「判別分析」の「計算的な方法」の一部に修正を加えた計算アルゴリズムを開発し、この「計算的な方法」が有効であることを確認した。これは「回帰分析」に対する「計算的な方法」の開発に相当する。

遺伝的アルゴリズムを利用した「判別分析」などの「計算的な方法」の開発によって、この方法を利用した数学的な「印象の推定モデル」の構築は、対象のカラー画像の特徴量(連続量)でも、仕様や特徴など(離散量)でも、また、その対象に対する特定の印象の有無(離散量)でも、印象の程度・大きさ(連続量)でも、それらの間の対応づけ(モデル化)をすることができるようになった。この新しい方法の開発によって、入力(説明変数)の数がかかなり“多く”、これらと出力(被説明変数)との間の対応づけがかかなり“複雑”である場合などにも、“実務に耐えられる”対応づけをすることができるようになった。

目 次

概 要	1
目 次	4
図表目次	6
1 章 「印象の推定」と「印象の推定モデル」	8
1－1 節 「印象」の検討の枠組み	9
1－2 節 「印象を扱う仕事」と「印象の推定」	15
1－3 節 「印象の推定」と「印象の推定モデル」	23
1－4 節 いままでの「印象の推定モデル」の構築	30
1－4－1 項 「感性情報処理」での研究	31
1－4－2 項 「印象の工学」での研究	38
2 章 遺伝的アルゴリズムを利用した計算的な「判別分析」の方法	48
2－1 節 「印象の推定モデル」としての「判別分析」	49
2－2 節 解析的な「判別分析」の方法の問題点	51
2－3 節 遺伝的アルゴリズムを利用した計算的な「判別分析」の方法	55
2－3－1 項 「遺伝的アルゴリズム」のレビュー	55
2－3－2 項 計算的な「判別分析」の方法	58
2－4 節 例題による「解析的な方法」と「計算的な方法」の解の比較	66
2－5 節 「第2目的関数」による解の改善	76
2－6 節 「感度分析」による解の改善	78
2－7 節 「判別式の形式」の自由な選択	82
2－8 節 「判別得点の分布」を考慮した判別分析	88
2－9 節 複雑な問題に対応できる「2段階判別方式」の方法	92

2-10節	複雑な問題に対応できる「二重判別方式」の方法	97
2-11節	「重判別分析」への適用の方法	102
2-12節	「計算的な判別分析」の特徴のまとめ	107
2-13節	ビジネスバッグのカラー画像に対する「印象の推定モデル」の構築の試み	113
3章	「遺伝的アルゴリズムを利用した計算的な『判別分析』の方法」の拡張	128
3-1節	「論理関係の判別分析」への適用の方法	129
3-1-1項	「印象の推定モデル」としての「論理関係の判別分析」	129
3-1-2項	「基本的な論理関係」の判別分析	131
3-1-3項	「排他的論理和」の判別分析	135
3-1-4項	「任意の論理関係」の判別分析	138
3-2節	「回帰分析」への適用の方法	141
3-2-1項	「印象の推定モデル」としての「回帰分析」	141
3-2-2項	「解析的な回帰分析」の方法の問題点	142
3-2-3項	「計算的な単回帰分析」の方法と計算の例示	144
3-2-4項	「計算的な重回帰分析」の方法と計算の例示	148
4章	「印象の推定」と「印象の推定モデル」の今後の研究	151
文 献		155
謝 辞		161

図表目次

図 1-4-1	「自由申告法」の調査シートの例	39
図 1-4-2	「印象の推定モデル」の概念の1つ	39
図 1-4-3	「同時関係」の概念のベン図による説明	41
図 1-4-4	「同時関係」の確かさの違い	41
図 1-4-5	「二重の同時関係」の概念のベン図による説明	43
図 1-4-6	「条件付きの同時関係」の概念のベン図による説明	43
表 1-4-1	ビジネスバッグの「同時関係」の例	44
表 1-4-2	ビジネスバッグの「二重の同時関係」の例	45
表 1-4-3	ビジネスバッグの「条件付きの同時関係」の例	46
図 2-1-1	「印象の推定モデル」としての「判別分析」	49
図 2-1-2	「判別分析」の論理	50
図 2-2-1	解析的な「判別分析」の目的関数	51
図 2-3-1	「遺伝的アルゴリズム」による解の改善	58
図 2-3-2	「遺伝的アルゴリズム」の処理の流れ	59
図 2-3-3	「優先生き残り」と「ランダム生き残り」の説明	61
図 2-3-4	「ランダム交叉」の説明	61
図 2-3-5	「トーナメント」の説明	63
図 2-3-6	「突然変異」の説明	64
表 2-4-1	あやめ科の3種類の植物の4変量観測値	67
表 2-4-2	「解析的な方法」で求めた例題の解	69
図 2-4-1	「計算的な方法」による計算のプロセスの例示	72
表 2-4-3	「計算的な方法」で求めた例題の2つの解	73
図 2-5-1	2つの集団を“完全に判別できる”場合の「解の自由度」	76
図 2-6-1	「遺伝的アルゴリズム」への「感度分析」の取り入れ	79
表 2-6-1	解の「感度分析」の例示	81
図 2-7-1	「2次式」の判別式による判別分析の概念	82
図 2-7-2	「2次式」の判別式による遺伝的アルゴリズムの計算	83
表 2-7-1	「2次式」と「“V型”の絶対値式」による判別分析の結果	85
図 2-7-3	「“V型”の絶対値式」の判別式による判別分析の概念	86
図 2-8-1	判別得点の「分布」を考慮した判別分析	89
表 2-8-1	「100%値」と「90%値」の解の比較	91
図 2-9-1	「2段階判別方式」の方法の説明	93
表 2-9-1	「2段階判別方式」による判別分析の結果	95
図 2-10-1	「二重判別方式」の方法の説明	97
図 2-10-2	「二重判別方式」の処理の流れ	98

表 2-10-1	「二重判別方式」による判別分析の結果	1 0 0
図 2-10-3	2つの判別式の組み合わせによる「判別のしかた」の違い	1 0 1
図 2-11-1	「重判別分析」への適用	1 0 2
図 2-11-2	「重判別分析」の処理の流れ	1 0 3
表 2-11-1	3つの群の「重判別分析」の結果(群の並び方を指定しない場合)	1 0 5
表 2-11-2	3つの群の「重判別分析」の結果(“BAC”の順序を指定した場合)	1 0 6
図 2-13-1	「局所自己相関特徴」の性質	1 1 4
図 2-13-2	「アタッシュケース」と「ショルダーバッグ」のカラー画像	1 1 6
表 2-13-1	「アタッシュケース」と「ショルダーバッグ」の カラー画像の特徴量と判別分析の結果	1 1 7
図 2-13-3	「沢山入りそうな」ビジネスバッグと 「沢山は入らなそう」ビジネスバッグのカラー画像	1 2 2
表 2-13-2	「沢山入りそうな」ビジネスバッグと 「沢山は入らなそう」ビジネスバッグの判別分析の結果	1 2 3
図 2-13-4	「高そうな」と「安そうな」のビジネスバッグのカラー画像	1 2 5
表 2-13-3	「高そうな」と「安そうな」の判別分析の結果	1 2 6
表 3-1-1	「論理和」の真理値表と判別分析の結果	1 3 1
表 3-1-2	「論理積」の真理値表と判別分析の結果	1 3 2
表 3-1-3	「否定」の真理値表と判別分析の結果	1 3 3
表 3-1-4	「否定論理和」の真理値表と判別分析の結果	1 3 3
表 3-1-5	「否定論理積」の真理値表と判別分析の結果	1 3 4
表 3-1-6	2変数の「排他的論理和」の真理値表と判別分析の結果	1 3 5
表 3-1-7	3変数の「排他的論理和」の真理値表と判別分析の結果	1 3 6
表 3-1-8	「3変数の論理関係の例題」の真理値表と判別分析の結果	1 3 8
表 3-1-9	「4変数の論理関係の例題」の真理値表と判別分析の結果	1 4 0
図 3-2-1	「回帰分析」の基本的なやり方	1 4 1
表 3-2-1	「回帰分析」の例題	1 4 4
表 3-2-2	問題1の「解析的な回帰分析」の結果	1 4 4
表 3-2-3	問題1で「誤差の絶対値和を最小にした場合」の結果	1 4 6
表 3-2-4	問題1で「誤差の最大値を最小にした場合」の結果	1 4 6
表 3-2-5	問題1で「推定値が観測値を下回らない場合」と 「推定値が観測値を上回らない場合」の結果	1 4 7
表 3-2-6	問題2の「解析的な回帰分析」の結果	1 4 8
表 3-2-7	問題2で「誤差の絶対値和を最小にした場合」の結果	1 4 9
表 3-2-8	問題2で「誤差の最大値を最小にした場合」の結果	1 4 9
表 3-2-9	問題2で「推定値が観測値を下回らない場合」の結果と 推定値が観測値を上回らない場合」の結果	1 5 0

1 章 「印象の推定」と「印象の推定モデル」

私たちは、私たちの暮らしや仕事で接するありとあらゆる対象に対して「印象」というものを“感じる”。また、「印象」というものを“持っている”。そして、私たちは、この“感じた”あるいは“持っている”「印象」によって、その対象あるいはそれに類似したものに對する態度や行動を決めている。したがって、私たちは、それを明に意識しているかどうかは別にして、私たちの毎日の暮らしや仕事のなかで、日常的に「印象」を扱っている、ということができる。

このうち、私たちに取り組んでいる「『印象』を扱う仕事」、つまり、私たちが“感じる”あるいは“持っている”「印象」に関連した、あるいは、私たちがその仕事について「印象」の果たす役割を“意識した”問題の発見や解決のうち、「印象の推定」という“機能的な”意味での要素的な仕事は、仕事の対象に対する“最終的な”態度や行動を決める（つまり判断する）ための情報である「印象」を導き出すための、とくに重要な仕事と考えることができる。また、この「印象の推定」という“機能的な”意味での要素的な仕事を含めて、「印象を扱う仕事」は、私たちが心理としてまた主観として感じるあるいは持っている「印象」を“専門的に”扱ってきた「心理学」の実際の問題への利用・応用ということができ、心理学の知識や経験の活用によって、仕事の成果の品質を上げ、また、そのプロセスの効率を上げることが期待できる。

この章では、この論文の主題である「印象の推定」について、まずは、私たちにとって「印象」とはなんであるかについて検討し、これに続いて、「『印象』を扱う仕事」、具体的にはこれに含まれる“機能的な”あるいは“プロセス的な”意味での要素的な仕事とその構成と内容を検討する。そして、“機能的な”意味での要素的な仕事の1つである「印象の推定」の内容とこれを考えるための枠組みを検討し、併せて、「印象の推定」の実際の仕事に利用するときの具体的な“道具立て”（からくり）である「印象の推定モデル」の構築の検討のための枠組みを検討する。そして、いままでに「感性情報処理」などの分野で試みられた「印象の推定モデル」の構築をレビューし、その問題点を検討する。

1-1 節 「印象」の検討の枠組み

私たちは、いろいろな対象に対して「印象」というものを“感じる”。たとえば生活用品、工業製品、建築物、造形物などのモノ（ハードウェア）に対しては、「大きい／小さい」「重そう／軽そう」「面白い／平凡だ」「派手だ／地味だ」「高そう／安そう」「便利そう／使いにくそう」などの印象を“感じる”。大きさ、形状、構造、色彩・配色、模様、材質など、モノのデザインの構成要素とその組み合わせに対してだけでなく、モノの機能・性能、重さ、使い勝手、価格、あるいは、「どんな人が使いそう／そうではない」「こんな場所に似合う／似合わない」など、さらに、社会的、文化的、環境的な側面からの価値・評価についての印象も“感じる”。

そのモノを見ただけでもそのモノのことを聞いただけでも、また、それを一寸使っただけでも長く使ったあとでも、私たちは、そのモノに関する情報（つまり知識や経験）とそのモノから新たに獲得した（つまり認識した）情報から、そのモノに対する「印象」を作り出して“感じる”のである。そして、その“感じた”「印象」によって、私たちは、そのモノに対して、「好きだ／嫌いだ」「欲しい／欲しくない」「買う／買わない」「使う／使わない」などの態度や行動などを決めるのである。

私たちは、空間や景観、写真や絵画、文字や筆跡などに対しても、それぞれの「印象」を“感じる”。“目に見える”ものだけでなく、言葉や文章、声や音、匂いや触感など、“目には見えない”ものに対しても、「印象」を“感じる”のである。物語を読んだり聞いたりすれば、「面白い作品だ」とか「感動的な話だ」などの「印象」を“感じる”。「つまらない話」だから途中で読むのをやめてしまう、というのも「印象」を“感じて”採った行動の結果である。歌や音楽を聴いたり、映画や芝居を見た場合も、私たちは、“感じた”「印象」によって、あれこれと想像してイメージを膨らませたり、気分をよくしたり、ある場合には、気分を悪くすることもあるのである。

また、仕事でも遊びなどでも、他人^{ひと}との付き合いには、互いに感じる「印象」や感情というものがある。その他人^{ひと}に対して感じる「印象」や感情によって、その他人^{ひと}との付き合い方や接し方が変わってくるのである。実際には付き合ったり接したことがない他人^{ひと}でも、その他人^{ひと}のした仕事の成果やその他人^{ひと}が現在している仕事のやり方を見れば、また、その他人^{ひと}の生き方の話を見たり聞いたりすれば、あるいは、その他人^{ひと}の姿をちらっと見ただけでも、私

私たちは、その他人^{ひと}に対してなにかの「印象」を感じ、ときには感心したり感動したり、尊敬したり馬鹿にしたり、笑ってみたり怒ってみたり、好きになったり嫌いになったり、ということもあるのである。

私たちが生活する上で、家族・親戚や近所・友達との付き合い、仕事や勉強の仲間、趣味や遊びの仲間、あるいは、企業や行政などから提供されるものを含めて、“社会”から受けている「サービス」は、物理的には“目には見えない”が、これらに対しても、私たちは「有用だ／無用だ」「分かりやすい／分かりにくい」「あたたかい／冷たい」などの「印象」を感じる。また、人の組織や集団、企業や都市、地域、国などといったものも、物理的には“存在しない”あるいは存在したとしても（少なくともその全体／全部は）“目には見えない”ことが多いのであるが、これらに対しても、私たちは、これらに対する知識や経験などこれらに関して新たに得た（つまり認識した）情報から、それぞれの「印象」を感じるのである。

私たちは、その対象が“目に見える”ものであれ“目には見えない”ものであれ、あるいは、実際に“存在する”ものであれ実際には“存在しない”ものであれ、さらには、実際に体験したものであれただ見たり聞いたり読んだりしただけのものであれ、私たちがそれをそれと認識したあるいは意識したすべての対象に対して、私たちの知識や経験などを通じて、また、私たちが新たに得た（つまり認識した）情報を通じて、なにかの「印象」や感情を感じるのである。そして、その感じた「印象」や感情で、その対象に対する態度や行動を決めるのである。

こういったことを考えると、「印象」は、私たちの生活や活動のありとあらゆる側面で存在し、また、ありとあらゆる側面で大きな役割を果たしている、ということが分かってくる。私たちが、なにかの対象に対して特別な「印象」を感じないのは、それが目に見えない、手で触れられない、耳に聞こえない、知らない、経験したことがない、頭で理解できないなどのために、私たちがその対象をそれと認識あるいは意識していない場合なのである。

では、私たちが感じる「印象」とはどんなものであろうか。国語辞典では、「印象」とは「見たり聞いたりしたときに対象物が人間の心に与える感じ。感銘。心に残っていること。」などと説明されている（文献1 以下「国語辞典」での説明について同様である）。また、「印象」の類義語である「感じ」は「感覚的にとらえたそのものの特徴。雰囲気。」などと、「イメージ」は「心の中に思い浮かべる姿や情景。心象。形象。イマージュ。」などと説明され

ている。同様に、「印象」の類義語である「心象」「感触」「気分」「フィーリング」「雰囲気」「ムード」などについても、「心で感じる」ことや「感覚的にとらえる」ことなどが共通に説明されている。

しかし、私たちは、自分にとってなんの利害や興味・関心もない対象に対しては、ごくごく一般的な「印象」は感じるとしても、特別な「印象」は感じない。私たちは、自分にとってなにかの利害なり興味・関心がある対象やその側面に対してだけ、なにかの特別の「印象」を感じるのである。

筆者は、こういったことから、「印象」とは「その対象が自分の態度や行動にとってどうであるかの事前あるいは事後の多元的な評価の結果」と考えるのがよい、と思っている。つまり、「印象」とは「その対象に対する評価の結果」と考えるのである。私たちがそれを明に意識しているかどうかは別にして、私たちは自分がなにかの利害や興味・関心を持っている対象に対する態度や行動を決めるために、その対象が自分にとってどうであるかという評価をするのである。その対象に接する前であれば、その対象に接した場合に、その対象は自分にとってどうであろうかという“事前”の評価をするであろう。これは「予想」や「予測」といってもよい。あるいは、実際にその対象に接した後であれば、その対象に接した結果として、その対象が自分にとってどうであったか、という“事後”の評価をするであろう。また、同様の対象に、次に接したときにはどうするかという“事前”の評価をするであろう。

“事前”であれ“事後”であれ、その対象に対する評価の結果によって、私たちはその対象に対する態度や行動を決めるのである。そして、私たちがその対象に対してする“事前”あるいは“事後”の評価の結果が「印象」である、と考えられるのである。私たちは、「印象」によってその対象に対する態度や行動を決めたあと、その決定がもたらす自分にとっての結果が、自分にとって「よかった」であれば、“事後”の評価としてのその対象に対する「印象」を、学習的によりよいものに変化させるであろうし、その決定の結果が、自分にとって「よくなかった」であれば、“事後”のその対象に対する「印象」を、学習的により悪いものに変化させるのである。

また、この解釈のなかで、「自分にとってどうか」というのは、その対象が私たちにとってなんの利害も興味・関心もないといった場合もあるであろうし、私たちにとって直接の利害や興味・関心につながるといった場合もあるであろう。その対象が私たちにとってなんの利害も興味・関心もない場合には、その対象に対する「印象」は薄くまたごく一般的なもの

にとどまるであろう。しかし、私たちがその対象に直接あるいは間接の利害や興味・関心がある場合には、その対象に対する「印象」はより具体的でより詳細なものに、また、ある場合にはより感情的なものになるであろう。テレビや新聞などで報道された“遠く”の事件と“身近”で実際に起こった同じような事件、あるいは、“他人”^{たにん}に関することと自分あるいは自分の“家族”や“友人・知人”に関する同じようなことでは、その「印象」が大きく異なっていることから、そうであることが分かるであろう。

さて、私たちがどこかでカラスを見た場合、通常、私たちは、「カラスだ」「黒い」などのごくごく一般的な「印象」を超えて、そのカラスに対する特別な「印象」を持つことはないであろう。それは、私たちがカラスに関する特別な知識や経験を持っていないからである。私たちは、知識、経験、判断基準、感情などのすべてを含めて、すでに持っている情報を基（手がかり）にして、その対象に関する新たな情報を獲得して（つまり認識して）、すでに持っている情報とこの新たに獲得した情報の両方を使って、その対象に対する「印象」を作りだす、つまり「印象」を感じるのである。したがって、私たちは、その対象に関連して持っている情報が少なければ、その対象に関する新たな情報を獲得することも少なく、それに対して、私たちは、ごくごく一般的な「印象」を感じたとしても、特別の「印象」は感じないのである。

ちなみに、なにかの対象に対する私たちの利害や興味・関心がより大きければ、その結果として、その対象に対する私たちの知識や経験はより多くなるであろう。反対に、その対象に対する私たちの利害や興味・関心がより小さければ、その対象に対する私たちの知識や経験もより少なくなるであろう。また、私たちのその対象に対する知識や経験がより多ければ、私たちのその対象に対する利害や興味・関心もより大きくなるであろうし、知識や経験がより少なければ、利害や興味・関心もより小さくなるであろう。一般的に言って、利害・興味・関心と知識・経験とは、こういった関係を持っているのである。

さらに、私たちは、その対象に対して多くの知識や経験を持っているものについては、その知識や経験を基（手がかり）に、あるいは、その知識や経験によって（頭のなかに）作られたなにかの基準（たとえばプロトタイプのようなもの）によって、その対象あるいはそれに似た対象に対して、ある程度は“絶対的な”「印象」を感じるであろう。しかし、私たちがその対象に対して、多くの知識や経験を持っていない場合、つまり、私たちがその対象あるいはそれに似た対象をよく知らない場合には、私たちは、その対象に対して“絶対的な”

「印象」を感じることは少なく、私たち自身あるいは他人が選択したなにかほかの対象と、
私たち自身が知識や経験を持っている側面について比較をすることによって、“相対的な”
「印象」を感じることが多い。一般的にいて、“絶対的に”感じる「印象」は、より確か
でより大きくより強いものであることが多く、“相対的に”感じる「印象」は、より変わり
やすくより小さくより薄いものであることが多いようである。

さて、私たちがなにかのモノやサービスを購入したり利用したりする場合、当然のことな
がら、その仕様、機能・性能、価格、デザインなど、そのモノやサービスの複数の側面や要
素に利害や興味・関心があるであろう。私たちの対象に対する利害や興味・関心は、一般的
にいて“多元的”なのである。私たちは、1つの対象の複数のつまり多元的な側面に利害
や興味・関心を持っているのである。そして、私たちのその“多元的な”利害や興味・関心
に対応して、私たちがその対象に対して、また、利害や興味・関心にしたがって感じる「印
象」もまた“多元的”になるのである。

私たちは、その対象の部分や側面、ほかの対象との関係、機能・性能、価値などに対して
は、それらの利害や興味・関心に対応した“部分的な”「印象」を感じ、「好き／嫌い」「欲
しい／欲しくない」など、その対象に対する態度や行動の決定に対しては、これに対応した
“全体的な”あるいは“総合的な”「印象」を感じるのである。一般的にいて、“部分的
な”「印象」が、直接的あるいは個別的なものであるのに対して、“全体的な”あるいは“総
合的な”「印象」は、間接的であり、（その対象に対する態度や行動の決定に関連している、
という意味で）結論的な「印象」ということができるかもしれない。あるいは、これは“態
度的な”や“行動的な”と表現した方がよいかもしれない。

私たちのいろいろなつまり多元的な利害や興味・関心と「印象」は、互いに独立している
わけではない。たとえば、あるモノの機能・性能やデザインに対して「素晴らしい」という
「印象」があれば、そのモノの価格は多少高くてもそのモノに対して「高い」という「印象」
は感じないであろう。しかし、そのモノの機能・性能やデザインが「普通のもの」という「印
象」を感じた場合には、これに対応した価格でなければ、そのモノは「高い」と感じるであ
ろう。私たちのいろいろな利害や興味・関心は、互いに“関連”し、また、互いに“依存”
しているのである。このため、私たちの多元的な「印象」も互いに“関連”し、また、互い
に“依存”しているのである。

さらに、私たちがなにかの対象に対して「印象」を感じる、その基になる知識や経験など

は、必ずしも確かなものばかりではない。なかには、あいまいなもの、いい加減なもの、ときには、誤ったものもあるであろう。ある知識や経験は正しくまた確かであったが、その前提条件が変化して、そのときには正しくなくまた確かではなくなった、ということもあるであろう。したがって、これらの情報を基にした「印象」は、「印象」の間の関連性などを含めて、その基となるこれらの情報が正しくまた確かであるということを前提にした“条件付き”のものである、と考えることができる。あるいは、私たちのある対象に対する興味と関心そのものが、これらの情報に関して“条件付き”のものであるともいえるのである。

ところで、当然のことながら、私たちが“感じた”「印象」は、私たちの頭のなかの記憶に残っている。この記憶が、頭のなかに残っている、つまり“持っている”「印象」なのである。私たちは、同じような状況で同じような対象に出会った場合、この“持っている”「印象」を基準（手がかり）にして、その対象に対する「印象」を作り出して（呼び出して）“感じる”のである。あるいは、私たちがその対象に対して“感じる”「印象」は、私たちがその対象あるいはそれと似た対象に対して“持っている”「印象」の影響を受けるのである。当然のことながら、“感じた”「印象」が大きくあるいは強ければ、“記憶に残る”「印象」も大きく強く、“感じた”「印象」が小さくあるいは弱ければ、“記憶に残る”「印象」も小さく弱いのである。また、“記憶に残っている”「印象」は、時間が経てば、次第に小さくあるいは薄くなっていく、という一般的な性格を持っているのである。

「印象」は、私たちがいろいろな対象に対して、頭や心で“心理的に”そして“主観的に”感じる、あるいは、“心理的として”そして“主観的として”頭や心のなかに持っているもので、一般的に、その性質として、「あいまいさ」「いい加減さ」「不安定さ」「変わりやすさ」など、あるいは、時や場所、人や状況などに応じた「違い」が挙げられている。問題の発見や解決をはかる「『印象』を扱う仕事」を行うに当たっては、これら「印象」の性質を十分に考慮して、また、“人の心理や主観を扱う”専門的な科学である「心理学」の知識と経験を利用して、仕事の結果として得られる成果の品質とその仕事のプロセスの効率を向上する、という取り組みの態度が必要であるといえよう。

1-2節 「印象を扱う仕事」と「印象の推定」

前節の繰り返しになるが、私たちは、私たちの暮らしや仕事のありとあらゆる対象に対して、なにかの「印象」を感じ、また、なにかの「印象」を持ち、これによって、それぞれの対象に対する態度や行動を決めている。その意味で、私たちが感じまた持つ「印象」は、私たちの暮らしや仕事のなかで、非常に大きな役割を果たしている、ということができる。また、私たちは、それを明に意識しているかいなかは別にして、毎日の暮らしや仕事のなかで、日常的に「印象」を扱っている、ということができる。とくに、私たちがしているいろいろな仕事は、それがどんな仕事であったとしても、なにかの形で「印象」に関連し、また、なにかの形で「印象」を扱っていることが多いのである。

しかし、「印象」は、私たちが頭や心で“心理的・主観的に”感じあるいは頭や心のなかで“心理的・主観的に”持っているもので、一般的にいつて、「人によって違う」「ときや状況などによって違う」「あいまいである」「変わりやすい」、あるいは、（“目には見えない”などの理由から、）「捉えどころがない」「よく分からない」などといった性格を持っている。このため、「物質科学」などが扱っている“人の心理や主観を含まない”対象に比べて、必ずしもその扱い方が簡単であるあるいは明確であるとはいえないところがあるのである。人が作り出すハードウェア、ソフトウェア、サービスを含めて、これらを「モノ」（アーティファクト）と呼ぶことにすると、“人の心理や主観を含まない”対象（モノ）に関する「『モノ』を扱う仕事」に比べて、“人の心理や主観を含んだ”対象に関する「『印象』を扱う仕事」はむずかしいとされる点が多いのである。

“人の心理や主観を含まない”対象に対する「『モノ』を扱う仕事」に対しては、私たちは、これにしたがって行えば、その仕事を進める方法や手続きについては、ある範囲内ではあるが、“客観的”“合理的”と主張できるものを用意しており、それにしたがって導き出した仕事の結果は（これも、ある範囲内ではあるが）正しくまた有効であると主張することができる。しかし、“人の心理や主観を含んだ”対象に対する「『印象』を扱う仕事」に対しては、（「客観的」「合理的」という言葉の意味の解釈のしかたにもよるが、）私たちは、一部の場合を除いてまだそのような方法や手続きを手にしておらず、このため、必ずしも仕事の結果が正しくまた有効であるとは主張できない部分があるのである。

もちろん、「『モノ』を扱う仕事」の場合にも、その方法や手続きなどの選択が仕事の目

的に対して“適切”でなければ、その方法や手続きなどが“客観的な”ものであったとしても、その仕事の結果が正しくまた有効である、とは必ずしも主張できないであろう。(ちなみに、筆者は、「『印象を扱う仕事』のための工学」として新しい工学として「印象の工学」を提唱し研究している。「印象の工学」の目的は、「印象を扱う仕事」の成果の品質とプロセスの効率を改善・向上することで、このため、「印象を扱う仕事」に適用される方法やプロセスをすべて明らかにしてその妥当性を問う、というやり方の採用を提案している(文献2)。)

さて、私たちがしている「『印象』を扱う仕事」とはどんなものなのであろうか。それはどんな内容を含んでいるのであろうか。

たとえばモノ(ハードウェア)の製品を考えてみると、私たちは、そのモノに対して、具体的には、そのモノの物理的なデザイン、機能・性能、価格、あるいは、そのモノの利用の側面に対して、「機能感」「高級感」「使いやすさ」「便利さ」「快適さ」などの「印象」を感じる。また、その全体に対しても、「欲しい」「使いたい」「買いたい」などの「印象」を感じる。これらの「印象」は、利用者が“感じる”「満足感」(あるいは「期待感」)ということもできよう。そして、私たちは、これらの「印象」あるいは「満足感」を通して、そのモノに対する態度や行動を決めているのである。具体的には、そのモノを「買う／買わない」、そのモノを「使う／使わない」、そのモノを「大切にする／しない」などといったことを決めているのである。

その意味で、企業が利用者(つまりは顧客)にモノを提供する、つまり、企画、設計、製造、販売などのすべてのプロセスを含めて、利用者にモノを提供する仕事は、企業の側から考えて、「印象を扱う仕事」ということができるのである。企業は、モノの提供を通じて、利用者にモノと企業の「印象」を提供し、これによって売り上げを上げているのだ、といってもよいのである。

近年は、モノの機能・性能が向上し、利用者の“機能・性能的な”要求のほとんどが満たされてしまい、モノの価値が、機能・性能などよりも、デザインなど、利用者の心理的な「満足感」つまりは利用者の感じる「印象」によるようになってきている。こういった傾向にしたがって、企業が行う「モノづくり」は、その機能や性能の作り込みを基本とすることは当然のことではあるが、これに加えてあるいはこれと平行して、そのモノに「印象」を作り込んでいく「印象づくり」といった仕事になっているのである。この意味からも、「モノづくり」は「印象づくり」つまりは「印象を扱う仕事」といえるのである。

さらに、利用者は、「モノ」に対して「印象」を感じると同様に、この「サービス」に対しても「印象」を感じるであろう。その意味で、いままでに、「モノづくり」について述べたことは、「サービス」についても当てはまるであろう。通常、企業の利用者への「モノ」の提供は、「モノ」の利用に関する情報の提供や修理などの「サービス」を含んでおり、利用者は、企業から提供される“すべて”に対して、なにかの「印象」を感じるのである。

当然のことながら、「印象を扱う仕事」は、仕事の内容や成果に「印象」が関係しており、仕事に取り組む人たちが、「印象」が“人が心理的・主観的に”感じるまた持つものであることを意識して、このためになにかの工夫をすることによって、仕事の成果の品質やプロセスの効率を上げよう、としているものをいうのである。たとえば「モノづくり」の場合、モノに「印象を作り込む」ということが“意識されて”行われている仕事は、「印象を扱う仕事」であるといえようが、こういったことをまったく“意識せずに”行われている「モノづくり」は、たとえそのモノの印象が仕事の結果に大きな役割を果たしているとしても、（少なくとも積極的には、）「印象を扱う仕事」とはいえないであろう。

これらのことは、（人が他人のために作りだす）^{ひと}「モノ」（アーティファクト）としてのソフトウェアやサービスなどについても同様である。とくに、そのモノの「モノづくり」のうち、モノと利用者との接点である「ヒューマンインターフェイス」づくりは、代表的な「印象を扱う仕事」といえる。たとえば、利用者にとって、モノの機能や内容が複雑で面倒な場合、利用者は、そのモノとの接し方をむずかしいと感じるであろう。そういう印象を持つであろう。その結果、利用者は、そのモノへの取り組みあるいは利用を躊躇するかもしれない。しかし、これに工夫を加えて、機能や内容を簡単なものにあるいは直観的に理解しやすいものにすれば、利用者のそのモノに対する印象は違ったものとなり、利用者はそのモノに簡単に取り組むあるいは利用することができるようになるであろう。

また、ヒューマンインターフェイスがモノの中身と大きく異なっている場合、利用者は、ヒューマンインターフェイスから感じる「印象」から判断つまり推定したモノの中身を“誤る”ことになり、そのモノを使わなくなるか、あるいは、大きな違和感を感じるであろう。（ちなみに、筆者は、「印象の工学」の研究のなかで、ヒューマンインターフェイスに関する「印象の設計」のもっとも基本的な原則の1つは、「オネストインプレッション」(honest impression) つまり「正直な印象」づくりである、と考えている。）

また、いわゆる「企業イメージ」に関する仕事も、代表的な「印象を扱う仕事」といえる

であろう。企業イメージがよければ、その企業に対して市場にたとえば技術的な意味での“信頼感”つまりそういう「印象」があれば、その企業の製品の売りに役に立つであろうし、また、その企業の人材の確保にも役に立つであろう。また、その企業が取引先を確保する場合にも、金融市場から資金を調達したりする場合にも、よい企業イメージはプラスに働くであろう。企業は、こういった理由で、自分の企業イメージを上げようとするのである。最近、企業が「CRM」(customer relation management)に注目をしているのも、こういったことのためといえよう。「企業イメージ」には、製品、技術、販売、宣伝、広告、社会活動など、企業の活動のすべてが“直接に”あるいは“間接に”関係しているのである。その意味で、「企業イメージ」の視点から、企業の活動のすべてが「印象を扱う仕事」といえるのである。

さて、「印象を扱う仕事」は、どのような内容からなっているのだろうか。まずは、「印象を扱う仕事」を、その“プロセス”に対応して分解すると、「印象の計画」「印象の設計」「印象の実現・製作」「印象の評価」「印象の管理」といった要素的な仕事からなっていることが分かる。たとえば「モノづくり」つまり「(モノの)印象づくり」を例にとると、「印象の計画」ではそのモノについて目標とする印象を選択・設定し、「印象の設計」ではそのモノにその印象を作り込む方法を具体化し、「印象の実現・製作」でその印象をモノの形や機能などとして実現し、「印象の評価」でそのモノにその印象が実現されたかどうかを確認し、「印象の管理」ではそのモノのその印象を維持・管理するのである。

「モノづくり」つまり「(モノの)印象づくり」は、モノに必要な印象を作り込むという仕事で、なにかの対象が実現されたときに、その対象が利用者などに与える印象を評価する、というのは“逆向き”に、その対象が利用者を与える印象を選択・設定して、利用者にその印象を与えることができるように、ある対象の仕様を実現するという“創造的な”仕事である、ともいえる。この対比として、モノに目標とする印象を作り込むのではなく、モノを作っていったらこういう印象になってしまった、という“順方向”の仕事は、一部の場合を除いて、“創造的な”仕事とはいえないであろう。

また、「印象を扱う仕事」を、それに含まれる“機能”に対応して分解すると、「印象の計測」「印象の分析」「印象の推定」などといった要素的な仕事を含んでいることが分かる。

「印象の計測」では、仕事の目的つまり必要に応じてその対象の印象を「はかる」、つまりは、仕事の目的に応じて、その対象の印象について、必要かつ十分な新しい情報を得るのである。また、「印象の分析」では、仕事の目的つまり必要に応じて、なにかの仮説と方法を

前提として、印象に性質や構造などに関する、これもまた必要かつ十分な新しい情報を得るのである。さらに、「印象の推定」は、仕事の目的つまり必要に応じて、また、なにかの仮説と方法を前提として、“すでに分かっている”ことを根拠にして、その対象の“まだ分からない”印象を推し量るのである。

当然のことながら、これらの要素的な仕事のそれぞれには、“人が心理・主観として”感じまた持つ「印象」を専門的に扱ってきた「心理学」の知識と経験が役に立つことになる。その意味で、「印象を扱う仕事」は、心理学と実務問題（あるいは工学）の出会い、実務問題からの心理学への挑戦、あるいは、心理学の知識と経験の実務への応用といえる。

さて、「印象を扱う仕事」の“機能的な”意味での要素的な仕事のなかで、「印象の推定」は非常に大切な仕事である。ある仕事では、私たちは、まず「印象の計測」を行い、その結果に対して「印象の分析」を行う。そして、その結果に基づいて、扱っている対象に対する態度や行動を決めている。しかし、ほかの多くの仕事では、私たちはまず「印象の計測」を行い、その結果に基づいて「印象の分析」を行い、その結果に基づいて「印象の推定」を行う。そして、この結果に基づいて扱っている対象に対する態度や行動を決めているのである。

私たちは、仕事の対象について、なにかの方法と手続きなどにしたがって、ある解を選択し、その解に対応する「印象の推定」を行なう。そして、得られた結果が“期待に近い”もので、“受け入れられる”範囲内にあれば、その対象についてその解を採用するのである。しかし、得られた結果が“期待に近い”ものではなく、“受け入れられる”範囲内になければ、その解を導き出した方法やプロセスの一部あるいは全部を変更することによって、得られる結果が“期待に近く”また“受け入れられる”範囲内のものとなる解とすることができるかいなかを検討することになるのである。

一般的に言って、「印象を扱う仕事」は、それに含まれている“機能的な”意味での要素的な仕事である「印象の計測」や「印象の分析」をしても、それだけでは仕事の結果としての態度や行動の決定に結びつかないことが多い。「印象の計測」や「印象の分析」の結果から、さらに、その方法が“直接的”であるか“間接的”であるかを別にして、「印象の推定」を行って得られた結果によって、扱っている対象に対する態度や行動を決めるという場合が多いのである。そのように考えると、「印象を扱う仕事」を完結するためには、「印象の推定」が必要である、あるいは、「印象を扱う仕事」は、どんな形であるにしても、必ず「印象の推定」を含んでいる、といえるのである。また、「印象の推定」は、「印象を扱う仕事」

の結論を出すための、あるいは、（結果の判断は除いて）目的としている仕事である、といえるのである。

ではここで、「印象の推定」の「推定」とはどんなことか、どんな内容を含んだものか、あるいは、含むべきであるか、を考えてみよう。国語辞典では、「推定」とは「はっきりとは分からないことをいろいろな根拠をもとに、あれこれ考えて決めること。」などと説明されている。（あるいは、「統計で、ある母集団から取り出された標本をもとにその母集団の平均・分散などを算出すること。」とも説明されている。）また、「推定」の類義語である「推測」は「ある事柄に基づいて、おしはかって考えること。推量。」と、「推量」は「おしはかること。推測。」と、「おしはかる（推し量る、推し測る）」は「ある事柄をもとにして他の事柄の見当をつける。推測する。推量する。」と、「見当」は「未知の事柄について立てた見込み。予想。」などと説明されている。

筆者は、（上記の説明を参考に、）「印象を扱う仕事」の要素的な仕事の1つである「印象の推定」もまた、目的と成果を持った仕事である、という立場から、「推定」とは「なにかの根拠をもとに、（仕事の目的の達成のために）必要かつ十分な“まだ分からない”情報を推し量るつまり得ること」と解釈するのがよい、と考えている。ここで、「仕事」とは、達成すべき目的があり、資金や時間などの資源を投入して達成する成果があり、これが、目的達成の立場から評価をされる、という意味である。

以下は、この解釈の内容をブレイクダウンして、「印象の推定」という仕事の目的を“有効に”達成しようとするための検討しなければならない注意点を整理したものである。ここで、「有効に」とは、仕事の成果の品質とプロセスの効率を上げる、つまり、仕事の目的に沿って必要かつ十分な品質の成果を得、かつ、仕事の目的に沿って必要かつ十分な効率のプロセスを実現しよう、という意味である。

- どの対象のどの印象に興味・関心があるのか。
- その印象はなんのために推定するのか。
- 得られた結果は誰がどう使うのか。
- その印象ですでに分かっていることはなにか。
- その印象でまだ分からないことはなにか。
- どんな結果を想定して推定するのか。
- （その目的を達成するために）その印象を推定するにはどんな方法やモデルがあるか。

●選択した方法やモデルには、どんな仮説や前提条件があるか。

●推定の根拠とするデータはどのように得られたか。

これらを少しく説明しておく。まずは、「『どの対象』の『どの印象』に興味・関心があるのか」と「その印象は『なんのために』推定するのか」である。「印象の推定」という仕事の目的を明らかにし、その仕事の対象をより具体的に明らかにすること、つまり、「なに」を「なんのために」やるのか、「なにが知りたいのか」を明らかにすることが大切である、ということである。仕事の目的や内容があいまいであったり、これらに不明なところがあったりした場合には、「印象の推定」という仕事の計画や設計にもあいまいさや不明なところが残し、したがって、実行や評価にもあいまいさや不明さが残ってしまう、ということである。これらは「なにをどこまで知りたいのか」といいかえることもできよう。

次は、「得られた結果は『誰』が『どう使う』のか」である。「印象の推定」という仕事の結果は、たとえば発売予定の製品のデザイン案の選択に使う場合もあるだろう。製品の生産量を決めるために使う場合もあるであろう。あるいは、なにかの判断や意思決定に“直接”使うこともあるだろうし、これらのためにはただの参考情報である場合もあるだろう。いずれにしても、「印象の推定」という仕事の結果を、誰がどのように使うかによって、「印象の推定」のしかた、つまり、方法や手続きの選択、予算や時間やスタッフの投入は違ってくるのである。時間や予算やスタッフを割いて、「印象の推定」という仕事をするのであるから、その成果は、投入した資源に見合ったものでなければならないのである。

この注意点も、「なにをどこまで知りたいのか」といいかえることができよう。ここで、「知りたいこと」は「必要かつ十分な情報」ということであり、仕事の目的に応じて、数値的な結果であれば、平均値であったり中央値（メディアン）であったり最頻値（モード）であったり、最大値や最小値であったり90%値（パーセンタイル）であったり、分布であれば、バラツキ（分散）であったりその形状であったり、時間的な変化であれば、傾向であったり周期であったりするものである。

その次は、「その印象で『すでに分かっていること』はなにか」である。「すでに分かっていること」はあらためて推定する必要はない。「すでに分かっていること」をあらためて推定してみても、それは時間と予算とスタッフの無駄遣いになってしまう。もちろん、「まだ“はっきりとは”分からないこと」を確認のために推定しよう、ということはあるであろうが、これは“確認”という新しい情報を得ているのである。

同様に、「その印象で『まだ分からないこと』はなにか」と「どんな結果を想定して推定するのか」である。「まだ分からないこと」はなんだということが分かっているから、推定をする必要があり、また、「こんな結果」になりそうだななどということが分かっているから、推定ができるのであって、「なにがなんだか分からない」では、推定の必要も推定のしようもないのである。その意味で、「すでに分かっていること」「まだ分からないこと」そして「知りたいこと」の整理と区別は大切なのである。

最後は、「（その目的を達成するために）その印象を推定するにはどんな方法があるか」と「選択した方法にはどんな仮説や前提条件があるか」である。「印象の推定」の方法（あるいはモデル）にはいろいろなものがあるが、このなかから仕事の目的や性格に合ったものを選択しよう、ということである。選択した方法は、「印象の推定」で得られる結果の“根拠”の1つであり、仕事の目的や性格に合ったものでなければ、得られる結果は、仕事の結果として受け入れられるものにはならないのである。また、選択した方法には仮説や前提条件があり、「印象の推定」で得られた結果は、これらになりたつという範囲内で受け入れられるのである。選択した方法の仮説や前提条件がなりたたなければ、得られる結果は受け入れられるものとはならないのである。

また、「推定の根拠とする『データ』はどのように得られたか」である。「印象の推定」の方法やモデルが、仕事の目的つまり必要性に合ったものであり、その処理に、高度な統計手法やスーパーコンピューターなどを使ったとしても、これに使うつまり推定の根拠とするデータが、仕事の目的を満たすために“必要かつ十分な”ものでなければ、得られる結果は不十分なものである！というわけである。このため、推定の根拠として使うデータがどのようにして得られたものであるかを明らかにすることが大切である！ということである。これは、キャッチフレーズの的にいえば、「G I G O」つまり garbage in garbage out！（ゴミを入れればゴミが出てくる！）にならないように！ということであろう。別のいい方をすれば、「印象の分析」や「印象の推定」などがする「記号の処理」よりも、「印象の計測」がする「問題の記号化」の方がはるかに大切である！ということでもある。

いずれにしても、「印象の推定」という仕事は、「なにかの根拠を基にして仕事の目的の達成のために必要かつ十分な“まだ分からない”情報を得ること」という視点から、上記の枠組みで考えることで、実際に扱う問題つまり問題発見や問題解決（の目的に応じた）成果の“品質”を上げ、また、仕事のプロセスの“効率”を上げることができる、と筆者は考えている。

1-3節 「印象の推定」と「印象の推定モデル」

前節でも述べたが、「印象の推定」という“機能的な”意味での要素的な仕事の内容は、“すでに分かっている”情報（つまり知識や経験）から導かれた（推定のための）「方法と手続き」によって、“すでに分かっている”データ（つまり入力）から“まだ分かっていない”「印象」（つまり出力）を推し量ることである。

このうち、“すでに分かっている”情報から導かれた（推定のための）「方法と手続き」が、“すでに分かっている”データつまり「入力」と“まだ分からない”ことつまり「出力」の間の「対応づけ」あるいは「モデル」と呼ばれるものである。これは「印象の推定モデル」と呼んでもよい。「入力」は説明変数、「出力」は被説明変数とも呼ばれる。あるいは、この「印象の推定モデル」は、入力と出力の間の「対応づけ」に関する「説明」とこの説明のための「仮説や前提条件」ということもできる。

繰り返しになるが、「モデル」に対応する“推定のための”「方法と手続き」は、“すでに分かっている”情報つまり知識や経験から導かれたものである。（一般に、知識や経験とはそういうものなのであるが、）この“すでに分かっている”情報は、なにかの仮説や条件がなりたつ範囲内で“有効”なのであって、これらの仮説や条件が異なったあるいは変わった場合を含めて、あらゆる場合にそのままなりたつというわけではないのである。したがって、入力と出力の「対応づけ」の説明も、この“すでに分かっている”情報が“有効な”範囲内で“有効”なのである。これは非常に大切なことである。

この“すでに分かっている”情報が有効な範囲を決めるのが、「仮説と前提条件」というものである。統計学を例にとると、たとえば扱うデータは正規分布にしたがっているとか、比較する2つの集団は分散や共分散が等しいとか、2つの変量は互いに独立であるとか、そういったことを「仮説や前提条件」として、必要な概念の定義をし、これに基づいた必要な統計的な計算をして、仕事の目的つまり必要性や性格を満たす解を導き出して、その説明をするのである。

したがって、「対応づけ」の「モデル」の仮説や前提条件とこれに基づいた説明が、仕事の目的や性格に照らして“妥当な”ものであり“受け入れられる”ものであれば、この仮説・前提条件を基にした説明にしたがって、“すでに分かっている”データを根拠に“まだ分からない”こと（つまり「印象」）を出力つまり推定することができるのである。その推定の

結果は、仕事の目的や性格に照らして「モデル」の仮説・前提条件とこれに基づいた説明が妥当な範囲内で、そして、（推定の根拠となる）“すでに分かっている”データが確かな範囲内で、確かであり有効であるというわけである。当然のことながら、これらのいずれかが妥当なあるいは確かな範囲を超えた場合には、推定の結果は、確かなものではなく有効なものでもないのである。

さて、「印象の推定」を行う説明と仮説・条件に基づいた「対応づけ」の「モデル」つまり「印象の推定モデル」は、入力（つまり説明変数）と出力（つまり被説明変数）の種類、および、「対応づけ」の方法と手続きによって、いくつかのタイプに分類できる。

まずは「入力」と「出力」の種類による分類である。「入力」と「出力」の種類として、アイテムの有無、カテゴリーや分類パターンのコード、数値などが挙げられる。このうち、「アイテム」は、なにかの仕様や特徴の有無のこと（離散量：0 or 1）であり、「カテゴリー／分類パターン」は、対象の分類の結果のこと（離散量：1, 2, 3,）である。また、「数値」（連続量）で表されるものには、1つの数値で表されるものと、複数の数値の組み合わせ（アレイ）で表されるものがある。さらに、これらには確定的なものと、そうではないものがある。ちなみに、一般的に、数値は“量的な”ものを表し、アイテムの有無、あるいは、カテゴリーや分類パターンのコードは“質的な”ものを表すといえる。

「印象の推定モデル」の入力と出力の検討には、こういった整理が必要であり、とくにその結果の妥当性の評価には、たとえば「入力がどのように得られたものであるか」などの検討事項を含めて、入力データの“妥当性”の検討が必要となる。ちなみに、前の1－2節で、「印象を扱う仕事」の“機能的な”意味での要素的な仕事の1つである「印象の計測」と「印象の推定」の仕事の大事な注意点として、「G I G Oにならないように！」を挙げたが、これはここで述べた「入力がどのように得られたものであるか」と同じことなのである。

次は、「対応づけ」の「方法と手続き」である。「対応づけ」の方法と手続きには、言葉によるもの、論理によるもの、数学によるもの、あるいは、それらの組み合わせによるものがある。これらはそれぞれ「言語モデル」「論理モデル」「数式モデル」「組み合わせモデル」などと呼ばれている。このうち「言語モデル」は、言葉という方法で入力と出力の対応づけを表現したモデルであり、「論理モデル」は、論理という方法で入力と出力の対応づけを表現したモデルである。また、「数学モデル」は、数学的な方法で入力と出力の対応づけを表現したモデルであり、入力データを数学的な方法で処理して、必要な結果を出力するの

である。当然のことながら、入力も出力も、数値である必要はなく、アイテムの有無、あるいは、カテゴリーや分類パターンコードであってもよい。

「対応づけ」つまり「モデル」の「説明」のしかたには、大別して2つのタイプがある。たとえば、“すでに分かっている”情報つまり経験として、ある入力が大きくなると出力も大きくなり、入力が小さくなると出力も小さくなる、ということが分かっている場合、つまり、経験的に、そういった「対応づけ」ができることが分かっている場合には、入力と出力を、たとえば1次式や2次式などの適当な関係式で対応づけ、そのよいパラメーターを見つけることによって、「対応づけ」のよりよいつまり不具合のより少ない説明としよう、というやり方である。これは「パラメーター依存型」の方法論である。ここで、対応づけの根拠とした「経験」とは、（なにを「理論」や「法則」と呼ぶか、という問題はあるが、）とくになにかの理論や法則などに基づいたものである必要はない。

たとえば「線型回帰分析モデル」では、入力（説明変数）と出力（被説明変数）の対応づけつまり回帰式の形式として1次式を選択し、ほかの方法で観測されたなにかのデータから、この回帰式のよりよいつまり不具合の少ないパラメーターを決める。回帰式の形式および決められたパラメーターの値には、“積極的な”根拠はないのではあるが、ほかの方法で観測されたなにかのデータに対して、よりよいつまり不具合の少ない説明ができる、という理由が、この対応づけの妥当性の“消極的な”根拠となる。これには、統計的な方法が用いられるのが普通である。

このほか、「ニューラルネットワークモデル」などの方法論も、入力と出力の「対応づけ」に“積極的な”根拠があるわけではないので、このタイプの方法論といえる。しかし、この統計的な方法は、実は、扱う対象の集団が“同質な”ものであることを前提とした方法であるので、“異質な”対象に対しても有効であるとは“積極的には”いえないことに注意が必要である。これについては、あとで1-3-1項や2-1-3節でも述べる。

これに対して、“すでに分かっている”情報つまり知識として、ある入力が大きくなると出力も大きくなり、入力が小さくなると出力も小さくなる、ということが分かっている場合、つまり、（ここでも、なにを「知識」や「理論」と呼ぶか、という問題はあるが、）知識や理論として、そういった「対応づけ」ができることが分かっている場合には、入力と出力を、その知識に基づいた関係で対応づけた説明としよう、というやり方である。これは「構造依存型」の方法論である。ここで、「知識」とは、経験的な法則あるいは理論的な法則など、

観察例の豊富なそしてその構造に関して豊富な分析に基づいた知識ということである。これに対して、上記の「経験」とは、相対的に観察例の少ない知識、ということができる。

このタイプの方法論の根拠は、「パラメーター依存型」の方法論のそれに比較して、より“積極的な”ものということができ、その結果の妥当性の検討は、この「知識」の妥当性、つまり、この知識がその対象の「対応づけ」に適用できるのか、この知識の有効な範囲はどこまでか、といったことを検討することと同じである。

ちなみに、入力と出力の「対応づけ」に対して、そのすべてになにかの理屈が付けられるか、というと、これは大変にむずかしいことである。これは、“すでに分かっている”情報である知識にしても経験にしても、それらの妥当性の判断は相対的にしかできないからである。同様に、その「対応づけ」の方法が、“パラメーター依存型”であるか“構造依存型”であるかの区別もまたむずかしい。これは、なにかの方法論が“パラメーター依存型”であるか“構造依存型”であるかの判断も、相対的なものであるからである。大切なことは、「印象の推定モデル」の構築のために選択した方法がどのような根拠に基づいたものかを明らかにしておき、これによって、その方法とその方法が導き出す結果が持っている限界を明らかにしておくことといえる。（ちなみに、2章で説明する「判別分析」を利用した「印象の推定モデル」は、基本的には“パラメーター依存型”の方法論である。）

さて、1-1節でも述べたように、「印象」は、物質科学が扱う自然科学的な対象とは違った性質を持っている。原則として、自然科学的な対象は、（ある程度あるいはある範囲内では「客観的」ともいえる）共通の定義、共通の基準や測度などによって（ある程度あるいはある範囲内では）機械的に「はかる」ことができ、直接には「はかる」ことができない場合にも、その方法が間接的あるいは代替的であるとしても、論理的な方法で「はかる」ことができる。これに対して、「印象」は、人が心理としてまた主観として“感じる”あるいは“持つ”ものであり、原則として、（「客観的」という言葉の意味の解釈にもよるが、）客観的な方法では「はかる」ことはできず、また、一般的に、印象には、人やときや状況や場所などによって違う、時間的な変化するといった性質があるため、これらの性質を考慮して、また、前提条件を明らかにして、論理的に「はかる」ことをし、論理的に「推定」をしなければならないのである。

「印象の推定モデル」を構築するときには、こういったことも注意点の1つとして考えておくことが必要であるが、現在までのところ、そこまでの対応を必要とする“実務”の問題

の必要性は提起されておらず、このためもあって、「印象の推定」ではそこまでの対応はなされていないのが現状といえよう。今後は、「感性情報処理」や「感性工学」などへの興味や関心が増すにつれ、こういった必要性（需要）も喚起されてくるかもしれない。しかし、現在までのところでは、「印象の推定モデル」の構築に関するこういったきめの細かい対応よりも、「印象の推定モデル」を構築するときに採用される方法の仮説や前提条件が、実務の問題として“妥当”つまり“受け入れられる”かいなか、そして、その方法がその問題の目的つまり必要性を“満たす”ことができるかどうかの検討の方が大切である、と考えられる。

ちなみに、2章で取り上げる「判別分析を利用した『印象の推定モデル』」は、比較的“ポピュラーな”対象に対する“平均的な”回答者の反応を推定するためのモデルである。回答者は、ある程度の知識や経験を持っており、これらの対象に対して、ある程度は、“絶対的な”印象を感じる事が想定されている。したがって、このモデルは、回答者に、特別な知識や経験を必要とする特別な印象は対象にしていない、ということに注意が必要である。

さて、数学的な「印象の推定モデル」は、数学的な手段や方法によって記述された論理的な対応づけである。その入力（説明変数）と出力（被説明変数）としては、それぞれ数値的、アイテム的あるいはカテゴリー的なものが考えられる。このうち、数値的なものは“量的な”入力あるいは結果、そして、アイテム的あるいはカテゴリー的なものは“質的な”入力あるいは結果といえる。入力と出力が、それぞれ数値的なものであるか、あるいは、アイテム的あるいはカテゴリー的なものであるかによって、「印象の推定モデル」のタイプには、いろいろなものが考えられる。

つまり、「印象の推定モデル」のタイプとして、第1は数値的な対象の仕様や特徴あるいはその組み合わせを入力して、数値的な印象の大きさを出力するモデル、第2は数値的な対象の仕様や特徴を入力して、アイテム的あるいはカテゴリー的な印象の大きさを出力するモデル、第3はアイテム的あるいはカテゴリー的な対象の仕様や特徴を入力して、数値的な印象の大きさを出力するモデル、第4はアイテム的あるいはカテゴリー的な対象の仕様や特徴を入力して、アイテム的あるいはカテゴリー的な印象の大きさを出力するモデル、そして、第5にそれらを組み合わせたモデルなどが考えられる。

言語的なモデルや論理的なモデルではなく、この数学的な「印象の推定モデル」を構築する目的つまり必要性は、数値的な入力から数値的な出力を得たい場合、あるいは、数値的な

入力からアイテム的あるいはカテゴリー的な出力を得たい場合が考えられる。また、入力と出力のそのほかの組み合わせを含めて、入力と出力の対応づけが“複雑”である場合、その対応づけの計算の手間を効率化したい、といった場合もあるだろう。たとえば、対応づけの方法のなかに、非線型な要素や交互作用、あるいは、時間的な遅れなどが含まれている場合は、対応づけは“複雑”で、人の計算能力や勘だけでは対応できない。あるいは、対応づけが状況に依存するつまり人やときや状況や場所などの違いがある場合も対応づけは“複雑”で、人の力だけではうまく対応づけができない。

こういったときには、数学的なモデルを構築して、数学的な方法と手続きによって、簡単に対応づけをするのである。さらに、この対応づけのモデルを使って、システムを「機械化」あるいは「自動化」するなどの目的で、繰り返し使う場合、数学的な手段によるモデルが使われる。「印象の推定モデル」を構築あるいは利用する場合には、こういったそれぞれのタイプのモデルの特徴、および、「印象の推定モデル」を構築・利用する目的つまり必要性を明らかにしておくことが必要であろう。

最後に、「印象の推定」に関する大切な話題の1つとして、「逆推定」に触れておく。たとえば製品のデザインが決まり、対象者と状況が決まれば、（その印象の再現性や安定性などを別にすれば、）その印象は決まる、と考えてよいであろう。その対象者が、そのデザインをどう感じるかが決まるといってもよい。これを推し量るのが「推定」である。これに対して、なにかの印象が決まり、対象者と状況が決まった場合、これに対応した製品のデザイン案を決める、というのがやり方がある。これが「逆推定」である。この場合は「設計」といってよいかもしれない。一般に、「推定」を「説明」や「印象」に対応するものと考えたとすると、「逆推定」は「想像」や「創造」に対応するものと考えてよい。

一般に、「推定」は、多入力から少出力への“単純な”メカニズムによっていると考えられ、比較的簡単に解を見つけることができる。これに対して、「逆推定」は、（不完全な）少入力から多出力への“複雑な”メカニズムによっているようで、容易には解を見つけることができない。この「逆推定」は、いわゆる「逆問題」といわれるもので、一般的に、解空間がごくごく単調な場合を除いて、解くことができない（文献3）。

人工知能で、「AならばBである」という知識（プロダクションルール）を使って、AからBを推定することを「前向き推論」といい、BからAを推定することを「後向き推論」というが、これと対比すると理解しやすいかもしれない。人工知能の例では、「AならばBである」という知識を

使って、AからBを推定することは、「AならばBである」という知識の確実さの範囲で可能であろう。しかし、BからAを推定するには、本来ならば「BならばAである」という知識が必要で、「AならばBである」という知識から、BからAを推定するには、論理的な無理がある。また、無理をするとしても、「BならばAである」という可能性も考えられる、といった程度の推定でしかない。また、「AならばBである」という知識のほかに、「A'ならばBである」「A''ならばBである」などという知識や経験が存在する可能性があれば、Bからの「逆推定」の解としては、Aのほかに、A'やA''なども逆推定しなければならないのである。

そういった意味で、「印象の“逆”推定」は、ある印象を実現する製品のデザインの部分的な解を、ある確かさの範囲内で導くことはできるかもしれないが、得られる部分的な解は解の全体を網羅していないことを知っておかなければならないのである。ちなみに、通常の順方向の推定にしても、ある製品のデザイン案に対する印象は、人やときや場所や状況などによって違うことがあるので、1つの印象の有無や程度に対応できるとは限らず、このような場合には、順方向の推定も、原理的にできる、というものではないことになる。

いずれにしても、「印象の推定」「印象の“逆”推定」とともに、印象の性質や推定の方法などから考えて、得られる結果は“その程度のもの”であることを知っておくことが必要であろう。確かにある程度の「推定」や「逆推定」は、いろいろな方法や根拠をもとにできるであろうが、私たち人間がするようなきめの細かい、あるいは、ときや状況などに応じた「推定」や「逆推定」は、現在および近い将来のうちには、原理的な意味を含めてできそうにはない、と考えてよさそうである。

1-4節 いままでの「印象の推定モデル」の構築

この節では、2章の「遺伝的アルゴリズムを利用した計算的な『判別分析』の方法」の説明に先立って、いままで、数学的な「印象の推定モデル」がどのような方法で構築されてきたか、そして、それらはどんな問題点をかかえていたかを簡単にまとめておく。

「人工知能」「知識情報処理」の研究に続く「感性情報処理」や「感性工学」の研究への興味や関心が高まるにつれて、人の“感性”に基づいたあるいは“好み”に合わせることをおもな狙いとした「感性検索」や「感性設計」などの研究が行われるようになった。ちなみに、人の“感性”や“好み”は、確かに存在し、人が扱う対象にとって大切な役割を果たしていることは分かっていたのではあるが、その“あいまいさ”あるいは人やときや場所や状況などによる“違い”つまりはデータとしての“不安定さ”のため、その扱い方がむずかしく、「感性情報処理」や「感性工学」以前の研究では、できるだけ取り除かれてきた。

しかし、「感性情報処理」や「感性工学」の課題である「感性検索」や「感性設計」などの研究では、モノやソフトウェアあるいはサービスに人の“感性”や“好み”を積極的に取り入れようという意図の下で、人の感性や好みに“合わせて”、対象の仕様を検索するあるいは設計するなどのための方法が模索・研究されている。これらの試みのなかで、人の感性や好みに“合わせる”ように、つまり、対象に対して目標とする「印象」が感じられるように、対象の仕様を選択しあるいは設計するなどのためには、その対象に対して目標とする印象を感じるかいなかを推定・評価をすることが必要で、このために、いろいろな形で、数学的な「印象の推定モデル」が構築されてきた。

ここでは、代表的なものとして、「感性情報処理」での研究と「印象の工学」での研究を簡単に紹介する。（「印象の工学」とは、筆者が提唱・研究している「『印象を扱う仕事』のための工学」で、そこで適用される方法や仕事のプロセスをすべて明らかにしその妥当性に検討を加えることによって、その成果の品質とプロセスの効率を改善・向上しよう、という意図を持った新しい工学で、「印象の推定」は、その重要な研究課題の1つである（文献2）。）

なお、「感性情報処理」や「印象の工学」の研究は、日本で提案され、現在までのところ、おもに日本で研究されており、欧米の研究者はその成果に関心を示してはいるが、まだ目立った研究成果は見当たらない。欧米では、“経験”によって、感性的な設計コンセプトを物理的な設計仕様などにブレークダウン（具体化）していく「感性設計」の方法は行なわれているものの、「感性検索」や「感性設計」などに関連した、数学的な「印象の推定モデル」の構築は、欧米ではほとんど行われていない、といってよい。

1-4-1 項 「感性情報処理」での研究

「感性情報処理」の1つである「感性工学」とは、「感性と工学とを結び付ける技術」のことであり、「人間の感性を分析しそれを商品の設計に取り込むことにより、人に喜びと満足をもたらす商品づくりを工学的に行なう分野」のことである（文献4～6）。「感性工学」を提唱した広島大学の長町三生は、彼の初期の研究の中で、希望する感性（つまり「印象」）が感じられるモノ（製品）のデザインを支援するエキスパート（専門家）システムを構築し、このなかで、モノのデザインの構成要素とそのモノのデザインに対する印象の対応づけ、つまり、数学的な「印象の推定モデル」を構築する方法を提案している。

この方法では、まず対象商品のデザイン（設計）を構成するアイテム／カテゴリーを抽出する。ここで、デザインのアイテムとは、「高さ」「幅（直径）」「カラー」など、デザイン上のある特性のユニットのことであり、カテゴリーとは、各アイテム内の小さな分類のことで、「高さ」などのような定量的な特性と「模様」などの定性的な特性が含まれている。

次に、この対象商品の印象を表す「感性ワード（語）」を抽出・整理する。感性ワードとは印象語のことで、ここでは形容詞を採用している。感性ワードは当初600個くらいを収集し、これを因子分析などの方法によって^{ふる}篩いにかけて、これと意味的な対応によって20～30対に縮小する。そして、この“縮小された”感性ワードを評価語にした「SD法」（semantic differential method）（文献7～12）を利用して、対象商品のデザインの印象を評価する。

ちなみに、「SD法」は、イリノイ大学のC・E・オズグッドが“情緒的意味”（affective meaning）を計測するために開発した方法で、現在は、印象を計測するための代表的な方法の1つとなっている。この方法では、評価対象（刺激）に対する印象を、対立する評価語（尺度または基準）の対の〇〇と××を用意して、回答者に「とても〇〇」から「とても××」までの間で評点してもらうことによって、印象を計測する。評価語の対は、複数用意される。また、評価語の対ではなく、1つの評価語〇〇について、「まったく〇〇」から「まったく〇〇でない」までの間で評価評点させる方法もある。「SD法」は、計測のためのガイダンスや質問文を目的に応じて丹念に作成しなくても、適切な評価語さえ用意すれば、簡単に実施できるのが最大の利点である。

対象商品のデザインを構成するアイテム／カテゴリーを説明変数に、感性ワードを評価語にした「SD法」による評価結果を被説明変数にして、「数量化理論1類」によって分析する。この結果、得られた対応づけが、数学的な「印象の推定モデル」となる。具体的には、感性ワードとアイテム／カテゴリーの間の偏相関係数の値が大きいものを“説明力が大きい”

対応づけ（関係）として拾い、これらを代表的な知識表現の1つであるプロダクションルールと同じ形式の「もし～ならば～」(if～then～)という形の「印象の推定モデル」として採用する。

長町は、このモデルを“逆”読みして（「ならば～」から「もし～」を推定して）、「SD法」による評価結果、つまり、感性ワードに対象商品のデザインを構成するアイテム／カテゴリーを対応づけることによって、たとえば“おしゃれな”ペンシルには「赤系の色で2色、樹脂で被覆された標準径、独立のクリップ、段差なしの蛇腹型」などのデザイン要素が対応づけられる、などと結論づけている。ちなみに、長町はこの方法を「感性工学手法2類」と呼び、「生活者が抱いている製品に対するイメージや感性、具体的な製品設計要素に翻訳する技術」と説明しているが、このことから考えると、長町は、「逆推定」を目的にした「印象の推定モデル」の構築を意図しているようである。

（ちなみに、長町が提案する「感性工学手法1類」とは、数学的なモデルを利用するのではなく、“経験的な”知見に基づいて、目標とする製品コンセプトあるいは印象を、人が次第に細かいコンセプトや印象に分解あるいは翻訳していき、徐々にコンセプトや印象をそれらを実現する具体的な製品の設計仕様とその組み合わせに落とし込んでいく、というやり方である。また、「感性工学手法3類」とは、数学的なモデルを“直接に”利用して、ある程度ではあるが“機械的に”目標とする製品コンセプトや印象を具体的な製品の設計仕様とその組み合わせに翻訳していく、というやり方である。）

長町は、この方法のデモンストレーションのための事例として、住宅設計システムの「H U L I S」(Human Living System)や女性のファッションデザインシステムの「F A I M S」(Fashion Image System)などを構築している。前者は、住宅を外観、外溝、玄関、和室、洋間、キッチン、風呂の7つの成分から構成されるものとして、この構成要素の順に、入力された感性ワード（つまり印象語）に対応した設計つまり“翻訳”をしていく。このシステムによって、従来のように、間取りから住宅設計をするのではなく、「こんな家に住みたい」という顧客の“感性”や“イメージ”に合わせて、住宅設計をすることができるのがポイントである。

また、後者は、女子大生のスーツとワンピースのデータベースを用意し、これらと感性ワードの対応づけを行なって、感性ワードに合った春夏秋冬の服装を提案するシステムである。このシステムは、女性の体型を身長、胸囲、腸骨稜角の3つのデータで表して（女子大生約400名を対象にした51項目の生体計測の結果、この3つのデータではほかのすべての計測値が再現できるそうである）、これらによって、それぞれの被験者を10種類の体型に分類して、それぞれの女性の体型に合わせて、また、「背が高くてかわいい」といったイメージ（つまり「印

象」)に合わせて、服装のデザインを提案し、部分的に気に入らないところがあれば、コンピューターグラフィックスの技術で修正・変更することができるのがポイントである。

長町の研究は、「感性工学」の目標を具体的に示し、また、その実現のための方法論を具体的にデモンストレーションしたという意味で、非常に大きな意義がある。しかし、長町自身が述べているように、この方法で得られた結果(対応づけ)は、必ずしも妥当なものばかりではないようである。

つまり、長町は、「数量化理論1類は、各アイテム/カテゴリーを独立性のあるものと考え、変数の線型和として構成しますが、感性そのものは非線型性という性質をもっているの

で、このような統計処理を行うことは適切ではありません。したがって、モザイク的にデザインの部品をまとめて商品表現することも適切とはいえません。」と述べており、この問題点を解決するため、「ファジィ数量化理論1類やニューラルネットワークモデルや遺伝的アルゴリズムを用いて、非線型性の性質をそのまま維持しながら処理することを試みています。」と述べている(文献4)。

1-3節でも述べたように、筆者は、「印象の推定モデル」の構築にいずれの方法論を適用するにしても、それが“パラメーター依存型”のものである場合には、その結果の妥当性の検証には、統計的な方法による信頼度の評価のほかに、なにかほかの方法あるいは知識や経験による検証が必要であり、1つの方法論だけで、“妥当な”(つまり“受け入れることができる”)対応づけが得られることは少ない、と考えている。あとで2-13節などでも述べるが、「統計的な方法」による結果の妥当性の検証あるいは結果の推定能力の評価は、分析する対象と推定する対象のデータが“同質”であることを前提とした方法であり、対象のデータのなかに“異質な”ものが含まれていた場合には、妥当性の検証あるいは結果の推定能力の評価はむずかしい。さらには、対象のデータで、なにが“同質”でなにが“異質”であるかの判断もむずかしいことである。

また、対象に対して感じる「印象」からその対象の「仕様や特徴」を推定する「逆推定」は、対象の仕様や特徴と対象に対して感じる(であろう)印象の間に、ごくごく“単調な”対応関係がなりたつ場合にのみ可能であって、一般には、もしできたとしても解空間のごく一部の対応づけしかできないので、この点でも問題があるのではないかと、とも考えている。また、(この方法が前提としている)モノのデザインの構成要素の抽出を“網羅的に”行うのは、その実行上かなりむずかしいことで、当然のことながら、抽出に漏れた構成要素に関しては推定ができない、という問題点も残る。(ちなみに、2章で説明する「判別分析」を利用した「印象の推定モデル」は、カラー画像の画像特徴量の入力に対して、印象の有無の出力を対応づけるた

め、入力としてのモノのデザイン要素を“網羅的に”抽出しなければならない、という問題点からは解放されている。)

このほか、いわゆる「感性検索」や「マルチメディア検索」などという名称で、対象の画像の特徴量あるいは仕様や特徴と、印象の大きさあるいは印象の有無を対応づけ、印象語に一致するあるいは類似するという意味で、対応する対象を検索するなどの試みは少なくない。以下、その代表的なものを挙げておく。

通商産業省・電子技術総合研究所の加藤俊一らは、「正準相関分析」などの方法によって、絵画のカラー画像としての特徴量（局所高次自己相関特徴（これについては2-13節で説明する）など）と“視覚的な”印象語の大きさ（程度）の関係を対応づけ、印象語からの絵画データベースの検索（加藤らはこのシステムを「電子美術館」と呼んでいる）を試みている。この対応づけは、「印象派」の絵画に限定されたものであるが、これによって、ある程度ではあるが、“類似画”の検索や“未知”の絵画の検索も可能になった、と報告している（文献14、15）。

これに類似したものとして、NTTの木本晴夫は、画像の配色パターン（色相、彩度、明度の割合）と感性語（印象語）を対応づけ、感性語による花やペイントブラシ画像の検索を試みて、その結果が有効であった、と報告している（文献16）。立命館大学の八村広三郎らは、「領域分割法」によって、画像の画面構成を表す情報を抽出し、これとこれに対応する印象語を対応づけることによって、印象語による“静物画”の画像検索を試み、試みは有効であった、と報告している（文献17）。

また、日立製作所の北原義典らは、マルチメディアプレゼンテーションのコンテンツの色彩およびバックグラウンド音楽と（雰囲気やイメージなどを表す）形容詞（印象語）とを対応づけ、マルチメディアプレゼンテーションのコンテンツの作成を支援するシステムを構築して、対応づけは有効であった、と報告している（文献18）。これらの試みは、いずれも機械的に計算された入力データ（画像の特徴量）と「SD法」などで計測された印象の大きさの間の対応づけを行なったもので、その方法は統計的な分析によるルールの生成によるものである。

また、静岡大学の原田将治らは、椅子のカラー画像を対象に、画像解析から得られる特徴量（たとえば椅子の面積の割合、輪郭線の長さなど）と客観的な基準から得られる椅子の形状的な特徴量（たとえば肘掛けの有無、脚の数など）を合わせた28個の特徴量と、これに対して「SD法」で計測された印象の大きさを「重回帰分析」で対応づけ、この対応づけを利用して、印象による椅子のカラー画像の検索システムを試作・評価して、この結果が充分

に有効であった（124個のデータのうちの62個のデータに対して得られた重回帰式が、残りの62個のデータに対しても十分に有効であったそうである。）、と報告している（文献19）。

さらに、慶応義塾大学の萩原将文は、「ファジィルール」の方法を用いて、カラーポスターの配色とフォントとその「見やすさ」の関係を対応づけ、また、「遺伝的アルゴリズム」を利用して、カラーポスターの配色とフォントの変更案を生成することによって、ユーザーの感性を反映させたカラーポスターの作成を支援するシステムを構築し、その有効性を確認した、と報告している（文献20、21）。

そして、関西大学の野村康雄らは、「ニューラルネットワーク」による「非線型判別分析」を利用して、音楽の楽譜から抽出した特徴と印象の大きさを対応づけ、人が感じる印象による楽曲の分類を試み、この対応づけによって、“未知”の楽曲でも高い精度で分類できた、と報告している（文献22）。「ニューラルネットワーク」を利用した試みとしては、ほかに、鳥取大学の白木渡らが、橋梁の景観評価に関して、橋梁美を構成する4つの要素（形式美、サイコベクトル、環境との調和、環境との色調和）と、3段階評価をした橋梁美（個性美、好感度、調和美、優美さ、存在感）の関係を“不十分ながら”ニューラルネットワークで対応づけることができた、と報告している（文献23）。

「統計分析によるルールの生成」「重回帰分析」「遺伝的アルゴリズム」「ニューラルネットワーク」などを利用したこれらの「対応づけ」つまり「モデル化」の試みは、いずれも（1-3節で説明した）“パラメーター依存型”の方法によるもので、得られた対応づけを、（分析の対象には含まれなかった）“未知”の対象のデータに適用するためには、この対応づけがなにかほかの方法あるいは知識や経験などで裏付けられているなどの合理的な根拠に乏しい、という共通の問題点を持っている。

あとで2-13節などでも述べるが、対象とするデータの（全部ではなく）ある部分について分析をして得られた解が残りのデータに対して適用できた、というのは、単にそのデータが（ほかのデータから離れている、という意味で）“特異な”つまり“異質な”データを含んでいなかった、あるいは、データが“同質”であった、ということの意味するに過ぎない。したがって、あるデータについて分析をして得られた解が、“未知”の対象のデータに対して有効であると主張するためには、そのデータが分析をしたデータと“同質な”ものであることの根拠を示すか、あるいは、分析で得られた対応づけが、“未知”のデータに対しても有効であることの根拠を、分析に採用した方法とは別のなにかの方法で示すことが必要なのである。

そういった意味で、これらの「対応づけ」つまり「モデル化」は、確かに統計的な意味で

(ある範囲内では) 確かな結果であると主張することもできようが、残念ながら、これらの結果を“未知”の対象に対して適用できると“積極的に”主張するだけ合理的な根拠を持っていない、というのが本当のところであろう。したがって、これらの試みは、適用されている方法論の技術的な意味での面白さを別にすれば、基本的に、長町が提案した方法を超えてはおらず、これと大きな違いはないといってよい。また、同様の意味で、数学的な「印象の推定モデル」の構築の方法には、まだ“標準的な”ものといったものが見いだされていない、というのが現状といってよい、と筆者は考えている。

筆者は、1-2節と1-3節で、「印象を扱う仕事」の“機能的な”意味での要素的な仕事の1つである「印象の計測」、つまり、「仕事の目的の達成のために、その対象に関して、必要かつ十分な情報を得る」ための仕事の注意点として、「G I G Oにならないように!」を挙げた。以上に説明したいろいろで試みでは、「印象の推定」に先立つ「印象の計測」の方法として、おもに「S D法」が採用されているが、この方法によって、「仕事の目的の達成に必要なかつ十分な情報を得る」ことができている、というのがこれらの試みの結果が有効である、最大の前提条件であろう。

この立場から考えると、「感性情報処理」の分野で行なわれている「印象の推定モデル」の構築の試みでは、必ずしもこういった検討が充分に行なわれていないように思われる。文献7~12は、印象の測定法に関するものであり、「印象の計測」を行なうにあたっては、これらを参照し、実際に適用する方法の性格によって、「仕事の目的に必要なかつ十分な情報を得る」ことができるものかを検討しておくことが大切であろう。

ちなみに、「S D法」の問題点として、金沢工業大学の神宮英夫は文献12で、意味尺度の選定が適切かどうか、被験者の意味尺度への評定が両極の間に連続的变化を示すかどうか、被験者が複数の意味尺度に対して同程度の評定ができるかどうか、共通の因子構造が存在するかどうか、などを挙げている。また、以下は、筆者が文献13で整理した「S D法」の問題点である。「印象の計測」における必要な問題点の例の1つとして、これらの問題点をクリアして、「仕事の目的の達成に必要なかつ十分な情報を得る」ことができているかを、充分に検討しておく必要がある。

第1に、評価語が示されない側面については、回答がないので、評価つまり計測ができない。この問題点の解決のためには、評価したいすべての側面について、評価語を用意しておくことが考えられるが、このようにして、評価語を増やすと、今度は回答者(被験者)の負担が大きくなり、結果的に信頼できる回答が得られなくなってしまう。さらに、可能性のあるすべての側面について詳細に検討して、厳選した評価語を用意することも考えられるが、あらかじめ結果のすべてのスペクトラムが予測でき、詳細な検討ができれば、そもそも計測の必要がない、ということもある。

第2は、SD法による計測には、評価語の対を使うことが多いが、たとえば「きれいな」の対語は「汚い」「見にくい」「きれいでない」のいずれか、といった問題もある。評価語つまり言葉には、複数の意味が含まれているのが普通であり、対語を選択することで、元の「きれいな」の意味が決まってくることもあるが、回答者が計測者（調査者）の意図するように、評価語の意味を理解してくれるかどうか問題であろう。「きれいでない」と「きれいではない」のように、表現のしかたのわずかな違いで、評価語の意味やニュアンスが違ってくる、ということもある。

第3は、評価語には、形容詞や形容動詞が使われる場合が多い。しかし、例として、モノ（製品）のデザインを印象を考えると、「重そうな」「高級そうな」「しっかりした」などの形容詞的な印象のほかに、「学校の先生が持つような」「パーティー向き」「大事な書類が入っていそうな」など、人とモノのかかわり合い（いわゆる「5W1H」）に関連した、形容詞だけでは表せない印象も多い。日本語には、形容詞や形容動詞が少なく、このため、漢語や外来語の△△に「な／だ」を付けて、「△△な／だ」という形の形容動詞が作られてきたが、これを含めても形容詞はせいぜい8000～1万語しかない。

第4として、日常使っている言葉は、たとえば「きれい」と「美しい」、「大きい」と「重い」、「強い」と「硬い」などのように、意味的あるいはイメージ的に、なにかの関わりがあるのが普通で、互いにまったく独立ということは少ない。しかし、SD法では、原則として、評価語は、互いに独立であるという立場が取られ、こういった関係は無視されてしまうことが多い。

第5に、SD法では、回答者には、示された評価語には、必ず回答しなければならない。しかし、現実には、評価すべき対象に対しては、示された評価語は関係がない、といった場合もある。こういった場合にも、回答者は、必ず回答しなければならないので、結局、よくわからないという理由で、「どちらでもない」に付けることになるが、これは示された評価語の対の中間の意味である「どちらともいえない」とは、まったく違った内容であろう。「無回答」という選択ができれば、少しく事情は違って来ようが、この区別ができないことも、大きな問題点の1つであろう。

このほか、SD法の教科書では、評価語の選択で、「適切な反意語のないものは避ける」「あいまいな言葉は避ける」などといったことが勧められているが、評価語の選択やガイダンスが恣意的ならば、結果もまた恣意的であることを忘れてはならない。また、SD法に限ったことではないが、評価語つまり言葉という表現手段だけで、印象の評価ができるか、といった本質的な問題もある。

1-4-2 項 「印象の工学」での研究

「印象の工学」とは、筆者が提唱・研究している「『印象を扱う仕事』のための工学」で、私たちが“心理としてまた主観として”感じるまた持っている「印象」を専門的に扱ってきた「心理学」の知識や経験を利用して、そこで適用される方法や仕事のプロセスを明らかにしその妥当性や合理性に検討を加えることによって、その成果の品質とプロセスの効率を上げる（改善・向上する）ことを達成しよう、という意図を持った新しい工学である（文献2）。具体的には、この目的に沿って、「印象を扱う仕事」の“機能的”あるいは“プロセス的な”意味での要素的な仕事である「印象の計測」「印象の分析」「印象の推定」「印象の計画・設計」「印象の管理」「仕事のプロセス管理」などの要素的な技術を研究開発しており、「印象の推定」は、その重要な“機能的な”意味での要素的な技術の1つである。

筆者は、この研究の一環として、「デザイン支援システム」の研究の中で、対象の仕様や特徴など（離散量）と“希望する”印象（離散量）との間のより確かな（これは“構造的な”あるいは“安定的な”といってもよい）対応づけ、つまり、数学的な「印象の推定モデル」を見つけ出し構築するため、「同時関係分析法」と呼ぶ方法を開発した（文献25、24）。

さて、この方法は、モノ（製品）や空間のデザインに対する印象を例にとると、最初の対象のデザインの構成要素を整理するところは、（1-4-1項で説明した）「感性工学」での長町らの方法とほぼ同じである。したがって、これに関連した問題点もほぼ同様である。それぞれの製品や空間のデザインに対する印象は、「自由申告法」という方法（図1-4-1）（文献13、24）によって調査をする。

「自由申告法」の調査シートの中心部には、対象（刺激）の写真（カラー画像）が貼ってある。この調査シートを利用して、被験者（回答者）にいくつかの対象の写真（最小のガイドンスで）見せて、まずはその仕様や特徴を回答の表現のしかたにとくに制限を加えない“自由な”表現形式で申告して（つまり書いて）もらう。そして、このプロセスで対象（刺激）を十分に観察してもらったあとに、同じ対象に対する印象を、これもまた特別な制限を加えない“自由な”表現形式で申告してもらう。調査のやり方にもよるが、通常は、1つの対象に対して6～7語の説明語と5～6語の印象語が申告される。その結果は、自然言語処理の技術を利用して構築した「電子シソーラス」（同義語・類義語辞書）を使い、また、コンピューターを“支援的に”あるいは“半自動的に”使って、説明語と印象語の意味の解釈をしてコー

ド化する。

ちなみに、「SD法」では、回答にガイダンスが設けられているので、ガイダンスに挙げた側面については“漏れなく”回答を得ることができる。しかし、当然のことながら、ガイダンスに挙げられていない側面については回答が得られない。これに対して、「自由申告法」では、基本的には“最少”のガイダンスしか設けないために、原則として、回答者が“意識した”仕様や特徴や“感じた”印象はすべて申告される。しかし、たまたまそれに気がつかなかった、適当な言葉が浮かばなかった、それはいうまでもないことであると考えた、あるいは、すでに類似のことをいってしまったので敢えて述べなかったなどの理由で、申告されないものもある（文献13）。こういった問題点を防ぐために、事前に、計測の目的をしっかりと説明したり、計測に使うほかの刺激を見せたり、ということをすることもある。

このカバンを説明してください

■このカバンで気がついた特徴をできるだけ多く、下の()内に簡単に書いてください。
このカバンを見ていない人に説明しようとするときなどに用いる言葉です。

1 (革製)	2 (茶色の色)
3 (カギ付)	4 (高級感がある)
5 (書類入れ)	6 (あまり物が多い)
7 (やや重い)	8 (シンプル)
9 (スリッパ型)	10 (固い)
11 (持ち手が革)	12 (持ち手が曲線)
13 (タック付)	14 (横の仕切りがある)
15 ()	16 ()



このカバンからどんな印象を受けますか

■このカバンから受けたあなたの印象をできるだけ多く、下の()内に簡単に書いて下さい。

1 (シンプル)	2 (高級感)
3 (色が目立つ)	4 (シンプルで使いやすい)
5 (書類入れ)	6 (持ち手が少ない)
7 (重さ)	8 (モダン)
9 (丈夫)	10 (若々しい)
11 (持ち手が目立つ)	12 (持ち手が太い)
13 (ランゲルなデザイン)	14 ()
15 ()	16 ()

右の欄には何も記入しないで下さい

実験番号 _____ 写真番号 _____

図 1-4-1 「自由申告法」の調査シートの例

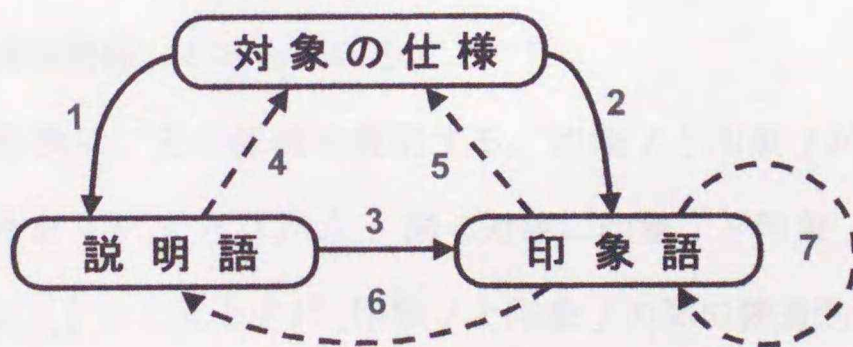


図 1-4-2 「印象の推定モデル」の概念の1つ

そして、この調査（計測）で、後述する“ある”仮説をもとにして、図 1-4-2 に示す枠組みによって対応関係を見つけ出そうとするのである。

この図の中で、実線の 1 は「対象を見て（あるいは対象の仕様や特徴を聞いて）、説明できる」、2 は「対象を見て（あるいは対象の仕様や特徴を聞いて）、印象を感じる」、3 は「対象の説明を聞いて、印象を感じる」、破線の 4 は「対象の説明を聞いて、対象（の仕様や特徴）を想像する」、5 は「対象に対する印象を聞いて、対象（の仕様や特徴）を想像する」、6 は「対象に対する印象を聞いて、説明をする」、7 は「対象に対する印象を聞いて、ほかに“同時に”感じる印象を想像する」を意味する。

実線の 1、2、3 は、“前向き”つまり順方向の推定であり、「説明」あるいは「印象」といえる。破線の 4、5、6 は、“後向き”つまり逆方向の推定つまり「逆推定」であり、あるものは「想像」、あるものは「創造」といえるかもしれない。（1－3 節でも説明したように、「逆推定」はいわゆる「逆問題」であり、一般的には解くことができないのであるが、ここでは、調査データの範囲内で“機械的に”対応づけている。）破線の 7 は、“同レベル”の「印象」つまり「共感覚」の推定である。

このモデルの枠組みにしたがって、「対象の仕様」「説明」「印象」の間で相互に翻訳することができる“確かな”対応関係を見つけだせば、この対応関係を「印象の推定モデル」として使うことができる。「同時関係分析法」は対応関係を見つけ出すための方法で、ある仮説の下で「対象の仕様」「説明」「印象」の間に“隠れている”構造的な対応関係を見つけ出すことができる。前にも述べたように、当然のことながら、「対象の仕様」「説明」「印象」の間の対応関係が“構造依存的”であるか“パラメーター依存的”であるかは、相対的なものである。ここで、「構造的」といっているのは、1－4－1 項で述べたいままでに試みられた対象とそれに対する印象の間の「対応関係」に比較して、ということである。

この仮説とは、「“同じ”回答者が“同じ”対象に対して“同時に”した回答の間には、構造的な対応関係がある可能性が高い」というものである。そして、これらの“隠れている”構造的な対応関係を「同時関係」と呼んでいる。

印象同士の対応関係を例に、その原理を説明する。印象 i と印象 j が申告された対象がそれぞれ全体に閉める割合を $[p_i]$ と $[p_j]$ 、同じ対象に印象 i と印象 j の両方が同時に申告されたものの割合を $[p_{ij}]$ としたときに、印象 i と印象 j の間の構造的な対応関係によって、印象 i が申告されると同時に印象 j も申告される」、つまり、「 i ならば j である」（ $i \rightarrow j$ ）

ことが確かである程度 $[r_{i \rightarrow j}]$ は、(1)式で表すことができる(図1-4-3)。

$$(1) \quad r_{i \rightarrow j} = \max \left(\frac{p_{ij}}{p_i} - p_j, 0 \right) \quad (i \neq j \text{ のとき})$$

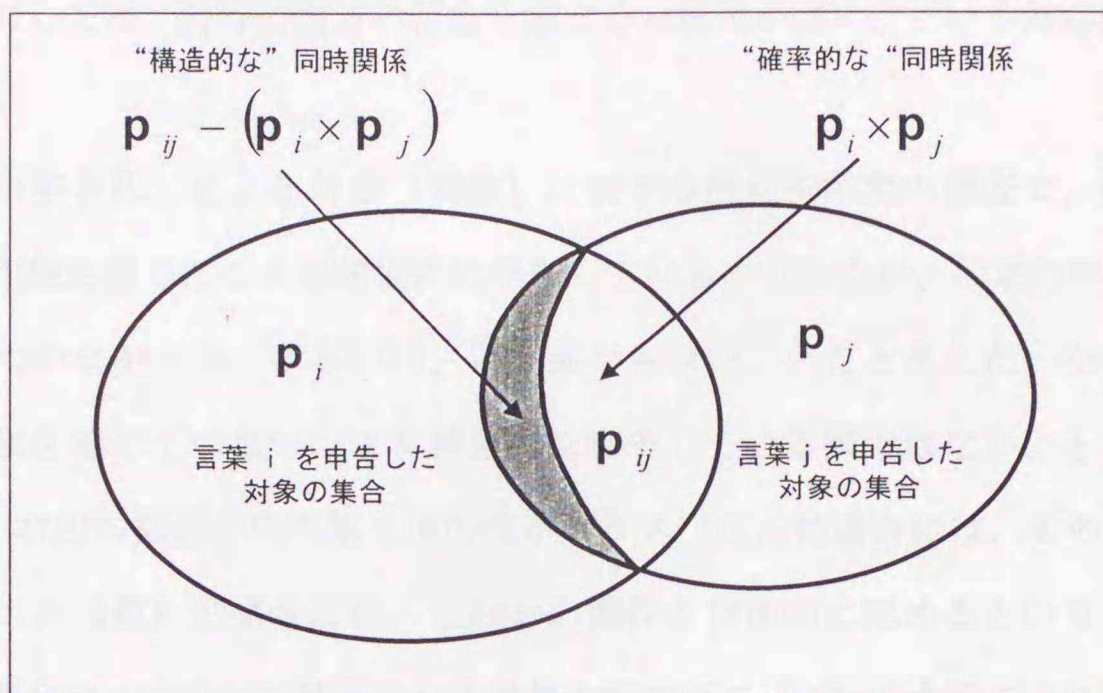


図1-4-3 「同時関係」の概念のベン図による説明

ここで、(1)式の右辺の括弧の中の第1項の前半の $[p_{ij}/p_i]$ は、印象 i が申告されたという条件の下での印象 j の申告される割合であり、第1項の後半の $[p_j]$ は、こういった条件なしに、印象 j が申告される全体的な割合である。印象 i と印象 j の間に“隠れている”「同時関係」の確かさの程度は、 i という条件付きの下での j の申告された割合から条件なしの場合の割合を引くことによって求められる、ということである。

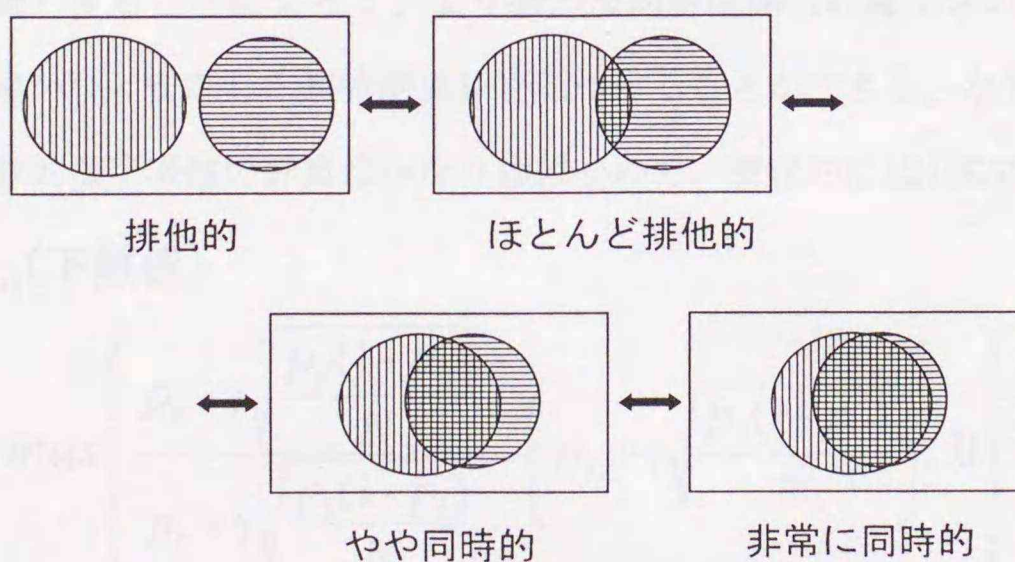


図1-4-4 「同時関係」の確かさの違い

図 1-4-4 で、印象 i を感じる対象の集団と印象 j を感じる対象の集団の間でまったく「重なり」がない場合には、2つの印象は互いに“排他的”である。また、「重なり」が少しくあっても“ほとんど排他的”である。これに対して、印象 i を感じる対象の集団と印象 j を感じる対象の集団の間で「重なり」が大きいときには、2つの印象は“同時的”つまり一方の印象を感じる対象に対しては、同時に他方の印象を感じる可能性が高い、という対応関係が存在する、と考えられる。

なお、「自由申告法」による対象（刺激）に対する説明や印象の調査で、印象の申告“あり”は、その印象を感じたことを積極的に申告しているのであるが、印象の申告“なし”は、たまたま気がつかなかった、あるいは、いうまでもないことだと考えた、などの場合があるので、その印象を感じていないことを積極的に申告している訳ではない、と考えられる。そこで、(1)式の右辺の括弧の中の第1項の値がプラス（正）の場合には、そのままよいが、この値がマイナス（負）の場合には、この対応関係を積極的に認めるというわけにはいかないので、この場合は、この式で計算される結果の値をゼロ（0）とすることになっている。

ちなみに、(1)式の値 $[r_{i \rightarrow j}]$ は、 i ならば j であることがどのくらい確かであることを示す値で、プロダクションルールの確信度（certainty factor, CF）に相当する値である。相対的な意味で、値が大きいほどこの対応関係は確かであり、値が小さいほどこの対応関係は確かではない、と考えることができる。

また、(1)式は、「同時関係」の確かさの程度の最尤値であるので、実際には、調査した対象の全体の数を $[N]$ 、統計的な推定の過誤の危険率に対する標準正規得点を $[\gamma]$ として、「 i ならば j である」ことが確かである割合の統計的な下限値（実際にはその近似値）を、(2)式で計算（評価）することによって、より確かな同時関係（の確かさの程度を評価し、その結果として、確かな／安定した同時関係）を見つけることができる。なお、(1)式で表される $[r_{i \rightarrow j}]$ の統計的な下限値の計算はかなり複雑なので、便宜的に(2)式で代用している。

(2) $r_{i \rightarrow j}$ (下限値)

$$\doteq \max \left(\frac{p_{ij} - \gamma \sqrt{\frac{p_{ij}(1-p_{ij})}{N}}}{p_i + \gamma \sqrt{\frac{p_i(1-p_i)}{N}}} - \left(p_j + \gamma \sqrt{\frac{p_j(1-p_j)}{N}} \right), 0 \right) \quad (i \neq j \text{ のとき})$$

さらに、二重の同時関係、つまり、「 i かつ j ならば k である」((3)式と図 1-4-5)、あ

るいは、条件付きの同時関係、つまり、ある特定の条件Aの下で*i*ならば*j*である」((4)式と図 1-4-6) などのように、条件項や結果項をより多く重ねたり条件を付けたりすることによって、これに対応するより確かな同時関係を見つけ出すこともできる。

$$(3) \quad r_{ij \rightarrow k} = \max \left(\frac{p_{ijk}}{p_{ij}} - p_k, 0 \right) \quad (i \neq j \neq k \neq i \text{ のとき})$$

$$(4) \quad r_{i \rightarrow j} = \max_{(i,j \in A)} \left(\frac{p_{ij}}{p_i} - p_j, 0 \right) \quad (i \neq j \text{ のとき})$$

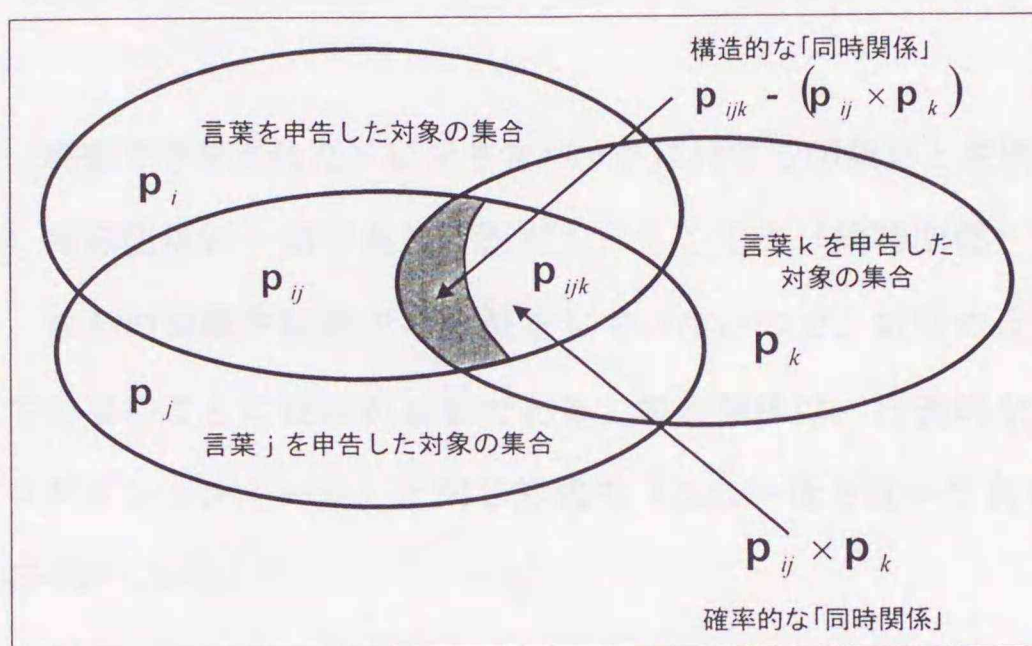


図 1-4-5 「二重の同時関係」の概念のベン図による説明

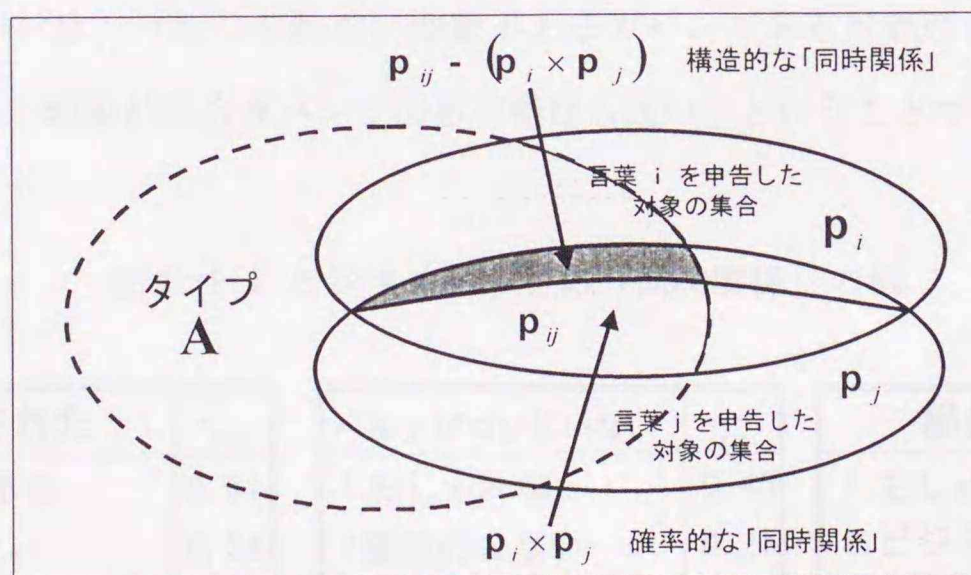


図 1-4-6 「条件付きの同時関係」の概念のベン図による説明

また、便宜的に、分析者が抽出した対象の仕様（カテゴリー／アイテム）や特徴の有無（離散量）を、「印象の計測」で回答者が申告した回答（説明語や印象語）の有無（離散量）と同

様の性格のデータとみなすことによって、印象同士の関係の分析と同様に、対象の仕様と印象の申告の間に“隠れている”（とみなすことができる）対応関係を見つけ出すこともできる。この場合は、これらの関係は、「特定の仕様ならば特定の印象を感じるで（あろう）」という「印象の推定モデル」として使うことができる。

例として、200枚のビジネスバッグ（カバン）のカラー画像（写真）を刺激に、「自由申告法」の方法と手続きにしたがって、それぞれ8名ずつの回答者（被験者）の回答を調査した結果をこの方法で分析した。回答者の数は、延べでは1600名であったが、個々の回答者には5枚の対象に対して回答をしてもらったので、実際には320名の回答者の協力を得た。

表1-4-1は、（調査で申告された）ビジネスバッグに対する印象語と説明語の間で見つかった“確からしい”対応関係の一部である。ただし、ここでの「同時関係」の確かさの値には最尤値を採用し、ほかの知識や経験での検証をしていないので、結果の全部が全部納得できるものにはなっていないことに注意が必要である。対応関係は、代表的な知識表現の方法の1つである「プロダクションルール」と同じ形式の「もし～ならば～である」（if～then～）という形式で対応づけている。

この中で、たとえば「都会的な」印象があるカバンであれば、そのカバンは、「おしゃれな」印象があるカバンである可能性が高く、また、それは「ビジネスマンの」という印象もあるカバンである可能性も高い、ということである。あるいは、「ベルトがついていない」と説明されたカバンは、「おしゃれな」印象があるカバンである可能性が高く、また、それは「飽きのこない」印象があるカバンである可能性も高い、ということである。

表 1-4-1 ビジネスバッグの「同時関係」の例

洗練された	$r_{i \rightarrow j}$	ベルトがついていない	$r_{i \rightarrow j}$	都会的な	$r_{i \rightarrow j}$
1. 把手のある	0.51	1. おしゃれな	0.45	1. おしゃれな	0.32
2. グリーン	0.34	2. 飽きのこない	0.21	2. ビジネスマンの	0.23
3. 薄い	0.30	3. シンプルな	0.20	3. ブランド物の	0.20
4. 書類の入る	0.27	4. 丈夫そう	0.17	4. シンプルな	0.20
5. ベルトのついた	0.24	5. 都会的	0.17	5. 高級そうな	0.20
6. グレー	0.19	6. 出し入れが面倒そう	0.17	6. 高そうな	0.19
印象語 → 説明語		説明語 → 印象語		印象語 → 印象語	

表 1-4-2 ビジネスバッグの「二重の同時関係」の例

カジュアルな(a) × たくさん入りそうな(b)				重そうな(a) × 高そうな(b)			
確かさの程度	$r_{ab \rightarrow j}$	$r_{a \rightarrow j}$	$r_{b \rightarrow j}$	確かさの程度	$r_{ab \rightarrow j}$	$r_{a \rightarrow j}$	$r_{b \rightarrow j}$
1. 学生っぽい	0.15	0.07	0.04	1. 革製の	0.15	0.07	0.04
2. 若者向け	0.14	0.10	0.00	2. 茶色の	0.14	0.10	0.00
3. 便利そうな	0.11	0.05	0.01	3. カギつき	0.11	0.05	0.01
4. 使いやすいそうな	0.10	0.06	0.00	4. 止め金がある	0.10	0.06	0.00
5. 実用的な	0.10	0.02	0.23	5. アタッシュケース	0.10	0.02	0.23

印象語 × 印象語 → 印象語

印象語 × 印象語 → 説明語

同様に、「二重の同時関係」を利用すると、たとえば「カジュアルな」印象があるカバンでありかつ「沢山はいりそうな」印象があるカバンであれば、それは「学生っぽい」印象があるカバンである可能性が高く、また、それは「若者向けの」という印象があるカバンである可能性が高い。これらの分析結果は十分に納得できるもの、といえる。

さらには、「条件付きの同時関係」を利用して、「アタッシュケース」「ブリーフケース」「クラッチバッグ」「ダレスバッグ」「ショルダーバッグ」というカバンの種類（タイプ）ごとに、それぞれの同時関係の確かさの程度を評価してみると、表 1-4-3 の結果が得られた。このなかで、「丈夫そうな」という印象があるカバンは「重そうな」という印象があるカバンである可能性が高い、という対応関係は、カバンの種類には関係なくなりつつあることがわかる。また、「固そうな」という印象がある可能性が高いこともわかる。これに対して、「古めかしい」という印象があるカバンならば「おしゃれな」「高級そうな」「汚れやすい」「重そうな」「年配者向けの」という印象があるカバンである、という対応関係は、カバンの種類が違ふと必ずしもなりたないことがわかる。

表 1-4-3 ビジネスバッグの「条件付きの同時関係」の例

アタッシュケース	ブリーフケース	クラッチバッグ	ダレスバッグ	ショルダーバッグ
丈夫そうな →	丈夫そうな →	丈夫そうな →	丈夫そうな →	丈夫そうな →
1. 重そうな	1. 重そうな	1. 重そうな	1. 重そうな	1. 重そうな
2. 固そうな	2. 固そうな	2. 高級そうな	2. 固そうな	2. 固そうな
3. デザインが良い	3. 長持ちしそう	3. 男性用の	3. 長持ちしそう	3. 便利そう
4. シンプルな	4. オソドックスな	4. シンプルな	4. たくさん入りそう	4. 使いやすいそう
ビジネスマンの →	ビジネスマンの →	ビジネスマンの →	ビジネスマンの →	ビジネスマンの →
1. 固そうな	1. 固そうな	1. 固そうな	1. 固そうな	1. 固そうな
2. 男性用の	2. 男性用の	2. 男性用の	2. 男性用の	2. 丈夫そう
3. 高級そう	3. 高級そう	3. 安っぽい	3. おしゃれな	3. 出張用の
4. 仕事用の	4. デザインが良い	4. あまり入らなさ	4. 新しい感覚の	4. たくさん入りそう
古めかしい →	古めかしい →	古めかしい →	古めかしい →	古めかしい →
1. おしゃれな	1. 高級そう	1. 汚れやすい	1. 重そう	1. 年配者向け
2. 丈夫そう	2. 固そう	2. 固そう	2. 出張用の	2. 高そう
3. 持ちやすそう	3. 丈夫そう	3. 地味な	3. 旅行用の	3. 重そう
4. 通勤用の	4. 大人用の	4. シンプルな	4. たくさん入りそう	4. 男性用の
機能的な →	機能的な →	機能的な →	機能的な →	機能的な →
1. たくさん入りそう	1. シンプルな	1. 華やかな	1. 長持ちしそう	1. 地味な
2. ビジネスマンの	2. 男性用の	2. 通勤用の	2. 出張用の	2. センスの悪い
3. 出張用の	3. 使いやすい	3. 高級そう	3. 学生っぽい	3. 重そう
4. 仕事用の	4. デザインが良い	4. 飽きのこない	4. 柔らかそう	4. 使いやすい

いずれにしても、いろいろな「同時関係」の分析を行って、その結果が妥当であるかいかを、ほかの知識や経験あるいはその確認のための調査などによって検討し、妥当で受け入れられるものを“確かな”対応関係として受け入れて、「印象の推定モデル」として使おう、というのが「同時関係分析法」の基本的なやり方である。

ちなみに、この方法の検証の1つとして、この方法で見つけ出された（単純なおよび二重の）対応関係を提示して、これが受け入れられるものかいかについて、一般の回答者に対してアンケート調査を行った。結果は、同時関係の分析で“確か”とされた関係が“受け入れられ”、分析で“そうではない”とされた関係は“否定された”という意味で、概ね（90%以上の）肯定的な結果が得られた。しかし、なかには、同時関係の分析では“確か”とされたものが“受け入れられない”、あるいは、分析で“そうではない”とされたものが“受け入れられる”とされたものもいくつか（10%以下）見られた。したがって、この方法で見つけ出された対応関係を意味のある対応関係として“機械的に”採用することはできず、なにかの方法や知識・経験あるいはなにかの別の調査によって、その妥当性を検証してから適用をする、という手続き（プロセス）が必要であることをあらためて確認した（文献26）。

以上の説明では、対象に対して感じる「印象の計測」には、「自由申告法」の適用を前提としていた。しかし、たとえば、対象 k に対して、印象 i と印象 j を感じる程度（大きさ）をそれぞれ $[a_i]$ $[a_j]$ とし、印象 i と印象 j の申告の割合をそれぞれ $[p_i]$ $[p_j]$ 、印

象 i と印象 j の“同時”申告の割合を $[p_{ij}]$ とすると、(5)式と(6)式のように解釈することによって、印象の“有無”ではなく、印象の“程度”（大きさ）を前提とした「印象の計測」の方法、たとえば「SD法」などにも適用することができる。

$$(5) \quad p_i = \frac{\sum_{\forall k} a_{ik}}{m} \quad p_j = \frac{\sum_{\forall k} a_{jk}}{m} \quad p_{ij} = \frac{\sum_{\forall k} a_{ik} a_{jk}}{m}$$

$$(6) \quad r_{i \rightarrow j} = \frac{p_{ij}}{p_i} - p_j \quad (i \neq j \text{ のとき})$$

「同時関係分析法」の問題点を挙げれば、この方法は、基本的に、対象については、その仕様や特徴などを、専門家（エキスパート）が整理しなければならない点で、これは分析に大きな手間が必要ということになる。また、仕様や特徴などは“網羅的に”挙げることであればよいが、この整理に漏れてしまった場合には、分析の結果の有効性は、整理に入った仕様と特徴などの範囲に限られる、ということになる。

なお、この方法は、大量のデータの中から、未知の、価値ある関係を見つけ出す「データマイニング」にも適用できる可能性があり、テスト的に、辞書や新聞に含まれているキーワードの間の関連性に適用することができている（文献27、28）。

また、「プロダクトデザイン」の過程で、デザイナーがイメージやコンセプトを構想したときに、そのイメージやコンセプトに該当するモノや空間のデザインにはどんなものがあるかを参考情報として検索する「コンテンツと印象による画像検索システム」などに実用化している。具体的には、モノや空間のカラー画像には、その内容（コンテンツ）と印象が登録されており、コンテンツやイメージ／コンセプトが入力されると、それに直接該当するもののほか、コンテンツの場合には、たとえば「OA機器」を「パソコン」「ワープロ」「ファックス」「電話機」などと解釈したり、あるいは、イメージやコンセプトの場合には、同時関係分析法で見つけ出された対応関係（つまりモデル）で、ほかのイメージ／コンセプトを推定して、これに該当する対象のカラー画像を検索してくる（文献29）。

2章 遺伝的アルゴリズムを利用した計算的な「判別分析」の方法

この章では、数学的な「印象の推定モデル」を構築するための確実で有力な方法論として、「判別分析」を取り上げる。ここでいう数学的な「印象の推定モデル」とは、対象のカラー画像の特徴量（連続量）とその対象に対する特定の印象の有無（離散量）の「対応づけ」のことである。

さて、「判別分析」には、すでに「解析的な方法」が開発されており広く普及している。しかし、この「解析的な方法」は、ほかの多変量解析の方法と同様に、たとえば変量の値の連続性やその分布の正規性、変量相互の間の独立性、目的関数の微分可能性など、いろいろな仮説や条件を前提として構築されている。しかし、私たちが実際に分析したい問題とデータは、この「解析的な方法」が前提としている仮説や条件を満たしていない場合もあり、この場合には、「解析的な方法」で求められた解（判別式と判別値）は実際の問題とデータに対しては必ずしも有効ではない、といった可能性が考えられる。

この章では、まず「判別分析の目的」を中心に「解析的な方法」の問題点を検討し、その解決のために、実際の問題に対してより“有効な”あるいはより“判別力の高い”解を求めることを目的に新たに開発した「計算的な判別分析」の方法と計算アルゴリズムを説明する。この「計算的な方法」は、実際の問題が要求する「判別分析の目的」などにしたがって、「遺伝的アルゴリズム」を利用して“計算的に”解を見つけだすもので、対応づける対象の属性値の分布や性質などには特別な仮説や条件は必要なく、また、その対応づけには非線型性や交互作用などを含んだ、かなり複雑な問題にも対応できるのが特徴である（文献30）。

2-1節 「印象の推定モデル」としての「判別分析」

たとえば対象のカラー画像の局所自己相関特徴（これについては2-1-3節で説明する）など、その対象のカラー画像を代表する特徴量（連続量）とその対象のカラー画像に対して申告された特定の印象の有無（離散量）の「対応づけ」つまり「モデル」ができれば、“任意”の対象のカラー画像に対しても、その特徴量を計算してその「モデル」に入力することによって、その対象のカラー画像に対する特定の印象の有無、つまり、その対象のカラー画像に対して（平均的な）回答者が特定の印象を感じるかいないかを推定することができる（図2-1-1）。「判別分析」は、この「対応づけ」つまり「印象の推定モデル」の構築のための有力な方法の1つである（文献31、32）。

ちなみに、1-3節でも述べたが、ここで取り上げる「印象の推定モデル」は、比較的にポピュラーな対象に対する“平均的な”回答者が感じる印象を推定するモデルである。回答者は、その対象に対してある程度の知識や経験を持っており、その対象に対して、ある程度は“絶対的な”印象を感じるものが想定されている。したがって、この「印象の推定モデル」は、“特別な”知識や経験を必要とする“特別な”印象は対象にしていないことに注意が必要である。

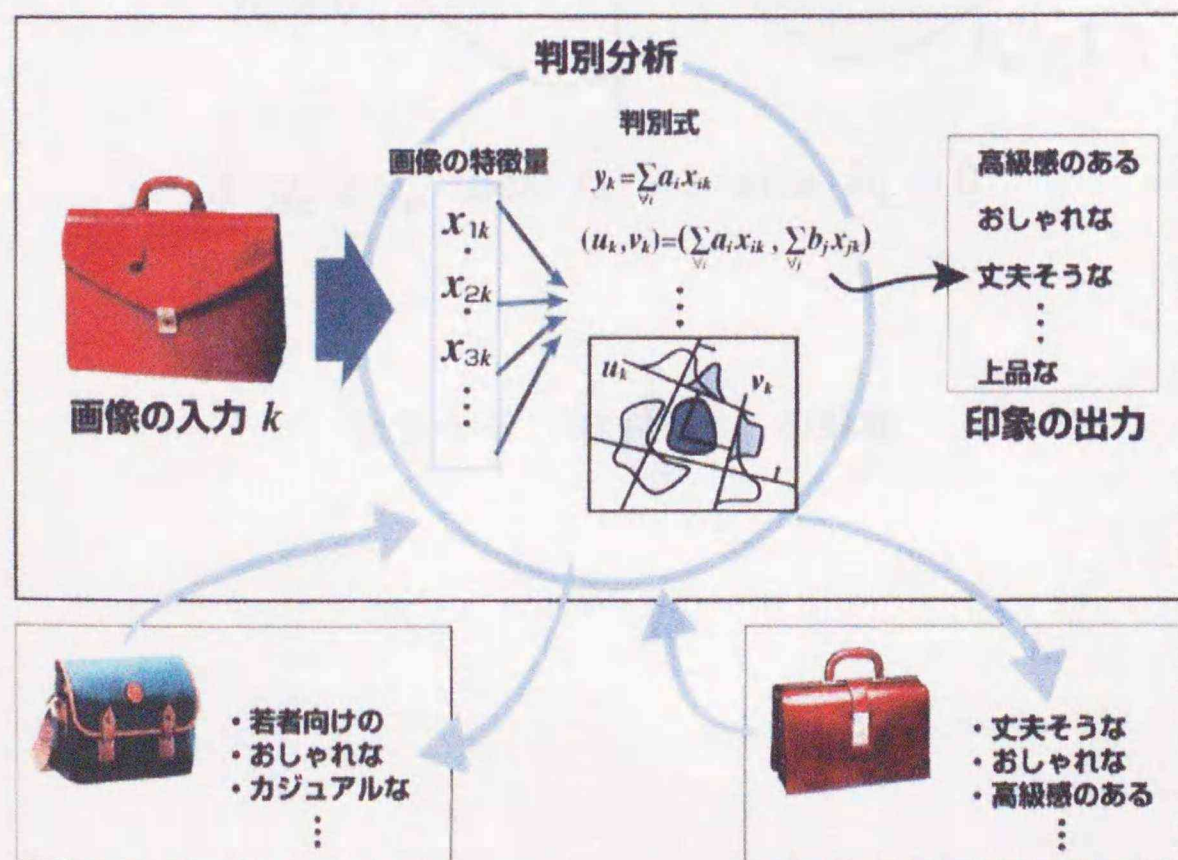


図2-1-1 「印象の推定モデル」としての「判別分析」

具体的には、対象 $[k]$ ($k=1,2,\dots,K$) を代表する複数の属性値（たとえばカラー画像の特微量） $[x_{ik}]$ ($i=1,2,\dots,m$) を（「分析の目的」に応じて選択された形式の）「判別式」（判別関数）によって1つの値つまり「判別得点 $[f_k]$ 」に変換し、その値が一定の値つまり「判別値 $[f_0]$ 」より大きいかいなかに対応して、その対象が「特定の特性を持った対象 ($h_k=1$) の集団」に属するか、あるいは、「そうではない対象 ($h_k=0$) の集団」に属するかを判別することができる。「特定の特性の有無」を「特定の印象の有無」で置き換えれば、対象 $[k]$ を代表する“連続的な”属性値 $[x_{ik}]$ を判別式によって判別得点 $[f_k]$ に変換し、その値が判別値 $[f_0]$ より大きいかいなかに対応して、その対象に対して“平均的な”回答者が特定の印象を感じるかいか ($h_k=1 \text{ or } 0$) を推定することができる（図 2-1-2）。

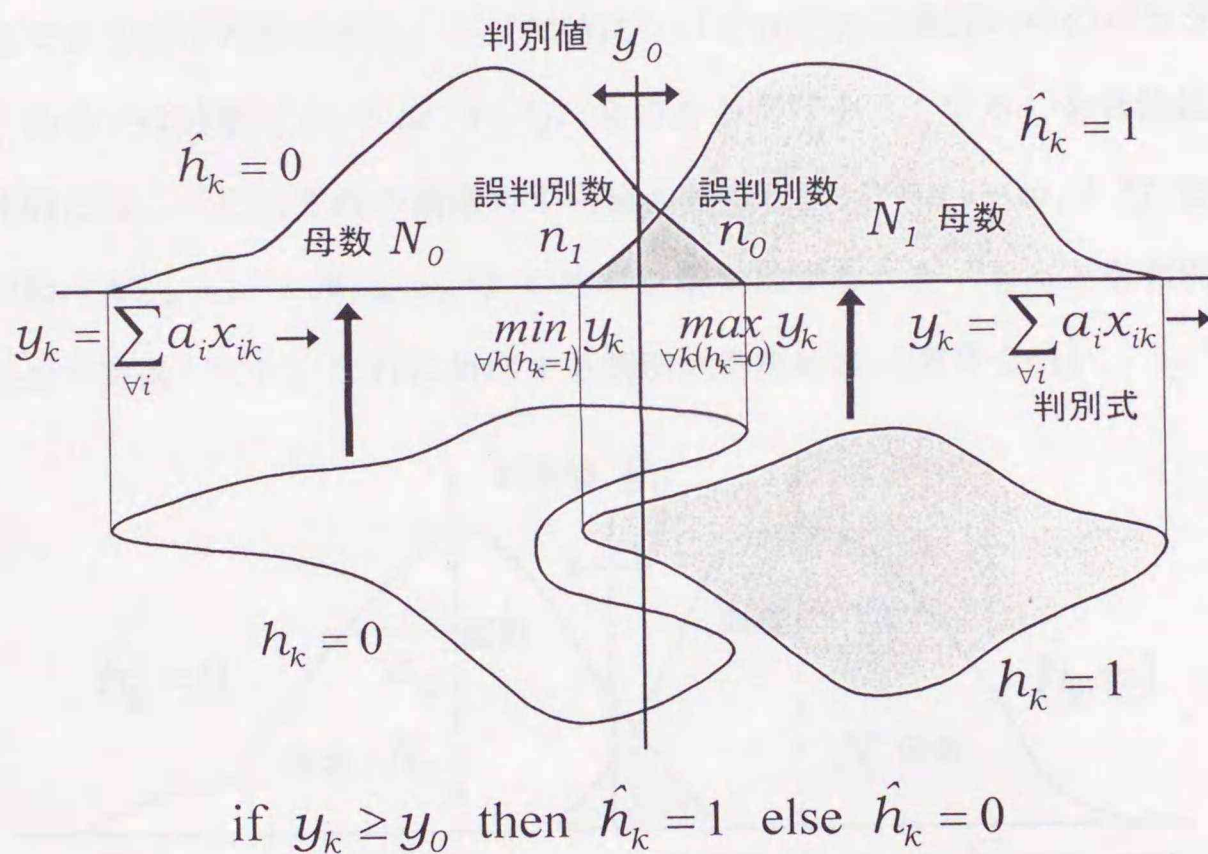
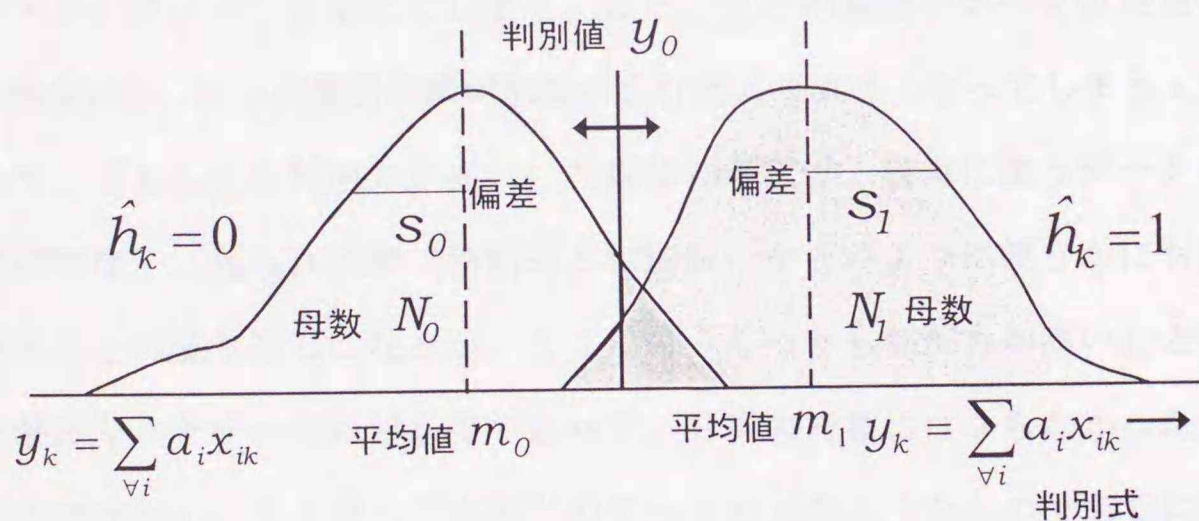


図 2-1-2 「判別分析」の論理

2-2節 解析的な「判別分析」の方法の問題点

「判別分析」には、すでに「解析的な方法」が開発されており広く利用されている。この「解析的な方法」では、判別すべき「特定の特性を持った対象 ($h_k=1$) の集団」と「そうでない対象 ($h_k=0$) の集団」の間で、それぞれの対象 $[k]$ ($k=1,2,\dots,K$) の複数の“連続的な”属性値 $[x_{ik}]$ ($i=1,2,\dots,m$) を判別式で変換した判別得点 $[f_k]$ について、「群間平方和と総平方和の比率（つまり相関比）を最大にする」、あるいは、これと等価な「群間平方和と群内平方和の比率を最大にする」ことを「もっとも判別力が高い」と定義して、これに対応する判別式のパラメーター $[a_i]$ ($i=1,2,\dots,m$) を求める（文献31、32）。

別のいい方をすれば、対象の判別得点の分布に対して、「2つの集団の間の平均的な距離 $[m_1 - m_0]$ をできるだけ大きくする」と“同時に”「それぞれの集団の中のバラツキ（偏差 $[s_1 \text{ or } s_0]$ あるいは分散 $[s_1^2 \text{ or } s_0^2]$ ）をできるだけ小さくする」を目的に、判別式を決める。具体的には、「それぞれの集団の平均値の差の自乗 $[(m_1 - m_0)^2]$ とそれぞれの集団の分散の和 $[N_1 * s_1^2 + N_0 * s_0^2]$ の比率を最大にする」を「もっとも判別力が高い」と考えて（つまり定義して）、これに対応する判別式を決める（図2-2-1）。



$$\begin{aligned}
 & |m_0 - m_1| \rightarrow \text{最大化} \\
 \text{and } & N_0 \times s_0^2 + N_1 \times s_1^2 \rightarrow \text{最小化} \\
 \Rightarrow & \frac{(m_0 - m_1)^2}{N_0 \times s_0^2 + N_1 \times s_1^2} \rightarrow \text{最大化}
 \end{aligned}$$

図2-2-1 解析的な「判別分析」での目的関数

しかし、判別分析を利用した「印象の推定モデル」を構築する立場から考えると、「もっとも判別力が高い」の意味の解釈をこのように“間接的に”するのではなく、もっと“直接的に”、たとえば「誤判別数の合計 $[n_1 + n_0]$ がもっとも少ない」あるいは「誤判別率 $[(n_1 + n_0) / (N_1 + N_0)]$ がもっとも小さい」などと解釈する立場も考えられる。この“直接的な”解釈の立場からは、“間接的な”解釈によって導かれた解は必ずしも満足できるものではないであろう。ちなみに、「誤判別数の合計 $[n_1 + n_0]$ がもっとも少ない」あるいは「誤判別率 $[(n_1 + n_0) / (N_1 + N_0)]$ がもっとも少ない」を表す目的関数は、“離散的な”値をとるために微分ができない、したがって、“解析的には”対応することができないことに注意が必要である。

実は、「解析的な方法」では、2つの集団の属性値のデータは、その分散共分散行列が互いに等しく、また、正規分布にしたがっていることを前提としている。分析する2つの集団のデータに、この前提条件がなりたっていれば、「解析的な方法」が目的としている「相関比を最大にする」を満たす解は、「誤判別数の合計を最小にする」解、より正確にいうと「誤判別を生じる確率の合計を最小にする」解と一致する。しかし、2つの集団のデータの分散共分散行列が等しくない場合には、2つの集団の誤判別数の一方が大きく他方が小さくなるという“アンバランス”を生じてしまう。また、2つの集団のデータが正規分布にしたがっていない場合には、2つの集団の誤判別数の合計はより大きくなってしまう（文献31、32）。

ところで、「もっとも判別力が高い」の意味の解釈は、現実には扱うデータの性質によって異なるだけでなく、得られた解（判別式と判別値）をどのように使うかにも関係している。前に述べたことの繰り返しになるが、ここでの「もっとも判別力が高い」とは、あくまでも分析の対象となったデータに対してであって、分析の対象にはならなかった“未知”のデータに対してではない。ここで、“未知”のデータがどのようなものであるにせよ、得られた解の“未知”のデータに対する「判別力」つまり「判別のよさ」は、なにか別の方法で評価されなければならないであろう。得られた解を“未知”のデータに対して適用する場合には、そのデータが分析したデータと“同質”であることを、なにかの方法で確認しておくことが必要であろう。また、その場合、得られた解の「判別力」は分析で評価されたそれを下回ることであろう。「得られた解をどのように使うか」というのは、そのような意味を含んでいるのである。

たとえば「特定の特性を持った対象 ($h_k=1$) の集団」と「そうではない対象 ($h_k=0$) の

集団」の間で、対象の数やその属性値の分布や性質に違いが少なく、また、得られる解を適用範囲をとくに意識しないで使う場合には、「誤判別数の合計を最小にする」のが適当であろう。この場合は判別分析の目的関数は(1)式で表される。

(1) $n_0 + n_1 \rightarrow$ 最小化

また、2つの集団の間で、対象の数やその属性値の分布や性質に違いが少なく、得られた解を2つの集団に対してほぼ同じ割合で適用したいという場合には、2つの集団のどちらか一方に対してだけ誤判別数が多いということがないように、「2つの集団の誤判別数のバランスをとる」ことが必要であろう。この場合の目的関数は、「2つの集団の誤判別数のうちより大きなものの値を最小にする」が適当であろう。具体的には(2)式で表すのが適当であろう。

(2) $\max(n_0, n_1) \rightarrow$ 最小化

さらに、2つの集団の間で、対象の数の違いが大きい場合には、誤判別数についてバランスをとるのではなく、「誤判別の割合(率)についてバランスをとる」ことが必要であろう。つまり、この場合の目的関数は、「2つの集団の誤判別率のうちより大きいものの値を最小にする」、つまり、具体的には、(3)式で表すのが適当であろう。

(3) $\max\left(\frac{n_0}{N_0}, \frac{n_1}{N_1}\right) \rightarrow$ 最小化

あるいは、もっと直接的に、「もっとも判別力が高い」は、「『特定の特性を持った対象 ($h_k=1$) の集団』と『そうではない対象 ($h_k=0$) の集団』の間の最短距離を最大にする」、つまり、「特定の特性を持った対象 ($h_k=1$) の集団」の判別得点の最小値と「そうではない対象 ($h_k=0$) の集団」の判別得点の最大値の差を最大にする、という方法も考えられる。この場合の目的関数は、(4)式で表される。

(4) $100\% \min_{\forall k(h_k=1)} y_k - 100\% \max_{\forall k(h_k=0)} y_k \rightarrow$ 最大化

さらに、対象の属性値あるいは判別得点の分布に(ほかの対象から飛び離れた)“特異な”つまり“異質な”点が含まれている可能性を考慮して、「『特定の特性を持った対象 ($h_k=1$)

の集団』の（上から90%の対象を含む）90%最小値（パーセンタイル）と『そうではない対象（ $h_k=0$ ）の集団』の（下から90%の対象を含む）90%最大値（パーセンタイル）の間の距離を最大にする」という方法も考えられる。この場合の目的関数は、(5)式で表される。（この目的関数については2-8節で詳細に説明する。）

$$(5) \quad 90\% \min_{\forall k(h_k=1)} y_k - 90\% \max_{\forall k(h_k=0)} y_k \rightarrow \text{最大化}$$

また、1-4-2項で説明した、最小限のガイダンスの下で、対象（刺激）に対する説明と印象を自由な申告語（説明語と印象語）で申告させる「自由申告法」による調査データを考えてみよう。この場合、特定の申告語（説明語と印象語）の「申告あり」については、その回答者は確かにその対象が特定の特性を持っていること（ $h_k=1$ ）を“積極的に”肯定していると考えてもよいが、「申告なし」については、その回答者は必ずしもその対象がその特性を持っていない（ $h_k=0$ ）と“積極的に”述べた訳ではなく、たまたまそれに気がつかなかった、いうまでもないことと考えた、あるいは、すでに類似のことをいってしまったので敢えて述べなかった、などのために“積極的には”述べなかった可能性が考えられる。

こういった調査データの間には、回答者の申告意思の積極性に違いが見られる。こういった場合には、対象が特定の特性を持っていないこと（ $h_k=0$ ）の意味が、その特性を持っていること（ $h_k=1$ ）ほどには積極的な意味を持たないので、たとえば「誤判別数を一定の割合だけ“割り引いて”評価をする」といった考慮や工夫も必要であろう（(6)式）。

$$(6) \quad (1-\alpha)n_0 + n_1 \rightarrow \text{最小化} \quad (0 \leq \alpha \leq 1)$$

実際に判別分析を適用する問題では、2つの集団に属する対象の数 [N_1 or N_0] に違いがあるだけでなく、その対象の属性値 [x_{ik}] などの性質にも違いがある場合が多い。また、その属性値の分布も、必ずしも（数学的な取り扱いが簡単な）正規分布などにしがっていることは少ない。したがって、実際の問題では、取り扱う対象の数 [N_1 or N_0] やその対象の属性値 [x_{ik}] の性質などを明らかにすることが必要であり、また、得られる解（判別式と判別値）をどのように使うかを明らかにしておくことが必要である。これらの意味の解釈は、

「判別分析の目的」の設定に密接に関係しており、これを実際に判別分析を適用する問題からの要請にしたがって行うことが大切である。

2-3節 遺伝的アルゴリズムを利用した計算的な「判別分析」の方法

では、「判別分析」を適用する実際の問題の必要にしたがって、「もっとも判別力の高い」解を見つけだすために開発した、遺伝的アルゴリズム（文献33～35）を利用した計算的な「判別分析」の方法と計算アルゴリズム（以下、「計算的な判別分析」あるいはより簡単に「計算的な方法」という）を説明する。

2-3-1項 「遺伝的アルゴリズム」のレビュー

まずはその準備として、「計算的な方法」の“道具”である「遺伝的アルゴリズム」つまり「GA」（genetic algorithm）を簡単に説明しておく。この方法は、1975年に、生物の進化のメカニズムをヒントにした「最適化問題」のアルゴリズムの1つとして、ミシガン大学のJ・H・ホランドが提案したもので、単純にいうと、以下のような方法である。

つまり、生物には「遺伝」と呼ばれるメカニズムがあり、交配や結婚などによって、生物の特徴を、具体的には「遺伝子」の並びである「染色体」を、親から子に、子から孫へ、孫から曾孫へ伝え、たとえばより環境条件に適したものが生き残る「適者生存の原理」などにしたがって全体として“進化”していく。いわゆる「ダーウィンの進化論」である。このプロセスをヒントに、解のそれぞれのパラメーターを「遺伝子」、そのパラメーターの並び（配列）を「染色体」とみなして、これに（遺伝や進化の実態である）交配や結婚などに相当する数学的な操作（これを「遺伝的操作」と呼ぶ）を加えてパラメーターの値とパラメーターの並びに変化を与え、設定した目的に対してより適合した染色体が生き残るように、つまり、“適者”が生き残るように選択する。このプロセスを繰り返して、最終的にもっとも目的に適合した（実際には、できるだけ目的に適合した）染色体つまり解を見つけだそう、というのである。

ちなみに、ここで説明した“単純な”方法は、「単純GA」（simple GA, SGA）と呼ばれており、実際の問題への適用では、うまくいく場合と必ずしもそうではない場合がある。うまくいかない場合の原因の1つとして、相対的な意味で“よい”解が存在すると、解の探索が充分でないにも関わらず、その解によって集団が占領されてしまう「過剰収束」があるが、一般的には、実際の問題への適用は、それぞれの問題の内容と性格に応じて、それぞれの問題に合った少しく“複雑な”

工夫が必要とされることが多いようである。

この方法の最大の特徴は、うまく定式化ができない問題でも、解の候補の「リストアップ」とその目的への適合度の「評価」さえできれば、この方法が適用できることである。また、もう1つの特徴として、たとえばシミュレーティッドアニーリング（焼き鈍し）法や山登り法など、ほかの多くの最適化法の計算アルゴリズムが解の“局所的な”探索の積み重ねという方法に基づいており、“局所的な”最適解を導き出してしまいう危険があるのに対して、この方法は、「多点探索」という方法で解の“多様性”を確保することによってこの危険を避けることができ、“大局的な”探索に優れている、とされている。

ちなみに、「遺伝的アルゴリズム」への関心の高まりは、1985年から2年おきに「遺伝的アルゴリズム国際会議」（International Conference on Genetic Algorithm, ICGA）が開催されるようになり、また、サンタフェ研究所のC・ラングトンによって「人工生命」という“野心的な”研究領域が出現したことによる。また、ニューラルネットワークの最適化など、“学習機能”を実現するためのアルゴリズムとして注目を集め始めたのは、1989年ころからである。日本で、遺伝的アルゴリズムへの関心が高まったのは、1990年ころからのようである。

現在は、「遺伝的アルゴリズム」の研究は、以下の課題に対して行なわれている。つまり、応用事例の研究として、「組合せ最適化問題」の1つである「スケジューリング」では、ジョブショップスケジューリング問題（文献36～38）のほか、仮設足場の配置計画（文献39）、バス仕業ダイヤの作成（文献40）など、数多くの事例研究が試みられ、“実務的に”よい結果が得られているようである。また、同じく「組合せ最適化問題」の1つである「配置問題」では、VLSI（超大型集積回路）のレイアウト（文献41、42）、布地の裁断問題などのほか、パソコンやワークステーションのディスプレイのユーザーインターフェイスの研究（文献43～45）が行なわれている。ちなみに、VLSIのレイアウトは「グラフ分割問題」である。VLSIのレイアウトでは、レイアウトの総面積や配線の長さなどが評価の基準であり、よい結果が得られているようであるが、ユーザーインターフェイスでは、配置の対称性や操作性などの“心理学的な”要素が評価基準であり、また、これには動的な変化があるため、まだまだ研究の余地が大きいようである。

「組合せ問題」では、ほかに「遺伝子情報解析」、つまり、遺伝子情報からの分類規則の抽出で、共通の祖先や機能を持つ遺伝子の配列に、進化的に保存されているモチーフ（パターン）の抽出などへの適用が試みられ、よい結果が得られているようである（文献46～49）。また、ネットワーク上のデータトラフィックの回線割り付けやルーティングの方式設計などを含めた「ネットワーク

設計」などへの適用の試みも多く、これも実務的な成果を上げているようである（文献50～52）。

さらに、ぼけやノイズによって劣化した画像からできるだけ元の画像に近い画像を求めようとする「画像復元」も、“よい成果を期待できる”遺伝的アルゴリズムの有力な応用例になっている（文献53～55）。そして、類似の研究として、三次元局面のメッシュの生成（文献56）、データの当てはめ（文献57、58）なども試みられている。

また、「多目的最適化問題」への適用では、計算アルゴリズム上の研究（文献59～61）のほか、スケジューリング問題への適用など（文献62、63）などが試みられている。このほか、機械翻訳手法への適用（文献64）も試みられている。これらについても、実務的な意味での効果を上げる事例が現れているようである。

さらに「規則学習」への適用では、「分類システム」（classifier system）に対するミシガンモデル（文献65、66）やピッツバーグモデル（文献67）のほか、分子系統樹の作成など（文献68～70）が試みられているが、この分野は、一部の例を除いて、まだ実際の問題に対して効果を上げる、という段階になってはいないようである。

次に、遺伝的アルゴリズムの「性能向上」を目的にした研究として、遺伝的アルゴリズムの持つ高度な並列性を利用して解の“質”（これは“最適性”である）と解を見つけ出すまでの“効率”を上げようとする「並列分散GA」の研究（文献71～74）、染色体の長さを可変長として（染色体の長さが固定の）単純GAの持つ非線型問題への弱さを克服しようとする「メッシーGA」（messy GA, MGA）と呼ばれる研究（文献75）、あるいは、遺伝的アルゴリズムと局所的な探索に優れたほかの最適化法と組み合わせて解の“質”の向上をはかろうとする「ハイブリッドGA」の研究（文献76～78）も進められている。ちなみに、応用事例の研究では、なにかの形で、ほかの最適化法を“補完的に”使うのが一般的な傾向である。

そのほか、「適者生存の原理」を基礎としたダーウィンの進化論ではなく、J・ラマルクが提唱した「ラマルク進化論」（つまり「用不用説」と「獲得形質の遺伝」）などをヒントに、より効率的な計算を実現しようとする試みも行なわれている（文献79～82）。

さらに、1985年、スタンフォード大学のJ・N・コザは、プログラムの論理に対して遺伝的アルゴリズムを適用し、“進化的に”最適なプログラムを見つけ出す「遺伝的プログラミング」

（genetic programming, GP）を提案し、人工知能の問題解決、ロボット、分子生物学などを含む、実際的な問題を対象に研究が進められている（文献83、84）。

2-3-2 項 計算的な「判別分析」の方法

さて、本題に戻ろう。「判別分析」にこの遺伝的アルゴリズムを適用する場合、そして、解つまり判別式の形式として、(1)式に示す1次式を採用した場合、判別式のパラメーター $[a_i]$ ($i=1, 2, \dots, m$)を「遺伝子」、そのパラメーターの並び(配列) $[a_1, a_2, \dots, a_m]$ を「染色体」とみなす。判別式のパラメーターが決まれば「判別式」は確定し、その判別式を使ってそれぞれの対象 $[k]$ ($k=1, 2, \dots, K$)の(連続的な)属性値 $[x_{ik}]$ を1つの値つまり「判別得点 $[y_k]$ 」に変換することができる。そして、この判別得点の分布に対して、たとえば「判別分析の目的」が「誤判別数の合計 $[n_1 + n_0]$ を最小にする」の場合には、これをもっともよく満たす「判別値 $[y_0]$ 」を見つけることができる。その結果として、それぞれの判別式(解)の目的関数の値つまり判別分析の目的への「適合度」(この場合には「誤判別数の合計」である)が決まる(図2-3-1)。

$$(1) f_k = \sum_{\forall i} a_i x_{ik} \quad (-1 \leq a_i \leq +1)$$

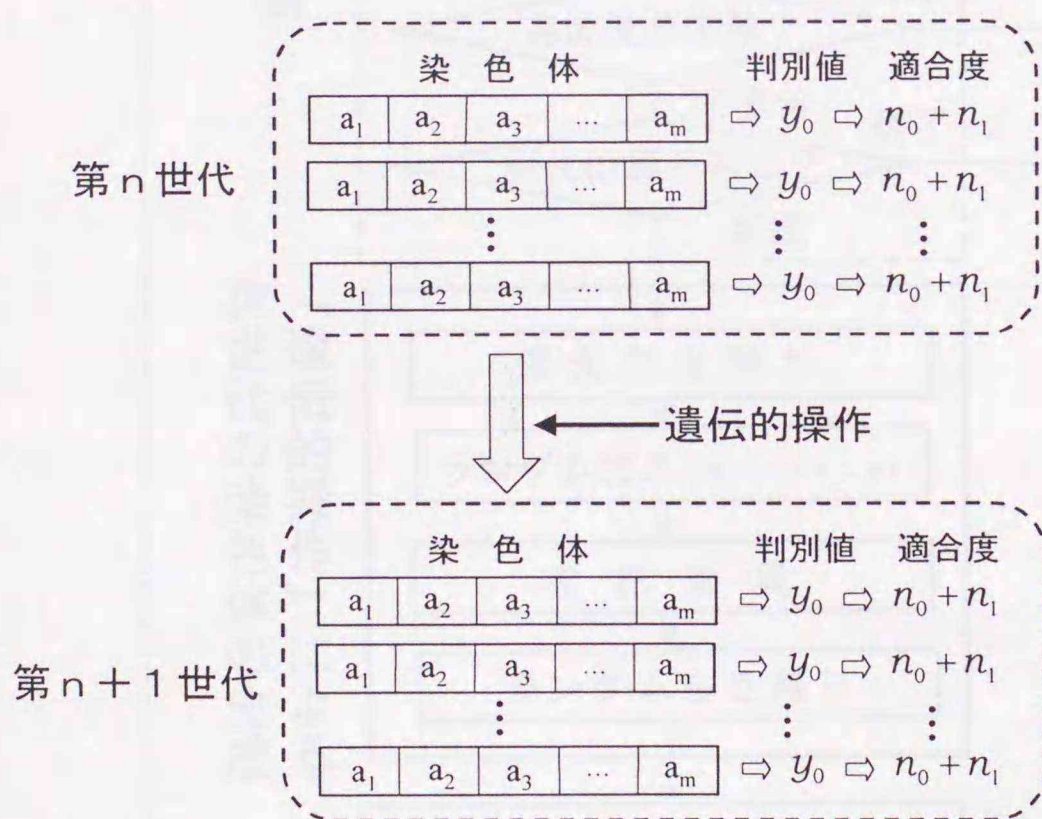


図2-3-1 「遺伝的アルゴリズム」による解の改善

具体的には、1つの世代について“複数”の染色体を用意し、これらに(染色体に変化を与えるための数学的な操作である)「遺伝的操作」を加えて、目的への適合度を改善する可

能性がある新しい“複数”の染色体を生成し、この中からより適合度の大きい染色体を選択して次の世代に生き残らせる。このプロセスを繰り返し世代が進むにしたがって、1つの世代の“複数”の染色体の集まりを全体としてより適合度の大きい染色体を含んだものに進化させていくことを期待するのである（図2-3-2）。

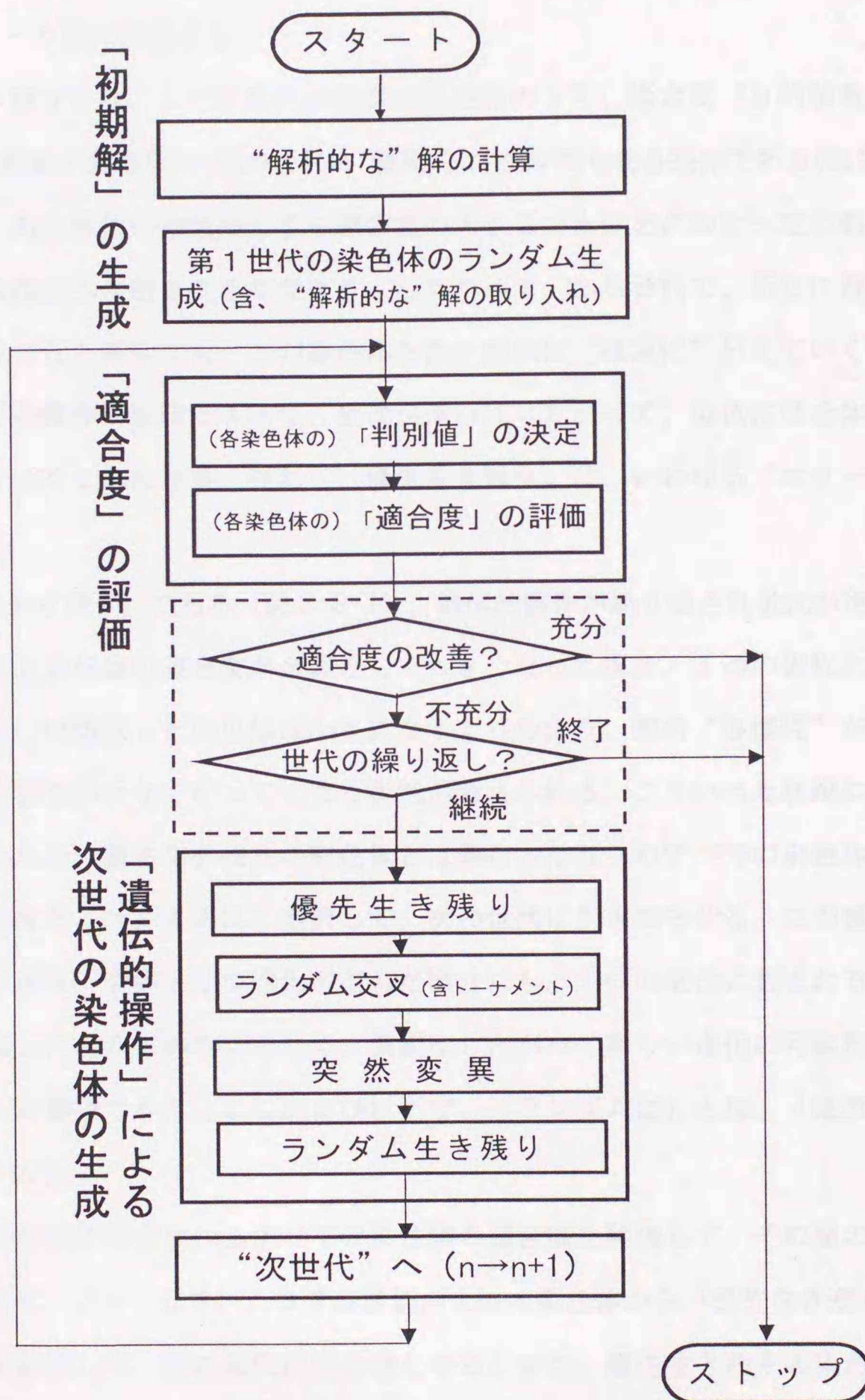


図2-3-2 「遺伝的アルゴリズム」の処理の流れ

この「計算的な方法」では、遺伝的操作として「優先生き残り」「ランダム生き残り」「ランダム交叉」「トーナメント」「突然変異」の5つを採用した。このほか、代表的な遺伝的操作として「逆位」などがあるが、この問題に対する効果が不明であるので採用しなかった。以下、これらの操作の内容を具体的に説明する。ちなみに、これらの遺伝的操作は“離散的”であり、必ずしも最適な解にうまくたどりつける数学的な保証はない。これを解決するための効果的な対策は2-6節で説明する。

第1の「優先生き残り」は、1つの世代の複数の染色体のうち、適合度（目的関数の値）の大きい順に一定の個数の染色体を、次の世代に“優先的に”生き残らせる操作である（図2-3-3）。この操作によって、ある世代の染色体のうち適合度の大きさに上位を占めた一定の個数の染色体は、その次の世代に生き残ることができる。したがって、ある世代で、新たに適合度の大きい染色体が生成された場合にも、この染色体を次の世代に“確実に”伝えていくことができる。つまり、この操作の採用によって、世代が進むにしたがって、染色体は全体として“確実に”進化していくことになる。なお、「優先生き残り」は、いわゆる「エリート保存戦略」と同じである。

第2は「ランダム生き残り」である（図2-3-3）。遺伝的操作が繰り返され世代が進んで、1つの世代に含まれる染色体の適合度が全体として大きくなった場合、1つの世代に含まれる染色体が同じあるいは類似した染色体ばかりになって（つまり、解の“多様性”が失われて）、次の進化の可能性が少なくなってしまう状況が考えられる。こういった状況になってしまうことを避けるため、優先生き残りの染色体とは異なった残りのすべての染色体の中から一定の個数の染色体を“ランダムに”選択して、次の世代に生き残らせる。この操作の採用によって、世代が進み、全体として進化が進んだ場合にも、1つの世代に含まれる染色体は同じあるいは類似したものばかりではなく、異質にしたがって新しい進化の可能性を持った染色体を含むことが期待できる。ここおよび以下で、「ランダムに」とは、「乱数を利用する」という意味である。

具体的には、1つの世代に含まれるすべての染色体の適合度を評価して、その値の大きさの順に並べて（つまり、ソートして）、まずは評価が上位の染色体から「優先生き残りの数」に相当する染色体を選択して、次の世代に生き残らせる。また、優先生き残りの染色体とは異なった残りのすべての染色体の中から、「ランダム生き残りの数」に相当する染色体を“ランダムに”選択して、これらも次の世代に生き残らせる。「優先生き残りの数」や「ランダ

ム生き残りの数」は、遺伝的アルゴリズムの（計算の）“戦略”として、分析者が設定する。

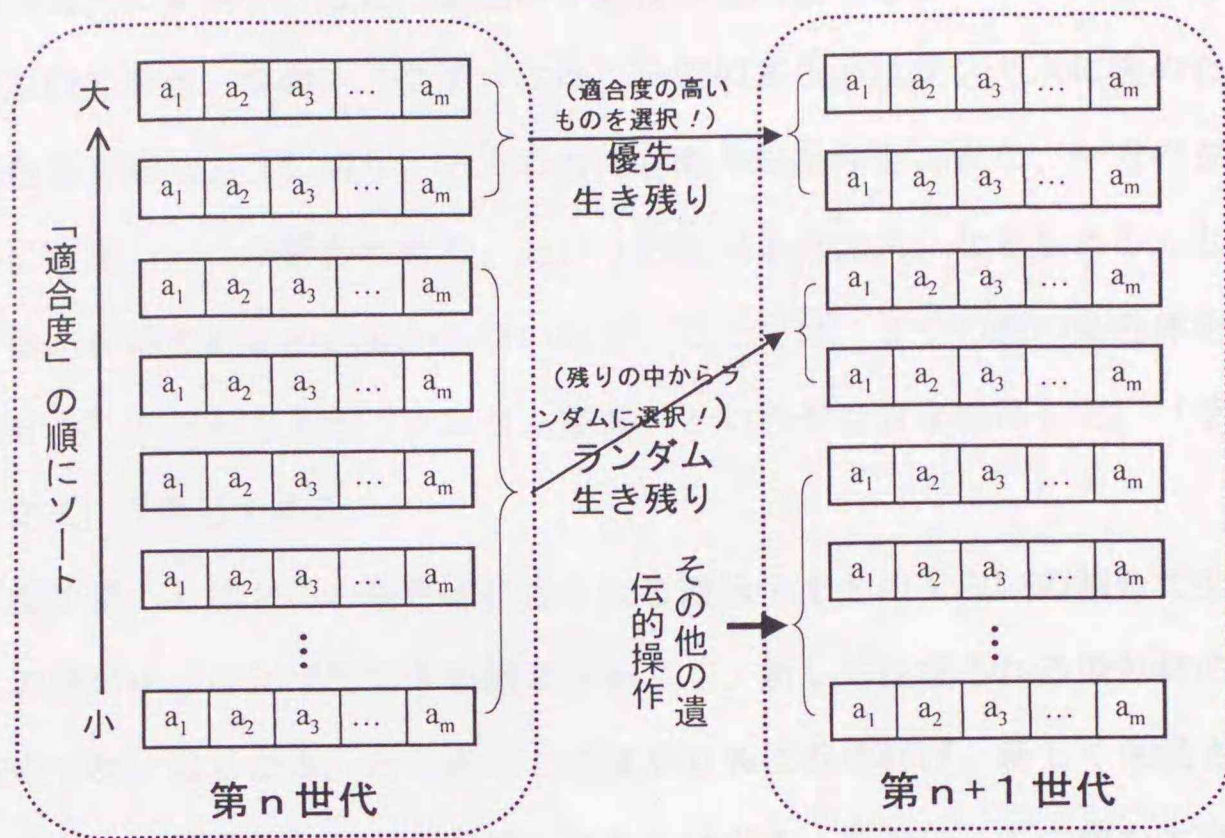


図 2-3-3 「優先生き残り」と「ランダム生き残り」の説明

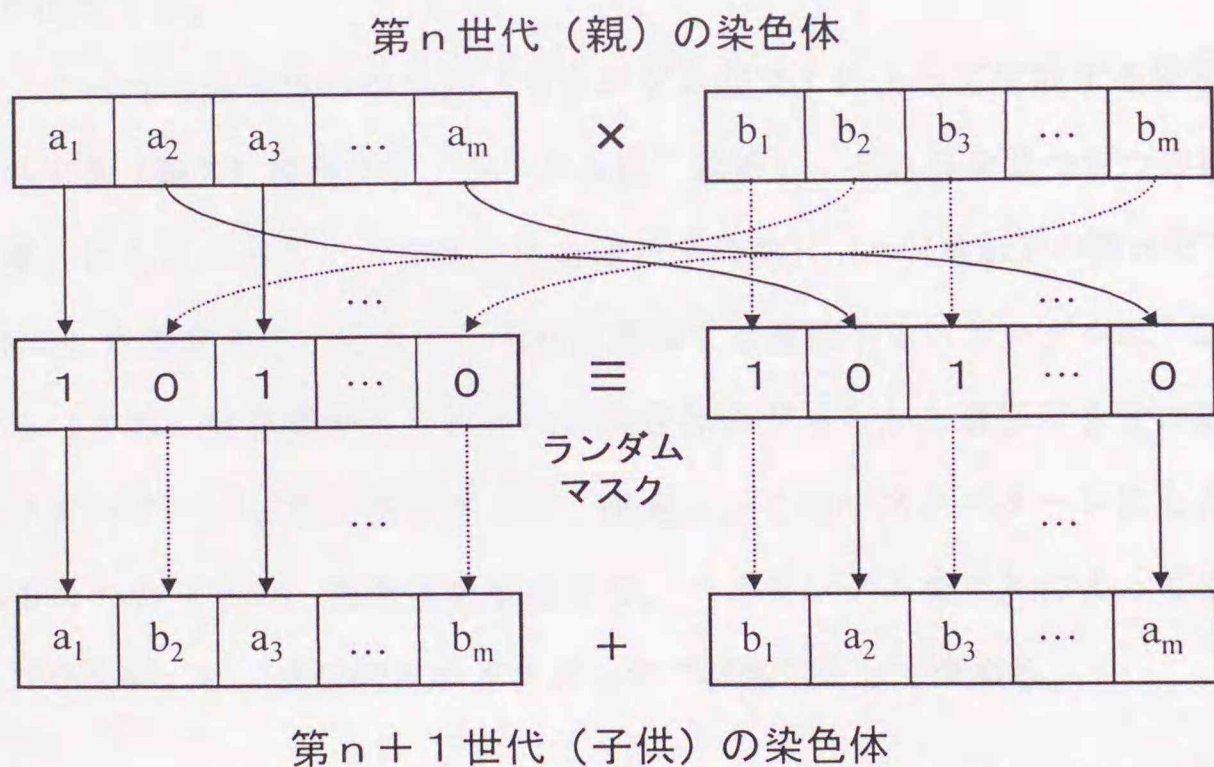


図 2-3-4 「ランダム交叉」の説明

第3は「ランダム交叉」である（図 2-3-4）。この操作では、1つの世代に含まれる染色体の中から2つ（1組）の染色体を“ランダムに”選択し、これを交叉率に対応して“ランダムに”交叉させて、適合度を改善する可能性のある2つの染色体を生成する。1組の男女

の結婚から、この男女の遺伝子を受け継いだ2人の子供を誕生させることに相当するもので、2つの染色体の交叉によって、もとの染色体の適合度を改善するよりよい染色体の生成（誕生）を期待するのである。なお、「交叉」には、一定のあるいはランダムに決めた長さで、2つの親の染色体を同じように切り、一方の右片と他方の左片を、また、一方の左片と他方の右片を互いに交換してつなぎ合わせる、という方法（1点交叉）などもある。しかし、そのいずれの方法が有効であるかが分からないので、ここでは、2つの親の染色体をもっともよく“かき混ぜる”ことができる「ランダム交叉」というやり方を採用した。「ランダム交叉」は「一様交叉」と同じである。

ここで、「交叉率」とは2つの染色体に含まれる遺伝子をどのくらいの割合で交叉させるかを表し、この値が小さければ交叉する割合が小さく、新しく生成される染色体の遺伝子はもとの染色体のそれに近くなる。たとえばこの値が0%に近ければ、新しく生成される染色体はもとのいずれかの染色体とかなり類似したものになる。反対に、この値が大きくたとえば100%ならば、新しく生成される染色体の遺伝子はもとの2つ染色体のどちらから受け継ぐかはまったく任意となり、もとの2つの染色体のいずれとも違った新しい染色体が生成される可能性が出てくる。

具体的には、すべての染色体の中から「『ランダム交叉』によって生成する染色体の数」（偶数）に相当する（親の）染色体を“ランダムに”選択し、それらを2つずつの組とする。また、「交叉率」にしたがって、「交叉できる遺伝子の数」（＝（遺伝子の数×交叉率）の値を整数化したもの）を決定する。そして、それに相当する遺伝子を“ランダムに”選択して、新しく生成する（子の）染色体がもとの2つの染色体のどちらから遺伝子を受け継ぐべきかを定める「マスクパターン」を“ランダムに”決定し、このマスクパターンにしたがって新しく生成される2つの（子の）染色体を生成する。「『ランダム交叉』によって生成する染色体の数」と「交叉率」は、遺伝的アルゴリズムの“戦略”として決める。

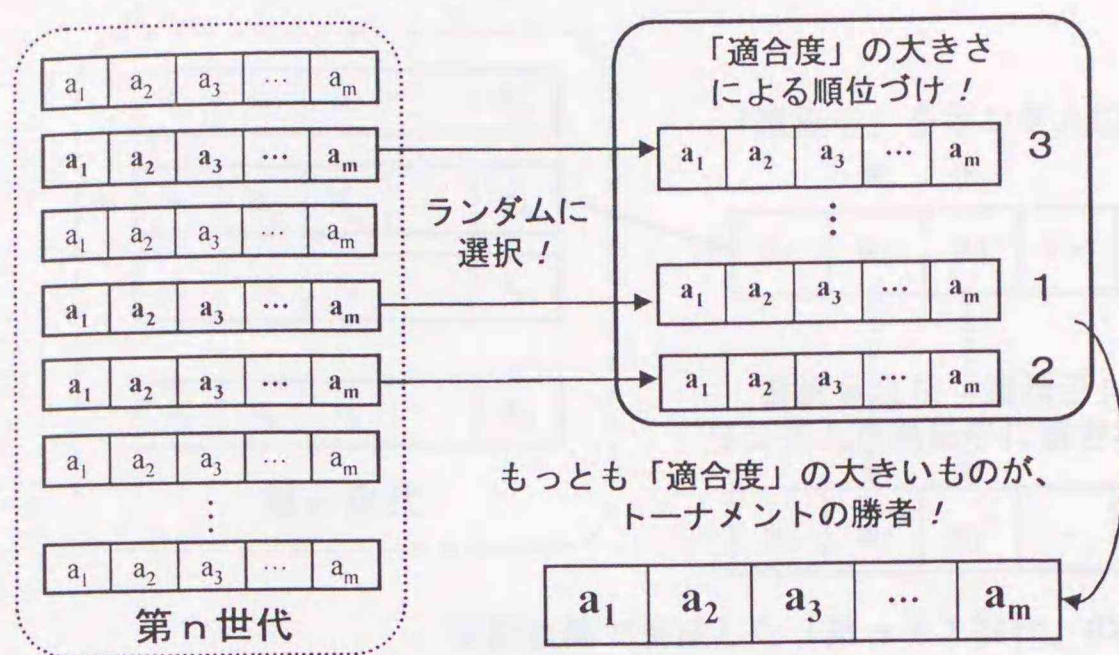


図 2-3-5 「トーナメント」の説明

第4は「トーナメント」である（図 2-3-5）。これは、複数の染色体を競わせてよりよい染色体を選抜しようという操作である。「ランダム交叉」では、親に相当するもとの世代の染色体を2つ（1組）選択するが、このときに、この操作を行って適合度のより大きい染色体を選抜してこれを親に相当するもとの染色体としよう、とするのである。親に相当するもとの世代の染色体の適合度が大きくなれば、「ランダム交叉」によって新しく生成される染色体の適合度も大きくなることが期待できる、というわけである。

具体的には、すべての染色体の中から、「トーナメント数」に相当する染色体を“ランダムに”選択する。そして、これらの染色体の適合度を評価して、その中でもっともよい染色体を選抜して、「ランダム交叉」の親に相当するもとの染色体とする。「トーナメント数」は、遺伝的アルゴリズムの“戦略”として決める。

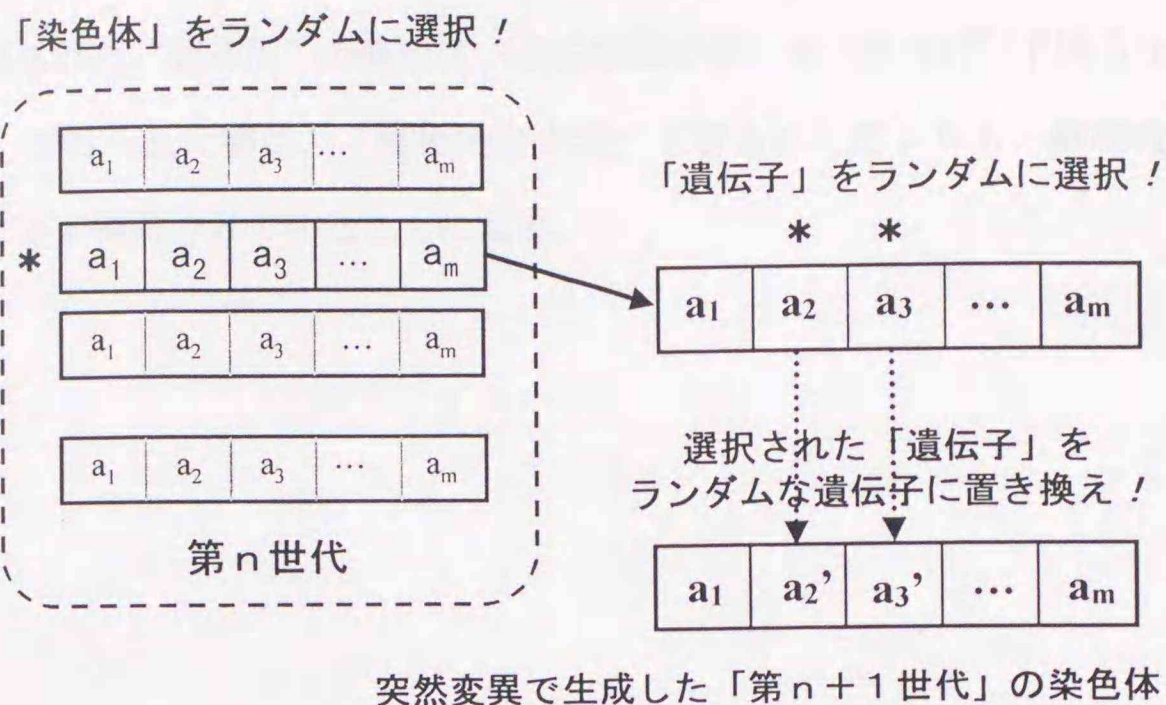


図 2-3-6 「突然変異」の説明

第5の「突然変異」は、染色体に含まれるいくつかの遺伝子に“ランダムに”変異を与え、適合度を改善する新しい可能性を作ろうとする操作である（図2-3-6）。この操作によって、突然変異を与える前の染色体に比べて、適合度がより大きい染色体が生成される可能性を期待するのである。「突然変異率」は、染色体に含まれる遺伝子にどのくらいの割合で変異を生じさせるかで、この値が小さければ、もとの染色体とはあまり変わらないつまり似たものが生成される。反対に、この値が大きければ、もとの染色体とは大きく違った新しい染色体が生成される可能性が高くなる。

具体的には、1つの世代に含まれるすべての染色体の中から、「『突然変異』により生成する染色体の数」に相当する染色体を“ランダムに”選択する。次に、「突然変異率」にしたがって、「『突然変異』が生じる遺伝子の数」（＝（遺伝子の数×突然変異率）の値を整数化したもの）を決定し、それぞれの染色体の遺伝子のうち「突然変異」が生じる遺伝子を“ランダムに”決定する。そして、それぞれの遺伝子を“ランダムに”変更する。「『突然変異』により生成する染色体の数」と「突然変異率」は、遺伝的アルゴリズムの“戦略”として決める。

ちなみに、「初期解」つまり「第1世代の染色体」は、基本的に“ランダムに”生成する。しかし、一般的に、「解析的な方法」で得られた“最適解”も“適合度の大きい”つまり“判別力の高い”解であるので、これを第1世代の複数の染色体のうちの1つとして取り入れておけば、（遺伝的アルゴリズムの）繰り返し計算の収束の速度を速めることができ、計算の効率化をすることができる（図2-3-2）。この場合、「計算的な方法」で得られる解は、「解

析的な方法」で得られた“最適解”の適合度（目的関数の値）を“絶対に”下回らないので、
「計算的な方法」で得られた解は、「解析的な方法」で得られた解よりも、論理的に“優れた”ものであることが保証されていることになる。

2-4節 例題による「解析的な方法」と「計算的な方法」の解の比較

表 2-4-1 は、英国の統計学者の R・A・フィッシャーが提示し、多くの教科書などでも判別分析の代表的な例題として使われている「あやめ科の 3 種類の植物の 4 変量観測値」のデータである（文献 85）。このデータは、「イリス・サトーサ（A 群）」と「イリス・ヴェルシコロール（B 群）」と「イリス・ヴィルジニカ（C 群）」の 50 件ずつのサンプルについて、がく（萼）の長さ $[x_1]$ 、がくの幅 $[x_2]$ 、花弁の長さ $[x_3]$ 、花弁の幅 $[x_4]$ の 4 つの変量を測定した結果である。

このうち、「イリス・サトーサ（A 群）」は、花弁の長さ $[x_3]$ と花弁の幅 $[x_4]$ の値の分布が、ほかの 2 つの群（B 群と C 群）のサンプルのそれとはまったく重なりがなく、これらとは容易に判別できることが分かっているので、「イリス・ヴェルシコロール（B 群）」と「イリス・ヴィルジニカ（C 群）」の 2 つの群の判別を、従来の「解析的な方法」と新たに開発した「計算的な方法」で解いてみて、得られた解を比較した。

表 2-4-1 あやめ科の 3 種類の植物の 4 変量観測値

##	イリス・セトーサ (A群)				イリス・ヴェルシコロール (B群)				イリス・ヴィルジニカ (C群)			
	がく の長さ	がく の幅	花弁 の長さ	花弁 の幅	がく の長さ	がく の幅	花弁 の長さ	花弁 の幅	がく の長さ	がく の幅	花弁 の長さ	花弁 の幅
	X ₁	X ₂	X ₃	X ₄	X ₁	X ₂	X ₃	X ₄	X ₁	X ₂	X ₃	X ₄
01	5.1	3.5	1.4	0.2	7.0	3.2	4.7	1.4	6.3	3.3	6.0	2.5
02	4.9	3.0	1.4	0.2	6.4	3.2	4.5	1.5	5.8	2.7	5.1	1.9
03	4.7	3.2	1.3	0.2	6.9	3.1	4.9	1.5	7.1	3.0	5.9	2.1
04	4.6	3.1	1.5	0.2	5.5	2.3	4.0	1.3	6.3	2.9	5.6	1.8
05	5.0	3.6	1.4	0.2	6.5	2.8	4.6	1.5	6.5	3.0	5.8	2.2
06	5.4	3.9	1.7	0.4	5.7	2.8	4.5	1.3	7.6	3.0	6.6	2.1
07	4.6	3.4	1.4	0.3	6.3	3.3	4.7	1.6	4.9	2.5	4.5	1.7
08	5.0	3.4	1.5	0.2	4.9	2.4	3.3	1.0	7.3	2.9	6.3	1.8
09	4.4	2.9	1.4	0.2	6.6	2.9	4.6	1.3	6.7	2.5	5.8	1.8
10	4.9	3.1	1.5	0.1	5.2	2.7	3.9	1.4	7.2	3.6	6.1	2.5
11	5.4	3.7	1.5	0.2	5.0	2.0	3.5	1.0	6.5	3.2	5.1	2.0
12	4.8	3.4	1.6	0.2	5.9	3.0	4.2	1.5	6.4	2.7	5.3	1.9
13	4.8	3.0	1.4	0.1	6.0	2.2	4.0	1.0	6.8	3.0	5.5	2.1
14	4.3	3.0	1.1	0.1	6.1	2.9	4.7	1.4	5.7	2.5	5.0	2.0
15	5.8	4.0	1.2	0.2	5.6	2.9	3.6	1.3	5.8	2.8	5.1	2.4
16	5.7	4.4	1.5	0.4	6.7	3.1	4.4	1.4	6.4	3.2	5.3	2.3
17	5.4	3.9	1.3	0.4	5.6	3.0	4.5	1.5	6.5	3.0	5.5	1.8
18	5.1	3.5	1.4	0.3	5.8	2.7	4.1	1.0	7.7	3.8	6.7	2.2
19	5.7	3.8	1.7	0.3	6.2	2.2	4.5	1.5	7.7	2.6	6.9	2.3
20	5.1	3.8	1.5	0.3	5.6	2.5	3.9	1.1	6.0	2.2	5.0	1.5
21	5.4	3.4	1.7	0.2	5.9	3.2	4.8	1.8	6.9	3.2	5.7	2.3
22	5.1	3.7	1.5	0.4	6.1	2.8	4.0	1.3	5.6	2.8	4.9	2.0
23	4.6	3.6	1.0	0.2	6.3	2.5	4.9	1.5	7.7	2.8	6.7	2.0
24	5.1	3.3	1.7	0.5	6.1	2.8	4.7	1.2	6.3	2.7	4.9	1.8
25	4.8	3.4	1.9	0.2	6.4	2.9	4.3	1.3	6.7	3.3	5.7	2.1
26	5.0	3.0	1.6	0.2	6.6	3.0	4.4	1.4	7.2	3.2	6.0	1.8
27	5.0	3.4	1.6	0.4	6.8	2.8	4.8	1.4	6.2	2.8	4.8	1.8
28	5.2	3.5	1.5	0.2	6.7	3.0	5.0	1.7	6.1	3.0	4.9	1.8
29	5.2	3.4	1.4	0.2	6.0	2.9	4.5	1.5	6.4	2.8	5.6	2.1
30	4.7	3.2	1.6	0.2	5.7	2.6	3.5	1.0	7.2	3.0	5.8	1.6
31	4.8	3.1	1.6	0.2	5.5	2.4	3.8	1.1	7.4	2.8	6.1	1.9
32	5.4	3.4	1.5	0.4	5.5	2.4	3.7	1.0	7.9	3.8	6.4	2.0
33	5.2	4.1	1.5	0.1	5.8	2.7	3.9	1.2	6.4	2.8	5.6	2.2
34	5.5	4.2	1.4	0.2	6.0	2.7	5.1	1.6	6.3	2.8	5.1	1.5
35	4.9	3.1	1.5	0.2	5.4	3.0	4.5	1.5	6.1	2.6	5.6	1.4
36	5.0	3.2	1.2	0.2	6.0	3.4	4.5	1.6	7.7	3.0	6.1	2.3
37	5.5	3.5	1.3	0.2	6.7	3.1	4.7	1.5	6.3	3.4	5.6	2.4
38	4.9	3.6	1.4	0.1	6.3	2.3	4.4	1.3	6.4	3.1	5.5	1.8
39	4.4	3.0	1.3	0.2	5.6	3.0	4.1	1.3	6.0	3.0	4.8	1.8
40	5.1	3.4	1.5	0.2	5.5	2.5	4.0	1.3	6.9	3.1	5.4	2.1
41	5.0	3.5	1.3	0.3	5.5	2.6	4.4	1.2	6.7	3.1	5.6	2.4
42	4.5	2.3	1.3	0.3	6.1	3.0	4.6	1.4	6.9	3.1	5.1	2.3
43	4.4	3.2	1.3	0.2	5.8	2.6	4.0	1.2	5.8	2.7	5.1	1.9
44	5.0	3.5	1.6	0.6	5.0	2.3	3.3	1.0	6.8	3.2	5.9	2.3
45	5.1	3.8	1.9	0.4	5.6	2.7	4.2	1.3	6.7	3.3	5.7	2.5
46	4.8	3.0	1.4	0.3	5.7	3.0	4.2	1.2	6.7	3.0	5.2	2.3
47	5.1	3.8	1.6	0.2	5.7	2.9	4.2	1.3	6.3	2.5	5.0	1.9
48	4.6	3.2	1.4	0.2	6.2	2.9	4.3	1.3	6.5	3.0	5.2	2.0
49	5.3	3.7	1.5	0.2	5.1	2.5	3.0	1.1	6.2	3.4	5.4	2.3
50	5.0	3.3	1.4	0.2	5.7	2.8	4.1	1.3	5.9	3.0	5.1	1.8
最大値	5.80	4.40	1.90	0.60	7.00	3.40	5.10	1.80	7.90	3.80	6.90	2.50
平均値	5.01	3.43	1.46	0.25	5.94	2.77	4.26	1.33	6.59	2.97	5.55	2.03
最小値	4.30	2.30	1.00	0.10	4.90	2.00	3.00	1.00	4.90	2.20	4.50	1.40

(1)式と(2)式および表 2-4-2 は、判別式の形式を 1 次式とし、「相関比（つまり群間平方和と群内平方和の比率）を最大化する」ことを目的として、「解析的な方法」で求めた解（判別式 $[f_k]$ と判別値 $[f_0]$ ）である。（以下、表中で、もともとのデータは“非斜字”で、計算した結果は“斜字”で示してある。）

この解では、B群の誤判別数 $[n_1]$ は 2 個、C群の誤判別数 $[n_0]$ は 1 個で、誤判別数の合計 $[n_0 + n_1]$ は 3 個であった。表中、（判別値の選択によって）誤判別を生じる可能性があるデータは“枠つき・太字”の数値で、（選択した判別値で）誤判別を生じたデータは“網かけ・太字”の数値で示してある。なお、B群の判別得点 $[f_k]$ の最小値 $[-18.965]$ と C群の判別得点 $[f_k]$ の最大値 $[-16.102]$ の間で丹念に探し、判別値 $[f_0]$ を $[-17.678 \sim -16.918]$ の間の値とすると、B群の誤判別数 $[n_1]$ は 1 個、C群の誤判別数 $[n_0]$ は 1 個で、誤判別数の合計 $[n_0 + n_1]$ は 2 個にできることが分かる。（このほか、(1)式の判別式と(2)式の判別値の「+」を「-」に、「-」を「+」に直した、B群とC群の並びの順序を逆転した“もう 1 つ”の解がある。これは以降の説明でも同様である。）

$$(1) \quad f_k = 3.556 x_{1k} + 5.579 x_{2k} - 6.970 x_{3k} - 12.386 x_{4k}$$

$$(2) \quad \text{if } f_k \geq 16.663(= f_0) \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

表 2-4-2 「解析的な方法」で求めた例題の解

イリス・ヴェルシコロール(B群)						イリス・ヴィルジニカ (C群)					
##	がく の長さ	がく の幅	花弁 の長さ	花弁 の幅	判別 得点	がく の長さ	がく の幅	花弁 の長さ	花弁 の幅	判別 得点	
	x_1	x_2	x_3	x_4	f		x_1	x_2	x_3	x_4	f
01	7.0	3.2	4.7	1.4	-7.355	6.3	3.3	6.0	2.5	-31.972	
02	6.4	3.2	4.5	1.5	-9.333	5.8	2.7	5.1	1.9	-23.392	
03	6.9	3.1	4.9	1.5	-10.901	7.1	3.0	5.9	2.1	-25.149	
04	5.5	2.3	4.0	1.3	-11.592	6.3	2.9	5.6	1.8	-22.745	
05	6.5	2.8	4.6	1.5	-11.906	6.5	3.0	5.8	2.2	-27.824	
06	5.7	2.8	4.5	1.3	-11.576	7.6	3.0	6.6	2.1	-28.250	
07	6.3	3.3	4.7	1.6	-11.763	4.9	2.5	4.5	1.7	-21.049	
08	4.9	2.4	3.3	1.0	-4.573	7.3	2.9	6.3	1.8	-24.068	
09	6.6	2.9	4.6	1.3	-8.515	6.7	2.5	5.8	1.8	-24.948	
10	5.2	2.7	3.9	1.4	-10.969	7.2	3.6	6.1	2.5	-27.794	
11	5.0	2.0	3.5	1.0	-7.843	6.5	3.2	5.1	2.0	-19.352	
12	5.9	3.0	4.2	1.5	-10.136	6.4	2.7	5.3	1.9	-22.653	
13	6.0	2.2	4.0	1.0	-6.656	6.8	3.0	5.5	2.1	-23.428	
14	6.1	2.9	4.7	1.4	-12.229	5.7	2.5	5.0	2.0	-25.405	
15	5.6	2.9	3.6	1.3	-5.101	5.8	2.8	5.1	2.4	-29.027	
16	6.7	3.1	4.4	1.4	-6.888	6.4	3.2	5.3	2.3	-24.818	
17	5.6	3.0	4.5	1.5	-13.293	6.5	3.0	5.5	1.8	-20.779	
18	5.8	2.7	4.1	1.0	-5.275	7.7	3.8	6.7	2.2	-25.367	
19	6.2	2.2	4.5	1.5	-15.623	7.7	2.6	6.9	2.3	-34.694	
20	5.6	2.5	3.9	1.1	-6.947	6.0	2.2	5.0	1.5	-19.819	
21	5.9	3.2	4.8	1.8	-16.918	6.9	3.2	5.7	2.3	-25.828	
22	6.1	2.8	4.0	1.3	-6.669	5.6	2.8	4.9	2.0	-23.390	
23	6.3	2.5	4.9	1.5	-16.382	7.7	2.8	6.7	2.0	-28.469	
24	6.1	2.8	4.7	1.2	-10.309	6.3	2.7	4.9	1.8	-18.982	
25	6.4	2.9	4.3	1.3	-7.135	6.7	3.3	5.7	2.1	-23.504	
26	6.6	3.0	4.4	1.4	-7.802	7.2	3.2	6.0	1.8	-20.659	
27	6.8	2.8	4.8	1.4	-10.994	6.2	2.8	4.8	1.8	-18.082	
28	6.7	3.0	5.0	1.7	-15.344	6.1	3.0	4.9	1.8	-18.019	
29	6.0	2.9	4.5	1.5	-12.429	6.4	2.8	5.6	2.1	-26.663	
30	5.7	2.6	3.5	1.0	-2.006	7.2	3.0	5.8	1.6	-17.903	
31	5.5	2.4	3.8	1.1	-7.163	7.4	2.8	6.1	1.9	-24.115	
32	5.5	2.4	3.7	1.0	-5.227	7.9	3.8	6.4	2.0	-20.087	
33	5.8	2.7	3.9	1.2	-6.358	6.4	2.8	5.6	2.2	-27.902	
34	6.0	2.7	5.1	1.6	-18.965	6.3	2.8	5.1	1.5	-16.102	
35	5.4	3.0	4.5	1.5	-14.005	6.1	2.6	5.6	1.4	-20.175	
36	6.0	3.4	4.5	1.6	-10.878	7.7	3.0	6.1	2.3	-26.887	
37	6.7	3.1	4.7	1.5	-10.218	6.3	3.4	5.6	2.4	-27.387	
38	6.3	2.3	4.4	1.3	-11.535	6.4	3.1	5.5	1.8	-20.577	
39	5.6	3.0	4.1	1.3	-8.028	6.0	3.0	4.8	1.8	-17.678	
40	5.5	2.5	4.0	1.3	-10.476	6.9	3.1	5.4	2.1	-21.817	
41	5.5	2.6	4.4	1.2	-11.468	6.7	3.1	5.6	2.4	-27.638	
42	6.1	3.0	4.6	1.4	-10.974	6.9	3.1	5.1	2.3	-22.204	
43	5.8	2.6	4.0	1.2	-7.613	5.8	2.7	5.1	1.9	-23.392	
44	5.0	2.3	3.3	1.0	-4.775	6.8	3.2	5.9	2.3	-27.577	
45	5.6	2.7	4.2	1.3	-10.399	6.7	3.3	5.7	2.5	-28.458	
46	5.7	3.0	4.2	1.2	-7.131	6.7	3.0	5.2	2.3	-24.170	
47	5.7	2.9	4.2	1.3	-8.928	6.3	2.5	5.0	1.9	-22.033	
48	6.2	2.9	4.3	1.3	-7.847	6.5	3.0	5.2	2.0	-21.165	
49	5.1	2.5	3.0	1.1	-2.452	6.2	3.4	5.4	2.3	-25.110	
50	5.7	2.8	4.1	1.3	-8.788	5.9	3.0	5.1	1.8	-20.124	
最大値B -2.006						最大値C -16.102					
最小値B -18.965						最小値C -34.694					

次に、同じ問題を、「解析的な方法」による解と同様に、「相関比を最大にする」ことを目的にして、「遺伝的アルゴリズムを利用した計算的な『判別分析』の方法」（以下、「計算的な判別分析」あるいは「計算的な方法」という）によって解いてみた。遺伝的アルゴリズムの“戦略”は、1つの世代での染色体の数は20個、「優先生き残り」の染色体の数は2個、「ランダム交叉」により生成される染色体の数は10個（5組）、交叉率は100%つまり交叉する遺伝子の数は4個のうちの4個（つまり全部）、交叉の親を決める「トーナメント」の数は2個（つまり2つの染色体のうちよりよいものを勝者とする）、「突然変異」により生成される染色体の数は5個、突然変異の割合は25%つまり突然変異をする遺伝子の数は4個のうちの1個、そして、「ランダム生き残り」の染色体の数は3個とした。

図 2-4-1 は、「計算的な方法」を適用した計算プロセスの例示である。つまり、第1世代の20個の染色体の遺伝子（つまりパラメーター）は、「解析的な方法」で求めた最適解の取り入れはせず、まったく“ランダムに”生成した。したがって、この例示の解の（「分析の目的」への）適合度は必ずしも高くない。図中のそれぞれの遺伝子は $[0]$ から $[20000]$ の数値で表し、遺伝子 $[0]$ はパラメーターの値 $[-1.00000]$ に、遺伝子 $[20000]$ はパラメーターの値 $[+1.00000]$ に、その間には10000分の1ずつのきざみの値に対応させた。生成した20個の染色体（判別式のパラメーターの並び）を表 2-4-2 のデータに適用して、相関比の値と重なり数（判別得点の値が、B群の最小値より大きいC群の対象の数と、C群の最大値より小さいB群の対象の数の合計）を計算した。ちなみに、第1世代の染色体の番号は $[\#nn]$ 、第2世代の染色体の番号は $[\ast nn]$ と表記してある。

まずは、第1世代の染色体を目的関数の値の大きい順に並べて、「優先生き残り」として、上位の2個の染色体（ $[\#15]$ と $[\#04]$ ）を、次の第2世代（ $[\ast 01]$ と $[\ast 02]$ ）に生き残らせた。ちなみに、この例では、作成と説明の便宜のため、適合度の大きさとして、相関比と重なり数の値の両方を“参考に”あるいは“恣意的に”使っている。このため、必ずしも“機械的に”この2つの値のいずれかが大きい染色体が適合度が高いとは評価していないので注意が必要である。

次は、「トーナメント」と「ランダム交叉」である。“ランダムに”選択した2つの染色体 $[\#06]$ と $[\#09]$ のトーナメントでは評価が高い染色体 $[\#06]$ が勝ち残り、 $[\#20]$ と $[\#06]$ のトーナメントでは $[\#20]$ が勝ち残った。勝ち残った2つの染色体 $[\#06]$ と $[\#20]$ の遺伝子は、“ランダムに”生成したマスクパターン $[0100]$ で交叉させて、2つの新しい

染色体 $[*03]$ と $[*04]$ を生成した。ちなみに、 $[*03]$ は $[\#06]$ をもとに、 $[*04]$ は $[\#020]$ をもとに生成した。「トーナメント」はこのあと 8 回繰り返し、8 個つまり 4 組の親の染色体の組み合わせを作って、「ランダム交叉」によって 8 個の新しい染色体 ($[*05]$ $[*06]$ $[*07]$ $[*08]$ $[*09]$ $[*10]$ $[*11]$ $[*12]$) を生成した。

次は「突然変異」である。すべての染色体の中から“ランダムに” 5 個の染色体 ($[\#11]$ $[*05]$ $[\#19]$ $[\#08]$ $[\#12]$) を選択し、そのそれぞれについて“ランダムに” 選択した遺伝子つまりパラメーター 1 個を“ランダムな” 値に置き換えて、新しい染色体 5 個 ($[*13]$ $[*14]$ $[*15]$ $[*16]$ $[*17]$) を生成した。 $[\#11]$ は第 4 パラメーターの $[60990]$ が $[14981]$ に変異した。 $[\#08]$ と $[\#12]$ はたまたま変異せず、そのまま次の世代の染色体 ($[*16]$ と $[*17]$) になった。

最後は「ランダム生き残り」である。優先生き残りをした $[\#15]$ と $[\#04]$ を除いた残りの 18 個の染色体の中から、ランダムに選択した $[\#12]$ $[\#08]$ $[\#07]$ の 3 個は、そのまま第 2 世代 ($[*18]$ $[*19]$ $[*20]$) に生き残らせた。

第 1 世代でもっとも評価が高かった (つまり「重なり数」が小さかった) 染色体 $[\#15]$ は、第 2 世代でもっとも評価が高かったが、第 1 世代で 2 番目に評価が高く優先生き残りをした染色体 $[\#04]$ は、第 2 世代では 5 番目の評価となり、遺伝的操作で新たに生成された染色体が 2 ~ 4 番目を占めた。

第 2 世代の 20 個の染色体についても、同じ手順と方法で「優先生き残り」「トーナメント」「ランダム交叉」「突然変異」「ランダム生き残り」の操作を加え、それ以降の世代についても同様の操作を繰り返した。このプロセスは約 200 世代で解の改善が見られなくなった。当然のことながら、1 つの世代に染色体の数を増やせば、より少ない世代数で、目的関数の値の大きいつまり適合度のより高い染色体を得ることができる。

【第1世代】

染色体#	遺伝子1	遺伝子2	遺伝子3	遺伝子4	1-相関比	重なり数	遺伝的操作へ
#01	047662	095802	132938	171604	0.60830	100	→ トーナメント
#02	121002	097390	166586	083164	0.51943	100	→ トーナメント
#03	023569	042537	178765	171935	0.64961	100	→ ランダム交叉 (*11/*12)
#04	059395	032635	063339	183013	0.09414	92	→ 優先生き残り (*02) & ランダム交叉 (*05/*06)
#05	136938	045375	132068	021108	0.00017	93	→ 突然変異 (*14)
#06	167312	086731	155462	191503	0.59101	100	→ ランダム交叉 (*03/*04) & トーナメント
#07	012772	006744	144022	125655	0.00011	98	→ ランダム生き残り (*20)
#08	089788	159279	044974	193418	0.00002	91	→ 突然変異 (*16) & ランダム交叉 (*09/*10) & ランダム生き残り (*19)
#09	164648	042613	113442	080257	0.17608	99	→ トーナメント
#10	161618	047203	151056	059576	0.31097	100	→ ランダム交叉 (*07/*08)
#11	176532	180474	096666	060990	0.09094	100	→ 突然変異 (*13)
#12	195673	157085	049703	090108	0.00052	99	→ 突然変異 (*17) & ランダム交叉 (*09/*10) & ランダム生き残り (*18)
#13	177101	019605	116428	068768	0.12423	99	→ トーナメント
#14	093563	061070	173750	168593	0.73643	100	→ ランダム交叉 (*11/*12)
#15	122939	144229	020340	045143	0.75341	10	→ 優先生き残り (*01)
#16	102436	022544	175175	194439	0.73878	100	→ トーナメント
#17	170049	140107	090158	128132	0.29470	100	→ *なし*
#18	065410	181059	147125	025624	0.00352	100	→ ランダム交叉 (*05/*06)
#19	107463	070720	088144	130803	0.09062	98	→ トーナメント & 突然変異 (*15)
#20	007188	093737	119731	127171	0.03637	96	→ ランダム交叉 (*03/*04/*07/*08)

【ランダム交叉】

染色体#	遺伝子1	遺伝子2	遺伝子3	遺伝子4	染色体#	遺伝子1	遺伝子2	遺伝子3	遺伝子4
#06>#09	167312	086731	155462	191503	#20>#06	007188	093737	119731	127171
⇒ *03	167312	093737	155462	191503	⇒ *04	007188	086731	119731	127171
#18>#01	065410	181059	147125	025624	#04>#16	059395	032635	063339	183013
⇒ *05	059395	032635	147125	183013	⇒ *06	065410	181059	063339	025624
#10>#19	161618	047203	151056	059576	#20>#13	007188	093737	119731	127171
⇒ *07	161618	047203	119731	059576	⇒ *08	007188	093737	151056	127171
#12>#09	195673	157085	049703	090108	#08>#13	089788	159279	044974	193418
⇒ *09	089788	157085	049703	090108	⇒ *10	195673	159279	044974	193418
#14>#09	093563	061070	173750	168593	#03>#02	023569	042537	178765	171935
⇒ *11	093563	061070	173750	171935	⇒ *12	023569	142537	178765	168593

【突然変異】

染色体#	遺伝子1	遺伝子2	遺伝子3	遺伝子4	⇒ 染色体#	遺伝子1	遺伝子2	遺伝子3	遺伝子4
#11	176532	180474	096666	060990	⇒ *13	176532	180474	096666	014981
#05	136938	045375	132068	021108	⇒ *14	198114	045375	132068	021108
#19	107463	070720	088144	130803	⇒ *15	107463	070720	088144	093670
#08	089788	159279	044974	193418	⇒ *16	089788	159279	044974	193418
#12	195673	157085	049703	090108	⇒ *17	195673	157085	049703	090108

【第2世代】

染色体#	遺伝子1	遺伝子2	遺伝子3	遺伝子4	1-相関比	重なり数	遺伝的操作から
*01	122939	144229	020340	045143	0.75341	10	← 優先生き残り (#15)
*02	059395	032635	063339	183013	0.09414	92	← 優先生き残り (#04)
*03	167312	093737	155462	191503	0.58683	100	← ランダム交叉 (#06+#20)
*04	007188	086731	119731	127171	0.04060	96	← ランダム交叉 (#06+#20)
*05	059395	032635	147125	183013	0.69728	100	← ランダム交叉 (#18+#04)
*06	065410	181059	063339	025624	0.66017	25	← ランダム交叉 (#18+#04)
*07	161618	047203	119731	059576	0.11200	99	← ランダム交叉 (#10+#20)
*08	007188	093737	151056	127171	0.09817	100	← ランダム交叉 (#10+#20)
*09	089788	157085	049703	090108	0.60074	28	← ランダム交叉 (#12+#08)
*10	195673	159279	044974	193418	0.25764	100	← ランダム交叉 (#12+#08)
*11	093563	061070	173750	171935	0.73803	100	← ランダム交叉 (#14+#03)
*12	023569	142537	178765	168593	0.67989	100	← ランダム交叉 (#14+#03)
*13	176532	180474	096666	014981	0.00058	99	← 突然変異 (#11)
*14	198114	045375	132068	021108	0.09459	99	← 突然変異 (#05)
*15	107463	070720	088144	093670	0.44804	69	← 突然変異 (#19)
*16	089788	159279	044974	193418	0.00002	91	← 突然変異 (#08)
*17	195673	157085	049703	090108	0.00052	99	← 突然変異 (#12)
*18	195673	157085	049703	090108	0.00052	99	← ランダム生き残り (#12)
*19	089788	159279	044974	193418	0.00002	91	← ランダム生き残り (#08)
*20	012772	006744	144022	125655	0.00011	98	← ランダム生き残り (#07)

図 2-4-1 「計算的な方法」の計算のプロセスの例示

表 2-4-3 「計算的な方法」で求めた例題の 2 つの解

イリス・ヴェルシコロール (B 群)							イリス・ヴィルジニカ (C 群)						
	がく の長さ	がく の幅	花卉 の長さ	花卉 の幅	判別 得点	判別 得点	がく 長さ	がく の幅	花卉 の長さ	花卉 の幅	判別 得点	判別 得点	
##	x_1	x_2	x_3	x_4	f	g	x_1	x_2	x_3	x_4	f	g	
01	7.0	3.2	4.7	1.4	-0.580	-2.446	6.3	3.3	6.0	2.5	-2.522	-4.106	
02	6.4	3.2	4.5	1.5	-0.736	-2.465	5.8	2.7	5.1	1.9	-1.845	-3.297	
03	6.9	3.1	4.9	1.5	-0.860	-2.674	7.1	3.0	5.9	2.1	-1.983	-3.724	
04	5.5	2.3	4.0	1.3	-0.914	-2.301	6.3	2.9	5.6	1.8	-1.794	-3.431	
05	6.5	2.8	4.6	1.5	-0.939	-2.598	6.5	3.0	5.8	2.2	-2.194	-3.804	
06	5.7	2.8	4.5	1.3	-0.913	-2.480	7.6	3.0	6.6	2.1	-2.228	-4.096	
07	6.3	3.3	4.7	1.6	-0.928	-2.650	4.9	2.5	4.5	1.7	-1.660	-2.926	
08	4.9	2.4	3.3	1.0	-0.361	-1.700	7.3	2.9	6.3	1.8	-1.898	-3.747	
09	6.6	2.9	4.6	1.3	-0.672	-2.419	6.7	2.5	5.8	1.8	-1.968	-3.592	
10	5.2	2.7	3.9	1.4	-0.865	-2.264	7.2	3.6	6.1	2.5	-2.192	-4.004	
11	5.0	2.0	3.5	1.0	-0.619	-1.894	6.5	3.2	5.1	2.0	-1.526	-3.187	
12	5.9	3.0	4.2	1.5	-0.799	-2.379	6.4	2.7	5.3	1.9	-1.787	-3.352	
13	6.0	2.2	4.0	1.0	-0.525	-2.046	6.8	3.0	5.5	2.1	-1.848	-3.514	
14	6.1	2.9	4.7	1.4	-0.964	-2.610	5.7	2.5	5.0	2.0	-2.004	-3.363	
15	5.6	2.9	3.6	1.3	-0.402	-1.920	5.8	2.8	5.1	2.4	-2.289	-3.644	
16	6.7	3.1	4.4	1.4	-0.543	-2.317	6.4	3.2	5.3	2.3	-1.957	-3.541	
17	5.6	3.0	4.5	1.5	-1.048	-2.596	6.5	3.0	5.5	1.8	-1.639	-3.327	
18	5.8	2.7	4.1	1.0	-0.416	-2.025	7.7	3.8	6.7	2.2	-2.001	-4.052	
19	6.2	2.2	4.5	1.5	-1.232	-2.697	7.7	2.6	6.9	2.3	-2.736	-4.499	
20	5.6	2.5	3.9	1.1	-0.548	-2.040	6.0	2.2	5.0	1.5	-1.563	-3.024	
21	5.9	3.2	4.8	1.8	-1.334	-2.924	6.9	3.2	5.7	2.3	-2.037	-3.730	
22	6.1	2.8	4.0	1.3	-0.526	-2.130	5.6	2.8	4.9	2.0	-1.845	-3.250	
23	6.3	2.5	4.9	1.5	-1.292	-2.867	7.7	2.8	6.7	2.0	-2.245	-4.114	
24	6.1	2.8	4.7	1.2	-0.813	-2.484	6.3	2.7	4.9	1.8	-1.497	-3.045	
25	6.4	2.9	4.3	1.3	-0.563	-2.258	6.7	3.3	5.7	2.1	-1.854	-3.584	
26	6.6	3.0	4.4	1.4	-0.615	-2.350	7.2	3.2	6.0	1.8	-1.629	-3.512	
27	6.8	2.8	4.8	1.4	-0.867	-2.613	6.2	2.8	4.8	1.8	-1.426	-2.975	
28	6.7	3.0	5.0	1.7	-1.210	-2.925	6.1	3.0	4.9	1.8	-1.421	-3.005	
29	6.0	2.9	4.5	1.5	-0.980	-2.572	6.4	2.8	5.6	2.1	-2.103	-3.661	
30	5.7	2.6	3.5	1.0	-0.158	-1.690	7.2	3.0	5.8	1.6	-1.412	-3.285	
31	5.5	2.4	3.8	1.1	-0.565	-2.011	7.4	2.8	6.1	1.9	-1.902	-3.708	
32	5.5	2.4	3.7	1.0	-0.412	-1.877	7.9	3.8	6.4	2.0	-1.584	-3.699	
33	5.8	2.7	3.9	1.2	-0.501	-2.050	6.4	2.8	5.6	2.2	-2.201	-3.735	
34	6.0	2.7	5.1	1.6	-1.496	-3.054	6.3	2.8	5.1	1.5	-1.270	-2.926	
35	5.4	3.0	4.5	1.5	-1.104	-2.618	6.1	2.6	5.6	1.4	-1.591	-3.222	
36	6.0	3.4	4.5	1.6	-0.858	-2.541	7.7	3.0	6.1	2.3	-2.121	-3.926	
37	6.7	3.1	4.7	1.5	-0.806	-2.574	6.3	3.4	5.6	2.4	-2.160	-3.767	
38	6.3	2.3	4.4	1.3	-0.910	-2.456	6.4	3.1	5.5	1.8	-1.623	-3.317	
39	5.6	3.0	4.1	1.3	-0.633	-2.205	6.0	3.0	4.8	1.8	-1.394	-2.955	
40	5.5	2.5	4.0	1.3	-0.826	-2.260	6.9	3.1	5.4	2.1	-1.721	-3.420	
41	5.5	2.6	4.4	1.2	-0.904	-2.409	6.7	3.1	5.6	2.4	-2.180	-3.785	
42	6.1	3.0	4.6	1.4	-0.865	-2.528	6.9	3.1	5.1	2.3	-1.751	-3.384	
43	5.8	2.6	4.0	1.2	-0.600	-2.132	5.8	2.7	5.1	1.9	-1.845	-3.297	
44	5.0	2.3	3.3	1.0	-0.377	-1.709	6.8	3.2	5.9	2.3	-2.175	-3.863	
45	5.6	2.7	4.2	1.3	-0.820	-2.329	6.7	3.3	5.7	2.5	-2.244	-3.878	
46	5.7	3.0	4.2	1.2	-0.562	-2.181	6.7	3.0	5.2	2.3	-1.906	-3.489	
47	5.7	2.9	4.2	1.3	-0.704	-2.276	6.3	2.5	5.0	1.9	-1.738	-3.222	
48	6.2	2.9	4.3	1.3	-0.619	-2.281	6.5	3.0	5.2	2.0	-1.669	-3.290	
49	5.1	2.5	3.0	1.1	-0.193	-1.546	6.2	3.4	5.4	2.3	-1.980	-3.583	
50	5.7	2.8	4.1	1.3	-0.693	-2.235	5.9	3.0	5.1	1.8	-1.587	-3.149	
最大値 B					-0.158	-1.546	最大値 C					-1.270	-2.926
最小値 B					-1.496	-3.054	最小値 C					-2.736	-4.499

結果は、(3)式と(4)式および表 2-4-3 の $[f]$ であり、この解は、「解析的な方法」で求めた解 ((1)式と(2)式および表 2-4-2) と定数倍 (約 0.08 倍) であることを除いてほぼ同じであった。(この結果は、「計算的な方法」で得られる解が、「解析的な方法」で得られる解に十分に“近い”レベル、つまり、“設定された目的を十分に満たしている”ものである、という証左でもある。) また、この解では、B群のサンプルの判別得点 $[f_k]$ の最小値 $[-1.496]$ とC群のサンプルの判別得点 $[f_k]$ の最大値 $[-1.270]$ の間で、「誤判別数の合計を最小にする」判別値 $[f_0]$ を探すと、 $[-1.394 \sim -1.334]$ の間の値で、B群の誤判別数 $[n_1]$ は1個、C群の誤判別数 $[n_0]$ は1個で、誤判別数の合計 $[n_0 + n_1]$ は2個とすることができた。ちなみに、(3)式の判別値 $[f_0 = 1.364]$ は、この範囲の中点を選んである。

$$(3) \quad f_k = 0.28047 x_{1k} + 0.44003 x_{2k} - 0.54971 x_{3k} - 0.97695 x_{4k}$$

$$(4) \quad \text{if } f_k \geq 1.364 (= f_0) \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

次に、「誤判別数の合計 $[n_1 + n_0]$ を最小にする」ことを目的に、「計算的な方法」を適用した。遺伝的アルゴリズムの手順とプロセスと計算の“戦略”は、上記とほぼ同様である。結果を(5)式と(6)式および表 2-4-3 の $[g]$ の項に示す。これによると、B群の誤判別数 $[n_1]$ は1個、C群の誤判別数 $[n_0]$ は0個で、誤判別数 $[n_0 + n_1]$ の合計は1個であった。したがって、「誤判別数の合計を最小にする」という立場からすると、この解は、「相関比を最大にする」ことを目的としている「解析的な方法」による解よりも“優れている”ことが分かる。

$$(5) \quad g_k = 0.11187 x_{1k} + 0.20957 x_{2k} - 0.61082 x_{3k} - 0.73500 x_{4k}$$

$$(6) \quad \text{if } g_k \geq 2.92573 (= g_0) \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

この結果は、当然といえば当然である。繰り返しになるが、「解析的な方法」で得られた解は、「相関比を最大にする」を目的として導き出した解であり、「誤判別数の合計を最小にする」を目的にした解ではないのである。これに対して、ここで適用した「計算的な方法」で得られた解は、「誤判別数の合計を最小にする」を目的として導き出した解であり、「相関比を最大にする」を目的とした解ではない。また、(前にも述べたように、) ここで適用した「計算的な方法」で得られた解は、(遺伝的アルゴリズムによる処理が充分には収束していないときに得られる) 単なる“実行可能解”ではなく、「解析的な方法」で得られる“最

適解”と同様なレベルで、十分に設定した目的を満たしている解なのである。

処理をする目的が違えば、当然のことながら、得られる結果も違ってくる。したがって、この場合、「誤判別数の合計を最小にする」という評価の立場からは、「誤判別数の合計を最小にする」を目的とした「計算的な方法」で得られた解は、（「誤判別数の合計を最小にする」とは“別”の）「相関比を最大にする」を目的としている「解析的な方法」で得られた解よりも、「誤判別数の合計を最小にする」という目的をより満たしているという意味で、論理的に“優れている”のである。

さらに、遺伝的アルゴリズムの初期解の1つに、「解析的な方法」で得られた解を取り入れれば、「計算的な方法」で得られる解は、その論理として、「解析的な方法」で得られる解よりも、“常に”かつ“必ず”優れているのである。この立場からは、「誤判別数の合計を最小にする」を目的とした「計算的な方法」は、「相関比を最大にする」を目的としている「解析的な方法」よりも、論理的に“優れた”解を導き出すことができるのである。

もちろん、「相関比を最大にする」という立場からは、「相関比を最大にする」を目的としている「解析的な方法」あるいは「相関比を最大にする」を目的とした「計算的な方法」で得られる解は、（「相関比を最大にする」とは“別”の）「誤判別数の合計を最小にする」を目的とした「計算的な方法」で得られる解よりも、「相関比を最大にする」という目的をより満たしているという意味で、論理的に“優れている”のである。同様に、判別の目的を、たとえば「誤判別数の合計を最小にする」として得られる解と「2つの集団の誤判別数をバランスさせて最小にする」として得られる解（2-2節の(2)式）は、論理的に互いに異なり、「誤判別数の合計を最小にする」の立場からは、前者の解が“優れた”ものであり、「2つの集団の誤判別数をバランスさせて最小にする」の立場からは、後者の解が“優れた”ものであることも、論理的に当然のことなのである。

再度の繰り返しになるが、判別分析の目的を“柔軟に”設定することができ、設定した目的に対応して、それぞれに異なった“最適”あるいはこれに近い解を得ることができるのが、遺伝的アルゴリズムを利用した「計算的な方法」の特徴の1つである。遺伝的アルゴリズムを利用した「計算的な方法」の問題点をあげるとすれば、それは得られた解が“十分に収束した”ものであるかいなかの判断であろう。実際のところ、この問題点に対しては、それぞれの問題で、解の改善の進行状況から“主観的に”判断をする以外には決定的な方法がない。そこで、収束をさせるのに“十分な”世代数の改善をすることで、得られる解に解の改善についての“充分さ”を与えるという方法を採用するのが普通の方法になっているようである。

2-5節 「第2目的関数」による解の改善

「判別分析の目的」として「誤判別数の合計を最小にする」（2-2節の(1)式）を選択した場合、この目的関数（誤判別数の合計）は“離散的”であることから、解（判別式のパラメーター）が多少異なっても目的関数（つまり適合度）の値はまったく同じである、といった状況が起こる。また、複数の異なった解の目的関数の値がまったく同じになり、解の収束の判断がつきにくいだけでなく、それらの解が本当によい、つまり「印象の推定モデル」として“判別力の高い”解であるかどうかの判断もつきにくい。図2-5-1の例では、2つの集団は、2つの集団を完全に判別する2本の破線の中のいずれの判別式であっても、“完全に”判別ができる、つまり、目的関数つまり「誤判別の数の合計」の値はゼロで、変わらない。

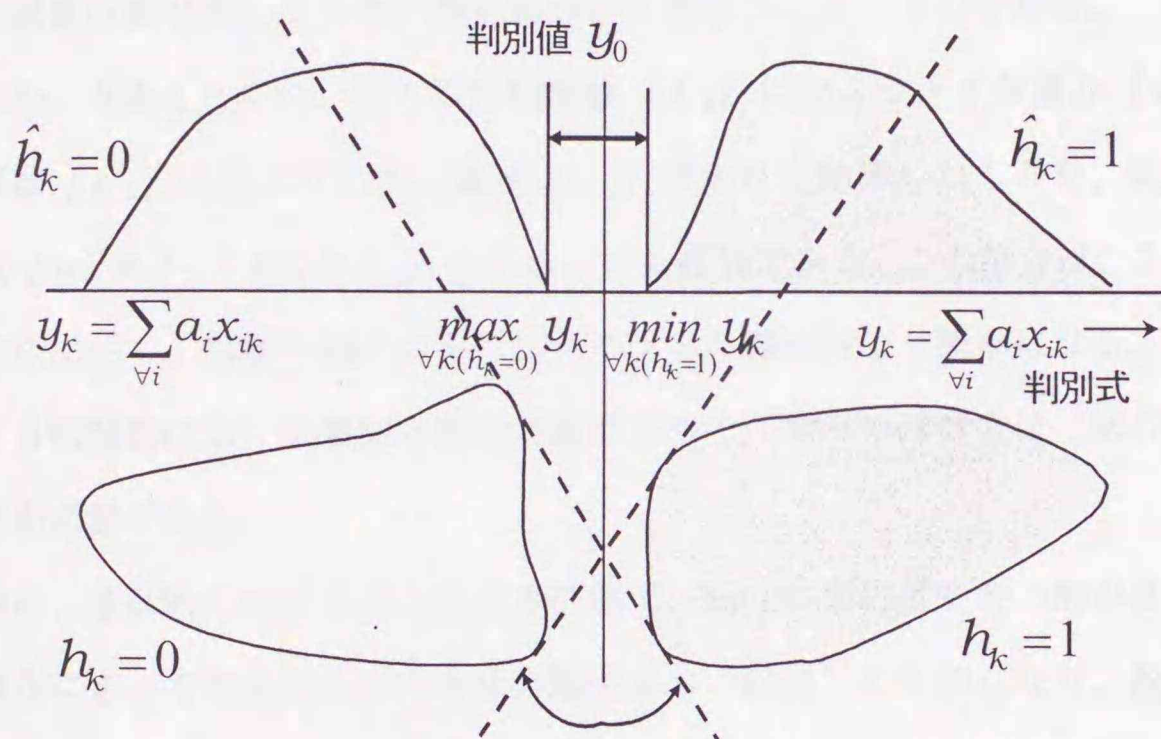


図2-5-1 2つの集団を“完全に判別できる”場合の「解の自由度」

このような場合には、第1の目的関数 $[f_k]$ のほかに第2の目的関数 $[g_k]$ を設け、第1の目的関数の値が同じ場合には、第2の目的関数の値を計算して、これによってより適合度（評価）の高い解を選択する、といった工夫が考えられる。ちなみに、第1の目的関数が“連続的”である場合には、複数の異なった解の目的関数の値がまったく同じになる可能性は小さいので、第2の目的関数を設ける意味は少ない。

具体的には、第1の目的関数 $[f_k]$ が「誤判別数の合計 $[n_1 + n_0]$ を最小にする」の場合、第2の目的関数 $[g_k]$ には、たとえば「『特定の特性を持った対象』の判別得点の最小値と『そうではない対象』の判別得点の最大値の差（つまり最短距離）を最大にする」を考える

ことができる。具体的には、(1)式で表すことができる。この場合には、「誤判別数の合計」が同じでも、「『特定の特性を持ったの対象』の判別得点の最小値と『そうではない対象』の判別得点の最大値の差」がより大きいものが“より判別力が高い”と評価されることになる。

$$(1) \quad \min_{\forall k(h_k=1)} f_k - \max_{\forall k(h_k=0)} f_k \rightarrow \text{最大化}$$

たとえば、判別式のパラメーターの選択によって、「特定の特性を持った対象の集団」と「そうではない対象の集団」が“完全に”判別できた場合には「誤判別数の合計」や「誤判別率」は $[0]$ になる。しかし、「誤判別数の合計」や「誤判別率」が $[0]$ になる解は沢山存在する可能性があるので、「2つの集団の間の最短距離をできるだけ大きくする」という第2の目的を設けることによって、できるだけ安定したつまり少しくらいパラメーターが変化しても、第1の目的関数の値が悪くならない解を見つけてやろう、というのである。

このほか、方法としては、第1の目的関数 $[f_k]$ にはより大きな重み $[w_1]$ をかけ、第2の目的関数 $[g_k]$ にはより小さな重み $[w_2]$ をかけて加算したものを、新たなそして唯一の目的関数 $[w_1 * f_k + w_2 * g_k]$ とすることも有効である。この場合は、2つの目的関数を設けた場合に比べて、処理の流れつまりプログラムの構成はより簡単になる。なお、第1目的関数と第2目的関数は互いに意味の次元が異なるので、重みの付け方は“試行錯誤的”であることに注意が必要である。

ちなみに、遺伝的アルゴリズムによる計算で、同一の適応度を持つ解が多数存在する場合、遺伝的操作によって生まれる次の世代の解全体も“同質”のものになり、解の収束速度が極端に低下する現象が起きることが多い。このような場合、一般には、解の適応度の評価関数になにかのバイアスを加える方法が有効であることが知られているが、上で説明した第2の目的関数を設ける方法は、この「バイアス」の1つといえることができよう。あるいはまた、この方法は、この問題の「多目的最適化問題」という立場からの1つの定式化ということができよう。

2-6節 「感度分析」による解の改善

いままでに説明した「ランダム交叉」や「突然変異」などの遺伝的操作は、遺伝子の交換や置換などによるもので、基本的に“離散的”ということができる。2-3節でも説明したように、これらの操作だけでは、たとえば解を少しだけ“連続的に”変化させたところによりよい解が存在する場合に、これにうまくたどりつける数学的な保証はない。一般に、遺伝的アルゴリズムは、“大局的な”探索には強く、最適解の周辺には早く近づくことができるが、“局所的な”探索には弱く、最適解そのものを見いだすことは苦手である、といわれている。その理由は、ここで述べたように、遺伝的操作が“離散的”である、ということなのである。

この問題点の解決には、「感度分析」の採用が有効である。つまり、“離散的な”遺伝的操作によって見つけられた解に対して、“連続的な”変化を与えることを試みて、その解がもとの解を改善できるかいなかを調べるのである。この場合、たとえば1次式の形式の判別式を採用したとして、1次式のそれぞれのパラメーターを“独立に”増加あるいは減少させて、解の適合度（たとえば、「誤判別数の合計」や「2つの集団の間の最短距離」など）が改善されるかいなかを検討するのである。解が改善される場合には、その解を新しい解として採用する。この方法によって、遺伝的操作によって見つけだされた解を少しだけ変化させたところによりよい解が存在する場合、これにうまくたどりつく可能性を期待するのである（図2-6-1）。

1次式のそれぞれのパラメーターの間に交互作用がある場合には、それぞれのパラメーターを“組み合わせ的に”増加あるいは減少させると、“独立に”増加あるいは減少させる場合とは異なった結果となる。しかし、“独立に”増加あるいは減少させる場合には、パラメーターの数のケースだけの計算で済むのに対して、“組み合わせ的に”増加あるいは減少させる場合には、パラメーターの数の自乗の数のケースだけの計算が必要になる。そこで、ここでは、パラメーターの増加あるいは減少させる「幅」を制限して、それぞれのパラメーターを“独立に”増加あるいは減少させる方法を選択した。

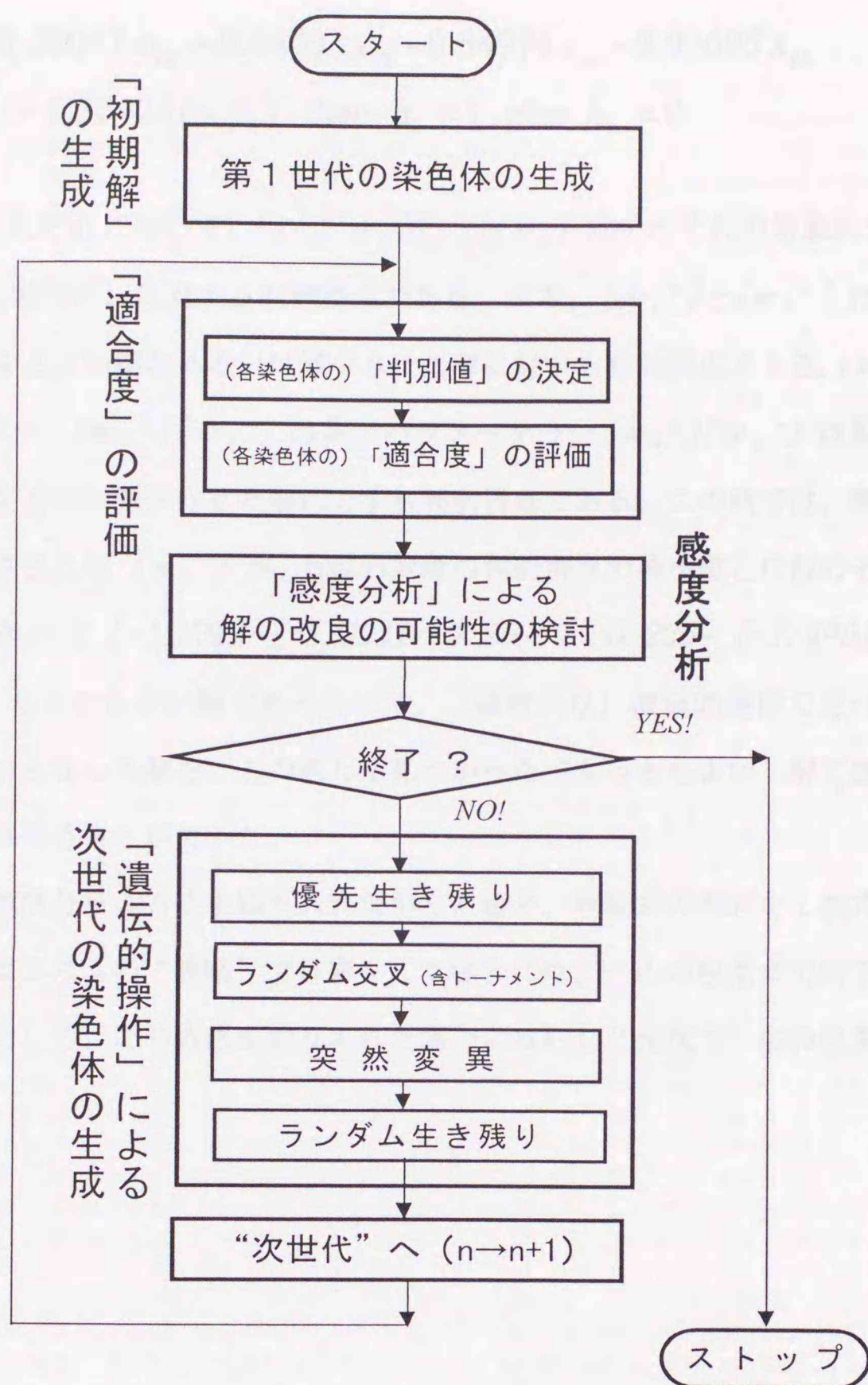


図 2-6-1 「遺伝的アルゴリズム」への「感度分析」の取り入れ

具体的な例示として、前と同じ「あやめ科の3種類の植物の4変量観測値」のB群とC群のデータに対して、計算が収束する以前の途中の解つまり判別式のパラメーターに対して20%の増加あるいは減少を加えて、解の適合度の改善の状況を検討した。表 2-6-1 は、1つの解((1)式と(2)式)に対して、「感度分析」を行った結果である。

$$(1) f_k = 0.28047 x_{1k} + 0.44003 x_{2k} - 0.54971 x_{3k} - 0.97695 x_{4k}$$

$$(2) \text{ if } f_k \geq 1.36423 (= f_0) \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

表中、2つの群それぞれに対して、 $[x_1][x_2][x_3][x_4]$ はそれぞれの対象のデータ、 $[f]$ は基準となる解（判別式）に対する判別得点である。また、 $[w_1^+]$ と $[w_1^-]$ は、それぞれ第1パラメーターを20%増加あるいは減少させた解に対する判別得点である。 $[w_2^+][w_2^-]$ は第2パラメーター、 $[w_3^+][w_3^-]$ は第3パラメーター、 $[w_4^+][w_4^-]$ は第4パラメーターを20%増加あるいは減少させた解に対する判別得点である。この例では、第1パラメーターを20%減少させた解 $[w_1^-]$ が、B群の対象の判別得点の最小値とC群のその最大値の差がもっとも大きい（ $[-1.832]-[-1.623]=-0.209 > -0.226 = [-1.496]-[-1.270]$ ）という意味で、もっともよい解であったので、（離散的な）遺伝的操作で見つけだされ、感度分析の「基準」となった解を、この新しく見つかった“もっともよい”解で置き換えて、よりよい解に進化させることができた。

ちなみに、この感度分析の方法を取り入れなかった場合、判別式の形式を1次式として、前と同じ遺伝的アルゴリズムの“戦略”で計算をした場合には、一応の収束まで約200世代の計算を要したのに対して、この方法を取り入れた場合には約20世代で一応の収束をさせることができた。

表 2-6-1 解の「感度分析」の例示

イリス・ヴェルシコロール (B群)														イリス・ヴィルジニカ (C群)													
パラメーターを増減したときの判別得点														パラメーターを増減したときの判別得点													
##	x ₁	x ₂	x ₃	x ₄	f	w ₁ ⁺	w ₁ ⁻	w ₂ ⁺	w ₂ ⁻	w ₃ ⁺	w ₃ ⁻	w ₄ ⁺	w ₄ ⁻	x ₁	x ₂	x ₃	x ₄	f	w ₁ ⁺	w ₁ ⁻	w ₂ ⁺	w ₂ ⁻	w ₃ ⁺	w ₃ ⁻	w ₄ ⁺	w ₄ ⁻	
01	70	32	47	14	-0.580	-0.187	-0.973	-0.298	-1.009	-2.279	-0.442	-0.979	-1.743	63	33	60	25	-2.522	-2.168	-2.875	-2.231	-1.130	-2.665	-0.321	-0.811	-2.175	
02	64	32	45	15	-0.736	-0.377	-1.095	-0.454	-0.941	-2.172	-0.414	-0.884	-1.702	58	27	51	19	-1.845	-1.520	-2.170	-1.607	-1.026	-2.319	-0.326	-0.804	-1.841	
03	69	31	49	15	-0.860	-0.473	-1.247	-0.587	-1.068	-2.366	-0.451	-0.999	-1.818	71	30	59	21	-1.983	-1.585	-2.382	-1.719	-1.094	-2.577	-0.271	-0.851	-1.997	
04	55	23	40	13	-0.914	-0.606	-1.223	-0.712	-0.726	-1.760	-0.197	-0.624	-1.333	63	29	56	18	-1.794	-1.440	-2.147	-1.539	-1.518	-2.931	-0.743	-1.346	-2.328	
05	65	28	46	15	-0.939	-0.574	-1.304	-0.693	-0.819	-2.025	-0.228	-0.717	-1.536	65	30	58	22	-2.194	-1.830	-2.559	-1.930	-1.124	-2.587	-0.321	-0.854	-2.054	
06	57	28	45	13	-0.913	-0.593	-1.233	-0.667	-1.346	-2.533	-0.775	-1.299	-2.008	76	30	66	21	-2.228	-1.802	-2.654	-1.964	-1.558	-3.177	-0.598	-1.315	-2.461	
07	63	33	47	16	-0.928	-0.574	-1.281	-0.637	-1.088	-2.369	-0.532	-1.014	-1.887	49	25	45	17	-1.660	-1.385	-1.935	-1.440	-1.020	-2.174	-0.416	-0.831	-1.759	
08	49	24	33	10	-0.361	-0.086	-0.635	-0.149	-0.759	-1.668	-0.378	-0.750	-1.296	73	29	63	18	-1.898	-1.489	-2.308	-1.643	-1.762	-3.312	-0.850	-1.590	-2.572	
09	66	29	46	13	-0.672	-0.301	-1.042	-0.416	-1.092	-2.309	-0.512	-1.056	-1.765	67	25	58	18	-1.968	-1.592	-2.343	-1.748	-1.362	-2.770	-0.504	-1.146	-2.128	
10	52	27	39	14	-0.865	-0.573	-1.157	-0.627	-0.799	-1.858	-0.334	-0.714	-1.478	72	36	61	25	-2.192	-1.788	-2.596	-1.875	-0.964	-2.551	-0.168	-0.677	-2.042	
11	50	20	35	10	-0.619	-0.338	-0.899	-0.443	-0.734	-1.638	-0.271	-0.682	-1.227	65	32	51	20	-1.526	-1.162	-1.891	-1.245	-0.801	-2.149	-0.156	-0.607	-1.699	
12	59	30	42	15	-0.799	-0.468	-1.130	-0.535	-0.780	-1.930	-0.289	-0.701	-1.519	64	27	53	19	-1.787	-1.428	-2.146	-1.549	-0.957	-2.289	-0.218	-0.735	-1.772	
13	60	22	40	10	-0.525	-0.188	-0.861	-0.331	-0.871	-1.894	-0.331	-0.840	-1.386	68	30	55	21	-1.848	-1.466	-2.229	-1.584	-0.835	-2.240	-0.091	-0.592	-1.738	
14	61	29	47	14	-0.964	-0.622	-1.307	-0.709	-1.273	-2.510	-0.673	-1.210	-1.974	57	25	50	20	-2.004	-1.684	-2.323	-1.784	-0.748	-1.999	-0.045	-0.477	-1.568	
15	56	29	36	13	-0.402	-0.088	-0.716	-0.147	-0.555	-1.577	-0.170	-0.519	-1.228	58	28	51	24	-2.289	-1.964	-2.615	-2.043	-0.387	-1.692	0.301	-0.040	-1.350	
16	67	31	44	14	-0.543	-0.167	-0.919	-0.270	-0.804	-2.004	-0.285	-0.763	-1.527	64	32	53	23	-1.957	-1.598	-2.316	-1.676	-0.631	-2.019	0.053	-0.355	-1.611	
17	56	30	45	15	-1.048	-0.734	-1.363	-0.784	-1.205	-2.414	-0.656	-1.126	-1.944	65	30	55	18	-1.639	-1.274	-2.003	-1.375	-1.377	-2.781	-0.632	-1.215	-2.198	
18	58	27	41	10	-0.416	-0.091	-0.741	-0.178	-1.276	-2.374	-0.772	-1.300	-1.846	77	38	67	22	-2.001	-1.569	-2.433	-1.666	-1.827	-3.554	-0.936	-1.645	-2.845	
19	62	22	45	15	-1.232	-0.884	-1.580	-1.039	-0.589	-1.710	0.048	-0.422	-1.240	77	26	69	23	-2.736	-2.304	-3.168	-2.507	-1.358	-2.992	-0.296	-1.017	-2.272	
20	56	25	39	11	-0.548	-0.234	-0.862	-0.328	-0.945	-1.982	-0.458	-0.919	-1.520	60	22	50	15	-1.563	-1.227	-1.900	-1.369	-1.166	-2.385	-0.431	-0.998	-1.817	
21	59	32	48	18	-1.334	-1.003	-1.665	-1.053	-1.045	-2.335	-0.459	-0.906	-1.888	69	32	57	23	-2.037	-1.650	-2.424	-1.755	-0.802	-2.267	-0.040	-0.526	-1.781	
22	61	28	40	13	-0.526	-0.184	-0.868	-0.280	-0.681	-1.771	-0.208	-0.635	-1.344	56	28	49	20	-1.845	-1.531	-2.159	-1.598	-0.826	-2.091	-0.176	-0.588	-1.679	
23	63	25	49	15	-1.292	-0.939	-1.645	-1.072	-1.068	-2.300	-0.385	-0.934	-1.752	77	28	67	20	-2.245	-1.813	-2.677	-1.999	-1.660	-3.277	-0.659	-1.422	-2.514	
24	61	28	47	12	-0.813	-0.471	-1.155	-0.567	-1.502	-2.728	-0.891	-1.482	-2.137	63	27	49	18	-1.497	-1.144	-1.850	-1.259	-0.747	-2.001	-0.086	-0.552	-1.535	
25	64	29	43	13	-0.563	-0.204	-0.922	-0.307	-0.887	-2.046	-0.365	-0.851	-1.560	67	33	57	21	-1.854	-1.478	-2.230	-1.563	-1.207	-2.683	-0.456	-0.997	-2.143	
26	66	30	44	14	-0.615	-0.245	-0.985	-0.351	-0.804	-1.993	-0.274	-0.752	-1.516	72	32	60	18	-1.629	-1.225	-2.033	-1.348	-1.645	-3.169	-0.825	-1.506	-2.488	
27	68	28	48	14	-0.867	-0.486	-1.248	-0.621	-1.019	-2.264	-0.389	-0.944	-1.708	62	28	48	18	-1.426	-1.078	-1.774	-1.180	-0.737	-1.983	-0.107	-0.554	-1.536	
28	67	30	50	17	-1.210	-0.834	-1.586	-0.946	-0.937	-2.243	-0.290	-0.803	-1.730	61	30	49	18	-1.421	-1.079	-1.763	-1.157	-0.967	-2.254	-0.339	-0.805	-1.788	
29	60	29	45	15	-0.980	-0.644	-1.317	-0.725	-0.985	-2.183	-0.425	-0.895	-1.713	64	28	56	21	-2.103	-1.744	-2.462	-1.856	-1.021	-2.423	-0.235	-0.756	-1.902	
30	57	26	35	10	-0.158	0.162	-0.478	0.071	-0.690	-1.660	-0.292	-0.703	-1.249	72	30	58	16	-1.412	-1.008	-1.816	-1.148	-1.635	-3.098	-0.831	-1.528	-2.401	
31	55	24	38	11	-0.565	-0.256	-0.873	-0.354	-0.847	-1.853	-0.368	-0.811	-1.411	74	28	61	19	-1.902	-1.487	-2.317	-1.655	-1.342	-2.842	-0.458	-1.132	-2.169	
32	55	24	37	10	-0.412	-0.104	-0.721	-0.201	-0.886	-1.873	-0.427	-0.877	-1.422	79	38	64	20	-1.584	-1.141	-2.027	-1.250	-1.719	-3.387	-0.886	-1.591	-2.682	
33	58	27	39	12	-0.501	-0.176	-0.827	-0.264	-0.808	-1.867	-0.343	-0.778	-1.432	64	28	56	22	-2.201	-1.842	-2.560	-1.954	-0.885	-2.287	-0.098	-0.592	-1.793	
34	60	27	51	16	-1.496	-1.159	-1.832	-1.258	-1.347	-2.640	-0.647	-1.207	-2.080	63	28	51	15	-1.270	-0.917	-1.623	-1.023	-1.395	-2.700	-0.707	-1.294	-2.112	
35	54	30	45	15	-1.104	-0.802	-1.407	-0.840	-1.293	-2.502	-0.744	-1.214	-2.032	61	26	56	14	-1.591	-1.249	-1.933	-1.362	-2.020	-3.400	-1.212	-1.924	-2.688	
36	60	34	45	16	-0.858	-0.521	-1.194	-0.559	-1.069	-2.322	-0.563	-1.006	-1.879	77	30	61	23	-2.121	-1.689	-2.552	-1.856	-0.753	-2.274	0.109	-0.455	-1.710	
37	67	31	47	15	-0.806	-0.430	-1.182	-0.533	-0.960	-2.220	-0.383	-0.892	-1.710	63	34	56	24	-2.160	-1.807	-2.513	-1.861	-0.920	-2.388	-0.199	-0.639	-1.948	
38	63	2																									

2-7節 「判別式の形式」の自由な選択

2-6節までの説明では、判別式の形式は「1次式」つまり「線型判別」であった。しかし、対象の複数の属性値 $[x_{ik}]$ とその対象が特定の特性を持つかないか $[h_k=1 \text{ or } 0]$ の対応づけに、「非線形性」や「交互作用」がある場合には、判別式の形式は1次式ではうまくはいかないことが多い。このような場合には、判別式の形式として「2次式」が使える（(1)式）。ちなみに、「解析的な方法」でも、2次式の形式の判別式を採用した場合の方法が開発され利用されている。（2つの集団のデータの分散共分散行列が等しい場合にも、マハラノビスの汎距離の近い方の集団に属すると判別する、という考え方に基づいて、「1次式」の形式の判別式を導いたのと同様に、これらが等しくない場合も、同じ考え方で、「2次式」の形式の判別式が導かれる（文献32）。）

$$(1) f_k = \sum_{\forall i} a_i x_{ik} + \sum_{\forall i} \sum_{\forall j} b_{ij} x_{ik} x_{jk} \quad (-1 \leq a_i, b_{ij} \leq +1)$$



$$y_k = \sum_{\forall i} a_i x_{ik} + \sum_{\forall i,j} b_{ij} x_{ik} x_{jk} \quad (-1 \leq a_i, b_{ij} \leq +1)$$

$$\text{if } y_k \geq y_0 \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

図 2-7-1 「2次式」の判別式による判別分析の概念

概念的に説明すると、「1次式」の形式の判別式による判別分析が、特定の特性の有無に関する2つの集団の対象を1次直線や1次平面で判別することに対応するとすれば、「2次式」の形式の判別式による判別分析は、2つの集団の対象を2次曲線や2次曲面で判別することになり、「非線形性」や「交互作用」により対応しやすい（図2-7-1）。ただし、「1次式」の形式の判別式による判別分析に比べて、「2次式」の形式の判別式による判別分析は、判別式のパラメーターの数が属性値の数の自乗の割合で増えるので、「計算的な方法」に必要な計算

の量や収束までの時間もこれに対応して増えることに注意が必要である。

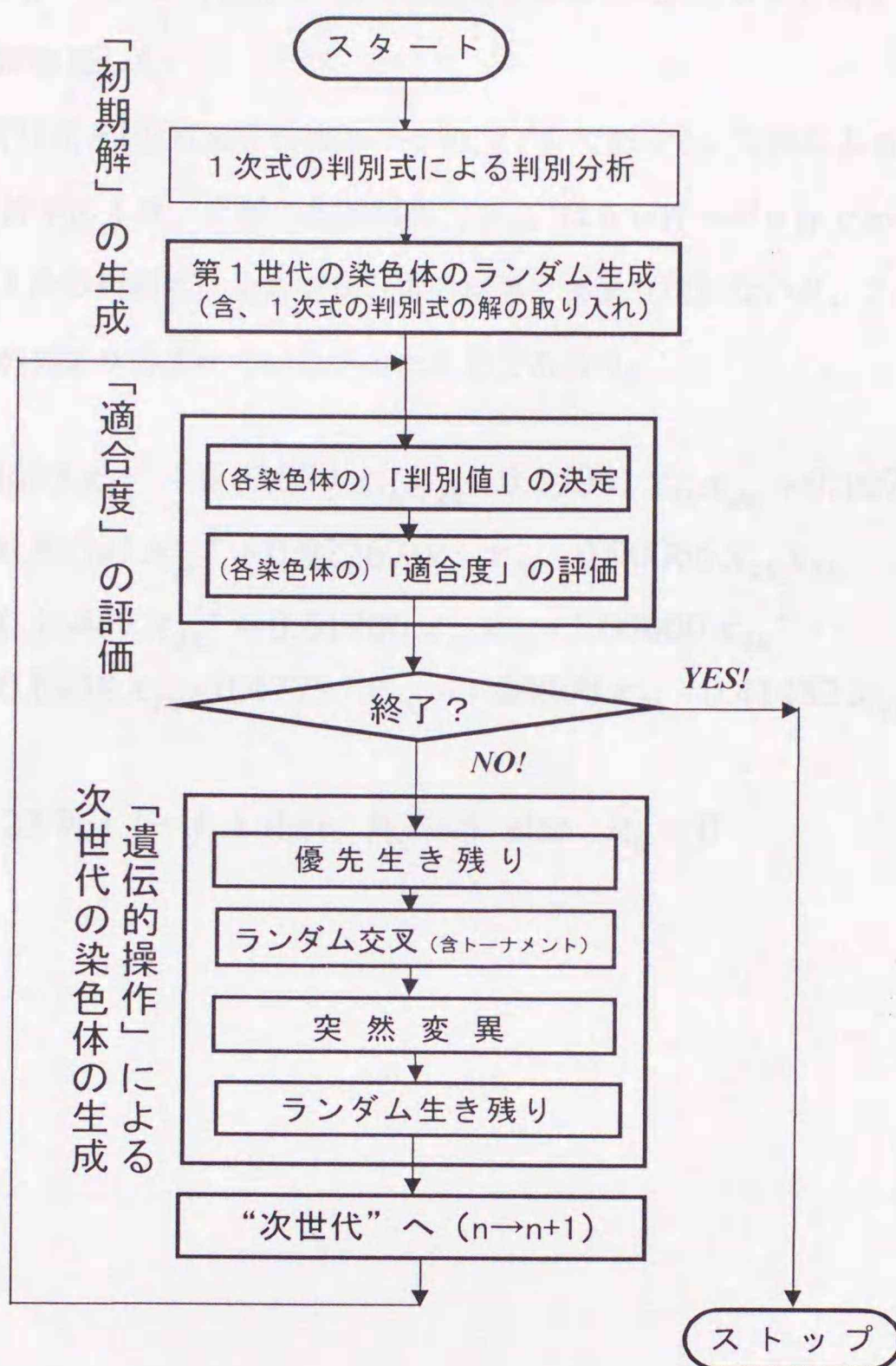


図 2-7-2 「2次式」の判別式による遺伝的アルゴリズムの計算

ちなみに、実際の処理では、まず「1次式」の判別式による判別分析を行って、その「もっとも判別力の高い」解を見つけだし、これを「2次式」の判別式の判別分析の初期解（第1世代の染色体）の1つに取り入れる（この場合、2次式の2次の項の係数の値は $[0]$ である）ことで、必要な計算の量や収束までの時間を少なくつまり効率化をすることができる（図 2-7-2）。

前と同じ「あやめ科の3種類の植物の4変量観測値」のB群とC群のデータについて、「2次式」の形式の判別式による判別分析を「計算的な方法」で行った。遺伝的アルゴリズムの“戦略”はいままでとほぼ同じで、1つの世代で用意する染色体を20個として、「優先生き残り」

「トーナメント」「ランダム交叉」「突然変異」「ランダム生き残り」の遺伝的操作を加えた場合、解は（「1次式」の形式の判別式による判別分析に要した約200世代を含めて）約1000世代までにほぼ収束した。

得られた結果は、(2)式と(3)式および表2-7-1の $[f]$ であった。これによると、B群の誤判別数 $[n_1]$ は50件中の1件、C群の誤判別数 $[n_0]$ は50件中の0件であった。この結果の誤判別の評価は1次の判別式による判別分析の結果とあまり変わらないが、これはデータに非線形性や交互作用があまり含まれていなかったためであろう。

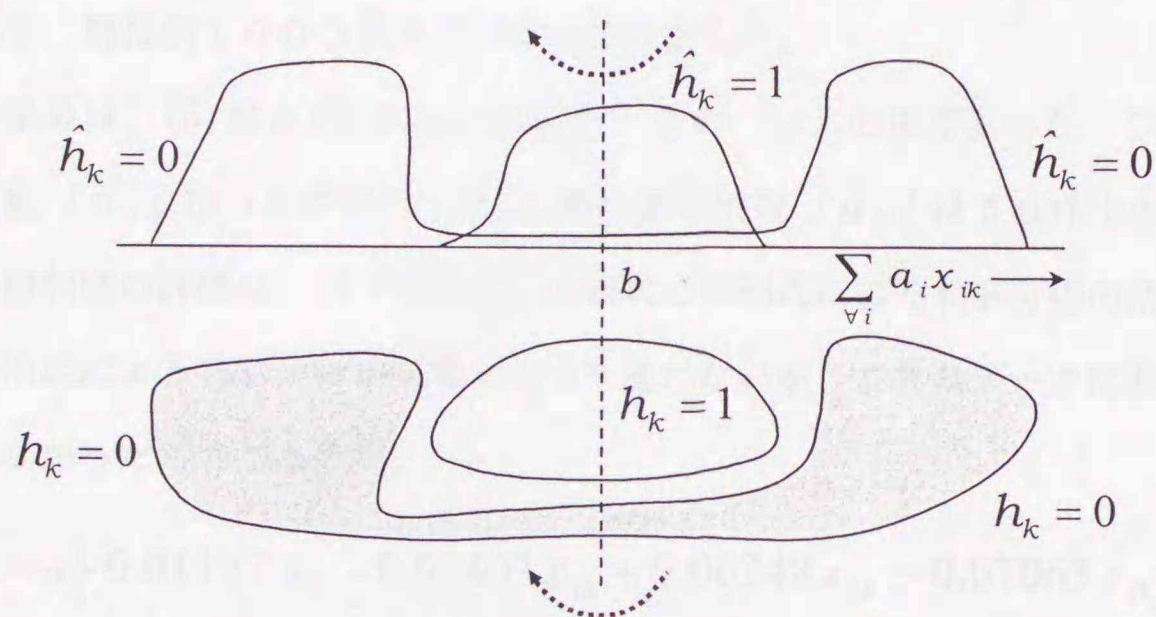
$$(2) \quad f_k = 0.14685 x_{1k}^2 + 0.63079 x_{1k} x_{2k} - 0.85061 x_{1k} x_{3k} + 0.10775 x_{1k} x_{4k} \\ - 0.92341 x_{2k}^2 + 0.92267 x_{2k} x_{3k} - 0.90066 x_{2k} x_{4k} \\ - 0.18407 x_{3k}^2 + 0.61706 x_{3k} x_{4k} - 1.00000 x_{4k}^2 \\ + 0.8938 x_{1k} - 0.47795 x_{2k} + 1.00000 x_{3k} - 0.41452 x_{4k}$$

$$(3) \quad \text{if } f_k \geq -23.961 (= f_0) \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

表 2-7-1 「2 次式」と「“V 型”の絶対値式」による判別分析の結果

イリス・ヴェルシコロール(B群)							イリス・ヴィルジニカ(C群)						
がく の長さ	がく の幅	花弁 の長さ	花弁 の幅	判別 得点	判別 得点		がく の長さ	がく の幅	花弁 の長さ	花弁 の幅	判別 得点	判別 得点	
##	X ₁	X ₂	X ₃	X ₄	f	g	X ₁	X ₂	X ₃	X ₄	f	g	
01	7.0	3.2	4.7	1.4	1.703	0.663	6.3	3.3	6.0	2.5	-5.299	0.497	
02	6.4	3.2	4.5	1.5	1.270	0.662	5.8	2.7	5.1	1.9	-1.982	0.576	
03	6.9	3.1	4.9	1.5	0.662	0.640	7.1	3.0	5.9	2.1	-4.031	0.534	
04	5.5	2.3	4.0	1.3	1.271	0.675	6.3	2.9	5.6	1.8	-2.576	0.562	
05	6.5	2.8	4.6	1.5	0.763	0.647	6.5	3.0	5.8	2.2	-4.081	0.526	
06	5.7	2.8	4.5	1.3	0.868	0.658	7.6	3.0	6.6	2.1	-6.544	0.496	
07	6.3	3.3	4.7	1.6	0.481	0.643	4.9	2.5	4.5	1.7	-0.714	0.613	
08	4.9	2.4	3.3	1.0	2.722	0.736	7.3	2.9	6.3	1.8	-4.949	0.530	
09	6.6	2.9	4.6	1.3	1.433	0.665	6.7	2.5	5.8	1.8	-4.148	0.545	
10	5.2	2.7	3.9	1.4	1.256	0.681	7.2	3.6	6.1	2.5	-5.138	0.509	
11	5.0	2.0	3.5	1.0	2.127	0.715	6.5	3.2	5.1	2.0	-1.540	0.590	
12	5.9	3.0	4.2	1.5	1.255	0.670	6.4	2.7	5.3	1.9	-2.365	0.571	
13	6.0	2.2	4.0	1.0	1.967	0.700	6.8	3.0	5.5	2.1	-2.916	0.556	
14	6.1	2.9	4.7	1.4	0.564	0.645	5.7	2.5	5.0	2.0	-2.324	0.570	
15	5.6	2.9	3.6	1.3	2.481	0.716	5.8	2.8	5.1	2.4	-3.605	0.543	
16	6.7	3.1	4.4	1.4	2.072	0.676	6.4	3.2	5.3	2.3	-3.135	0.555	
17	5.6	3.0	4.5	1.5	0.433	0.648	6.5	3.0	5.5	1.8	-2.127	0.574	
18	5.8	2.7	4.1	1.0	2.313	0.703	7.7	3.8	6.7	2.2	-5.377	0.503	
19	6.2	2.2	4.5	1.5	-0.197	0.635	7.7	2.6	6.9	2.3	-9.657	0.454	
20	5.6	2.5	3.9	1.1	2.194	0.702	6.0	2.2	5.0	1.5	-1.728	0.601	
21	5.9	3.2	4.8	1.8	-0.714	0.616	6.9	3.2	5.7	2.3	-3.838	0.535	
22	6.1	2.8	4.0	1.3	2.336	0.695	5.6	2.8	4.9	2.0	-1.865	0.582	
23	6.3	2.5	4.9	1.5	-0.707	0.618	7.7	2.8	6.7	2.0	-7.427	0.493	
24	6.1	2.8	4.7	1.2	0.841	0.657	6.3	2.7	4.9	1.8	-1.068	0.602	
25	6.4	2.9	4.3	1.3	2.041	0.682	6.7	3.3	5.7	2.1	-3.098	0.549	
26	6.6	3.0	4.4	1.4	1.865	0.673	7.2	3.2	6.0	1.8	-3.134	0.555	
27	6.8	2.8	4.8	1.4	0.636	0.645	6.2	2.8	4.8	1.8	-0.752	0.610	
28	6.7	3.0	5.0	1.7	-0.421	0.615	6.1	3.0	4.9	1.8	-0.856	0.607	
29	6.0	2.9	4.5	1.5	0.699	0.650	6.4	2.8	5.6	2.1	-3.565	0.540	
30	5.7	2.6	3.5	1.0	3.450	0.738	7.2	3.0	5.8	1.6	-2.469	0.577	
31	5.5	2.4	3.8	1.1	2.216	0.705	7.4	2.8	6.1	1.9	-4.720	0.534	
32	5.5	2.4	3.7	1.0	2.612	0.718	7.9	3.8	6.4	2.0	-3.724	0.538	
33	5.8	2.7	3.9	1.2	2.370	0.702	6.4	2.8	5.6	2.2	-3.875	0.533	
34	6.0	2.7	5.1	1.6	-1.176	0.600	6.3	2.8	5.1	1.5	-0.714	0.613	
35	5.4	3.0	4.5	1.5	0.285	0.645	6.1	2.6	5.6	1.4	-2.386	0.581	
36	6.0	3.4	4.5	1.6	0.529	0.655	7.7	3.0	6.1	2.3	-5.251	0.515	
37	6.7	3.1	4.7	1.5	1.029	0.650	6.3	3.4	5.6	2.4	-4.090	0.532	
38	6.3	2.3	4.4	1.3	0.736	0.660	6.4	3.1	5.5	1.8	-2.009	0.575	
39	5.6	3.0	4.1	1.3	1.651	0.687	6.0	3.0	4.8	1.8	-0.714	0.612	
40	5.5	2.5	4.0	1.3	1.487	0.680	6.9	3.1	5.4	2.1	-2.467	0.566	
41	5.5	2.6	4.4	1.2	0.953	0.664	6.7	3.1	5.6	2.4	-4.136	0.530	
42	6.1	3.0	4.6	1.4	0.890	0.654	6.9	3.1	5.1	2.3	-2.416	0.571	
43	5.8	2.6	4.0	1.2	2.065	0.693	5.8	2.7	5.1	1.9	-1.982	0.576	
44	5.0	2.3	3.3	1.0	2.791	0.735	6.8	3.2	5.9	2.3	-4.381	0.521	
45	5.6	2.7	4.2	1.3	1.330	0.673	6.7	3.3	5.7	2.5	-4.597	0.521	
46	5.7	3.0	4.2	1.2	1.822	0.689	6.7	3.0	5.2	2.3	-2.884	0.560	
47	5.7	2.9	4.2	1.3	1.543	0.679	6.3	2.5	5.0	1.9	-1.923	0.584	
48	6.2	2.9	4.3	1.3	1.829	0.679	6.5	3.0	5.2	2.0	-1.943	0.579	
49	5.1	2.5	3.0	1.1	3.230	0.753	6.2	3.4	5.4	2.3	-3.373	0.551	
50	5.7	2.8	4.1	1.3	1.687	0.683	5.9	3.0	5.1	1.8	-1.368	0.592	
最大値 B							最大値 C						
最小値 B							最小値 C						

次に、対象の属性値と特定の特性を持つかないかの対応づけに、「交互作用」はなく「非線型性」だけが問題となる場合には、「絶対値関数」を利用した判別式（以下、「絶対値式」という）が使える。これを概念的に説明すると、（「2次式」の場合の説明と同様に、）「1次式」による判別分析が、特定の特性の有無に関する2つの集団の対象を1次直線や1次平面で判別することに対応するとすれば、「絶対値式」による判別分析は、2つの集団の対象を折り返し直線や折り返し平面で判別することになり、非線型性により対応しやすい、と考えられる（図2-7-3）。



$$y_k = \left| \sum_{\forall i} a_i x_{ik} - b \right| \quad (-1 \leq a_i, b \leq +1)$$

$$\text{if } y_k \geq y_0 \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

図2-7-3 「“V型”の絶対値式」による判別分析の概念

「絶対値式」の形式の判別式のパラメーターの数は、普通の「1次式」とほぼ同じで、「2次式」の場合のように、判別式のパラメーターが属性値の数の自乗の割合で増えることはない。なお、絶対値関数の値は、正でも負でも同じであるので、解は1つだけではなく、解の正負を反転したものも同じ評価となることにも注意が必要である。また、判別すべき対象の並びの順序に応じて、判別式の値を非負（ ≥ 0 ）とするものと非正（ ≤ 0 ）とするものの2つが考えられる。当然のことながら、このいずれの形式を採用かによって解は異なる。

$$(4) f_k = \pm \left| \sum_{\forall i} a_i x_{ik} - b \right| \quad (-1 \leq a_i, b \leq +1)$$

前と同じ「あやめ科の3種類の植物の4変量観測値」のB群とC群のデータについて、この絶対値式の形式の判別式による判別分析を行った。遺伝的アルゴリズムの“戦略”はいままで例とまったく同じである。つまり、1つの世代で用意する染色体を20個として、「優先生き残り」「トーナメント」「ランダム交叉」「突然変異」「ランダム生き残り」の遺伝的操作を加えた場合、解は約1000世代までにほぼ収束した。

得られた結果は、(5)式と(6)式および表2-7-1の $[g]$ の項であった。これによると、B群の誤判別数 $[n_1]$ は50件中の1件、C群の誤判別数 $[n_0]$ は50件中の0件であった。この結果の誤判別の評価は、(「2次式」の形式の判別式による判別分析の結果と同様に、)1次式の判別式による判別分析の結果とあまり違わないが、これはデータに非線形性があまり含まれていなかったためであろう。

$$(5) f_k = + \left| -0.01197 x_{1k} - 0.02403 x_{2k} + 0.06348 x_{3k} + 0.07063 x_{4k} - 0.90000 \right|$$

$$(6) \text{ if } f \geq 0.614 (= f_0) \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

この絶対値式は、(4)式のように“V型”のもののほかに、(7)式のように、“W型”のものも考えられる。概念的にいうと、「“V型”の絶対値式」が折り目が1つの折り返し直線や平面であると考えたと、「“W型”の絶対値式」は折り目が2つの折り返し直線や平面、と考えられる。したがって、「“W型”の絶対値式」は、「“V型”の絶対値式」よりも複雑な非線形性に対応することができる。

$$(7) f_k = \pm \left| \left| \sum_{\forall i} a_i x_{ik} - b \right| - c \right| \quad (-1 \leq a_i, b, c \leq +1)$$

ちなみに、判別分析の「計算的な方法」は、対象の複数の属性値を1つの値に変換できる対応関係(関数関係)があれば、どんな形式の判別式でも採用することができる((8)式)。これは、この「計算的な方法」に“柔軟性”を与えている理由の1つである。

$$(8) f_k = F(x_{1k}, x_{2k}, \dots, x_{mk})$$

2-8節 「判別得点の分布」を考慮した判別分析

いままでに説明してきた判別分析の「計算的な方法」では、「2つの集団に含まれる対象の誤判別数の合計を最小にする」あるいは「(判別得点の値を尺度にした) 2つの集団の間の最短距離を最大にする」などの目的関数を採用してきた。これは、従来の「解析的な方法」が、判別式によって変換された判別得点の値の分布の「全体」に関心を持っているのに対して、「計算的な方法」のこれらの目的関数は、2つの集団の“重なり合っている”「部分」にだけ関心を持っていることを意味している。

つまり、この「計算的な方法」では、判別式で変換された判別得点の値について、判別式のパラメーター(係数)が多少変化しても、判別の結果には影響しない、つまり、2つの集団の“重なり”以外の「部分」には関心を持たず、もっぱら判別式のパラメーターの変化によって判別の結果が違ってくるところ、つまり、2つの集団の“重なりあった”「部分」にだけ関心を持っている、ということである。

しかし、分析の対象のなかに(少数の)“特異な”つまり“異質な”ものが含まれており、この“特異な”対象の変換された判別得点の値がほかのすべての対象のそれとは(分布の上で)かけ離れている、といった場合が考えられる。こういった場合には、“特異な”対象を含むすべての対象に対する判別分析の解と、“特異な”対象を除いたすべての対象に対する解とは、かなり違ったものになってしまう可能性がある。ちなみに、「解析的な方法」では、判別得点の分布の全体を考慮しているので、“特異な”対象の割合が少なければ、いずれの場合でも解はそれほど変わらない。

このような場合、“特異な”対象が含まれていることの影響を避けるためには、以下の工夫が有効である。つまり、判別分析の目的関数として、「(判別得点の値を尺度にした) 2つの集団の間の最短距離を最大にする」(以下、「100%値(パーセンタイル)」という)の代わりに、「(判別得点の値を尺度にした) 2つの集団の90%を含む“主要な”部分の間の最短距離を最大にする」(以下、「90%値(パーセンタイル)」という)を採用する(図2-8-1)。判別得点がより大きい対象の集団については、その集団に含まれる対象の上から90%の対象が含まれる判別得点の値と、判別得点がより小さい対象の集団については、その集団に含まれる対象の下から90%の対象が含まれる判別得点の値の差を最大にするのである。90%の値は、データの性格と分析の目的に応じて、95%でも80%でも75%でもよい。これは、

判別式で変換された値の分布を代表する程度によって判別分析を行なうか、という分析の“戦略”あるいは“方針”といえる。

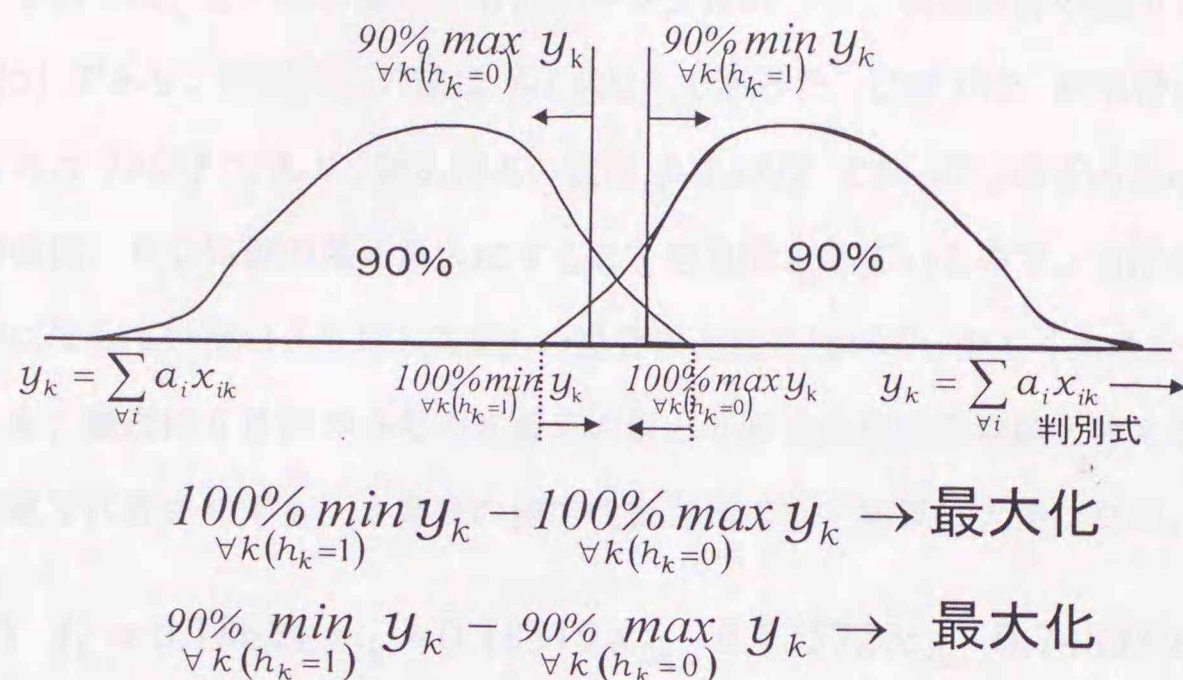


図 2-8-1 判別得点の「分布」を考慮した判別分析

あるいは、実務上の知恵・工夫としては、目的関数として、「100%値」と「90%値」の両方の解を求め、2つの解の間に大きな違いがなければ、そのいずれかの解を採用し、また、両者の間に大きな違いが認められる場合には、“特異な”対象が含まれている可能性があると考えて、後者（90%値）の解を採用する、という方法も考えられる。あるいは、「100%値」のときの解（判別式）で変換された値（判別得点）の分布を検討して、ほかの対象とはかけ離れた値をとる“特異な”対象が含まれていないどうかを検討する、というやり方も考えられる。

前と同じ「あやめ科の3種類の植物の4変量観測値」のB群とC群のデータについて、判別式の形式には1次式を採用して、「100%値」と「90%値」に対する解を求めてみた。遺伝的アルゴリズムの“戦略”はいままでと同じものとし、1つの世代で用意する染色体を20個として、「優先生き残り」「トーナメント」「ランダム交叉」「突然変異」「ランダム生き残り」の遺伝的操作を加えた場合、解は約1000世代までの間にほぼ収束した。

「100%値」の場合、得られた結果は、(1)式と(2)式および表2-8-1の[f]であった。この解の場合、B群のサンプルのうち、判別得点の値が下から6番目（上からは45番目）のものは[#35]であり、判別得点の値は[-3.059]であった。C群では、判別得点の値が上から6番目（下からは45番目）のものは[#24]であり、判別得点の値は[-3.481]であった。

表中、これらの値は、表で太字で示してある。両者の差は、 $[0.422]$ であった。

これに対して、「90%値」の場合には、得られた結果は、(3)式と(4)式および表 2-8-1 の $[g]$ であった。この解の場合、B群のサンプルのうち、判別得点の値が下から6番目のものは $[\#35]$ であり、判別得点の値は $[-1.902]$ であった。C群では、判別得点の値が上から6番目のものは $[\#07]$ であり、判別得点の値は $[-2.451]$ であった。両者の差は $[0.549]$ であった。この値は、90%値の差を最大にすることを目的としているので、当然のことながら、これを目的にはしていない「100%値」の場合のそれに比べて、少しく大きい。ちなみに、「90%値」を、厳密に50件のうちの5番目の値と6番目の値の平均値と考えることもできるが、5番目の値で代表させても、6番目の値で代表させても、結果に大差はない。

$$(1) f_k = 0.18621 x_{1k} + 0.18319 x_{2k} - 0.77572 x_{3k} - 0.74895 x_{4k}$$

$$(2) \text{ if } f_k \geq 3.39013 (= f_0) \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

$$(3) g_k = 0.24177 x_{1k} + 0.46253 x_{2k} - 0.69381 x_{3k} - 0.98207 x_{4k}$$

$$(4) \text{ if } g_k \geq 2.19330 (= g_0) \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

なお、これによると、B群の誤判別数 $[n_1]$ は50件中の1件、C群の誤判別数 $[n_0]$ は50件中の0件であった。この結果は、「100%値」に対する解とあまり変わらないが、これは、データに（他とはかけ離れた）“特異な”つまり“異質な”対象が含まれていなかったためであろう。

表 2-8-1 「100%値」と「90%値」の解の比較

イリス・ヴェルシコロール(B群)							イリス・ヴィルジニカ(C群)						
がく の長さ	がく の幅	花弁 の長さ	花弁 の幅	判別 得点	判別 得点		がく の長さ	がく の幅	花弁 の長さ	花弁 の幅	判別 得点	判別 得点	
##	X ₁	X ₂	X ₃	X ₄	f	g	X ₁	X ₂	X ₃	X ₄	f	g	
01	7.0	3.2	4.7	1.4	-2.805	-1.463	6.3	3.3	6.0	2.5	-4.749	-3.569	
02	6.4	3.2	4.5	1.5	-2.836	-1.568	5.8	2.7	5.1	1.9	-3.805	-2.753	
03	6.9	3.1	4.9	1.5	-3.072	-1.771	7.1	3.0	5.9	2.1	-4.278	-3.052	
04	5.5	2.3	4.0	1.3	-2.631	-1.658	6.3	2.9	5.6	1.8	-3.988	-2.789	
05	6.5	2.8	4.6	1.5	-2.968	-1.798	6.5	3.0	5.8	2.2	-4.387	-3.226	
06	5.7	2.8	4.5	1.3	-2.890	-1.726	7.6	3.0	6.6	2.1	-4.728	-3.416	
07	6.3	3.3	4.7	1.6	-3.067	-1.783	4.9	2.5	4.5	1.7	-3.394	-2.451	
08	4.9	2.4	3.3	1.0	-1.957	-0.977	7.3	2.9	6.3	1.8	-4.345	-3.032	
09	6.6	2.9	4.6	1.3	-2.782	-1.531	6.7	2.5	5.8	1.8	-4.142	-3.016	
10	5.2	2.7	3.9	1.4	-2.611	-1.575	7.2	3.6	6.1	2.5	-4.604	-3.282	
11	5.0	2.0	3.5	1.0	-2.167	-1.276	6.5	3.2	5.1	2.0	-3.657	-2.451	
12	5.9	3.0	4.2	1.5	-2.733	-1.573	6.4	2.7	5.3	1.9	-3.848	-2.747	
13	6.0	2.2	4.0	1.0	-2.332	-1.289	6.8	3.0	5.5	2.1	-4.023	-2.847	
14	6.1	2.9	4.7	1.4	-3.027	-1.820	5.7	2.5	5.0	2.0	-3.857	-2.899	
15	5.6	2.9	3.6	1.3	-2.192	-1.079	5.8	2.8	5.1	2.4	-4.161	-3.198	
16	6.7	3.1	4.4	1.4	-2.646	-1.374	6.4	3.2	5.3	2.3	-4.056	-2.909	
17	5.6	3.0	4.5	1.5	-3.022	-1.854	6.5	3.0	5.5	1.8	-3.855	-2.625	
18	5.8	2.7	4.1	1.0	-2.355	-1.176	7.7	3.8	6.7	2.2	-4.715	-3.190	
19	6.2	2.2	4.5	1.5	-3.057	-2.079	7.7	2.6	6.9	2.3	-5.165	-3.982	
20	5.6	2.5	3.9	1.1	-2.348	-1.276	6.0	2.2	5.0	1.5	-3.482	-2.474	
21	5.9	3.2	4.8	1.8	-3.387	-2.191	6.9	3.2	5.7	2.3	-4.273	-3.065	
22	6.1	2.8	4.0	1.3	-2.428	-1.282	5.6	2.8	4.9	2.0	-3.743	-2.715	
23	6.3	2.5	4.9	1.5	-3.293	-2.193	7.7	2.8	6.7	2.0	-4.748	-3.456	
24	6.1	2.8	4.7	1.2	-2.896	-1.670	6.3	2.7	4.9	1.8	-3.481	-2.395	
25	6.4	2.9	4.3	1.3	-2.586	-1.371	6.7	3.3	5.7	2.1	-4.142	-2.871	
26	6.6	3.0	4.4	1.4	-2.683	-1.444	7.2	3.2	6.0	1.8	-4.076	-2.710	
27	6.8	2.8	4.8	1.4	-2.993	-1.766	6.2	2.8	4.8	1.8	-3.404	-2.304	
28	6.7	3.0	5.0	1.7	-3.355	-2.131	6.1	3.0	4.9	1.8	-3.464	-2.305	
29	6.0	2.9	4.5	1.5	-2.966	-1.803	6.4	2.8	5.6	2.1	-4.212	-3.105	
30	5.7	2.6	3.5	1.0	-1.926	-0.830	7.2	3.0	5.8	1.6	-3.807	-2.467	
31	5.5	2.4	3.8	1.1	-2.308	-1.277	7.4	2.8	6.1	1.9	-4.264	-3.014	
32	5.5	2.4	3.7	1.0	-2.155	-1.109	7.9	3.8	6.4	2.0	-4.295	-2.737	
33	5.8	2.7	3.9	1.2	-2.349	-1.233	6.4	2.8	5.6	2.2	-4.287	-3.203	
34	6.0	2.7	5.1	1.6	-3.543	-2.410	6.3	2.8	5.1	1.5	-3.394	-2.193	
35	5.4	3.0	4.5	1.5	-3.059	-1.902	6.1	2.6	5.6	1.4	-3.780	-2.583	
36	6.0	3.4	4.5	1.6	-2.949	-1.670	7.7	3.0	6.1	2.3	-4.471	-3.242	
37	6.7	3.1	4.7	1.5	-2.954	-1.680	6.3	3.4	5.6	2.4	-4.346	-3.147	
38	6.3	2.3	4.4	1.3	-2.792	-1.742	6.4	3.1	5.5	1.8	-3.855	-2.603	
39	5.6	3.0	4.1	1.3	-2.562	-1.380	6.0	3.0	4.8	1.8	-3.405	-2.260	
40	5.5	2.5	4.0	1.3	-2.594	-1.566	6.9	3.1	5.4	2.1	-3.909	-2.707	
41	5.5	2.6	4.4	1.2	-2.811	-1.699	6.7	3.1	5.6	2.4	-4.326	-3.189	
42	6.1	3.0	4.6	1.4	-2.931	-1.704	6.9	3.1	5.1	2.3	-3.826	-2.695	
43	5.8	2.6	4.0	1.2	-2.445	-1.349	5.8	2.7	5.1	1.9	-3.805	-2.753	
44	5.0	2.3	3.3	1.0	-1.956	-0.999	6.8	3.2	5.9	2.3	-4.447	-3.228	
45	5.6	2.7	4.2	1.3	-2.694	-1.588	6.7	3.3	5.7	2.5	-4.442	-3.264	
46	5.7	3.0	4.2	1.2	-2.546	-1.327	6.7	3.0	5.2	2.3	-3.959	-2.859	
47	5.7	2.9	4.2	1.3	-2.639	-1.471	6.3	2.5	5.0	1.9	-3.671	-2.656	
48	6.2	2.9	4.3	1.3	-2.623	-1.420	6.5	3.0	5.2	2.0	-3.772	-2.613	
49	5.1	2.5	3.0	1.1	-1.743	-0.772	6.2	3.4	5.4	2.3	-4.134	-2.934	
50	5.7	2.8	4.1	1.3	-2.580	-1.448	5.9	3.0	5.1	1.8	-3.656	-2.492	
最小値 B							最大値 C						
上から 5 番目 B							下から 5 番目 C						
上から 6 番目 B							下から 6 番目 C						

2-9節 複雑な問題に対応できる「2段階判別方式」の方法

一般に、対象の属性値の数が多く、また、対象の数が多い場合には、ここで取り上げた例題（あやめ科の3種類の植物の4変量観測値）のようにきれいには判別できないことが多い。このような複雑な問題に対応するために、「2段階判別方式」を開発した。

この方法では、前に説明した通常の方法（以下で説明する「“2段階”判別方式」と対比すると、通常の方法は「“1段階”判別方式」といえる）で、まずは、もっとも判別力の大きい解、つまり、“第1”の判別式を求める。これは通常の判別分析の結果である。この場合、「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[u_1]$ が「そうではない対象の集団」の対象の判別得点の最大値 $[u_2]$ よりも大きければ、この解（判別式）でまったく誤判別は起こらず、この解で2つの集団の対象は“完全に”判別できる。

この解の判別値 $[u_0]$ は、「そうではない対象の集団」の対象の判別得点の最大値 $[u_2]$ と「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[u_1]$ の間のいずれの値でもよい。（特別な理由がなければ、推定モデルとしては、判別値 $[u_0]$ を、そうではない対象の集団」の対象の判別得点の最大値 $[u_2]$ と「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[u_1]$ の平均値（中点）とするのがよいであろう。）

しかし、「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[u_1]$ が「そうではない対象の集団」の対象の判別得点の最大値 $[u_2]$ よりも小さければ、判別値 $[u_0]$ をどのように選択しても、必ずどちらかの集団のどれかの対象で誤判別が生じてしまう。判別値 $[u_0]$ が「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[u_1]$ より小さければ、「特定の特性を持った対象の集団」には誤判別がまったく生じないが、「そうではない対象の集団」のうち、判別得点が判別値 $[u_0]$ を超える対象については誤判別が生じてしまう。反対に、判別値 $[u_0]$ が「そうではない対象の集団」の対象の判別得点の最大値 $[u_2]$ より大きければ、「そうではない対象の集団」には誤判別はまったく生じないが、「特定の特性を持った対象の集団」のうち、判別得点が判別値 $[u_0]$ を下回る対象について誤判別が生じてしまう。

こういったことから、「誤判別数の合計 $[n_0 + n_1]$ を最小にする」などの立場からは、判別値 $[u_0]$ は、「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[u_1]$ と「そうではない対象の集団」の対象の判別得点の最大値 $[u_2]$ の間で選択される。しかし、判別値をどのように選択しても、「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[u_1]$

$u_1]$ と「そうではない対象の集団」の対象の判別得点の最大値 $[u_2]$ 、あるいは、その一方は必ず誤判別となってしまうことに注意が必要である。以下、判別得点について、この2つの値の間にある範囲を「局所」と呼ぶことにする。

そこで、この場合、“第1”の判別式で誤判別が起こる可能性のある（つまり、「局所」に入っている）すべての対象について、具体的には、判別得点の値が、「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[u_1]$ と「そうではない対象の集団」の対象の判別得点の最大値 $[u_2]$ の間（含、両端）に入っているすべての対象について、“第1”の判別式とは別に、もっとも判別力の大きい“第2”の判別式を求めて、これによって“局所的に”判別をしよう、というのである（図 2-9-1）。実際には、「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[u_1]$ と「そうではない対象の集団」の対象の判別得点の最大値 $[u_2]$ の間だけでなく、その近傍を含んだ方が、「印象の推定モデル」としての判別力は高くなるであろう。なお、第2の段階でも、判別値の決定は、通常の判別分析の方法と同様である。

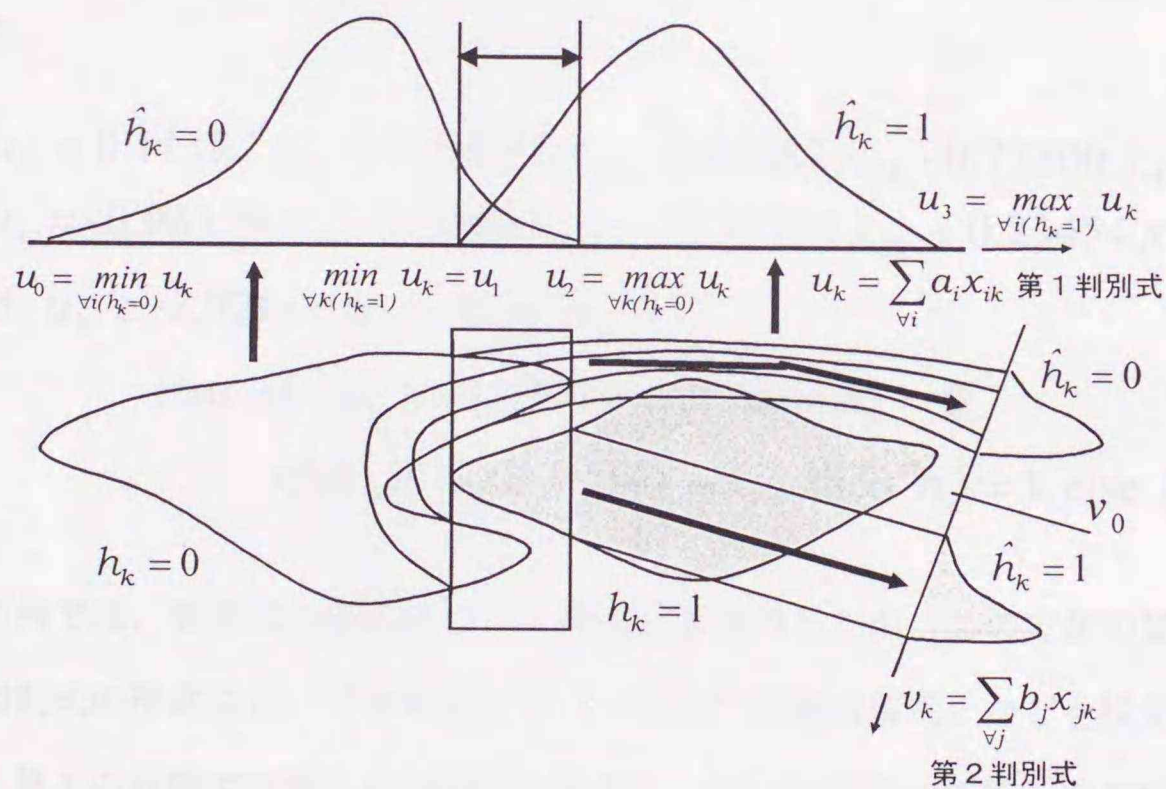


図 2-9-1 「2段階判別方式」の方法の説明

$$(1) \text{ if } u_k \geq u_2 \text{ then } \hat{h}_k = 1$$

$$\text{else if } u_k \leq u_1 \text{ then } \hat{h}_k = 0$$

$$\text{else if } v_k \geq v_0 \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

もう一度整理をすると、“第1”の判別式は、まずは、それぞれの対象が「局所」に入るかいないかを判別し、また、局所に入らなかった（つまり、明らかな判別ができた）対象に対しては、

いずれの集団に属するかを有効に判別する。“第1”の判別式で、局所に入ると判別された対象は、“第2”の判別式で有効な判別をする、という訳である（(1)式）。

前と同じ「あやめ科の3種類の植物の4変量観測値」のB群とC群のデータについて、「誤判別数の合計 $[n_0 + n_1]$ を最小にする」を目的に、この方法で解いた結果が、(2)式、(3)式、(4)式および表 2-9-1 である。“第1”の判別式 $[u_k]$ で判別得点を求めると、B群の対象の最小値 $[u_1 = -0.128]$ はC群の対象の最大値 $[u_2 = 0.000]$ より小さく、この間で判別値 $[u_0]$ をどのように設定しても“完全には”判別できないことがわかる。ちなみに、この間には、B群では1個(#34)、C群では7個(#07, #20, #24, #27, #28, #34, #39)の対象が含まれている。

そこで、この8個の対象について、「誤判別数の合計を最小にする」“第2”の判別式 $[v_k]$ を求めた結果が(3)式で、この判別式では、（「局所」に入った）B群とC群のすべての対象について、まったく誤りなく判別することができた。ちなみに、遺伝的アルゴリズムの“戦略”はいままでのもと同様として、約200世代以内で“完全に”判別できる解を見つけることができた。

$$(2) \quad u_k = 0.11187 x_{1k} + 0.20957 x_{2k} - 0.61082 x_{3k} - 0.73500 x_{4k}$$

$$(3) \quad v_k = -0.08359 x_{1k} - 0.03807 x_{2k} + 0.31436 x_{3k} + 0.23454 x_{4k}$$

$$(4) \quad \text{if } u_k \geq -2.926 (= u_2) \text{ then } \hat{h}_k = 1$$

$$\text{else if } u_k \leq -3.054 (= u_1) \text{ then } \hat{h}_k = 0$$

$$\text{else if } v_k \geq 1.356 (= v_0) \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

なお、この例では、判別式の形式には「1次式」を採用したが、この方法の場合にも、必要に応じて、判別式の形式には、「2次式」や「“V型”の絶対値式」などを採用することができる。また、第1の段階で採用した判別式の形式と、第2の段階で採用した判別式の形式が異なってもまったく問題はない。大切なことは、分析の対象としているデータの性質に合った判別式の形式を採用することである。第1の段階の判別式の形式として「1次式」を採用した結果、“完全には”判別できなかった場合、この節で説明した「2段階判別方式」を適用するという選択もよいが、その前に、第1段階の判別式の形式として、「2次式」あるいは（V型あるいはW型の）「絶対値式」などを採用した判別分析を行ってみることも必要かもしれない。判別式にいろいろな形式のものを採用してみる方法と、この「2段階判別方式」の方法の採用のいずれがよいかは、実際の問題を分析する分析者の“好み”の問題であろう。

表 2-9-1 「2 段階判別方式」による判別分析の結果

イリス・ヴェルシコロール(B群)							イリス・ヴィルジニカ(C群)						
がく の長さ	がく の幅	花卉 の長さ	花卉 の幅	判別 得点	判別 得点		がく の長さ	がく の幅	花卉 の長さ	花卉 の幅	判別 得点	判別 得点	
##	X ₁	X ₂	X ₃	X ₄	u	v	X ₁	X ₂	X ₃	X ₄	u	v	
01	7.0	3.2	4.7	1.4	-2.446	=====	6.3	3.3	6.0	2.5	-4.106	=====	
02	6.4	3.2	4.5	1.5	-2.465	=====	5.8	2.7	5.1	1.9	-3.297	=====	
03	6.9	3.1	4.9	1.5	-2.674	=====	7.1	3.0	5.9	2.1	-3.724	=====	
04	5.5	2.3	4.0	1.3	-2.301	=====	6.3	2.9	5.6	1.8	-3.431	=====	
05	6.5	2.8	4.6	1.5	-2.598	=====	6.5	3.0	5.8	2.2	-3.804	=====	
06	5.7	2.8	4.5	1.3	-2.480	=====	7.6	3.0	6.6	2.1	-4.096	=====	
07	6.3	3.3	4.7	1.6	-2.650	=====	4.9	2.5	4.5	1.7	-2.926	1.309	
08	4.9	2.4	3.3	1.0	-1.700	=====	7.3	2.9	6.3	1.8	-3.747	=====	
09	6.6	2.9	4.6	1.3	-2.419	=====	6.7	2.5	5.8	1.8	-3.592	=====	
10	5.2	2.7	3.9	1.4	-2.264	=====	7.2	3.6	6.1	2.5	-4.004	=====	
11	5.0	2.0	3.5	1.0	-1.894	=====	6.5	3.2	5.1	2.0	-3.187	=====	
12	5.9	3.0	4.2	1.5	-2.379	=====	6.4	2.7	5.3	1.9	-3.352	=====	
13	6.0	2.2	4.0	1.0	-2.046	=====	6.8	3.0	5.5	2.1	-3.514	=====	
14	6.1	2.9	4.7	1.4	-2.610	=====	5.7	2.5	5.0	2.0	-3.363	=====	
15	5.6	2.9	3.6	1.3	-1.920	=====	5.8	2.8	5.1	2.4	-3.644	=====	
16	6.7	3.1	4.4	1.4	-2.317	=====	6.4	3.2	5.3	2.3	-3.541	=====	
17	5.6	3.0	4.5	1.5	-2.596	=====	6.5	3.0	5.5	1.8	-3.327	=====	
18	5.8	2.7	4.1	1.0	-2.025	=====	7.7	3.8	6.7	2.2	-4.052	=====	
19	6.2	2.2	4.5	1.5	-2.697	=====	7.7	2.6	6.9	2.3	-4.499	=====	
20	5.6	2.5	3.9	1.1	-2.040	=====	6.0	2.2	5.0	1.5	-3.024	1.338	
21	5.9	3.2	4.8	1.8	-2.924	=====	6.9	3.2	5.7	2.3	-3.730	=====	
22	6.1	2.8	4.0	1.3	-2.130	=====	5.6	2.8	4.9	2.0	-3.250	=====	
23	6.3	2.5	4.9	1.5	-2.867	=====	7.7	2.8	6.7	2.0	-4.114	=====	
24	6.1	2.8	4.7	1.2	-2.484	=====	6.3	2.7	4.9	1.8	-3.045	1.333	
25	6.4	2.9	4.3	1.3	-2.258	=====	6.7	3.3	5.7	2.1	-3.584	=====	
26	6.6	3.0	4.4	1.4	-2.350	=====	7.2	3.2	6.0	1.8	-3.512	=====	
27	6.8	2.8	4.8	1.4	-2.613	=====	6.2	2.8	4.8	1.8	-2.975	1.306	
28	6.7	3.0	5.0	1.7	-2.925	=====	6.1	3.0	4.9	1.8	-3.005	1.338	
29	6.0	2.9	4.5	1.5	-2.572	=====	6.4	2.8	5.6	2.1	-3.661	=====	
30	5.7	2.6	3.5	1.0	-1.690	=====	7.2	3.0	5.8	1.6	-3.285	=====	
31	5.5	2.4	3.8	1.1	-2.011	=====	7.4	2.8	6.1	1.9	-3.708	=====	
32	5.5	2.4	3.7	1.0	-1.877	=====	7.9	3.8	6.4	2.0	-3.699	=====	
33	5.8	2.7	3.9	1.2	-2.050	=====	6.4	2.8	5.6	2.2	-3.735	=====	
34	6.0	2.7	5.1	1.6	-3.054	1.374	6.3	2.8	5.1	1.5	-2.926	1.322	
35	5.4	3.0	4.5	1.5	-2.618	=====	6.1	2.6	5.6	1.4	-3.222	=====	
36	6.0	3.4	4.5	1.6	-2.541	=====	7.7	3.0	6.1	2.3	-3.926	=====	
37	6.7	3.1	4.7	1.5	-2.574	=====	6.3	3.4	5.6	2.4	-3.767	=====	
38	6.3	2.3	4.4	1.3	-2.456	=====	6.4	3.1	5.5	1.8	-3.317	=====	
39	5.6	3.0	4.1	1.3	-2.205	=====	6.0	3.0	4.8	1.8	-2.955	1.315	
40	5.5	2.5	4.0	1.3	-2.260	=====	6.9	3.1	5.4	2.1	-3.420	=====	
41	5.5	2.6	4.4	1.2	-2.409	=====	6.7	3.1	5.6	2.4	-3.785	=====	
42	6.1	3.0	4.6	1.4	-2.528	=====	6.9	3.1	5.1	2.3	-3.384	=====	
43	5.8	2.6	4.0	1.2	-2.132	=====	5.8	2.7	5.1	1.9	-3.297	=====	
44	5.0	2.3	3.3	1.0	-1.709	=====	6.8	3.2	5.9	2.3	-3.863	=====	
45	5.6	2.7	4.2	1.3	-2.329	=====	6.7	3.3	5.7	2.5	-3.878	=====	
46	5.7	3.0	4.2	1.2	-2.181	=====	6.7	3.0	5.2	2.3	-3.489	=====	
47	5.7	2.9	4.2	1.3	-2.276	=====	6.3	2.5	5.0	1.9	-3.222	=====	
48	6.2	2.9	4.3	1.3	-2.281	=====	6.5	3.0	5.2	2.0	-3.290	=====	
49	5.1	2.5	3.0	1.1	-1.546	=====	6.2	3.4	5.4	2.3	-3.583	=====	
50	5.7	2.8	4.1	1.3	-2.235	=====	5.9	3.0	5.1	1.8	-3.149	=====	
最大値 -1.546 1.374							最大値 -2.926 1.338						
最小値 -3.054 1.374							最小値 -4.499 1.306						

この「2段階判別方式」の方法は、判別分析の方法に重要な示唆を与えている。たとえば、第1の段階では、“すべて”の対象に対して“もっとも判別力の高い”解（判別式）を求める。この解は、この解のパラメーターが多少違っていても、この解で判別した結果が十分に“有効である”対象にだけ適用し、それ“以外”の対象に対しては、この解では判別をせず（この解では判別をしない、という“判断”は、この解で求められる、それぞれの対象の判別得点の値による）、この解が十分に“有効である”対象を除いた“残りのすべて”の対象に対して、再度、新たな判別分析を行う。最初の判別分析の解を適用する範囲は、判別する対象の性質によって、また、判別の結果が十分に“有効である”と考えられるならば、対象のうちの30%でもあっても40%であっても50%であってもよい。

もう少し具体的にいうと、最初の解によって判別得点を計算しその値の大きさの順に、上から15%および下から15%の対象に適用すると、“すべて”の対象のうち30%に対しては最初の解が有効とされ、（最初の解の判別得点の値の大きさによって、そのように“判別”された）残りの70%の対象に対しては、別の解を求めて適用する、ということになる。

最初の解で判別した結果が“有効である”対象を除いた“残りのすべて”の対象についても同様で、この“残りのすべて”の対象に対して“もっとも判別力の高い”解を求めて、この解が“残りのすべて”の対象に対して、十分に“有効である”ならばそれでよいが、そうではない場合には、（前の場合とまったく同様に、）その解が十分に“有効である”対象についてだけ適用し、それ“以外”の対象に対しては、さらに別の判別分析をする、というようにして、結果的に“すべて”の対象について、十分に“有効な”（つまり満足に行く）解（判別式）を得る、という「“逐次的な”判別分析」という概念を導き出すことができる。ちなみに、この方法は、判別分析の「解析的な方法」の結果についても、まったく同様に適用することができる。

2-10節 複雑な問題に対応できる「二重判別方式」の方法

2-9節と同様に、対象の属性値の数や対象の数が多く、単純な判別分析ではきれいには判別できない複雑な問題に対応するため、もう1つの方法として「二重判別方式」を開発した。この方法では、2つの判別式（の組み合わせ）を用意し、この2つの判別式（の組み合わせ）を“同時に”使って、それぞれの対象がどちらの集団に属するかいなかを判別する。つまり、“第1”の判別式で変換した判別得点 $[u_k]$ が“第1”の判別値 $[u_0]$ を超え、同時に“第2”の判別式で変換した判別得点 $[v_k]$ も“第2”の判別値 $[v_0]$ を超えている場合に、その対象はある集団に属する ($h_k=1$) と判別し、そうではない場合は、その集団には属さない ($h_k=0$) と判別するのである（(1)式）（図2-10-1）。

$$(1) \text{ if } u_k \geq u_0 \text{ and } v_k \geq v_0 \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

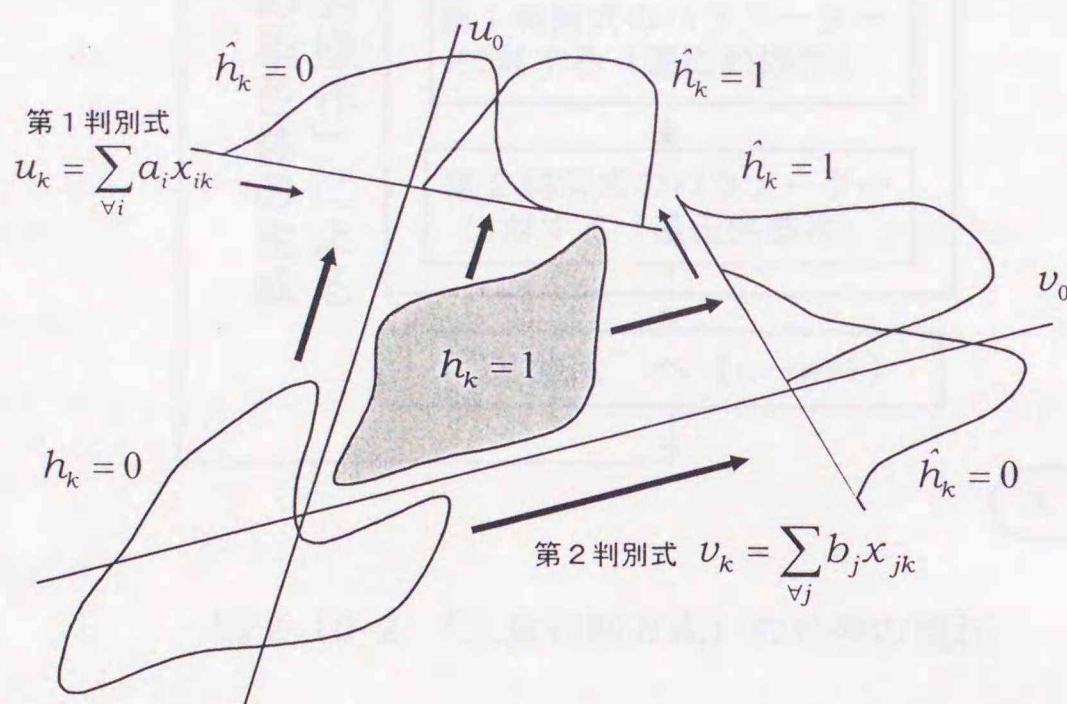


図 2-10-1 「二重判別方式」の方法の説明

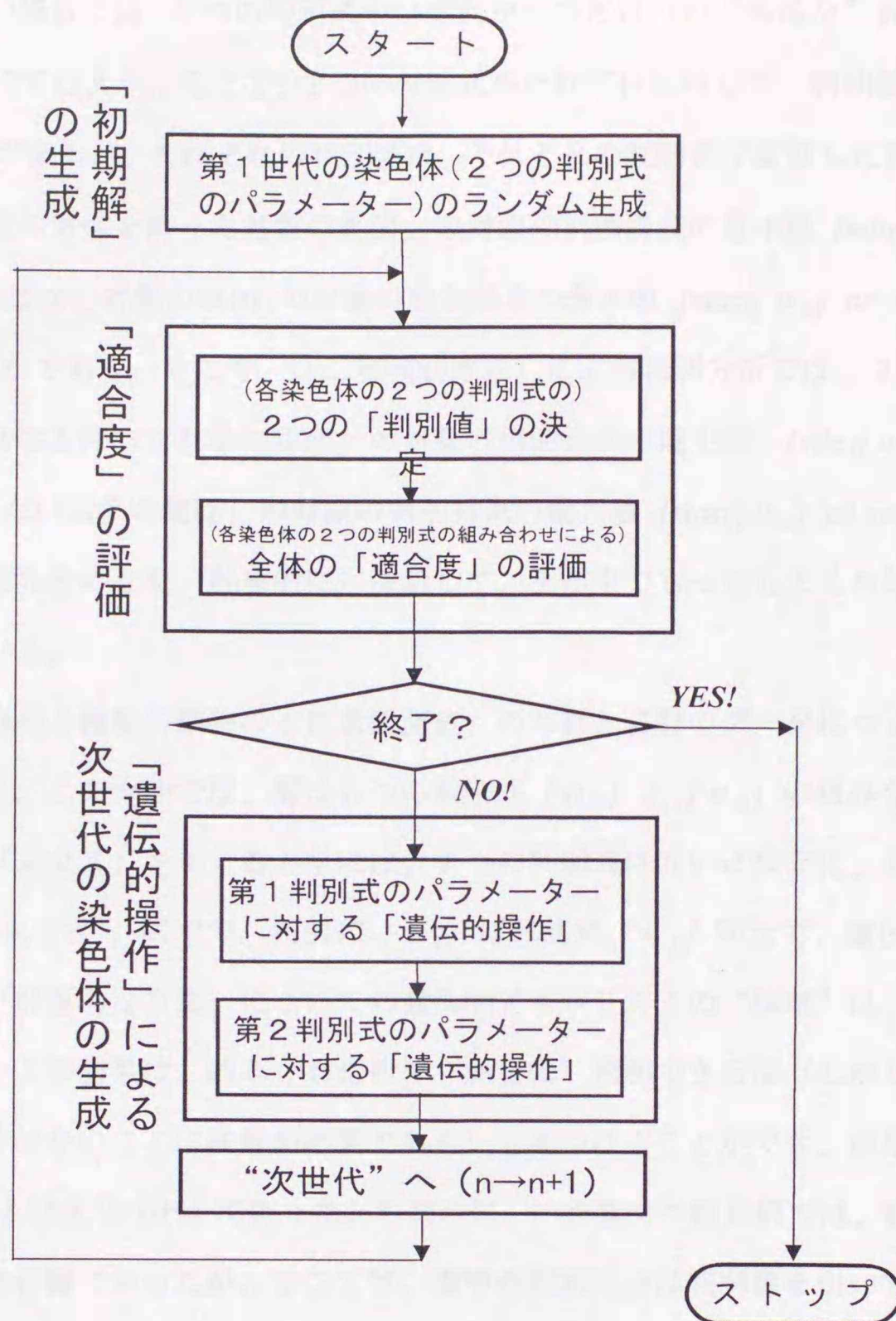


図 2-10-2 「二重判別方式」の処理の流れ

ちなみに、2つの判別式のいずれかについて、「特定の特性を持った対象 ($h_k=1$) の集団」の対象の判別得点の最小値 $[\min_l(u_k) \text{ or } \min_l(v_k)]$ が「そうではない対象 ($h_k=0$) の集団」の対象の判別得点の最大値 $[\max_0(u_k) \text{ or } \max_0(v_k)]$ よりも大きく、2つの集団が“完全に”分離できている場合には、誤判別は生じない。この場合は、(判別式が1つだけの) 単純な方式で判別ができるので、複雑な問題に対応できるこの「二重判別方式」を適用する必要はない。

しかし、2つの判別式のいずれについても、「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[\min_l(u_k) \text{ or } \min_l(v_k)]$ が「そうではない対象の集団」の対象の判別得点の最大値 $[\max_0(u_k) \text{ or } \max_0(v_k)]$ よりも小さく、(1つの判別式だけでは) 2つの集団が“完全

には”判別ができない場合には、2つの判別式のいずれか一方だけ（の“単純な”判別分析の解）では誤判別が生じてしまう。そこで、2つの判別式のそれぞれに対して、判別値 $[u_0 \text{ or } v_0]$ を選択する必要があるが、それぞれの判別値は、それぞれの判別式で変換した判別得点の分布に関して、「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[\min_l(u_k) \text{ or } \min_l(v_k)]$ と「そうではない対象の集団」の対象の判別得点の最大値 $[\max_0(u_k) \text{ or } \max_0(v_k)]$ の間にあることは確かである。そこで、「二重判別方式」による判別分析では、2つの判別式について、「特定の特性を持った対象の集団」の対象の判別得点の最小値 $[\min_l(u_k) \text{ or } \min_l(v_k)]$ と「そうではない対象の集団」の対象の判別得点の最大値 $[\max_0(u_k) \text{ or } \max_0(v_k)]$ の間の値のすべての組み合わせを（計算的に）探索して、その中でもっともよく目的関数を満たすものを選択している。

前と同じ「あやめ科の3種類の植物の4変量観測値」のB群とC群のデータについて、この方法による解を求めた。この方法では、解は2つの判別式 $[u_k]$ と $[v_k]$ の組み合わせであり、たとえば「ランダム交叉」を加えるときには、2つの判別式は互いに独立に、具体的には判別式 $[u_k]$ は判別式 $[u_k]$ 同士で、判別式 $[v_k]$ は判別式 $[v_k]$ 同士で、遺伝的操作を加えるようにした。「計算的な方法」についての遺伝的アルゴリズムの“戦略”は、いままでのものと同様である。この結果は、約600世代で“完全に”判別できる解（しかし、この解は必ずしも最適な解ではないことに注意が必要である）を見つけることができ、結果は、(1)式と(2)式と(3)式および表2-10-1であった。ちなみに、いままでの計算例では、判別得点は判別式で変換したままの値であったが、ここでは、表中の判別得点は判別値を引いてあることに注意が必要である。

$$(1) \quad u_k = 0.18531 x_{1k} + 0.18212 x_{2k} \\ - 0.76675 x_{3k} + 0.80984 x_{4k} + 1.0146865$$

$$(2) \quad v_k = 0.18531 x_{1k} + 0.18212 x_{2k} \\ - 0.12431 x_{3k} - 0.73113 x_{4k} + 0.2425055$$

$$(3) \quad \text{if } u_k \geq 0 \text{ and } v_k \geq 0 \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

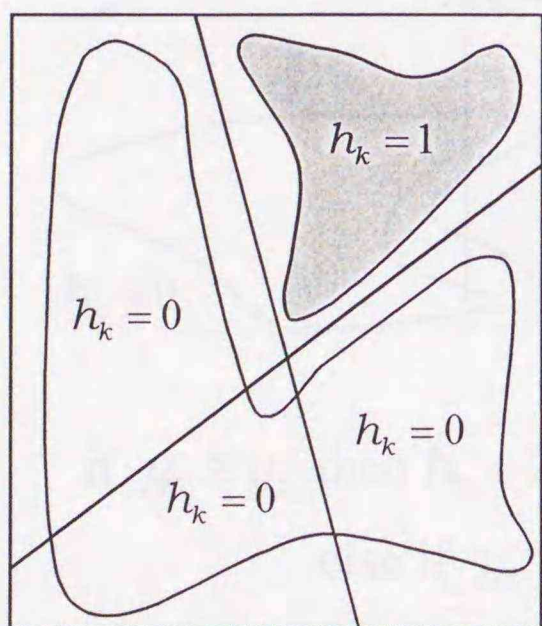
表 2-10-1 「二重判別方式」による判別分析の結果

イリス・ヴェルシコロール(B群)						イリス・ヴィルジニカ(C群)							
	がく の長さ	がく の幅	花弁 の長さ	花弁 の幅	判別 得点	判別 得点	がく の長さ	がく の幅	花弁 の長さ	花弁 の幅	判別 得点	判別 得点	
##	X ₁	X ₂	X ₃	X ₄	u	v	X ₁	X ₂	X ₃	X ₄	u	v	
01	7.0	3.2	4.7	1.4	0.425	0.515	6.3	3.3	6.0	2.5	0.207	-0.563	
02	6.4	3.2	4.5	1.5	0.548	0.355	5.8	2.7	5.1	1.9	0.209	-0.214	
03	6.9	3.1	4.9	1.5	0.316	0.380	7.1	3.0	5.9	2.1	0.054	-0.164	
04	5.5	2.3	4.0	1.3	0.439	0.233	6.3	2.9	5.6	1.8	-0.126	-0.074	
05	6.5	2.8	4.6	1.5	0.417	0.288	6.5	3.0	5.8	2.2	0.100	-0.336	
06	5.7	2.8	4.5	1.3	0.183	0.299	7.6	3.0	6.6	2.1	-0.390	-0.159	
07	6.3	3.3	4.7	1.6	0.475	0.257	4.9	2.5	4.5	1.7	0.304	-0.196	
08	4.9	2.4	3.3	1.0	0.639	0.446	7.3	2.9	6.3	1.8	-0.477	0.024	
09	6.6	2.9	4.6	1.3	0.292	0.471	6.7	2.5	5.8	1.8	-0.278	-0.098	
10	5.2	2.7	3.9	1.4	0.613	0.189	7.2	3.6	6.1	2.5	0.352	-0.354	
11	5.0	2.0	3.5	1.0	0.432	0.367	6.5	3.2	5.1	2.0	0.511	-0.066	
12	5.9	3.0	4.2	1.5	0.649	0.263	6.4	2.7	5.3	1.9	0.167	-0.128	
13	6.0	2.2	4.0	1.0	0.270	0.527	6.8	3.0	5.5	2.1	0.305	-0.170	
14	6.1	2.9	4.7	1.4	0.203	0.293	5.7	2.5	5.0	2.0	0.312	-0.330	
15	5.6	2.9	3.6	1.3	0.873	0.410	5.8	2.8	5.1	2.4	0.633	-0.561	
16	6.7	3.1	4.4	1.4	0.581	0.478	6.4	3.2	5.3	2.3	0.582	-0.329	
17	5.6	3.0	4.5	1.5	0.363	0.171	6.5	3.0	5.5	1.8	0.006	-0.006	
18	5.8	2.7	4.1	1.0	0.247	0.568	7.7	3.8	6.7	2.2	-0.222	-0.080	
19	6.2	2.2	4.5	1.5	0.329	0.136	7.7	2.6	6.9	2.3	-0.513	-0.396	
20	5.6	2.5	3.9	1.1	0.408	0.446	6.0	2.2	5.0	1.5	-0.092	0.037	
21	5.9	3.2	4.8	1.8	0.468	0.006	6.9	3.2	5.7	2.3	0.368	-0.286	
22	6.1	2.8	4.0	1.3	0.641	0.435	5.6	2.8	4.9	2.0	0.425	-0.281	
23	6.3	2.5	4.9	1.5	0.095	0.159	7.7	2.8	6.7	2.0	-0.566	-0.116	
24	6.1	2.8	4.7	1.2	0.023	0.421	6.3	2.7	4.9	1.8	0.375	-0.023	
25	6.4	2.9	4.3	1.3	0.485	0.472	6.7	3.3	5.7	2.1	0.187	-0.159	
26	6.6	3.0	4.4	1.4	0.544	0.441	7.2	3.2	6.0	1.8	-0.211	0.098	
27	6.8	2.8	4.8	1.4	0.238	0.392	6.2	2.8	4.8	1.8	0.451	-0.011	
28	6.7	3.0	5.0	1.7	0.346	0.166	6.1	3.0	4.9	1.8	0.392	-0.006	
29	6.0	2.9	4.5	1.5	0.419	0.226	6.4	2.8	5.6	2.1	0.117	-0.293	
30	5.7	2.6	3.5	1.0	0.671	0.606	7.2	3.0	5.8	1.6	-0.256	0.232	
31	5.5	2.4	3.8	1.1	0.448	0.422	7.4	2.8	6.1	1.9	-0.243	-0.024	
32	5.5	2.4	3.7	1.0	0.444	0.508	7.9	3.8	6.4	2.0	-0.117	0.141	
33	5.8	2.7	3.9	1.2	0.563	0.447	6.4	2.8	5.6	2.2	0.198	-0.366	
34	6.0	2.7	5.1	1.6	0.004	0.042	6.3	2.8	5.1	1.5	-0.004	0.189	
35	5.4	3.0	4.5	1.5	0.326	0.133	6.1	2.6	5.6	1.4	-0.541	0.127	
36	6.0	3.4	4.5	1.6	0.591	0.244	7.7	3.0	6.1	2.3	0.173	-0.224	
37	6.7	3.1	4.7	1.5	0.432	0.368	6.3	3.4	5.6	2.4	0.451	-0.422	
38	6.3	2.3	4.4	1.3	0.280	0.331	6.4	3.1	5.5	1.8	0.006	-0.007	
39	5.6	3.0	4.1	1.3	0.508	0.366	6.0	3.0	4.8	1.8	0.450	-0.012	
40	5.5	2.5	4.0	1.3	0.475	0.269	6.9	3.1	5.4	2.1	0.418	-0.121	
41	5.5	2.6	4.4	1.2	0.106	0.311	6.7	3.1	5.6	2.4	0.471	-0.402	
42	6.1	3.0	4.6	1.4	0.298	0.324	6.9	3.1	5.1	2.3	0.810	-0.230	
43	5.8	2.6	4.0	1.2	0.468	0.416	5.8	2.7	5.1	1.9	0.209	-0.214	
44	5.0	2.3	3.3	1.0	0.640	0.447	6.8	3.2	5.9	2.3	0.196	-0.330	
45	5.6	2.7	4.2	1.3	0.377	0.299	6.7	3.3	5.7	2.5	0.511	-0.451	
46	5.7	3.0	4.2	1.2	0.369	0.446	6.7	3.0	5.2	2.3	0.678	-0.298	
47	5.7	2.9	4.2	1.3	0.432	0.354	6.3	2.5	5.0	1.9	0.342	-0.145	
48	6.2	2.9	4.3	1.3	0.448	0.435	6.5	3.0	5.2	2.0	0.398	-0.115	
49	5.1	2.5	3.0	1.1	1.006	0.466	6.2	3.4	5.4	2.3	0.505	-0.342	
50	5.7	2.8	4.1	1.3	0.490	0.349	5.9	3.0	5.1	1.8	0.202	-0.068	
平均値					0.427	0.348	平均値					0.161	-0.166
最大値					1.006	0.606	最大値					0.810	0.232
最小値					0.004	0.006	最小値					-0.566	-0.563

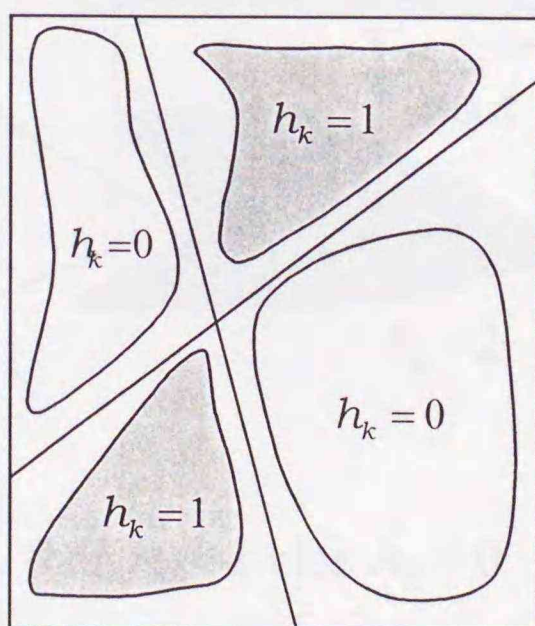
この解では、第1の判別式 $[u_k]$ と判別値 ((1)式) では、B群では50件のうち0個の誤判別が、C群では50件のうち36個の誤判別 (表中、 $[f_k < 0]$ のもの) があり、第2の判別式と判別値 ((2)式) では、B群では50件のうち0個の誤判別が、C群では50件のうち7個の誤判別 (表中、 $[g_k < 0]$ のもの) があった。しかし、これらの2つの判別式と判別値の組み合わせを“同時に”適用すると、B群とC群を合わせて100件のデータはまったく誤りなく判別することができた。

この「二重判別方式」の方法は、非線型性を持つデータに対応でき、また、判別式の形式に「1次式」を採用した場合、パラメーターの数は、1次式の判別式の2倍と比較的少なくて済むのがポイントである。

ちなみに、この方法は、判別の論理があらかじめ決まっていなくても適用できる。つまり、第1の判別式の判別得点 $[u_k]$ が第1の判別値 $[u_0]$ を超え、“同時に”第2の判別式の判別得点 $[v_k]$ も第2の判別値 $[v_0]$ を超えている場合、および、第1の判別式の判別得点 $[u_k]$ が第1の判別値 $[u_0]$ を下回り、“同時に”第2の判別式の判別得点 $[v_k]$ も第2の判別値 $[v_0]$ を下回る場合に、その対象はある集団に属する ($h_k=1$) と判別し、そうではない場合は、その集団には属さない ($h_k=0$) と判別する場合も考えられる (図2-10-3)。判別式の組み合わせに応じて、判別得点の平均値などによってそれぞれの領域に属する対象が、どちらの集団に属するかを決めて、いずれの「判別のしかた」にするかを定めるのである。この方法は、判別すべき対象が2つの“異なった”あるいは“類縁的な”対象群からなっている場合に有効であろう。同様の方法として、3つの判別式を“同時に”使って判別する「三重判別方式」なども考えられる。



if $u_k \geq u_0$ and $v_k \geq v_0$
then $\hat{h}_k = 1$ else $\hat{h}_k = 0$



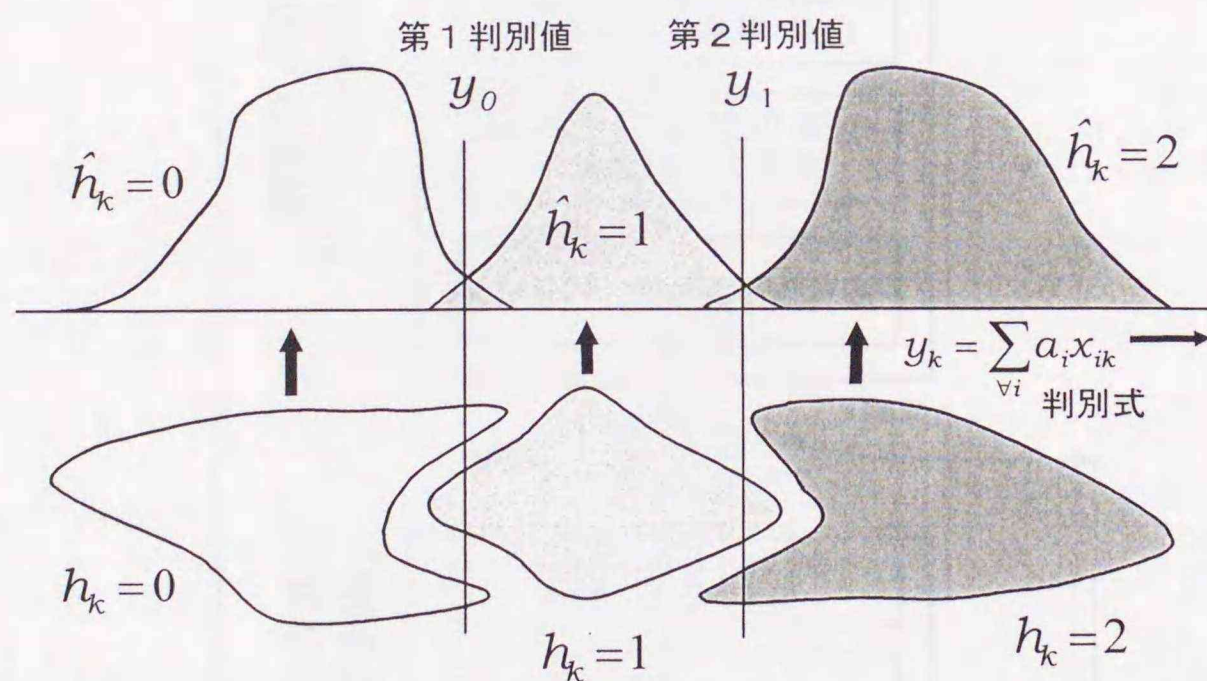
if $(u_k \geq u_0 \text{ and } v_k \geq v_0)$
or $(u_k \leq u_0 \text{ and } v_k \leq v_0)$
then $\hat{h}_k = 1$ else $\hat{h}_k = 0$

図2-10-3 2つの判別式の組み合わせによる「判別のしかた」の違い

2-11節 「重判別分析」への適用の方法

「計算的な判別分析」の方法は、ある対象が1つの集団に属するかどうかを判別するだけでなく、ある対象が3つ以上の集団のいずれに属するかを判別する「重判別分析」にも適用できる。（「重判別分析」は、「正準判別分析」あるいは「正準分析」ともいう。）この場合も、1つの判別式によってそれぞれの対象の複数の属性値を1つの判別得点に変換し、この値が複数の判別値で区切られるどの範囲に入っているかで、いずれの集団に属するかを判別する（図2-11-1）。

ある対象が3つ以上の集団のいずれに属するかを判別する「判別分析」には、集団の数に対応する“複数”の判別式を用意して、それぞれの対象について“複数”の判別得点を計算し、そのうちの最大の値を示した判別式に対応する集団に属する、と判別する方法もある。これは、「マハラノビスの汎距離の最小化」に基づく「線形判別分析」を一般化したものである。これに対して、この節で取り上げた、1つの判別式で計算した判別得点の値の大きさによって、どの集団に属するかを判別するやり方は、「相関比を最大にする」ことを目的にしたものを拡張したものといえる。



if $y_k \geq y_1$ then $\hat{h}_k = 2$
 else if $y_k \geq y_0$ then $\hat{h}_k = 1$ else $\hat{h}_k = 0$

図2-11-1 「重判別分析」への適用

遺伝的アルゴリズムによって判別式のパラメーターを見つけだす処理の流れは、普通の「計算的な判別分析」と同様であるが、この場合、それぞれの解の目的への適合度の評価のしかた

に若干の工夫が必要である。つまり、判別式のパラメーターつまり判別式が決まると、それぞれの集団に属する対象 $[k]$ の「判別得点 $[f_k]$ 」が計算できる。そして、それぞれの集団に属する対象の判別得点の平均値の大きさの順序で、それぞれの「集団の並びの順序」を決める。これは中央値やメディアンなどの値の大きさの順序によってもよい。あるいは、最大値や最小値などを含めて、これらに適当な重みを付けた計算的な代表値であってもよい。いずれにしても、それぞれの集団に属する対象の判別得点の分布を代表する値の大きさの順序にしたがって、それぞれの「集団の並びの順序」を決める。

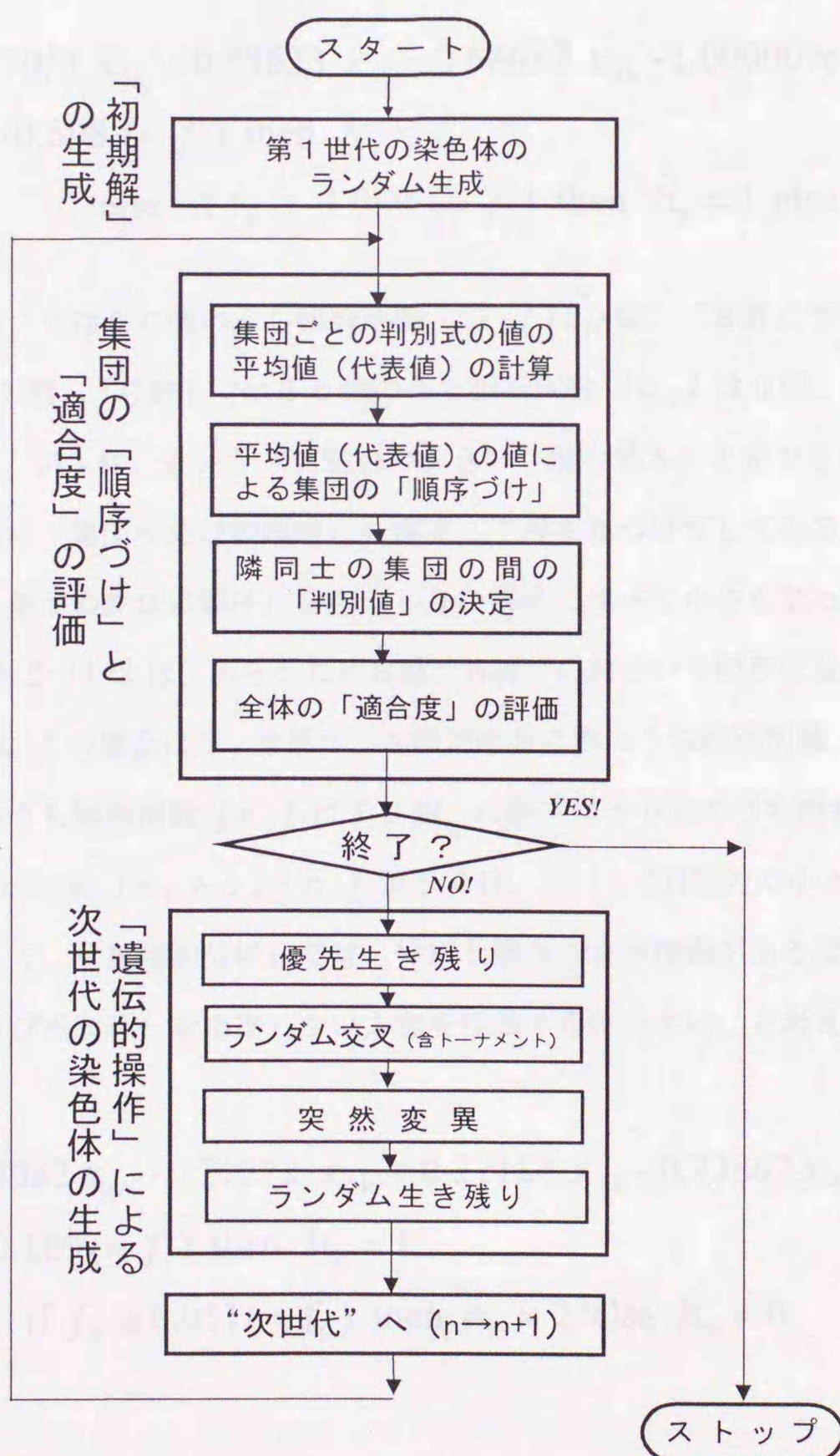


図 2-11-2 「重判別分析」の処理の流れ

そして、この「並びの順序」にしたがって、隣り合った集団の間の「判別値」 ($[y_1]$ と $[y_2]$) を決定して、この判別式の目的への適合度、たとえば「誤判別数の合計 $[n_0 + n_1 + n_2]$ 」などを評価する。なお、「集団の並びの順序」があらかじめ決まっている場合には、これにしたがって、隣り合った集団の間の判別値を見つけだして、それぞれの解の目的への適合度を評価すればよい (図 2-11-2)。

前と同じ「あやめ科の3種類の植物の4変量観測値」のA群、B群、C群の3つの集団のデータについて、この方法による「重判別分析」を行った。(1)式と(2)式および表 2-11-1 は、その結果である。

$$(1) \quad f_k = 0.13051 x_{1k} + 0.39623 x_{2k} - 0.68637 x_{3k} - 1.00000 x_{4k}$$

$$(2) \quad \text{if } f_k \geq -0.598 (= f_1) \text{ then } \hat{h}_k = 2$$

$$\text{else if } f_k \geq -3.069 (= f_0) \text{ then } \hat{h}_k = 1 \text{ else } \hat{h}_k = 0$$

結果として、「A群」では50個のうち誤判別数 $[n_0]$ は0個、「B群」では50個のうち誤判別数 $[n_1]$ は1個、「C群」では50個のうち誤判別数 $[n_2]$ は0個、誤判別の数の合計 $[n_0 + n_1 + n_2]$ が1件、という“判別力の大きい”解を得ることができた。

ちなみに、あらかじめ「集団の並びの順序」を指定して解を見つけだしてみると、その解の目的への適合度は、「集団の並びの順序」を指定しない場合に比べて小さくなってしまった。

(3)式と(4)式および表 2-11-2 は、あらかじめB群、A群、C群という順序に並べた場合の重判別分析の結果である。この場合には、結果は、A群では50個のうち誤判別数 $[n_0]$ は13個、B群では50個のうち誤判別数 $[n_1]$ は18個、C群では50個のうち誤判別数 $[n_2]$ は21個、誤判別の数の合計 $[n_0 + n_1 + n_2]$ が52件、という“判別力の小さい”解になってしまった。したがって、「重判別分析」では、特別な理由つまり根拠がある場合を除いて、あらかじめ「集団の並びの順序」を指定しない方法を採用するのがよい、と考えられる。

$$(3) \quad f_k = 0.14052 x_{1k} - 0.21272 x_{2k} + 0.22184 x_{3k} - 0.73867 x_{4k}$$

$$(4) \quad \text{if } f_k \geq 0.189 (= f_1) \text{ then } \hat{h}_k = 1$$

$$\text{else if } f_k \geq 0.011 (= f_0) \text{ then } \hat{h}_k = 2 \text{ else } \hat{h}_k = 0$$

表 2-11-1 3つの群の「重判別分析」の結果（群の並び方を指定しない場合）

#	イリス・セトーサ (A群)					イリス・ヴェルシコロール (B群)					イリス・ヴィルジニカ (C群)				
	x ₁	x ₂	x ₃	x ₄	得点	x ₁	x ₂	x ₃	x ₄	得点	x ₁	x ₂	x ₃	x ₄	得点
01	5.1	3.5	1.4	0.2	0.891	7.0	3.2	4.7	1.4	-2.444	6.3	3.3	6.0	25	-4.488
02	4.9	3.0	1.4	0.2	0.667	6.4	3.2	4.5	1.5	-2.485	5.8	2.7	5.1	19	-3.574
03	4.7	3.2	1.3	0.2	0.789	6.9	3.1	4.9	1.5	-2.734	7.1	3.0	5.9	21	-4.034
04	4.6	3.1	1.5	0.2	0.599	5.5	2.3	4.0	1.3	-2.416	6.3	2.9	5.6	18	-3.672
05	5.0	3.6	1.4	0.2	0.918	6.5	2.8	4.6	1.5	-2.700	6.5	3.0	5.8	22	-4.144
06	5.4	3.9	1.7	0.4	0.683	5.7	2.8	4.5	1.3	-2.535	7.6	3.0	6.6	21	-4.449
07	4.6	3.4	1.4	0.3	0.687	6.3	3.3	4.7	1.6	-2.696	4.9	2.5	4.5	17	-3.159
08	5.0	3.4	1.5	0.2	0.770	4.9	2.4	3.3	1.0	-1.675	7.3	2.9	6.3	18	-4.022
09	4.4	2.9	1.4	0.2	0.562	6.6	2.9	4.6	1.3	-2.447	6.7	2.5	5.8	18	-3.916
10	4.9	3.1	1.5	0.1	0.738	5.2	2.7	3.9	1.4	-2.328	7.2	3.6	6.1	25	-4.321
11	5.4	3.7	1.5	0.2	0.941	5.0	2.0	3.5	1.0	-1.957	6.5	3.2	5.1	20	-3.384
12	4.8	3.4	1.6	0.2	0.675	5.9	3.0	4.2	1.5	-2.424	6.4	2.7	5.3	19	-3.633
13	4.8	3.0	1.4	0.1	0.754	6.0	2.2	4.0	1.0	-2.091	6.8	3.0	5.5	21	-3.799
14	4.3	3.0	1.1	0.1	0.895	6.1	2.9	4.7	1.4	-2.681	5.7	2.5	5.0	20	-3.697
15	5.8	4.0	1.2	0.2	1.318	5.6	2.9	3.6	1.3	-1.891	5.8	2.8	5.1	24	-4.034
16	5.7	4.4	1.5	0.4	1.058	6.7	3.1	4.4	1.4	-2.317	6.4	3.2	5.3	23	-3.835
17	5.4	3.9	1.3	0.4	0.958	5.6	3.0	4.5	1.5	-2.669	6.5	3.0	5.5	18	-3.538
18	5.1	3.5	1.4	0.3	0.791	5.8	2.7	4.1	1.0	-1.987	7.7	3.8	6.7	22	-4.288
19	5.7	3.8	1.7	0.3	0.783	6.2	2.2	4.5	1.5	-2.908	7.7	2.6	6.9	23	-5.001
20	5.1	3.8	1.5	0.3	0.842	5.6	2.5	3.9	1.1	-2.055	6.0	2.2	5.0	15	-3.277
21	5.4	3.4	1.7	0.2	0.685	5.9	3.2	4.8	1.8	-3.057	6.9	3.2	5.7	23	-4.044
22	5.1	3.7	1.5	0.4	0.702	6.1	2.8	4.0	1.3	-2.140	5.6	2.8	4.9	20	-3.523
23	4.6	3.6	1.0	0.2	1.140	6.3	2.5	4.9	1.5	-3.050	7.7	2.8	6.7	20	-4.484
24	5.1	3.3	1.7	0.5	0.306	6.1	2.8	4.7	1.2	-2.520	6.3	2.7	4.9	18	-3.271
25	4.8	3.4	1.9	0.2	0.470	6.4	2.9	4.3	1.3	-2.267	6.7	3.3	5.7	21	-3.830
26	5.0	3.0	1.6	0.2	0.543	6.6	3.0	4.4	1.4	-2.370	7.2	3.2	6.0	18	-3.711
27	5.0	3.4	1.6	0.4	0.502	6.8	2.8	4.8	1.4	-2.698	6.2	2.8	4.8	18	-3.176
28	5.2	3.5	1.5	0.2	0.836	6.7	3.0	5.0	1.7	-3.069	6.1	3.0	4.9	18	-3.178
29	5.2	3.4	1.4	0.2	0.865	6.0	2.9	4.5	1.5	-2.657	6.4	2.8	5.6	21	-3.999
30	4.7	3.2	1.6	0.2	0.583	5.7	2.6	3.5	1.0	-1.628	7.2	3.0	5.8	16	-3.453
31	4.8	3.1	1.6	0.2	0.557	5.5	2.4	3.8	1.1	-2.039	7.4	2.8	6.1	19	-4.012
32	5.4	3.4	1.5	0.4	0.622	5.5	2.4	3.7	1.0	-1.871	7.9	3.8	6.4	20	-3.856
33	5.2	4.1	1.5	0.1	1.174	5.8	2.7	3.9	1.2	-2.050	6.4	2.8	5.6	22	-4.099
34	5.5	4.2	1.4	0.2	1.221	6.0	2.7	5.1	1.6	-3.248	6.3	2.8	5.1	15	-3.069
35	4.9	3.1	1.5	0.2	0.638	5.4	3.0	4.5	1.5	-2.695	6.1	2.6	5.6	14	-3.417
36	5.0	3.2	1.2	0.2	0.897	6.0	3.4	4.5	1.6	-2.558	7.7	3.0	6.1	23	-4.293
37	5.5	3.5	1.3	0.2	1.012	6.7	3.1	4.7	1.5	-2.623	6.3	3.4	5.6	24	-4.074
38	4.9	3.6	1.4	0.1	1.005	6.3	2.3	4.4	1.3	-2.586	6.4	3.1	5.5	18	-3.511
39	4.4	3.0	1.3	0.2	0.671	5.6	3.0	4.1	1.3	-2.195	6.0	3.0	4.8	18	-3.123
40	5.1	3.4	1.5	0.2	0.783	5.5	2.5	4.0	1.3	-2.337	6.9	3.1	5.4	21	-3.678
41	5.0	3.5	1.3	0.3	0.847	5.5	2.6	4.4	1.2	-2.472	6.7	3.1	5.6	24	-4.141
42	4.5	2.3	1.3	0.3	0.306	6.1	3.0	4.6	1.4	-2.573	6.9	3.1	5.1	23	-3.672
43	4.4	3.2	1.3	0.2	0.750	5.8	2.6	4.0	1.2	-2.158	5.8	2.7	5.1	19	-3.574
44	5.0	3.5	1.6	0.6	0.341	5.0	2.3	3.3	1.0	-1.701	6.8	3.2	5.9	23	-4.194
45	5.1	3.8	1.9	0.4	0.467	5.6	2.7	4.2	1.3	-2.382	6.7	3.3	5.7	25	-4.230
46	4.8	3.0	1.4	0.3	0.554	5.7	3.0	4.2	1.2	-2.150	6.7	3.0	5.2	23	-3.806
47	5.1	3.8	1.6	0.2	0.873	5.7	2.9	4.2	1.3	-2.290	6.3	2.5	5.0	19	-3.519
48	4.6	3.2	1.4	0.2	0.707	6.2	2.9	4.3	1.3	-2.293	6.5	3.0	5.2	20	-3.532
49	5.3	3.7	1.5	0.2	0.928	5.1	2.5	3.0	1.1	-1.503	6.2	3.4	5.4	23	-3.850
50	5.0	3.3	1.4	0.2	0.799	5.7	2.8	4.1	1.3	-2.261	5.9	3.0	5.1	18	-3.342
最大値 A					1.318	最大値 B					最大値 C				
最小値 A					0.306	最小値 B					最小値 C				

表 2-11-2 3つの群の「重判別分析」の結果（“BAC”の順序を指定した場合）

##	イリス・セトーサ (A群)					イリス・ヴェルシコロール (B群)					イリス・ヴィルジニカ (C群)				
	X ₁	X ₂	X ₃	X ₄	得点	X ₁	X ₂	X ₃	X ₄	得点	X ₁	X ₂	X ₃	X ₄	得点
01	5.1	3.5	1.4	0.2	0.135	7.0	3.2	4.7	1.4	0.311	6.3	3.3	6.0	2.5	-0.332
02	4.9	3.0	1.4	0.2	0.213	6.4	3.2	4.5	1.5	0.109	5.8	2.7	5.1	1.9	-0.031
03	4.7	3.2	1.3	0.2	0.120	6.9	3.1	4.9	1.5	0.289	7.1	3.0	5.9	2.1	0.117
04	4.6	3.1	1.5	0.2	0.172	5.5	2.3	4.0	1.3	0.211	6.3	2.9	5.6	1.8	0.181
05	5.0	3.6	1.4	0.2	0.100	6.5	2.8	4.6	1.5	0.230	6.5	3.0	5.8	2.2	-0.063
06	5.4	3.9	1.7	0.4	0.011	5.7	2.8	4.5	1.3	0.243	7.6	3.0	6.6	2.1	0.343
07	4.6	3.4	1.4	0.3	0.012	6.3	3.3	4.7	1.6	0.044	4.9	2.5	4.5	1.7	-0.101
08	5.0	3.4	1.5	0.2	0.164	4.9	2.4	3.3	1.0	0.171	7.3	2.9	6.3	1.8	0.477
09	4.4	2.9	1.4	0.2	0.164	6.6	2.9	4.6	1.3	0.371	6.7	2.5	5.8	1.8	0.367
10	4.9	3.1	1.5	0.1	0.288	5.2	2.7	3.9	1.4	-0.013	7.2	3.6	6.1	2.5	-0.247
11	5.4	3.7	1.5	0.2	0.157	5.0	2.0	3.5	1.0	0.315	6.5	3.2	5.1	2.0	-0.113
12	4.8	3.4	1.6	0.2	0.158	5.9	3.0	4.2	1.5	0.015	6.4	2.7	5.3	1.9	0.097
13	4.8	3.0	1.4	0.1	0.273	6.0	2.2	4.0	1.0	0.524	6.8	3.0	5.5	2.1	-0.014
14	4.3	3.0	1.1	0.1	0.136	6.1	2.9	4.7	1.4	0.249	5.7	2.5	5.0	2.0	-0.099
15	5.8	4.0	1.2	0.2	0.083	5.6	2.9	3.6	1.3	0.008	5.8	2.8	5.1	2.4	-0.422
16	5.7	4.4	1.5	0.4	-0.098	6.7	3.1	4.4	1.4	0.224	6.4	3.2	5.3	2.3	-0.305
17	5.4	3.9	1.3	0.4	-0.078	5.6	3.0	4.5	1.5	0.039	6.5	3.0	5.5	1.8	0.166
18	5.1	3.5	1.4	0.3	0.061	5.8	2.7	4.1	1.0	0.412	7.7	3.8	6.7	2.2	0.135
19	5.7	3.8	1.7	0.3	0.148	6.2	2.2	4.5	1.5	0.294	7.7	2.6	6.9	2.3	0.361
20	5.1	3.8	1.5	0.3	0.019	5.6	2.5	3.9	1.1	0.308	6.0	2.2	5.0	1.5	0.376
21	5.4	3.4	1.7	0.2	0.265	5.9	3.2	4.8	1.8	-0.116	6.9	3.2	5.7	2.3	-0.146
22	5.1	3.7	1.5	0.4	-0.033	6.1	2.8	4.0	1.3	0.189	5.6	2.8	4.9	2.0	-0.199
23	4.6	3.6	1.0	0.2	-0.045	6.3	2.5	4.9	1.5	0.332	7.7	2.8	6.7	2.0	0.495
24	5.1	3.3	1.7	0.5	0.022	6.1	2.8	4.7	1.2	0.418	6.3	2.7	4.9	1.8	0.068
25	4.8	3.4	1.9	0.2	0.225	6.4	2.9	4.3	1.3	0.276	6.7	3.3	5.7	2.1	-0.047
26	5.0	3.0	1.6	0.2	0.272	6.6	3.0	4.4	1.4	0.231	7.2	3.2	6.0	1.8	0.332
27	5.0	3.4	1.6	0.4	0.039	6.8	2.8	4.8	1.4	0.391	6.2	2.8	4.8	1.8	0.011
28	5.2	3.5	1.5	0.2	0.171	6.7	3.0	5.0	1.7	0.157	6.1	3.0	4.9	1.8	-0.024
29	5.2	3.4	1.4	0.2	0.170	6.0	2.9	4.5	1.5	0.117	6.4	2.8	5.6	2.1	-0.005
30	4.7	3.2	1.6	0.2	0.187	5.7	2.6	3.5	1.0	0.286	7.2	3.0	5.8	1.6	0.478
31	4.8	3.1	1.6	0.2	0.222	5.5	2.4	3.8	1.1	0.293	7.4	2.8	6.1	1.9	0.394
32	5.4	3.4	1.5	0.4	0.073	5.5	2.4	3.7	1.0	0.344	7.9	3.8	6.4	2.0	0.244
33	5.2	4.1	1.5	0.1	0.117	5.8	2.7	3.9	1.2	0.219	6.4	2.8	5.6	2.2	-0.079
34	5.5	4.2	1.4	0.2	0.042	6.0	2.7	5.1	1.6	0.218	6.3	2.8	5.1	1.5	0.313
35	4.9	3.1	1.5	0.2	0.214	5.4	3.0	4.5	1.5	0.011	6.1	2.6	5.6	1.4	0.512
36	5.0	3.2	1.2	0.2	0.140	6.0	3.4	4.5	1.6	-0.064	7.7	3.0	6.1	2.3	0.098
37	5.5	3.5	1.3	0.2	0.169	6.7	3.1	4.7	1.5	0.217	6.3	3.4	5.6	2.4	-0.368
38	4.9	3.6	1.4	0.1	0.159	6.3	2.3	4.4	1.3	0.412	6.4	3.1	5.5	1.8	0.130
39	4.4	3.0	1.3	0.2	0.121	5.6	3.0	4.1	1.3	0.098	6.0	3.0	4.8	1.8	-0.060
40	5.1	3.4	1.5	0.2	0.178	5.5	2.5	4.0	1.3	0.168	6.9	3.1	5.4	2.1	-0.043
41	5.0	3.5	1.3	0.3	0.025	5.5	2.6	4.4	1.2	0.309	6.7	3.1	5.6	2.4	-0.248
42	4.5	2.3	1.3	0.3	0.210	6.1	3.0	4.6	1.4	0.205	6.9	3.1	5.1	2.3	-0.257
43	4.4	3.2	1.3	0.2	0.078	5.8	2.6	4.0	1.2	0.263	5.8	2.7	5.1	1.9	-0.031
44	5.0	3.5	1.6	0.6	-0.130	5.0	2.3	3.3	1.0	0.207	6.8	3.2	5.9	2.3	-0.115
45	5.1	3.8	1.9	0.4	0.034	5.6	2.7	4.2	1.3	0.184	6.7	3.3	5.7	2.5	-0.343
46	4.8	3.0	1.4	0.3	0.125	5.7	3.0	4.2	1.2	0.208	6.7	3.0	5.2	2.3	-0.242
47	5.1	3.8	1.6	0.2	0.116	5.7	2.9	4.2	1.3	0.156	6.3	2.5	5.0	1.9	0.059
48	4.6	3.2	1.4	0.2	0.129	6.2	2.9	4.3	1.3	0.248	6.5	3.0	5.2	2.0	-0.049
49	5.3	3.7	1.5	0.2	0.143	5.1	2.5	3.0	1.1	0.038	6.2	3.4	5.4	2.3	-0.353
50	5.0	3.3	1.4	0.2	0.163	5.7	2.8	4.1	1.3	0.155	5.9	3.0	5.1	1.8	-0.007
平均値 A					0.117	平均値 B					平均値 C				
最大値 A					0.288	最大値 B					最大値 C				
最小値 A					-0.130	最小値 B					最小値 C				

2-1-2 節 「計算的な判別分析」の特徴のまとめ

この節では、いままでに説明した判別分析の「計算的な方法」の特徴の整理をする。整理をすることによって、この論文の主題である、数学的な「印象の推定モデル」を構築するための方法としての判別分析の「計算的な方法」の“特徴”つまり利点と問題点を明らかにし、また、同じ目的のためのほかの方法との“違い”を明らかにするためである。

しかし、この章で取り上げた複数の連続的な説明変数と1つの離散的な被説明変数を対応づける方法に限ったとしても、説明変数と被説明変数を対応づける「数学的なモデル」を構築するための方法は数多くあり、また、「数学的なモデル」を構築・利用する目的に応じて、ある同じ特徴が利点であるとも欠点であるとも評価されたりする。ここでは、判別分析の「計算的な方法」について、判別分析の「解析的な方法」との比較を中心とした違いを明らかにし、併せて、「ニューラルネットワークモデル」および「重回帰分析モデル」の方法との比較を試みる。

最初は、判別分析の「解析的な方法」と「計算的な方法」の違いの整理である。前節までの説明の繰り返しになるが、「判別分析」の論理は、分析をする対象の複数の属性値（連続値）を「判別式」によって1つの値つまり「判別得点」に変換し、この値の大きさによって、それぞれの対象がある集団に属するか、あるいは、ある集団には属さないつまりもう1つの別の集団に属するかを判別あるいは推定をしよう、という方法つまりモデルである。「印象の推定モデル」としては、対象のカラー画像の複数の特徴量を説明変数とし、そのカラー画像に対する特定の印象の有無などを被説明変数とした「判別分析」を行なおう、というのである。

この判別分析に対して、すでに開発され広く使われている「解析的な方法」は、「判別式の形式」として、1次式あるいは2次式を想定し、また、「判別力の高さ」（大きさ）つまり「判別のよさ」を「相関比を最大にすること」と定義して、解である判別式つまり判別式の係数（パラメーター）を求めるのである。この方法では、判別式の形式は、計算技術的な意味から1次式と2次式に限定されている。また、相関比を計算するので、分析をする“すべて”の対象に対して、同じ割合の考慮をすることを前提としている。

これに対して、新たに開発した「計算的な方法」は、「判別力の高さ」を実際の問題からの必要に対応して、どのようにでも定義をすることができる。これは「計算的な方法」の非常に大切な特徴である。「判別力の高さ」を（「解析的な方法」と同じ）「相関比を最大にする」と定義することも、また、「誤判別数の合計を最小にする」「（判別得点の値を尺度にした）

2つの集団に含まれる対象の間の最短距離を最大にする」「(判別得点の値を尺度にした) 2つの集団に含まれる対象の90%値の間の最短距離を最大にする」などの定義を“柔軟に”採用することができる。

たとえば「誤判別数の合計を最小にする」を採用した場合には、求めるべき解の関心は、対象の“すべて”ではなく、その解(つまり判別式)によって計算された判別得点の分布について、(判別すべき) 2つの集団が“重なる”あるいは“近接した”対象で、誤判別を生じる可能性があるものだけに限られる。判別得点の分布が“重ならない”あるいは“離れた”対象で、誤判別を生じる可能性がないものについては、ほとんど関心を持たない。

また、「2つの集団に含まれる対象の間の最短距離を最大にする」を採用した場合にも、求めるべき解の関心は、対象の“すべて”ではなく、一方の集団に属する対象のうちの判別得点が最大のものと他方の集団に属する対象のうちの判別得点が最小のもの、あるいは、これらに該当する可能性のある対象だけに限られ、これ“以外”の対象については、ほとんど関心を持たない。

あるいは、「判別力の高さ」として「2つの集団に含まれる対象の90%値の間の(判別得点の上での)最短距離を最大にする」を採用した場合には、求めるべき解の関心は、判別すべき2つの集団が“重なる”あるいは“近接した”対象だけではなく、一方の集団に属する対象のうちの判別得点が下から90%に当たるものと、他方の集団に属する対象のうちの判別得点が上から90%に当たるもの、あるいは、これらに該当する可能性のある対象に広がるが、この場合も、それ以外の対象についてはほとんど関心を持たない。

これに対して、「解析的な方法」では、「判別力の高さ」に対して、実際の問題の必要に応じて、こういった“柔軟な”定義を採用することはできず、求めるべき解は、分析すべき対象の“すべて”に等しく関心を持つのである。実際の問題への適用において、この違いは非常に大きい。

次は、「判別式の形式」の選択である。「解析的な方法」では、判別式の形式は、1次式あるいは2次式であり、このほかの形式は採用することができない。これは、“もっとも判別力が高い”解を求める方法として、「判別力の高さ」を表す関数が微分できる、という計算技術的な要請からの制約である。

これに対して、「計算的な方法」では、判別式の形式はどんなものであってもよい。1次式、2次式、3次式のほか、たとえば絶対値関数、四捨五入関数、剰余関数、周期関数など、必要

に応じて多種多様な関数を取り入れることも可能である。したがって、対象のデータに、複雑な交互作用や非線型性が含まれている場合にも、判別式の形式の工夫によって対応がしやすい。「判別力の高さ」を少しでも向上させたいという点で、この違いも非常に大きい。

さらに、2-2節の繰り返しになるが、「解析的な方法」では、(判別式の形式が1次式の)線型判別の場合、分析をする2つの集団の対象のデータのばらつきと相関は同じつまり分散共分散行列が等しく、また、正規分布にしていることが前提とされている。実際に分析するデータに、この前提条件がなりたっていれば、「相関比を最大にする」解と「誤判別数の合計を最小にする」解と一致する。しかし、たとえば分散共分散行列が等しくない場合には、2つの集団の誤判別数の一方が大きく他方が小さくなるという“アンバランス”を生じてしまう。また、2つの集団の対象のデータが正規分布にしていなかった場合には、2つの集団の誤判別数の合計はより大きくなってしまう。2-7節でも述べたように、分散共分散行列が等しくない場合には、「線型判別」は不適当であり、判別式の形式は「2次式」の採用が勧められ、判別の論理もマハラノビスの汎距離によるものがよい、といわれている。

これに対して、「計算的な方法」では、説明変数のデータにはなんの仮説も前提条件も必要としていない。たとえば、説明変数のデータの間に高度な相関関係や因果関係があった場合でも、この方法では、仕事の目的に応じて設定した「判別力の高さ」の値をできるだけ大きくする解を見つけ出すことができるのである。分析するデータが必ずしもきれいなものではない実際の問題への適用を考えると、この違いも非常に大きい。

また、「解析的な方法」は、分析すべき“すべて”の対象を“一時に”^{いちどき}判別しよう、というやり方である。これに対して、「計算的な方法」のうち、(データが交互作用や非線型性を含んだ)複雑な問題に対応するための「2段階判別方式」やこれを拡張した「“逐次的な”判別方式」は、分析の対象のなかから、まずはより確実に判別できる対象から判別していき、残った対象についても、より確実に判別できる対象から、“段階的に”あるいは“逐次的に”判別をしていく、という方法である。この違いは、実務的な立場からいって(も、また、解の「判別力の高さ」からいっても)非常に大きい。2-9節でも述べたように、「解析的な方法」にこういったやり方を持ち込むことで、「解析的な方法」に基づいた「2段階判別方式」や「“逐次的な”判別方式」も可能である。

また、3つ以上の集団に属する対象を判別する「重判別分析」についても、「解析的な方法」では、あらかじめ分析すべき集団の「並びの順序」を“決めて”おき、その順序を前提として

分析をし、たまたまよい結果が出ればよいが、そうではない場合には、別の「並びの順序」を“決めて”分析を試みるということが必要であった。これに対して、「計算的な方法」では、仮に解（つまり判別式）を決め、これによってそれぞれの集団に属する対象の判別得点の値を計算し、この結果に基づいて、「並びの順序」が（“決める”のではなく）“決まる”ので、分析者が集団の「並びの順序」を“決める”という煩雑な作業から開放されている。分析すべき集団の数が大きくなった場合、「並びの順序」の組み合わせの数は、階乗的に増えていくので、この違いも非常に大きい。

いずれにしても、判別分析を利用して、対象の複数の“連続的な”属性値に対して、特定の印象の有無などを対応づける「印象の推定モデル」を構築する場合、実際に存在する問題に対して、「解析的な方法」が持っているいろいろな制約を、「計算的な方法」によって“柔軟に”超えることができる、という点で優れている、ということができる。敢えて「計算的な方法」の欠点を挙げるとすれば、「計算的な方法」は、コンピューターによる処理を前提とした方法で、手間と時間が必要であり、また、得られる結果は必ずしも最適な解である保証がない、といったことであろうが、これらは、実行可能な解を導き出すことを必要としている実際の仕事の上ではほとんど問題にならない、と筆者は考えている。

「判別分析」と同様に、対象の複数の“連続的な”属性値に対して特定の印象の有無などを対応づける「印象の推定モデル」を構築する場合、「階層型ニューラルネットワーク」や「重回帰分析」が適用できる。では、次に、「階層型ニューラルネットワーク」と「重回帰分析」を利用した、数学的な「印象の推定モデル」との比較をしてみよう。（以下の比較は、「判別分析」を利用した数学的な「印象の推定モデル」との比較であって、「計算的な判別分析」を利用した数学的な「印象の推定モデル」との比較ではないことに注意が必要である。）

まずは、「パターン認識」などにも使われることが多い「階層型ニューラルネットワーク」による対応づけつまりモデルの構築の方法の1つである（文献86）。階層型ニューラルネットワークによっても、対象の複数の“連続的な”属性値に対して特定の印象の有無などを対応づける「印象の推定モデル」の構築ができる。前に挙げた「あやめ科の3種類の植物の4変量観測値」のデータを例にとると、この4種類のデータを入力とし、1つのB群であるかC群であるか（たとえばB群のときは出力の値が〔0. 0〕、C群のときは出力の値が〔1. 0〕とする）を出力とする階層型ニューラルネットワークに、例示したデータを教師データとして学習させ、“すべて”の教師データに対してまったく誤りがなくなるようになれば、この階層型

ニューラルネットワークは、学習した教師データの範囲で“完全な”判別ができる、といえる。

この場合、たとえば出力の値が $[0, 0 \sim 0.4]$ になったときにはB群とみなし、 $[0.6 \sim 1, 0]$ になったときにはC群とみなすのである。この階層型ニューラルネットワークによるモデルは、学習した教師データと“同質”のデータに対しては、判別つまり推定する能力も持っているであろう。（実は、この理由づけもなかなかむずかしいのである。2-13節でも同様の議論をするが、「階層型ニューラルネットワーク」によるモデルの構築は、「パラメーター依存型」の方法論である。したがって、階層型ニューラルネットワークで学習した教師データと“同質”のデータであれば、おそらく学習したニューラルネットワークによるモデルは、判別つまり推定する能力を持っている、と考えてもよいと思われる。しかし、対象のデータとして、なにが“同質”でなにが“異質”かという議論になると、これに対して、きちっとした答えを用意することはかなりむずかしいのである。）

階層型ニューラルネットワークは、“学習”する教師データのなかに複雑な交互作用や非線型性が含まれている場合にも、たとえば中間層（隠れ層）の素子の数を増やすことによって、学習しやすくなるのが特徴である。たとえば入力層、中間層、出力層の3層のニューラルネットワークモデルを想定した場合、原理的には、中間層の素子の数を必要なだけ増やせば、論理的にどのような入力と出力の対応づけも可能である。しかし、このモデルが与えられたデータを“メタ”のレベルでの学習しているか、という質問には、この対応づけだけからはうまく答えることができないであろう。

対象の複数の属性値を変換する方法が判別式であるか階層型ニューラルネットワークであるかは別にして、複数の属性値を1つの値に変換して、その値の大きさで、2つの集団のいずれに属するかを判別する、という意味では、「判別分析」による方法も「階層型ニューラルネットワーク」による方法も、本質的な違いはないである。判別分析の「計算的な方法」で得られる解は、必ずしも最適解ではなく、そこまでの改善で得られた実行可能な解である。これに対して、「階層型ニューラルネットワーク」で得られる解も、必ずしも最適解ではなく、そこまでの学習で得られた実行可能な解なのである。そういった意味でも、この2つの方法の間には、本質的に違いはないのである。ちなみに、「計算的な判別分析」で開発した「2段階判別方式」や「“逐次的”判別方式」は、階層型ニューラルネットワークの出力の値を判別得点と考えれば、まったく同様に適用できる。

「重回帰分析」によっても、上記と同様のことが可能である。「重回帰分析」は、「印象の

推定モデル」としては、「SD値」で計測した特定の印象の程度（大きさ）を説明するモデルとして使われる例が多いが、これを拡張して、特定の印象の有無などを説明するモデルとしても使うことができる。なお、「SD法」で結束した結果は、必ずしも量的尺度に基づくものではないので、連続的な被説明変数に対する方法である「重回帰分析」は、適当ではないかもしれない。上と同じ例を使うと、4種類のデータを説明変数とし、B群であるかC群であるか（たとえばB群のときは〔0. 0〕、C群のときは〔1. 0〕とする）を被説明変数として、（実は、この論理はかなりの無理があるのではあるが、）“もっともよい”回帰式を求め、この結果が十分に有意であれば、この回帰式は、分析したデータの範囲内で、十分な推定能力を持つ、と考えることができる。この場合、たとえば計算した被説明変数の値が〔0. 0～0. 4〕の間ならばB群とみなし、〔0. 6～1. 0〕ならばC群とみなすのである。

「階層型ニューラルネットワーク」によるモデルの構築と同様に、対象の複数の属性値を判別式で変換する代わりに、回帰式で変換したと考えれば、推定した被説明変数の値の大きさを、2つの集団のいずれに属するかを判別する、という論理は、「判別分析」による方法と本質的な差はないといってよいであろう。ただし、重回帰分析は、被説明変数も“連続的”であることを前提としているので、このモデルに対しては、この方法の適用には無理があるともいえる。また、（工夫ができないわけではないが、）重回帰分析は、論理が単純なだけに、分析する対象のデータに交互作用や非線型性が含まれている場合、対応がむずかしい。これらがこの方法の問題点といえるものであろう。（ちなみに、このあと3-2節で、回帰分析の「計算的な方法」を説明する。この方法によれば、分析すべき対象のデータに対する回帰式の「あてはまりのよさ」は、推定誤差の自乗和で評価されるだけでなく、実際の問題の必要に応じたいろいろな定義を採用することができる。）

2-1-3 節 ビジネスバッグのカラー画像に対する「印象の推定モデル」の構築の試み

この節では、「ビジネスバッグ」のカラー画像を対象に、そのタイプの分類およびそれに対して申告された2種類の印象の有無に対して、「計算的な判別分析」を利用して、カラー画像の特徴量とこれに対する特定の印象の有無などを対応付ける「印象の推定モデル」の構築を試みる。

「ビジネスバッグ」のカラー画像は縦480ドット×横704ドットのもので、R（赤色）成分、G（緑色）成分、B（青色）成分はそれぞれ8ビットで表し、 $(2^8=)$ 256階調の区別ができる。これらのカラー画像を代表する特徴量として、(1)式と(2)式などで表される「局所自己相関特徴量」($s_{rr}(\Delta x, \Delta y)$, $s_{rg}(\Delta x, \Delta y)$ など)を採用した。この特徴量は、カラー画像を“局所”つまり数ドットあるいは十数ドットの範囲で変位させた自己相関関数の値で、定常時系列データの特徴量として用いられる「“高次”局所自己相関特徴」(文献87)を(3乗以上の)“高次”ではなく“単純な”(つまり2乗の)相関に限定したものである。ここで、 $[x]$ と $[y]$ はカラー画像の横方向と縦方向の位置、 $[\Delta x]$ と $[\Delta y]$ は相関をとるカラー画像の色成分同士の横方向と縦方向の変位(ズレ)、 $[r(x, y)]$ $[g(x, y)]$ $[b(x, y)]$ はそれぞれ2次元のカラー画像のR（赤色）成分、G（緑色）成分、B（青色）成分の値である。

$$(1) s_{rr}(\Delta x, \Delta y) = \frac{\sum_{\forall x} \sum_{\forall y} [r(x, y) - \bar{r}] [r(x + \Delta x, y + \Delta y) - \bar{r}]}{\sum_{\forall x} \sum_{\forall y} [r(x, y) - \bar{r}]^2}$$

$$(2) s_{rg}(\Delta x, \Delta y) = \frac{\sum_{\forall x} \sum_{\forall y} [r(x, y) - \bar{r}] [g(x + \Delta x, y + \Delta y) - \bar{g}]}{\sqrt{\sum_{\forall x} \sum_{\forall y} [r(x, y) - \bar{r}]^2 \sum_{\forall x} \sum_{\forall y} [g(x, y) - \bar{g}]^2}}$$



図 2-13-1 「局所自己相関特徴」の性質

ちなみに、「“高次”局所自己相関特徴」は、画像の大きさや回転への独立性が高く、画像が大きくても小さくても、多少回転していても、その値は変わらないようであるが、「高次」ではない「局所自己相関特徴」は、対象の位置が違っていても同じ値という性質はあるものの、対象の大きさが違っていたり、回転していたりすると、この特徴量の値は少しく違ってしまふ(図 2-13-1)。このため、「印象の計測」および「印象の推定モデル」の構築のための) 判別分析の対象にする「ビジネスバッグ」のカラー画像は、できるだけ同じ大きさ、同じ向きにそろえたものを用意して、カラー画像の大きさや回転の影響がないようにした。

具体的には、ここで採用した局所の変位は、x 方向に [0~5 ドット]、y 方向に [-5~+5 ドット] である。(ちなみに、x 方向に [6 ドット]、y 方向に [11 ドット] の局所的な変位を与えている理由は、「ビジネスバッグ」のカラー画像の特徴が、横方向の変化よりも縦方向の変化の方が大きいと考えられるからである。また、y 方向の変位の [-5~+5 ドット] は、座標の取り方の問題で、これは [0~+10 ドット] であっても構わない。) この特徴量は、3 種類の色成分のうちの 2 つの成分の組み合わせ、つまり、RR、GG、BB、RG、GR、GB、BG、BR、RB の 9 種類があるが、変位が [0] の場合、たとえば R (赤色) と G (緑色) の組み合わせの特徴量 $[s_{rg}(\Delta x, 0) \text{ or } s_{rg}(0, \Delta y)]$ と G (緑色) と R (赤色) の組み合わせの特徴量 $[s_{gr}(\Delta x, 0) \text{ or } s_{gr}(0, \Delta y)]$ はまったく同じ値であるため、その一方のみを採用した。互いに異なる特徴量は全部で 546 個 (= 6 × 11 × 9 個 - 16 × 3 個) である。

後述するように、「ビジネスバッグ」のカラー画像の特徴量の組み合わせとしては、このうち、判別分析の対象とするカラー画像の違いによる“変化量大きい”14 個を選択した。本来ならば、より少ない数のまた“互いに独立な”特徴量を選択した方が、得られる解はより“洗練された”ものとなるかもしれないが、こうした特徴量の組み合わせの選択の方法は、

ここでの検討の主題ではない。ここでの検討の主題は、なにかの方法によって選択されたカラー画像の特徴量の組み合わせに対して、これらのカラー画像が与えられた2つの集団のいずれに属するかの判別について分析して、十分に“実用的な”解が得られるかいなかを検討することである。

また、前節までの説明で使った「あやめ科の3種類の植物の4変量観測値」によるB群とC群の判別では、B群50件とC群50件の合計100件の対象に対して、4つの変量での説明を試みた。これに対して、ここでの試みは、それぞれ20件ずつの2つの集団の対象の合計40件に対して、より多くの14個の変量での説明を試みようとしている。当然のことながら、説明変数の数が増えれば判別がしやすい。しかし、繰り返しになるが、ここでの試みの主題は、“もっとも適当な”説明変数の組み合わせを示すことではなく、判別分析の「計算的な方法」によって、十分に“実用的な”解が得られることを例示的に示すことなのである。

では、第1のケースとして、判別すべき「ビジネスバッグ」のカラー画像の40件を図2-13-2に示す。これらのカラー画像は、(1-4-2項で説明した)「自由申告法」による「印象の計測」で使用したもので、200件のカラー画像の中から、以下の判別分析の対象となる仕様(分類)のものを選択した。図中、カラー画像の下番号は、対象のサンプル番号である。最初の20件([#01]~[#20])は「アタッシュケース」であり、これにはソフトアタッシュとハードアタッシュの両方が含まれている。残りの20件([#1]~[#20])は「ショルダーバッグ」であり、これには、^{まち}襠の厚いものと薄いものの両方が含まれている。また、「アタッシュケース」「ショルダーバッグ」とともに、その材質としては、皮革、人工皮革、ビニールなどが含まれている。

これらの40件のビジネスバッグのカラー画像について、546個の局所自己相関特徴量を計算し、この中から、カラー画像の違いによって“特徴量の変化の大きい”14個の特徴量($GR(2,0)$, $RG(1,1)$, $RG(2,-1)$, $GR(1,2)$, $BG(1,-2)$, $GR(1,-5)$, $RR(2,-4)$, $GG(2,5)$, $BR(5,2)$, $BB(3,3)$, $GR(4,3)$, $BG(4,-3)$, $BR(5,-4)$, $GG(5,-5)$)を選択した(表2-13-1)。

たとえば、 $[RG(2,-1)]$ は、画像のR成分を基準にG成分をx方向に[2ドット]、y方向に[-1ドット]ずらして計算した相互相関係数の値である((2)式)。また、 $[GG(2,5)]$ は、画像のG成分自身をx方向に[2ドット]、y方向に[5ドット]ずらして計算した自己相関係数の値である((1)式)。表中、数値はスペースの関係で、たとえば $[-0.123456]$ は $[-.123456]$ のように、整数部第1位の[0]を省略して記してある。



図 2-13-2 「アタッシュケース」と「ショルダーバッグ」のカラー画像

表 2-13-1 「アタッシュケース」と「ショルダーバッグ」のカラー画像の特徴量と判別の結果

アタッシュケース

#	GR	RG	RG	GR	BG	GR	RR	GG	BR	BB	GR	BG	BR	GG	
	(2,0)	(1,1)	(2,-1)	(1,2)	(1,-2)	(1,-5)	(2,-4)	(2,5)	(5,2)	(3,3)	(4,3)	(4,-3)	(5,-4)	(5,-5)	得点
01	.529314	.526926	.525575	.522572	.519417	.512774	.514280	.512323	.516706	.512659	.516072	.512982	.509096	.509719	-.000156
02	.582291	.581223	.580619	.579048	.563168	.573852	.580171	.568473	.565326	.549905	.576317	.559898	.562226	.568465	-.000214
03	.509251	.507934	.507159	.506061	.496698	.497958	.511771	.491021	.495775	.490021	.501120	.491100	.490727	.488727	-.000175
04	.520865	.519217	.518707	.517358	.487976	.510964	.533068	.495099	.495907	.475069	.514349	.483398	.490922	.493664	-.000177
05	.526894	.525566	.524723	.522845	.520703	.513810	.517024	.512118	.517582	.514688	.516904	.515182	.512425	.510196	-.000205
06	.533414	.532484	.530867	.529472	.528514	.519532	.521154	.519569	.524200	.522737	.523234	.521591	.516544	.515995	-.000247
07	.532987	.531522	.530515	.528561	.522376	.519888	.523375	.517873	.520144	.514553	.523412	.517009	.514192	.516026	-.000220
08	.495587	.494340	.493693	.492622	.491154	.487462	.520635	.482882	.483311	.491307	.490418	.486295	.479519	.481073	-.000210
09	.524740	.523441	.522460	.520798	.520218	.511904	.513721	.510928	.515757	.514329	.514701	.514222	.510001	.508533	-.000174
10	.525753	.524762	.524119	.522892	.497272	.517629	.567942	.498356	.501535	.490277	.520389	.492226	.497265	.496487	-.000189
11	.526453	.525452	.524533	.522940	.520693	.515906	.516863	.516129	.517231	.514608	.518303	.516350	.512922	.514912	-.000235
12	.551254	.550023	.549145	.547439	.547821	.539924	.541583	.539649	.544594	.544135	.542999	.543252	.539781	.537785	-.000183
13	.511014	.509829	.509152	.507221	.500642	.499322	.504408	.496104	.497553	.493343	.501925	.495098	.492410	.494377	-.000236
14	.561708	.560539	.560138	.558847	.539495	.552543	.598536	.536559	.542986	.532774	.555912	.534508	.538455	.535020	-.000167
15	.497939	.496657	.495496	.493719	.492646	.483771	.485577	.483304	.486726	.485206	.486320	.485902	.480994	.480494	-.000193
16	.593259	.592098	.591840	.590265	.575835	.583926	.596352	.576014	.578807	.567448	.587194	.571918	.574470	.575011	-.000213
17	.503155	.502232	.500886	.499705	.495346	.491842	.494715	.490237	.492226	.488448	.494378	.489515	.486744	.487649	-.000227
18	.529753	.528817	.528340	.527007	.504966	.520358	.558628	.504005	.509428	.497863	.523944	.499472	.503843	.501940	-.000220
19	.560164	.559091	.557979	.556683	.551275	.547946	.552965	.545356	.549653	.544514	.551530	.546153	.544155	.543549	-.000248
20	.489235	.488636	.487209	.484618	.474200	.475394	.483512	.469536	.472203	.465498	.476969	.468009	.466783	.466882	-.000222
最大	.593259	.592098	.591840	.590265	.575835	.583926	.598536	.576014	.578807	.567448	.587194	.571918	.574470	.575011	-.000156
最小	.489235	.488636	.487209	.484618	.474200	.475394	.483512	.469536	.472203	.465498	.476969	.468009	.466783	.466882	-.000248

ショルダーバッグ

#	GR	RG	RG	GR	BG	GR	RR	GG	BR	BB	GR	BG	BR	GG	
	(2,0)	(1,1)	(2,-1)	(1,2)	(1,-2)	(1,-5)	(2,-4)	(2,5)	(5,2)	(3,3)	(4,3)	(4,-3)	(5,-4)	(5,-5)	得点
01	.534831	.533971	.532629	.531312	.524211	.523999	.525466	.523208	.519838	.513081	.526179	.518617	.515337	.521444	-.000258
02	.525924	.525191	.523445	.522577	.515025	.512963	.517402	.510624	.510910	.506534	.515339	.508391	.504805	.506981	-.000249
03	.514195	.513402	.512750	.511485	.490787	.504929	.519007	.496101	.492109	.478349	.507172	.486351	.487472	.495078	-.000252
04	.474747	.474635	.472854	.471767	.467256	.464126	.463661	.464220	.457689	.454663	.463878	.459628	.453443	.461168	-.000254
05	.569270	.568776	.567237	.566519	.564614	.558791	.560113	.558950	.560166	.559635	.561113	.558552	.554948	.554967	-.000254
06	.580588	.579708	.579085	.577821	.552769	.572290	.599625	.557257	.557093	.543764	.574809	.548493	.553117	.555961	-.000256
07	.559408	.558781	.557999	.556769	.535888	.550617	.583031	.535562	.539300	.530216	.553603	.531359	.535206	.533957	-.000273
08	.565733	.564622	.563961	.562353	.527324	.554756	.578995	.538091	.535300	.510328	.558464	.522718	.530096	.537257	-.000289
09	.517833	.517578	.515879	.515010	.493754	.508545	.518184	.502923	.486448	.477827	.508008	.487624	.483084	.500673	-.000335
10	.542522	.541174	.540374	.538570	.538429	.529857	.533364	.529322	.535710	.534639	.534383	.533823	.530062	.528136	-.000249
11	.534094	.533214	.532617	.531537	.477775	.527181	.559075	.503093	.489968	.454866	.529501	.473993	.486913	.502514	-.000322
12	.545857	.544660	.544088	.541910	.546655	.535864	.530605	.542045	.536894	.541845	.537428	.542369	.533409	.541436	-.000297
13	.568577	.567383	.566455	.564335	.561946	.555114	.557483	.554221	.558864	.555326	.558754	.556484	.552918	.552884	-.000275
14	.491945	.491348	.490229	.488913	.481847	.482662	.485378	.480923	.479083	.473853	.484518	.477840	.475467	.480142	-.000265
15	.496253	.495767	.494303	.493018	.489193	.484119	.487945	.481785	.485036	.482128	.486643	.482736	.479676	.479114	-.000274
16	.559555	.558333	.557654	.555782	.544649	.548093	.555757	.543891	.544789	.535778	.551843	.540080	.539749	.542872	-.000288
17	.605082	.603725	.603542	.601466	.587930	.593885	.604631	.588131	.589526	.580202	.597718	.583834	.584804	.586699	-.000270
18	.591705	.591604	.589712	.589369	.587844	.581218	.582874	.580032	.581452	.581238	.582306	.580813	.577280	.576726	-.000322
19	.470127	.469088	.467958	.465789	.465740	.457784	.457232	.459190	.457621	.458967	.459896	.459743	.452311	.457568	-.000261
20	.573030	.571837	.570894	.569056	.569035	.559836	.561482	.560411	.565091	.564320	.563860	.564154	.560003	.559132	-.000303
最大	.605082	.603725	.603542	.601466	.587930	.593885	.604631	.588131	.589526	.581238	.597718	.583834	.584804	.586699	-.000249
最小	.470127	.469088	.467958	.465789	.465740	.457784	.457232	.459190	.457621	.454663	.459896	.459628	.452311	.457568	-.000335

判別式の形式として「1次式」を採用して、「計算的な方法」を適用した。遺伝的アルゴリズムの“戦略”として、1つの世代に20個の染色体を用意し、初期解つまり第1世代の染色体の1つに、「解析的な方法」で得られた“最適解”を取り入れ、残りはランダムに生成した。遺伝的操作として「優先生き残り」「2個のトーナメントをした後のランダム交叉」「突然変異」「ランダム生き残り」を行い、それぞれ2個、10個、5個、3個を生成した。ランダム交叉の交叉率は100%、突然変異の変異率は25%である。計算は“もっともよい”解に改善が見られなくなった約400世代で打ち切った。

結果は、個々の特徴量を $[x_1] \dots [x_{14}]$ として、(3)式の判別式で表され、この解で40件の対象は“完全に”判別することができた。つまり、「アタッシュケース」の対象の判別得点の最小値は $[-0.000248]$ であり、「ショルダーバッグ」の対象の判別得点の最大値は $[-0.000249]$ となり、前者は後者よりも大きい。

(この結果では、説明変数のデータの有効桁数は6桁、判別式の係数の有効桁数は4桁で、判別得点の有効桁数は4桁である。この点から考えても、「アタッシュケース」の判別得点の最小値と「ショルダーバッグ」の判別得点の最大値の差は非常に小さい。このような場合には、前述した「2段階判別方式」を適用して、最初の判別式はこの判別式で“確実に”判別できる対象にだけ適用し、最初の判別式で変換された判別得点の上で、2つの集団の対象が“近接した”。つまりこの判別式では“確実な”判別ができない部分では、これに対応した“別”の判別式を用意して、これによって判別した方が、より“安定した”解を得ることができであろう。あるいは、この“対応づけ”に非線型性や交互作用が含まれていると考えられる場合には、判別式の形式に、2次式や絶対値式などを採用するなどの工夫を加えることによって、より“確実な”対応づけとすることもできるであろう。これは、このあとに述べる第2と第3のケースでもまったく同様である。)

$$(3) \quad f = 0.08220 x_1 - 0.08830 x_2 + 0.00897 x_3 - 0.04036 x_4 - 0.03563 x_5 \\ - 0.07210 x_6 - 0.00070 x_7 + 0.03633 x_8 + 0.02300 x_9 - 0.03385 x_{10} \\ - 0.02207 x_{11} + 0.10585 x_{12} - 0.02971 x_{13} - 0.07807 x_{14}$$

ところで、ある集団に“属する”対象に対して行なった判別分析の結果、得られた解がその集団の対象に“類似した”対象に対しても、その集団の対象に“属する”対象に対してと同様に適用できるかという評価は、大変にむずかしい問題を含んでいる。これは、判別分析を利用した「印象の推定モデル」の推定能力が、判別分析で分析した対象“以外”の対象にどのくらい適用できるかという評価をするということと同じである。

これに対して、分析すべき対象の集団のなかから“ランダムに”選択した対象の集団について複数回の分析を行い、これらの分析の結果がそれほど違っていない、したがって、その分析からは除かれた対象に対してもその分析で得られた解が適用できることをもって、その解は“安定的”である、と評価する考え方がある。

別のいい方をすれば、このやり方は、“部分的に”選択された対象から得られた解が選択されなかった対象あるいは対象の全体にも適用できるならば、分析した対象の集団のなかに、（ほかの対象とは特別に違ったという意味での）“異質な”対象が含まれていない、つまり、分析した対象の集団がすべて“同質”であった、と評価できるということである。J・W・テューキーの「ジャックナイフ法」（文献88）やB・エフロンの「ブートストラップ法」（文献89、90）などを含めて、「クロスバリデーション」（つまり、交叉妥当性の検証）の方法は、こういった考えに基づいているといえる。

しかし、なにが“同質”の対象であり、なにが“異質”の対象であるかを示すことは大変にむずかしいことである。筆者は、実は、こういった方法は、本質的には、解の未知の対象に対する推定能力の評価にはなっておらず、推定の対象となった対象が、分析に使われた対象と“同質”かいなかを評価しているに過ぎない、と考えている。ちなみに、M・G・ケンドールは、文献91のなかで、判別分析の方法の前提として、「グループは、はじめから決められている。異質の混入した対象を、どういう母集団またはグループに分けるのが最良かを見つけることは、ここで要求されている目的ではない。それらは、すでに決まっているのである。」と述べ、判別分析の目的が、必ずしも“未知”の対象の判別あるいは推定にあるのではない、と説明している。

上記の見解はさておき、(3)式の解が“安定した”ものであるかどうかの評価の1つとして、「アタッシュケース」と「ショルダーバッグ」それぞれのなかから“ランダムに”選択したn件ずつの対象の組み合わせについて、上記と同じ「計算的な方法」によって判別分析を行なって得られた解（判別式）が、選択されなかった対象に対して“有効”であるかどうかを検証する方法が考えられる。しかし、ここで“選択した”判別分析の目的は、「『アタッシュケース』のなかでの判別得点の最小値と『ショルダーバッグ』のなかでの判別得点の最大値の差を最大にする」であるので、分析のために“選択された”対象のなかに、「アタッシュケース」のなかで判別得点が最小である[#19]と「ショルダーバッグ」のなかで判別得点が最大である[#02]と[#10]が含まれていれば、これらの対象がほかの対象に“優先”して解のパ

ラメーターの決定に関与するため、解はあまり変化しない、と考えられる。

実際にいくつかのケースに対してこの方法を試みたが、(3)式の解はほとんど改善されなかった。また、「アタッシュケース」のなかで判別得点が最小である[#19]と「ショルダーバッグ」のなかで判別得点が最大である[#02]と[#10]が含まれていなかった場合でも、「アタッシュケース」のなかで判別得点が次に最小である[#06]、[#13]、[#11]、[#17]や、「ショルダーバッグ」のなかで判別得点が次に最大である[#03]、[#04]、[#05]が含まれていれば、これらの対象がほかの対象に“優先”して、解のパラメーターの決定に関与するので、多分、(3)式の解はそれほどの変化をしないであろう。

したがって、ここで“選択した”（「2つの集団の最短距離を最大にする」という）判別分析の目的は、“安定した”解を導き出すことができる、といえるのかもしれない。しかし、一方では、すべての対象を含んだ集団と、そこからその“一部”の対象を除いた集団とは、その対象の“重なり”が大きく、解を導き出す計算の“重なり”も大きいので、導き出される解もほとんど同じになるのは当然であり、この検討の結果は、必ずしも解の“安定性”を示すものではない、という主張も考えられる。

そこで、もう1つの方法として「感度分析」による評価を行なった。この評価の結果では、(3)式の解は、必ずしも“安定した”ものではない、と評価された。具体的には、(3)式の解（判別式）の14個のパラメーターのそれぞれ“独立に”（この変化が大きいか小さいかは議論のあるところであるが）上下に±1%変化させ、そのときに、2つの集団が“完全に判別できた”(3)式の解の「“重なり”に入った対象の数」（「アタッシュケース」のうちで「ショルダーバッグ」の判別得点の最大値を下回る対象の数と、「ショルダーバッグ」のうちで「アタッシュケース」の判別得点の最小値を上回る対象の数の合計）がどう変化するかを観察した。その結果として、「“重なり”に入った対象の数」は、40件中、最小は0件であり、平均は約10件、最大では（第12パラメーターを+1%増加させたときの）25件に増加してしまう。

したがって、この立場からは、(3)式の解は、そのパラメーターの±1%の変化について必ずしも“安定した”ものではない、と考えられる。ちなみに、判別値の選択のしかたにもよるが、通常、「誤判別数の合計」の値は「“重なり”に入った対象の数」の半分程度である。

繰り返しになるが、この場合、解のパラメーターを変化させても結果が変わらなかった対象については、(3)式の解にしたがい、結果が変わってしまった対象については、これらの対象に対して求めた“第2”の判別分析の解を適用する「“2段階”判別方式」の方法を適用

するなどすれば、より“安定した”解を導き出すことができる可能性があるであろう。あるいは、判別式の形式に工夫を加えることもよいであろう。また、この問題の場合、14個の説明変数のデータは必ずしも互いに独立したものではなく、また、その値の範囲も似通ったものであるので、「“1段階”判別方式」の方法で得られた(3)式の解は、「感度分析」という評価方法に対しては、この程度の“安定性”しか持っていないのだ、といえるかもしれない。

次に、第2のケースとして、判別すべきビジネスバッグのカラー画像を図2-13-3に示す。これらのカラー画像も、(1-4-2項で説明した)「自由申告法」による「印象の計測」で使用したもので、200件のカラー画像を1件当たり8名の回答者(延べ1600名)に回答してもらった結果の中から、当該の印象が申告されたものを選択した。最初の20件(#1~#20)は「沢山入りそうな」という印象が申告されたものであり、全般的な印象として^{まち}襷が厚いものが多く含まれている。残りの20件(#21~#40)は、「(「沢山入りそうな」とは“反対”の意味の)「沢山は入らなそうな」という印象が申告されたもので、これには襷が薄いものが多く含まれている。

これらの40件の画像についても、546個の「局所自己相関特徴量」を計算し、この中から、表2-13-2に示す“変化量の大きい”14個の特徴量($GR(1,0), RG(2,0), GG(0,4), RR(2,1), GB(4,-1), GB(1,-3), RB(2,-2), RB(5,-2), GG(3,-5), GB(3,-5), GR(5,-3), GR(4,4), BB(4,-4), GR(4,5)$)を選択した。

「計算的な方法」の遺伝的アルゴリズムの“戦略”は第1のケースの場合と同様である。計算は“もっともよい”解に改善が見られなくなった約600世代で打ち切った。結果の判別式は(4)式の通りで、表2-13-2に示すように、「沢山入りそうな」の対象の判別得点の最小値の $[-0.000042]$ は、「沢山は入らなそうな」の対象の判別得点の最大値の $[-0.00047]$ よりも大きい。したがって、「沢山入りそうな」と「沢山は入らなそうな」の印象は、(4)式の判別式で“完全に”判別することができた。ちなみに、この解の“安定性”の評価の結果も、第1のケースとほぼ同様であった。

$$(4) \quad f = 0.11725 x_1 - 0.07139 x_2 - 0.07939 x_3 - 0.00975 x_4 - 0.00188 x_5 \\ + 0.05676 x_6 - 0.05299 x_7 + 0.03966 x_8 + 0.04336 x_9 - 0.01446 x_{10} \\ - 0.06674 x_{11} - 0.02037 x_{12} - 0.01690 x_{13} + 0.07675 x_{14}$$



図 2-13-3 「沢山入りそうな」と「沢山は入らなそうな」ビジネスバッグのカラー画像

表 2-13-2 「沢山入りそうな」と「入らなそうな」ビジネスバッグのカラー画像の特徴量と判別の結果

沢山入りそうな

#	GR	RG	GG	RR	GB	GB	RB	RB	GG	GB	GR	GR	BB	GR	
	(1,0)	(2,0)	(0,4)	(2,1)	(4,-1)	(1,-3)	(2,-2)	(5,-2)	(3,-5)	(3,-5)	(5,-3)	(4,4)	(4,-4)	(4,5)	得点
01	.538786	.537582	.529037	.534771	.534198	.530737	.532462	.529109	.524502	.523709	.527249	.525221	.525408	.522316	-.000039
02	.476915	.474847	.470457	.470936	.464769	.464614	.463275	.457888	.464077	.456292	.462459	.461201	.451167	.458472	.000015
03	.569870	.568618	.558860	.566506	.562591	.559057	.561529	.558676	.553958	.551681	.558237	.555973	.551817	.553336	.000012
04	.605525	.605199	.590663	.611894	.588589	.585948	.591954	.589582	.587561	.580246	.597203	.595477	.576839	.593221	-.000019
05	.471085	.470172	.464264	.465507	.465887	.462949	.461075	.458334	.458994	.455951	.459213	.457335	.455712	.454815	-.000030
06	.575740	.575079	.565346	.571997	.568080	.565442	.567274	.563959	.560688	.557855	.563899	.562007	.557327	.559131	-.000019
07	.528074	.528053	.502785	.555781	.490429	.487924	.498002	.496237	.499625	.482599	.521130	.519679	.474467	.517795	-.000010
08	.505787	.503850	.492399	.503258	.494257	.492251	.495998	.489927	.486087	.483265	.490320	.488972	.482956	.485826	.000023
09	.614067	.614145	.582139	.645172	.573158	.568947	.587536	.585885	.578663	.562907	.607745	.605484	.553412	.603291	-.000026
10	.525037	.524116	.513674	.522031	.518333	.515046	.517479	.514055	.508376	.506977	.512598	.510508	.507990	.507382	-.000042
11	.588818	.589216	.563661	.614287	.552584	.549085	.560726	.559848	.561940	.544501	.584244	.582380	.535812	.580653	.000001
12	.540215	.538772	.528325	.536317	.533407	.531146	.533443	.528319	.523342	.522826	.525794	.523818	.524299	.520694	.000008
13	.498017	.497738	.463656	.548095	.447635	.445531	.456340	.454309	.460380	.439395	.492200	.490587	.431994	.488829	-.000029
14	.486151	.484769	.476499	.481809	.480263	.477172	.478619	.475302	.470844	.469457	.473502	.471754	.470181	.468768	-.000015
15	.532855	.531773	.518107	.532055	.523521	.520026	.525629	.519749	.513274	.511808	.518093	.516326	.512913	.513174	-.000036
16	.490866	.489613	.474367	.493794	.474282	.471043	.475918	.472029	.468709	.463056	.475976	.473964	.461337	.471202	.000047
17	.522267	.521223	.512235	.519053	.516138	.513077	.515129	.511483	.507438	.505522	.510749	.508964	.505944	.506169	-.000025
18	.527828	.525703	.515813	.527489	.513702	.512241	.514343	.510327	.509278	.503976	.513953	.512410	.501982	.509478	.000033
19	.523174	.521524	.511819	.518810	.513090	.510423	.512376	.507760	.505495	.501633	.508372	.506436	.499656	.503439	.000027
20	.514139	.512606	.505272	.509990	.507884	.505768	.506752	.503084	.498975	.497914	.501611	.499908	.498704	.497162	-.000042
最大	.614067	.614145	.590663	.645172	.588589	.585948	.591954	.589582	.587561	.580246	.607745	.605484	.576839	.603291	.000047
最小	.471085	.470172	.463656	.465507	.447635	.445531	.456340	.454309	.458994	.439395	.459213	.457335	.431994	.454815	-.000042

沢山は入らなそうな

#	GR	RG	GG	RR	GB	GB	RB	RB	GG	GB	GR	GR	BB	GR	
	(1,0)	(2,0)	(0,4)	(2,1)	(4,-1)	(1,-3)	(2,-2)	(5,-2)	(3,-5)	(3,-5)	(5,-3)	(4,4)	(4,-4)	(4,5)	得点
01	.562435	.561775	.540566	.603090	.540009	.536927	.543938	.542710	.536030	.530444	.555732	.554021	.529208	.552257	-.000155
02	.539081	.538247	.522386	.547586	.523439	.520058	.526386	.524045	.517438	.513180	.528653	.526139	.512152	.522976	-.000106
03	.516058	.514357	.506089	.512411	.508211	.506652	.508586	.503267	.499755	.498223	.502336	.500769	.498661	.497730	-.000063
04	.521291	.519779	.511334	.517066	.513645	.510703	.512240	.508175	.505232	.502174	.507689	.505524	.502322	.502234	-.000067
05	.595072	.594814	.582776	.595359	.586257	.582756	.587033	.583208	.578407	.575546	.584883	.583061	.575174	.580294	-.000102
06	.498256	.496895	.485556	.493524	.491235	.487710	.489728	.485033	.477954	.477853	.481194	.479020	.480042	.475362	-.000055
07	.481327	.480126	.470524	.479991	.471638	.469684	.473142	.469795	.463977	.460957	.469837	.467939	.461030	.464824	-.000095
08	.509821	.509364	.482921	.552862	.470980	.468996	.477863	.477045	.478920	.463353	.504455	.502949	.456903	.501448	-.000066
09	.549044	.548027	.530699	.571622	.527450	.524870	.530711	.528266	.526744	.519441	.543097	.541708	.517818	.540027	-.000082
10	.537861	.537379	.524071	.545211	.526257	.523277	.528586	.525795	.518389	.516154	.529276	.528178	.516350	.526246	-.000072
11	.556953	.556331	.547068	.553904	.552207	.548920	.550924	.547988	.542635	.541910	.546025	.544163	.543128	.541402	-.000048
12	.536483	.535983	.514549	.567818	.512839	.510968	.517763	.516086	.510151	.503763	.529263	.527486	.502314	.525391	-.000075
13	.502419	.502145	.474851	.545399	.467890	.465148	.473314	.471481	.472038	.459792	.496464	.494890	.455862	.493113	-.000069
14	.506934	.505874	.496760	.503172	.501991	.498191	.499645	.496531	.491236	.489700	.494501	.492156	.491086	.488750	-.000067
15	.572276	.572442	.541136	.614406	.532510	.528721	.543095	.542076	.538442	.523573	.566646	.564817	.516347	.563054	-.000048
16	.494045	.493600	.485165	.488804	.491185	.486885	.487553	.484055	.480821	.479548	.482218	.479887	.480589	.476950	-.000080
17	.590952	.590455	.581469	.589458	.586042	.582469	.585015	.582886	.578722	.577038	.582863	.580764	.577910	.578363	-.000055
18	.531132	.530796	.517229	.536587	.514239	.510786	.515177	.512960	.513987	.504619	.522388	.520280	.501006	.517719	-.000047
19	.516772	.516031	.504694	.519164	.506102	.504019	.507735	.502884	.500173	.497033	.506809	.505545	.496244	.503054	-.000080
20	.580294	.579787	.557441	.610600	.556039	.553140	.562437	.560129	.554105	.546896	.573680	.572040	.544760	.569967	-.000082
最大	.595072	.594814	.582776	.614406	.586257	.582756	.587033	.583208	.578722	.577038	.584883	.583061	.577910	.580294	-.000047
最小	.481327	.480126	.470524	.479991	.467890	.465148	.473142	.469795	.463977	.459792	.469837	.467939	.455862	.464824	-.000155

次に、第3のケースとして、判別すべきビジネスバッグのカラー画像を図2-13-4に示す。これらのカラー画像も、「自由申告法」による「印象の計測」で使用したもので、200件のカラー画像を1件当たり8名の回答者（延べ1600名）に回答してもらった結果の中から、当該の印象が申告されたものを選択した。最初の20件（#1～#20）は「高そうな」という印象が申告されたものであり、残りの20件（#21～#40）は（「高そうな」とは“反対”の意味の）「安そうな」という印象が申告されたものである。これらの40件の画像についても、546個の「局所自己相関特徴量」を計算し、この中から表2-13-3に示す“変化量の大きい”14個の特徴量（ $BG(2, 0), GG(5, 0), BG(5, 0), GG(0, 5), GB(2, 1), GB(1, 2), GG(1, 4), RR(2, -3), GR(2, -5), RB(3, -2), BG(3, -3), RR(3, -4), GR(4, -3), BB(4, -5)$ ）を選択した。

「計算的な方法」の遺伝的アルゴリズムの“戦略”は、第1と第2のケースと同様である。計算は“もっともよい”解の改善に見られなくなった約900世代で打ち切った。結果の判別式は(5)式の通りで、表2-13-3に示すように、「高そうな」の対象の判別得点の最小値の $[-0.000609]$ は、「安そうな」の対象の判別得点の最大値の $[-0.000468]$ よりも大きい。したがって、「高そうな」と「安そうな」の印象も、(5)式の判別式で“完全に”判別することができた。ちなみに、この解の“安定性”の評価の結果も、第1、第2のケースとほぼ同様であった。

$$(5) \quad f = -0.02463 x_1 - 0.14907 x_2 - 0.12629 x_3 + 0.27058 x_4 + 0.44046 x_5 \\ - 0.27994 x_6 - 0.21555 x_7 - 0.02573 x_8 - 0.17937 x_9 - 0.03026 x_{10} \\ + 0.12750 x_{11} + 0.01887 x_{12} + 0.23024 x_{13} - 0.05580 x_{14}$$



図 2-13-4 「高そうな」と「安そうな」ビジネスバッグのカラー画像

表 2-13-3 「高そうな」と「安そうな」ビジネスバッグのカラー画像の特徴量と判別の結果

高そうな

#	BG	GG	BG	GG	GB	GB	GG	RR	GR	RB	BG	RR	GR	BB	
	(2, 0)	(5, 0)	(5, 0)	(0, 5)	(2, 1)	(1, 2)	(1, 4)	(2, -3)	(2, -5)	(3, -2)	(3, -3)	(3, -4)	(4, -3)	(4, -5)	得点
01	.518631	.519884	.514632	.512849	.516480	.515159	.514682	.520797	.511566	.512955	.509620	.515807	.514825	.499131	.000547
02	.499146	.496392	.495549	.487480	.496777	.494978	.489823	.491532	.485678	.492995	.490344	.487472	.489565	.482779	.000508
03	.559227	.563682	.556731	.554941	.557083	.555406	.557299	.607396	.570338	.561959	.551951	.605645	.574040	.542037	.000600
04	.470566	.469725	.466513	.461549	.467961	.466410	.463309	.464089	.457655	.463761	.460528	.459483	.461717	.450482	.000536
05	.512080	.509888	.508865	.502473	.509586	.508386	.503962	.504657	.499449	.505705	.503279	.500919	.503141	.495925	.000496
06	.551639	.549326	.549338	.541217	.549394	.547775	.542919	.544020	.539171	.546175	.543911	.540960	.542963	.538326	.000501
07	.496832	.494504	.492701	.486034	.494091	.492533	.487929	.488847	.482477	.490167	.486927	.484465	.486892	.478254	.000598
08	.579161	.584928	.577532	.577273	.577287	.575896	.578738	.598220	.583418	.580274	.572381	.596257	.586978	.562101	.000522
09	.505091	.502303	.500832	.494346	.502566	.501287	.496657	.501381	.493134	.498667	.495366	.497354	.497023	.487575	.000495
10	.528108	.531931	.524849	.524239	.526029	.524434	.525936	.533593	.525007	.524299	.519366	.530182	.528500	.507994	.000576
11	.530265	.536499	.527655	.528626	.527944	.526526	.530386	.568480	.539114	.530352	.522049	.566323	.543239	.511200	.000602
12	.490137	.498047	.487204	.491099	.487659	.486409	.492515	.545659	.507219	.491740	.481949	.544078	.510517	.470463	.000609
13	.518046	.517905	.513772	.508583	.515359	.513388	.510716	.512145	.505980	.511076	.507465	.507551	.510013	.496510	.000556
14	.536268	.539252	.533040	.532572	.533987	.532919	.534069	.590311	.545748	.538884	.528949	.588673	.548716	.520227	.000558
15	.536964	.531170	.533449	.522441	.534509	.532920	.524500	.558813	.532188	.539250	.527439	.555947	.535969	.523675	.000539
16	.480083	.509886	.477702	.503178	.477605	.476497	.504237	.630546	.546129	.493106	.472674	.629798	.549614	.453665	.000549
17	.552751	.553805	.550489	.546786	.550987	.549632	.548484	.585329	.554935	.553567	.545241	.583024	.558332	.537970	.000558
18	.488451	.498386	.486459	.491407	.486909	.485437	.492960	.558272	.512248	.495340	.482351	.557271	.515103	.471549	.000583
19	.520835	.519297	.517644	.507185	.518267	.515870	.510266	.521847	.509870	.516874	.510865	.517655	.515435	.501261	.000580
20	.505643	.502207	.500877	.491725	.502609	.500516	.494735	.497626	.489761	.498409	.494795	.492446	.494655	.485045	.000570
最大	.579161	.584928	.577532	.577273	.577287	.575896	.578738	.630546	.583418	.580274	.572381	.629798	.586978	.562101	.000609
最小	.470566	.469725	.466513	.461549	.467961	.466410	.463309	.464089	.457655	.463761	.460528	.459483	.461717	.450482	.000495

安そうな

#	BG	GG	BG	GG	GB	GB	GG	RR	GR	RB	BG	RR	GR	BB	
	(2, 0)	(5, 0)	(5, 0)	(0, 5)	(2, 1)	(1, 2)	(1, 4)	(2, -3)	(2, -5)	(3, -2)	(3, -3)	(3, -4)	(4, -3)	(4, -5)	得点
01	.493756	.504224	.491557	.498063	.491751	.490583	.499084	.520897	.504068	.494247	.486971	.518343	.506914	.473816	.000416
02	.526755	.529687	.524949	.519442	.524314	.522427	.521411	.540547	.523851	.525986	.518488	.536330	.528938	.509219	.000430
03	.555991	.565122	.553891	.558700	.553799	.552736	.559796	.601139	.571638	.558423	.549018	.599539	.574572	.538698	.000458
04	.593212	.590653	.590302	.584628	.591259	.590579	.586382	.587357	.582915	.588550	.586343	.584576	.585923	.581387	.000456
05	.623484	.621474	.621417	.612309	.621281	.619531	.614266	.606938	.605863	.612929	.615648	.603749	.609801	.610107	.000420
06	.573063	.571948	.569102	.562184	.570533	.568688	.565190	.566039	.560591	.566064	.563643	.562186	.564578	.554380	.000421
07	.566388	.576908	.565277	.569695	.564317	.563041	.570771	.581792	.573382	.566128	.560093	.580168	.576603	.545994	.000423
08	.524432	.522218	.521194	.513461	.521705	.520053	.515524	.511796	.508797	.515626	.515536	.508237	.512877	.508373	.000437
09	.543618	.548245	.541586	.538408	.540453	.538914	.540123	.560231	.544314	.543540	.534727	.557251	.549278	.523256	.000395
10	.589933	.590123	.587239	.579748	.587210	.585207	.582534	.590355	.580750	.586382	.581086	.587005	.585556	.571967	.000426
11	.496997	.492799	.492654	.481994	.493874	.491470	.484079	.485703	.478714	.488319	.484920	.480598	.482963	.476505	.000468
12	.476247	.475606	.472981	.467418	.474048	.472635	.469541	.474307	.466362	.471926	.467232	.470217	.470364	.457722	.000458
13	.473848	.479522	.471420	.469789	.471261	.469238	.472003	.486572	.475068	.471191	.465519	.483957	.478869	.454108	.000310
14	.529799	.532145	.527352	.522596	.527546	.525508	.524883	.528942	.522224	.525024	.521487	.525711	.526420	.512207	.000406
15	.483959	.477019	.480052	.468524	.481846	.480264	.470971	.472573	.467090	.477950	.475420	.468640	.470554	.471909	.000452
16	.590810	.588453	.587604	.577613	.588312	.586282	.580404	.577718	.573339	.581692	.581233	.573309	.577746	.573433	.000350
17	.509056	.517134	.506960	.507828	.506689	.504859	.509941	.538797	.517464	.509065	.500924	.536336	.521751	.489366	.000410
18	.572387	.570002	.569204	.562329	.570322	.568795	.564470	.564484	.560080	.566139	.564169	.560874	.563007	.557602	.000417
19	.494964	.493297	.492035	.482210	.492257	.489901	.484729	.482768	.478815	.486814	.485403	.478979	.483575	.477586	.000382
20	.523273	.522318	.520579	.513032	.520867	.518905	.515159	.516226	.511241	.517032	.514540	.512794	.515048	.506858	.000392
最大	.623484	.621474	.621417	.612309	.621281	.619531	.614266	.606938	.605863	.612929	.615648	.603749	.609801	.610107	.000468
最小	.473848	.475606	.471420	.467418	.471261	.469238	.469541	.472573	.466362	.471191	.465519	.468640	.470364	.454108	.000310

以上の3つのケースの試みは、いずれも判別すべき対象は、それぞれ20件ずつ合計で40件という“小さな”規模での判別分析であった。また、結果としても、（解の「安定性」の意味をどう考えるかにもよるが、）「“1段階”判別方式」の方法の適用だけでは、あるいは、判別式の形式として1次式の選択だけでは、必ずしも“安定した”解を得ることができなかった、といえる。しかし、この試みで得られた解は、判別すべき対象の一部だけであるが、これらを“正しく”そして“安定して”判別することができ、たとえば（2-9節で説明した）「2段階判別方式」の方法を適用するなどによって、より“正しい”かつ“安定した”解とすることができる可能性を期待することができよう。

また、判別分析の対象とするビジネスバッグの「カラー画像」を増やし、また、これらに対して人が感じる「（特定の）印象の有無」などの調査データを増やすことによって、つまり、これによって、分析をする対象の集団のなかに、「特定の“同じ”印象」を感じる、できるだけ多くの“質の違う”、つまり、幅広いスペクトラムの対象を取り入れて、集団に含まれる対象の“質の違い”をできるだけ大きくすることによって、得られる判別分析モデルによる「印象の推定モデル」が、より多くの対象に対してより多くの（平均的な）人が考えるビジネスバッグのタイプの認識あるいは感じる（特定の）印象の有無を、幅広くカバーできるようになることが期待できるであろう。

さらに、カラー画像の特徴量（ここでは「局所自己相関特徴量」）の選択とそれらの組み合わせの検討によっても、より判別力の高い「印象の推定モデル」の構築が期待できるであろう。この節の初めでも述べたが、ここでの試みでは、546個の局所自己相関特徴量のなかから（カラー画像の違いによる“変化量の大きい”）14個の特徴量を選択し、これらを説明変数として判別分析を行ったが、これとは別に、より少ない数のまた互いに独立な特徴量の組み合わせを説明変数とする判別分析も可能であろう。この場合、ここで示した解よりも“洗練された”解が得られる可能性もあろう。このためには、特徴量の性質の分析と選択についての詳細な検討が必要であろう。いずれにしても、こうした方策を通じて、「計算的な方法」による判別分析を利用した「印象の推定モデル」の推定能力を向上させれば、“実務”の問題からの期待に対して十分に答えられる「印象の推定モデル」の構築の可能性を期待することができよう。

3章 「遺伝的アルゴリズムを利用した計算的な『判別分析』の方法」の拡張

2章で説明した、遺伝的アルゴリズムを利用した「計算的な判別分析」の方法とこれを利用した「印象の推定モデル」は、対象のカラー画像の“複数”の特徴量（連続量）とその対象に対する特定の印象の有無（離散量）を対応づけるものであった。この章では、まず、このモデルの入力つまりカラー画像の特徴量（連続量）のところに、物理的な仕様（アイテム／カテゴリー）の有無（離散量）を含むことができるように拡張する。具体的には、離散量同士の対応づけで、非線型性と交互作用の両方を含んだものの代表として「論理関係」を取り上げ、これに「計算的な判別分析」を適用する。これは「数量化理論2類」である。次に、このモデルの出力つまり特定の印象の有無（離散量）のところに、特定の印象の大きさ（連続量）を含むことができるように拡張する。これは「回帰分析」である。

2章で取り上げた「判別分析」と同様に、「数量化理論2類」と「回帰分析」にも、すでに「解析的な方法」が開発されており広く普及している。しかし、「判別分析」の方法と同様に、これらの「解析的な方法」も、計算技術的な意味合いから、必ずしも私たちが実際に分析したい問題の目的や性格に対する“柔軟性”は備えていない。

そこで、この章では、まず、代表的な離散量でありその対応関係に非線型性と交互作用を含む「論理関係」に、「計算的な判別分析」を適用して、この方法がこの問題にそのまま適用できることを例によって説明する。次に、「回帰分析」については、（「判別分析」の場合と同様に、）まず「分析の目的」を中心に「解析的な方法」の問題点を検討し、その解決のために、「計算的な判別分析」と同様の、遺伝的アルゴリズムを利用した「計算的な方法」が有効であることを例によって説明する。

ちなみに、これらの拡張によって、「印象の推定モデル」の構築の方法論としての“外的な基準のある”多変量解析の「計算的な方法」は、対応づけるべき説明変数、被説明変数ともに、連続量あるいは離散量のいずれであってもよく、基本的に同じ「方法」つまり定式化と遺伝的アルゴリズムを利用した計算アルゴリズムによって、実際に適用する問題からの要求に“柔軟に”対応することができるようになった。

3-1節 「論理関係の判別分析」への拡張

3-1-1項 「印象の推定モデル」としての「論理関係の判別分析」

2章で説明した「計算的な判別分析」は、対象の“連続的な”属性値（あるいは代表値）を判別式によって1つの値つまり「判別得点」に変換し、その値が一定の値つまり「判別値」よりも大きいかいなかによって、その対象が「ある特定の特性を持った集団」に属するかいなかを判別あるいは推定する方法であった。「印象の推定」という立場からいえば、対象の画像を代表する複数の“連続的な”特徴量に対して、その対象に対する特定の印象の有無（離散量）を対応づけ、これによって、任意の画像を代表する特徴量に対して、特定の印象の有無を判別あるいは推定するための方法であった。（ちなみに、1-4-2項で説明した「印象の工学」での数学的な「印象の推定モデル」の構築の方法の1つである「同時関係分析法」は、対象の物理的な仕様の有無や分類などの“離散的な”属性値あるいはその組み合わせと“離散的な”特定の印象の有無を対応づける方法であった。）

ここで、「計算的な判別分析」を利用して、対象の“離散的な”属性値、つまり対象の特定の仕様や特徴の有無やカテゴリーの分類など（たとえば、部品や部材の有無、形状や色彩の分類コードなど）を判別式によって1つの判別得点に変換し、その値が判別値よりも大きいかいなかによって、その対象がある特性を持っているかいかを判別する方法が考えられる。これは「数量化理論2類」である。もしこういったことができれば、対象の特定の仕様の有無や特定のカテゴリーの分類などの組み合わせに対して、特定の印象の有無を“計算的に”対応づけることができる。（ちなみに、1-4-2項で説明した「同時関係分析法」は、これらを“解析的に”分析し対応づける方法であった。）

この節では、離散的な値をとり、また、非線型性と交互作用の両方を含む代表的な変量であり対応関係でもある「論理変数と論理関係」に対して、「計算的な判別分析」の適用を試みる（文献92）。この適用がうまくいけば、「計算的な判別分析」は、対象の“連続的な”属性値と対象の特定の特性の有無を“形式的に”対応づけるだけでなく、対象の“離散的な”属性値あるいはその組み合わせと対象の特定の特性の有無を“論理的に”対応づける方法としての新しい可能性が期待できる。

（2-1節で示した）図2-1-1で、対象のカラー画像の特徴量を、対象の特定の仕様や特徴

の有無やカテゴリーの分類などに置き換えたものが、「“離散的な”属性値に対する『判別分析』を利用した『印象の推定モデル』」の概念である。この図では、対象の特定の仕様や特徴の有無やカテゴリーの分類などを入力すると、これに対応して特定の印象の有無が出力される。

さて、“論理的な”対応関係は、基本的にその判別分析で誤判別を生じることは許されないので、「計算的な判別分析」の目的関数は、「誤判別数の合計を最小にする」ではなく（というよりも、これは当然のこととして）、「“真” (*true*) と“偽” (*false*) の集団の間の判別得点の最短距離を最大にする」を採用する。具体的には、(1)式に示す「“真”の集団の判別得点の最小値と“偽”の集団の判別得点の最大値の差を最大にする」を採用する。これは2-5節の(1)式と同じである。あるいは、(2-5節の方法にしたがって、)「誤判別数の合計を最小にする」を第1の目的関数に、さらに、「“真”と“偽”の集団の間の最短距離を最大にする」を第2の目的関数に採用して、誤判別を最小にするだけでなく、さらにより明らかな判別ができるような目的関数を採用する。

また、論理関係は、基本的に非線型であるので、判別式の形式としては、(2)式で表される「1次式」のほかに、非線型性と交互作用に対応しやすい形式、つまり、(3)式の「2次式」、(4)式の「“V型”の絶対値式」、(5)式の「“W型”の絶対値式」（“V型”の折り返しが2回あるので“W型”となる）などを適用する。また、より複雑な論理関係の場合には、(2-9節で説明した)「2段階判別方式」や(2-10節で説明した)「二重判別方式」の方法を適用する。

$$(1) \min_{\forall k \in true} f_k - \max_{\forall k \in false} f_k \rightarrow \text{最大化}$$

$$(2) f_k = \sum_{\forall i} a_i x_{ik}$$

$$(3) f_k = \sum_{\forall i} a_i x_{ik} + \sum_{\forall i} \sum_{\forall j} b_{ij} x_{ik} x_{jk}$$

$$(4) f_k = \left| \sum_{\forall i} a_i x_{ik} + b \right| \quad \text{あるいは} \quad - \left| \sum_{\forall i} a_i x_{ik} + b \right|$$

$$(5) f_k = \left| \left| \sum_{\forall i} a_i x_{ik} + b \right| + c \right| \quad \text{あるいは} \quad - \left| \left| \sum_{\forall i} a_i x_{ik} + b \right| + c \right|$$

3-1-2 項 「基本的な論理関係」の判別分析

表 3-1-1(1)の $[a]$ $[b]$ $[z]$ は、2つの論理変数 $[a]$ と $[b]$ の「論理和」 $[z = a \cup b]$ の真理値表である。この論理関係の判別分析のための判別式の形式には、(6)式で表される1次式を適用する。

$$(6) \quad f = w_1 a + w_2 b \quad (-1 \leq w_1, w_2 \leq +1)$$

「計算的な判別分析」の遺伝的アルゴリズムの“戦略”として、1つの世代で用意する染色体を20個として、遺伝的操作として「優先生き残り」「2個のトーナメントをした後のランダム交叉」「突然変異」「ランダム生き残り」を行い、それぞれ2個、10個、6個、2個を生成する。ランダム交叉の交叉率は100%（つまり、2個のパラメーターの両方を交叉する）、突然変異の変異率は50%（つまり、2個のパラメーターの一方だけを変位させる）である。次の世代の染色体を生成したあと、それぞれ5%の感度分析によって解の改善をした。計算は約10～15世代で収束した。結果は(7)式と表 3-1-1(1)の通りである。つまり、判別得点の値 $[f]$ が、判別値 $[f_0 = 0.5]$ よりも大きい場合には“真”、小さい場合には“偽”と判別できる。

$$(7) \quad \text{if } f = a + b \geq 0.5 (= f_0) \text{ then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

表 3-1-1 「論理和」の真理値表と判別分析の結果

(1) 2変数の場合				(2) 3変数の場合				
a	b	f	z	a	b	c	f	z
0	0	0	0	0	0	0	0	0
0	1	1	1	0	0	1	1	1
1	0	1	1	0	1	0	1	1
1	1	2	1	0	1	1	2	1
				1	0	0	1	1
				1	0	1	2	1
				1	1	0	2	1
				1	1	1	3	1

3つの論理変数 $[a]$ $[b]$ $[c]$ の論理和 $[z = a \cup b \cup c]$ の場合の判別分析の結果は、(8)式と表 3-1-1(2)である。また、これらの結果から、一般的に、K個の論理変数 $[a_k]$ ($k = 1, 2, \dots, K$) の論理和 $[z = a_1 \cup a_2 \cup \dots \cup a_K]$ は、(9)式で判別できることが分かる。

$$(8) \quad \text{if } f = a + b + c \geq 0.5 (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

$$(9) \quad \text{if } f = \sum_{\forall k} a_k \geq 0.5 (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

表 3-1-2(1)の $[a] [b] [z]$ は、2つの論理変数 $[a]$ と $[b]$ の「論理積」 $[z = a \cap b]$ の真理値表である。この論理関係の判別分析にも、(6)式で表される1次式を適用する。「論理和」の場合と同様に、「計算的な判別分析」の遺伝的アルゴリズムの“戦略”として、1つの世代で用意する染色体を20個として、遺伝的操作として「優先生き残り」「2個のトーナメントをした後のランダム交叉」「突然変異」「ランダム生き残り」を行い、それぞれ2個、10個、6個、2個を生成した。ランダム交叉の交叉率は100%、突然変異の変異率は25%である。次の世代の染色体を生成したあと、それぞれ5%の感度分析によって解の改善をした。計算は約15～20世代で収束した。結果は(10)式と表 3-1-2(1)の通りである。

$$(10) \quad \text{if } f = a + b \geq 1.5 (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

表 3-1-2 「論理積」の真理値表と判別分析の結果

(1) 2変数の場合

a	b	f	z
0	0	0	0
0	1	1	0
1	0	1	0
1	1	2	1

(2) 3変数の場合

a	b	c	f	z
0	0	0	0	0
0	0	1	1	0
0	1	0	1	0
0	1	1	2	0
1	0	0	1	0
1	0	1	2	0
1	1	0	2	0
1	1	1	3	1

3つの論理変数 $[a] [b] [c]$ の論理積 $[z = a \cap b \cap c]$ の場合の判別分析の結果は、(11)式と表 3-1-2(2)である。また、これらの結果から、一般的に、K個の論理変数 $[a_k]$ ($k=1, 2, \dots, K$) の論理和 $[z = a_1 \cap a_2 \cap \dots \cap a_K]$ は、(12)式で判別できることが分かる。

$$(11) \quad \text{if } f = a + b + c \geq 2.5 (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

$$(12) \quad \text{if } f = \sum_{\forall k} a_k \geq K - 0.5 (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

また、論理変数 $[a]$ の「否定」 $[z = \neg a]$ は、(13)式で判別できる。

$$(13) \quad \text{if } f = -a \geq -0.5 (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

表 3-1-3 「否定」の真理値表と判別分析の結果

a	f	z
0	-0	1
0	-1	1

さらに、2つの論理変数 $[a] [b]$ の否定論理和「NOR」 $[z = \neg (a \cup b)]$ の判別分析の結果は、(14)式と表 3-1-4(1)である。3つの論理変数 $[a] [b] [c]$ の否定論理和 $[z = \neg (a \cup b \cup c)]$ の判別分析の結果は、(15)式と表 3-1-4(2)である。これらの結果から、一般的に、K個の論理変数 $[a_k] (k=1, 2, \dots, K)$ の否定論理和 $[z = \neg (a_1 \cup a_2 \cup \dots \cup a_K)]$ は、(16)式で判別できることが分かる。

$$(14) \quad \text{if } f = -a - b \geq -0.5 (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

$$(15) \quad \text{if } f = -a - b - c \geq -0.5 (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

$$(16) \quad \text{if } f = -\sum_{\forall k} a_k \geq -0.5 (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

表 3-1-4 「否定論理和」の真理値表と判別分析の結果

(1) 2変数の場合

a	b	f	z
0	0	-0	1
0	1	-1	0
1	0	-1	0
1	1	-2	0

(2) 3変数の場合

a	b	c	f	z
0	0	0	-0	1
0	0	1	-1	0
0	1	0	-1	0
0	1	1	-2	0
1	0	0	-1	0
1	0	1	-2	0
1	1	0	-2	0
1	1	1	-3	0

また、2つの論理変数 $[a] [b]$ の否定論理積「NAND」 $[z = \neg (a \cap b)]$ の判別分析の結果は、(17)式と表 3-1-5(1)である。3つの論理変数 $[a] [b] [c]$ の否定論理積 $[z = \neg (a \cap b \cap c)]$ の判別分析の結果は、(18)式と表 3-1-5(2)である。これらの結果から、一般的に、K個の論理変数 $[a_k] (k=1, 2, \dots, K)$ の否定論理積 $[z = \neg (a_1 \cap a_2 \cap \dots \cap a_K)]$ は、(19)式で判別できることが分かる。

$$(17) \quad \text{if } f = -a - b \geq -1.5 (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

$$(18) \quad \text{if } f = -a - b - c \geq -2.5 (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

$$(19) \quad \text{if } f = -\sum_{\forall k} a_k \geq 0.5 - K (= f_0) \quad \text{then } \hat{z} = 1 \quad \text{else } \hat{z} = 0$$

表 3-1-5 「否定論理積」の真理値表と判別分析の結果

(1) 2 変数の場合

a	b	f	z
0	0	-0	1
0	1	-1	1
1	0	-1	1
1	1	-2	0

(2) 3 変数の場合

a	b	c	f	z
0	0	0	-0	1
0	0	1	-1	1
0	1	0	-1	1
0	1	1	-2	1
1	0	0	-1	1
1	0	1	-2	1
1	1	0	-2	1
1	1	1	-3	0

3-1-3 項 「排他的論理和」の判別分析

表 3-1-6(1)は、2つの論理変数 $[a]$ $[b]$ の「排他的論理和」 $[z = a \text{ xor } b]$ の真理値表である。排他的論理和は、非線形性と交互作用を含んだ論理で、「1次式」の形式の判別式では判別できないので、非線形性と交互作用の両方を含むデータに有効な(20)式の「2次式」を採用した判別分析を試みる。

$$(20) \quad f = w_1 a + w_2 b + w_3 ab \quad (-1 \leq w_1, w_2, w_3 \leq +1)$$

1つの世代で用意する染色体を20個として、遺伝的操作として「優先生き残り」「2個のトーナメントをした後のランダム交叉」「突然変異」「ランダム生き残り」を行い、それぞれ2個、10個、6個、2個を生成した。ランダム交叉の交叉率は100%、突然変異の変異率は25%である。次の世代の染色体を生成したあと、それぞれ5%の感度分析によって解の改善をした。計算は50～100世代で収束した。計算の結果は(21)式と表 3-1-6(1)の通りである。つまり、判別得点の値 $[f]$ が、判別値 $[f_0=0.25]$ よりも大きい場合には“真”、小さい場合には“偽”と推定できる。

$$(21) \quad \text{if } f = 0.5a + 0.5b - ab \geq 0.25 (= f_0) \text{ then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

表 3-1-6 2変数の「排他的論理和」の真理値表と判別分析の結果

(1) 2次式((21)式)				(2) 絶対値式((23)式)				(3) 絶対値式((24)式)			
a	b	f	z	a	b	f	z	a	b	f	z
0	0	0.0	0	0	0	-1	0	0	0	0	0
0	1	0.5	1	0	1	-0	1	0	1	1	1
1	0	0.5	1	1	0	-0	1	1	0	1	1
1	1	0.0	0	1	1	-1	0	1	1	0	0

あるいは、判別式の形式として、非線形性を含むデータに有効な(22)式の「“V型”の絶対値式」を作用した判別分析を行なうと、(23)式と表 3-1-6(2)、および、(24)式と表 3-1-6(3)の解が得られた。ちなみに、遺伝的アルゴリズムによる計算は上記と同様であり、計算は20～30世代で収束した。

$$(22) \quad f = -|w_1 a + w_2 b + w_3| \quad \text{あるいは} \quad +|w_1 a + w_2 b + w_3|$$

$$(-1 \leq w_1, w_2, w_3 \leq +1)$$

$$(23) \quad \text{if } f = -|a + b - 1| \geq -0.5 (= f_0) \text{ then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

$$(24) \quad \text{if } f = |a - b| \geq 0.5 (= f_0) \text{ then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

3つの論理変数 $[a] [b] [c]$ のうち、1つの変数だけが“真”のときに、排他的論理和 $[z]$ が“真”となるように定義すると、解は(25)式と表 3-1-7(1)、あるいは、(26)式と表 3-1-7(2)のようになる。また、これらの結果から、論理変数 $[a_k] (k=1, 2, \dots, K)$ の数が多いときには、排他的論理和は、(27)式あるいは(28)式で判別できることがわかる。

表 3-1-7 3変数の「排他的論理和」の真理値表と判別分析の結果

(1) 2次式((25)式)					(2) 絶対値式((26)式)					(3) 2次式((29)式)					(4) 絶対値式((30)式)				
a	b	c	f	z	a	b	c	f	z	a	b	c	f	z	a	b	c	f	z
0	0	0	0.0	0	0	0	0	-1	0	0	0	0	0	0	0	0	0	-1.5	0
0	0	1	0.5	1	0	0	1	-0	1	0	0	1	1	1	0	0	1	-0.5	1
0	1	0	0.5	1	0	1	0	-0	1	0	1	0	1	1	0	1	0	-0.5	1
0	1	1	0.0	0	0	1	1	-1	0	0	1	1	1	1	0	1	1	-0.5	1
1	0	0	0.5	1	1	0	0	-0	1	1	0	0	1	1	1	0	0	-0.5	1
1	0	1	0.0	0	1	0	1	-1	0	1	0	1	1	1	1	0	1	-0.5	1
1	1	0	0.0	0	1	1	0	-1	0	1	1	0	1	1	1	1	0	-0.5	1
1	1	1	-1.5	0	1	1	1	-2	0	1	1	1	0	0	1	1	1	-1.5	0

$$(25) \quad \text{if } f = 0.5 a + 0.5 b + 0.5 c - ab - bc - ca \geq 0.25 (= f_0)$$

$$\text{then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

$$(26) \quad \text{if } f = -|a + b + c - 1| \geq -0.5 (= f_0) \text{ then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

$$(27) \quad \text{if } f = 0.5 \sum_{\forall k} a_k - \sum_{\forall k_1} \sum_{\forall k_2} a_{k_1} a_{k_2} \geq 0.25 (= f_0)$$

$$\text{then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

$$(28) \quad \text{if } f = -\left| \sum_{\forall k} a_k - 1 \right| \geq -0.5 (= f_0) \text{ then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

さらに、3つの論理変数 $[a] [b] [c]$ のうち、1つまたは2つの変数が“真”のときに、排他的論理和 $[z]$ が“真”となるように(拡張して)定義すると、解は(29)式と表 3-1-7(3)、あるいは、(30)式と表 3-1-7(4)のようになる。また、これらの結果から、論理変数 $[a_k] (k=1, 2, \dots, K)$ の数が多いときには、排他的論理和は、(31)式あるいは(32)式で判別できることがわかる。

$$(29) \quad \text{if } f = a + b + c - ab - bc - ca \geq 0.5 (= f_0)$$

$$\text{then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

$$(30) \quad \text{if } f = |a + b + c - 1.5| \geq -1.0 (= f_0) \text{ then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

$$(31) \quad \text{if } f = \sum_{\forall k} a_k - \sum_{\forall k_1} \sum_{\forall k_2} a_{k_1} a_{k_2} \geq 0.5 (= f_0) \text{ then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

$$(32) \quad \text{if } f = - \left| \sum_{\forall k} a_k - 1.5 \right| \geq -1.0 (= f_0) \text{ then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

3-1-4 項 任意の論理関係の判別分析

「計算的な判別分析」は、“任意の”論理関係についても適用できる。まずは第1の例題として、3つの論理変数 $[a]$ $[b]$ $[c]$ に関して、表 3-1-8 の $[a]$ $[b]$ $[c]$ $[z]$ で示される“与えられた”真理値表の論理関係について、同じ方法でその判別分析の解を求めてみよう。ちなみに、この論理関係は(33)式の論理式で表すことができる。

$$(33) \quad z = (\bar{a} \cap \bar{b}) \cup (a \cap b) \cup (\bar{a} \cap b \cap \bar{c})$$

遺伝的アルゴリズムの“戦略”として、1つの世代で用意する染色体を20個として、遺伝的操作として「優先生き残り」「2個のトーナメントをした後のランダム交叉」「突然変異」「ランダム生き残り」を行い、それぞれ2個、10個、6個、2個を生成した。ランダム交叉の交叉率は100%、突然変異の変異率は25%である。次の世代の染色体を生成したあと、それぞれ5%の感度分析によって解の改善をした。判別式の形式を「2次式」とした場合、“完全に”判別できてからしばらくたった約300世代までの計算で得られた解は(34)式と表 3-1-8(1)であった。判別式を“V型”の絶対値式とした場合も、約200世代までの計算で得られた解は(35)式と表 3-1-8(2)であった。

$$(34) \quad \text{if } f = -0.112a - 0.048b - 0.011c - 0.311a^2 + 0.046b^2 + 0.009c^2 \\ + 0.843ab - 0.418bc + 0.000ca \geq -0.213 (= f_0) \\ \text{then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

$$(35) \quad \text{if } f = |0.667a + 1.000b - 0.334c - 0.500| \geq 0.333 (= f_0) \\ \text{then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

表 3-1-8 「3変数の論理関係の例題」の真理値表と判別分析の結果

(1) 2次式による解((34)式) (2) 絶対値式による解((35)式)

a	b	c	f	z	a	b	c	f	z
0	0	0	-0.000	1	0	0	0	0.500	1
0	0	1	-0.002	1	0	0	1	0.834	1
0	1	0	-0.002	1	0	1	0	0.500	1
0	1	1	-0.422	0	0	1	1	0.166	0
1	0	0	-0.423	0	1	0	0	0.167	0
1	0	1	-0.425	0	1	0	1	0.167	0
1	1	0	0.418	1	1	1	0	0.167	1
1	1	1	-0.002	1	1	1	1	0.833	1

当然のことながら、もっと計算をつまり解の探索を続ければ、よりよい解が見つかる可能性もあるが、(34)式や(35)式は十分に実行可能な解であり、実用の上で十分に役に立つ解である、といえる。

次に、より複雑な例題として、表 3-1-9 の $[a] [b] [c] [d] [z]$ の真理値表で与えられる 4 つの論理変数 $[a][b][c][d]$ の論理関係について、同じ方法でその判別分析の解を求めてみよう。ちなみに、この論理関係は(36)式で表すことができる。

$$(36) \quad z = (\bar{a} \cap c \cap d) \cup (a \cap b \cap c) \cup (\bar{a} \cap b \cap \bar{c} \cap \bar{d}) \cup (a \cap \bar{b} \cap \overline{(c \cap d)})$$

結果として、判別式の形式に「1 次式」「2 次式」「“V 型”の絶対値式」の形式を採用しても、(1 段階だけの)単純な判別分析では、また、限られた世代数までの計算つまり解の探索では、“完全な”判別ができなかった。

そこで、(2-9 節で説明した)より複雑なデータに対応できる「2 段階判別方式」の方法を適用すると、“完全な”判別をすることができた。ちなみに、遺伝的アルゴリズムの“戦略”として、1 つの世代で用意する染色体を 20 個として、遺伝的操作として「優先生き残り」「2 個のトーナメントをした後のランダム交叉」「突然変異」「ランダム生き残り」を行い、それぞれ 2 個、10 個、6 個、2 個を生成した。ランダム交叉の交叉率は 100%、突然変異の変異率は 25% である。次の世代の染色体を生成したあと、それぞれ 5% の感度分析によって解の改善をした。第 1 段階の解は 1000 世代で、第 2 段階の解は 500 世代で打ち切った。いずれも、判別分析の解の改善が期待できそうになくなるまでに十分な世代数である。

第 1 段階の判別式は“V 型”の絶対値式((37)式)で、この判別式で計算した判別得点が(“真”の判別得点の最小値の) $[0.395]$ 以下の場合には“偽”、(“偽”の判別得点の最大値の) $[0.822]$ 以上の場合には“真”と推定される。16 件中の 7 件は、これで判別ができた。判別得点の値が $[0.395]$ と $[0.822]$ の間の場合には、第 1 段階の判別式では誤判別が生じる可能性があるので、第 2 段階の(1 次式の)判別式((38)式)で計算した判別得点が、 $[0.219]$ より小さい場合は“偽”、大きい場合は“真”と推定できる。第 1 段階で完全には判別できなかった 9 件は、第 2 段階のこの判別式で“完全に”判別できた。

$$(37) \quad f = |0.406a - 0.375b - 0.426c - 0.791d + 0.791|$$

$$(38) \quad g = -0.437a + 0.438b + 0.876c + 0.878d$$

$$(39) \quad \text{if } f \geq 0.822 \text{ then } \hat{z} = 1 \text{ else if } f \leq 0.395 \text{ then } \hat{z} = 0 \\ \text{else if } g \geq 0.219 \text{ then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

表 3-1-9 「4変数の論理関係の例題」の真理値表と判別分析の結果

a	b	c	d	f	g	z
0	0	0	0	0.7910	0.0000	0
0	0	0	1	0.0000	-----	0
0	0	1	0	0.3650	-----	0
0	0	1	1	0.4260	1.7540	1
0	1	0	0	0.4160	0.4380	1
0	1	0	1	0.3750	-----	0
0	1	1	0	0.0100	-----	0
0	1	1	1	0.8010	2.1920	1
1	0	0	0	1.1970	-----	1
1	0	0	1	0.4060	0.4410	1
1	0	1	0	0.7710	0.4390	1
1	0	1	1	0.0200	-----	0
1	1	0	0	0.8220	0.0010	0
1	1	0	1	0.0310	-----	0
1	1	1	0	0.3960	0.8770	1
1	1	1	1	0.3950	1.7550	1

当然のことながら、この場合も、もっと計算をつまり解の探索を続ければ、よりよい解が見つかる可能性もあるが、(37)～(39)式は十分に実行可能な解であり、実用の上で十分に役に立つ解である、といえる。

ちなみに、(なにかほかの方法で入力と出力の間の「対応づけ」に関して、)すでに論理的な関係が分かっている場合には、このような面倒な手間をかけて、論理関係を判別式に直す必要はないであろう。しかし、実際に解決しなければならない問題では、“論理的な”対応関係にせよ、“相関的な”関係にせよ、「対応関係」がはっきり分かっていることは少ない。筆者は、“連続的な”説明変数に対するのとまったく同じ方法で、“機械的に”判別分析（意味的には「数量化理論2類」である）の解を求めることによって、“離散的な”入力に対して、“離散的な”出力を対応づける「印象の推定モデル」を構築することができる、という意味は決して小さくはない、と考えている。

3-2節 「回帰分析」の適用の方法

3-2-1項 「印象の推定モデル」としての「回帰分析」

繰り返しになるが、「解析的な判別分析」の方法と計算アルゴリズムは、計算技術的な理由から、いろいろな仮説や条件を前提として、解析的な「微分」という方法によって“最適な”解を求めようとするものであった。これに対して、(2章で説明した)判別分析の「計算的な方法」では、できるだけ仮説や条件を前提とはせずに、「分析の目的」を実際の問題の要請に基づいて設定し、これに対応して“最適な”(実際には“よりよい実行可能な”)解を、遺伝的アルゴリズムを利用して“計算的に”見つけだすことができる。この「計算的な方法」は、ほかの統計分析の問題にも適用することができる。

「回帰分析」は、1つあるいは複数の独立(説明)変数の組み合わせと1つの従属(被説明)変数の対応関係を“もっともよく説明できる”回帰式を導き出し、これを利用して、観測した独立変数のデータから従属変数の値を推定しようという方法である。

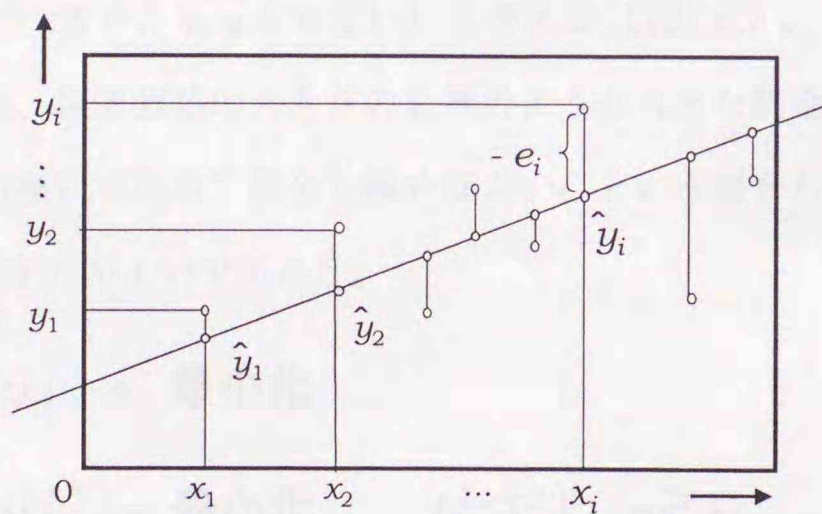


図 3-2-1 「回帰分析」の基本的なやり方

数学的な「印象の推定モデル」を構築しようとする立場では、たとえば、対象の画像の特徴量や物理的な仕様などを独立変数に、(「印象の計測」の方法として比較的ポピュラーである)「SD法」などの方法で測った印象を感じる程度(大きさつまり連続量)を従属変数として対応づけるモデルを考えることができる。(2-1節で示した)図 2-1-1 で、「判別分析」と「判別式」をそれぞれ「回帰分析」と「回帰式」に置き換えたものが、「『回帰分析』を利用した『印象の推定モデル』」の概念である。この図では、対象のカラー画像の特徴量を入力すると、(特定の印象の有無ではなく、)その印象を感じる程度(大きさ)が出力される。

3-2-2 項 「解析的な回帰分析」の方法の問題点

「判別分析」と同様に、「回帰分析」についても、「解析的な方法」と計算アルゴリズムが開発されており広く普及している。この「解析的な方法」では、観測データ (x_i, y_i) を“もっともよく説明できる”回帰式のパラメーター（回帰式が(1)式の場合には、 $[a]$ と $[b]$ ）は、「推定誤差 $[e_i]$ の平方和を最小にする」こと、つまり「最小自乗法」の原理と方法にしたがって求めることができる（(2)式）。この方法で求められた回帰式は、独立変数の値域とその近傍の範囲内で、ある統計的な信頼度の下に「推定」の目的に利用することができる。

$$(1) \quad \hat{y}_i = a x_i + b$$

$$(2) \quad \sum_{\forall i} e_i^2 = \sum_{\forall i} (\hat{y}_i - y_i)^2 \rightarrow \text{最小化}$$

この方法は、回帰式による推定値 $[\hat{y}_i]$ と実際の観測値 $[y_i]$ の差、つまり推定誤差 $[e_i (= \hat{y}_i - y_i)]$ に対して、その大きさの2乗の罰金を課して、その罰金の合計を最小にすることを“最適”（つまり、もっともよい）と考える（(2)式）。しかし、実際に回帰分析が適用される問題では、推定誤差の大きさの2乗のような大きな罰金を課する必要はなく、たとえば推定誤差の“絶対値に当たる”罰金を課せばよい、という場合も考えられる。この場合は、目的関数は(3)式で表すのがよいであろう。

$$(3) \quad \sum_{\forall i} |\hat{y}_i - y_i| \rightarrow \text{最小化}$$

$$(4) \quad \sum_{\forall i} |\hat{y}_i - y_i|^n \rightarrow \text{最小化} \quad (\text{ただし } n \geq 3)$$

$$(5) \quad \sum_{\forall i} f(\hat{y}_i - y_i) \rightarrow \text{最小化}$$

また、推定誤差の大きさの2乗のような小さな罰金では足りず、推定誤差の絶対値の n 乗 ($n \geq 3$) の罰金を課したい、という場合もあるであろう。この場合の目的関数は(4)式のようなであろう。あるいは、推定誤差の大きさに応じて、たとえば、段階的な罰金など、なにかの対応関係 $[f]$ で罰金の大きさを決める、といったことも考えられる。（この「対応関係」には、推定誤差が大きくなると、罰金も大きくなる“単調な”ものだけでなく、誤差の値の大きさに応じて、罰金の大きさが増減するものも含んでよいであろう。）この場合の目的関数は(5)式で表される。

また、「計算的な判別分析」でも、分析の対象のすべてに関心をもつのではなく、誤判別が生じる可能性がある対象にだけ関心を持つという問題の設定があったのと同様に、回帰分析を適用する問題でも、すべての対象の推定誤差に関心を持つのではなく、推定誤差の最大値だけに関心を持っている、つまり、罰金の合計ではなく、推定誤差の最大値を最小にしたい、といった場合もあるであろう。この場合は、目的関数は(6)式で表すのがよいであろう。

$$(6) \quad \max_{\forall i} |\hat{y}_i - y_i| \rightarrow \text{最小化}$$

あるいは、推定誤差がプラス側ならば受け入れられるが、マイナス側であることは許さない、という条件があることもある。この推定は“常に安全側に”ということである。この場合にも、(3)式～(6)式などの目的関数に加えて、(7)式の制約が付く。反対に、推定誤差がマイナス側でなければいけないという、反対の方向に“常に安全側に”の場合は、(3)式～(6)式などの各種の目的関数に加えて、(8)式の制約が付く。

$$(7) \quad \forall i \text{ に対して } \hat{y}_i - y_i \geq 0$$

$$(8) \quad \forall i \text{ に対して } \hat{y}_i - y_i \leq 0$$

また、推定誤差は、ある範囲内にあることが必要である、つまり、この範囲を超えた場合にはなにかの重大な危険が生じる可能性がある、といった制約条件も考えられる。この場合は確率的には小さくても、推定誤差が一定の値を超えてしまうことが絶対にないように、といった制約である。この条件は(9)式で表される。なお、データによっては、この条件を満足する解は存在しないこともある。

$$(9) \quad |\hat{y}_i - y_i| \leq \text{一定}$$

いずれにしても、こういった現実の問題で要求されるいろいろな条件や制約を満足する解は、従来の回帰分析の「解析的な方法」（以下、「解析的な回帰分析」という）では求めることはむずかしい。しかし、遺伝的アルゴリズムを利用した「計算的な方法」ならば、こういった目的や条件を満たす解を“計算的に”見つけ出すことができる。この回帰分析の「計算的な方法」（以下、「計算的な回帰分析」という）では、目的関数（つまり適合度の評価）に“任意”のものを選択できるだけでなく、回帰式の形式も、微分が可能な「1次式」や「2次式」だけではなく、微分ができない「絶対値関数」や「四捨五入関数」などを含んだ“任意”のものを採用することができる。ここで、「任意」とは、実際の問題の要求に合わせて、という意味である。

3-2-3 項 「計算的な単回帰分析」の方法と計算の例示

表 3-2-1 「回帰分析」の例題

#	国語	英語	数学	物理	I Q
i	x_i	y_i	u_i	v_i	z_i
01	73	57	61	48	84
02	74	59	70	48	94
03	86	63	79	52	116
04	74	59	65	48	88
05	80	60	55	42	92
06	78	59	85	58	112
07	86	61	57	48	100
08	82	59	83	52	100
09	87	64	70	52	108
10	80	59	75	52	106
平均値	80.00	60.00	70.00	50.00	100.00

表 3-2-1 は、「回帰分析」の簡単な例題である（文献 9 3）。

このうち、国語の成績 $[x_i]$ を独立変数に、英語の成績 $[y_i]$ を従属変数にした単純な回帰分析の問題（以下、「問題 1」という）を例にして、前項で説明したいろいろな目的を満足する回帰式を、遺伝的アルゴリズムを利用した「計算的な回帰分析」で見つけだしてみよう。
以下、回帰式の形式は 1 次式（(1) 式）を採用する。

$$(10) \quad \hat{y}_k = 0.34 x_k + 32.8$$

表 3-2-2 問題 1 の「解析的な単回帰分析」の結果

#	国語	英語	標準の 推定値	推定値 の誤差	誤差の 二乗
i	x_i	y_i	$\hat{y}_i$	e_i	e_i^2
01	73	57	57.62	0.62	0.38
02	74	59	57.96	-1.04	1.08
03	86	63	62.04	-0.96	0.92
04	74	59	57.96	-1.04	1.08
05	80	60	60.00	0.00	0.00
06	78	59	59.32	0.32	0.10
07	86	61	62.04	1.04	1.08
08	82	59	60.68	1.68	2.82
09	87	64	62.38	-1.62	2.62
10	80	59	60.00	1.00	1.00
平均値	80.00	60.00	60.00	0.00	11.10
			最大値	1.68	
			最小値	-1.62	

(10)式と表 3-2-2 は、「解析的な回帰分析」による解である。(以下、表中、もともとのデータは“非斜字”で、計算された値は“斜字”で表してある。また、最小値と最大値は“太字”で、(平均値ではなく)合計値は“網かけ字”で表してある。)

次は、「計算的な回帰分析」である。ここで、計算技術的な理由から、(1)式は、(11)式のように変形しておく。そして、 $[m]$ $[a']$ $[b']$ の3つのパラメーターを「遺伝子」とした「染色体」を想定し、これに繰り返し遺伝的操作を加えることによって、よりよい実行可能な解を見つけていく。ここで、遺伝的操作の計算を簡単にするために、 $[-1 \leq a' \leq +1]$ と $[-1 \leq b' \leq +1]$ という制約条件を付加しておく。この結果、 $[a]$ と $[b]$ が任意の値をとれるよう $[a = m a']$ と $[b = m b']$ となる、倍率パラメーターの $[m]$ を設ける。 $[m_0]$ は倍率パラメーターが採りうる最大値である。

$$(11) \quad \hat{y}_k = a x_k + b = m (a' x_k + b')$$

ここで、 $0 \leq m \leq m_0$, $-1 \leq a' \leq +1$, $-1 \leq b' \leq +1$

$$(12) \quad a = m a' \quad \text{and} \quad b = m b'$$

遺伝的アルゴリズムの“戦略”として、1つの世代に20個の染色体を用意し、初期解つまり第1世代の染色体は、ランダムに生成した。倍率パラメーターが採りうる最大値 $[m_0]$ は、「解析的な方法」で得られた最適解を基に設定した。遺伝的操作として「優先生き残り」「2個のトーナメントをした後のランダム交叉」「突然変異」「ランダム生き残り」を行い、それぞれ2個、10個、6個、2個を生成した。ランダム交叉の交叉率は100%、突然変異の変異率は25%である。次の世代の染色体を生成したあと、それぞれ10%の感度分析によって解の改善をした。計算は30～40世代で収束した。この結果、遺伝的アルゴリズムによる「計算的な回帰分析」で見つけ出した解は、「解析的な回帰分析」による結果とほぼ同じものであった。

次に、「推定誤差の絶対値和を最小にする」(3)式)を分析の目的にして、「計算的な回帰分析」を行なうと、(13)式と表 3-2-3 の結果(GA推定値、誤差、自乗)が得られた。この結果は、「解析的な方法」による「誤差の自乗和を最小にする」解とは(サンプルが少ないので違いも小さいが、それでも)若干違っているのがわかる。ちなみに、遺伝的アルゴリズムの計算は、上記と同じであるが、初期解つまり第1世代の染色体の1つには、「解析的な方法」で得られた最適解を採り入れ、残りは“ランダムに”生成した。計算は30～40世代で収束した。

$$(13) \hat{y}_k = 0.381 x_k + 29.437$$

表 3-2-3 問題 1 の「誤差の絶対値和を最小にした場合」の結果

#	国語	英語	標準の 推定値	誤差	自乗	GAの 推定値	誤差	自乗
i	x_i	y_i	${}_1y_i$	${}_1e_i$	${}_1e_i^2$	${}_2y_i$	${}_2e_i$	${}_2e_i^2$
01	73	57	57.62	0.62	0.38	57.25	0.25	0.06
02	74	59	57.96	-1.04	1.08	57.63	-1.37	1.87
03	86	63	62.04	-0.96	0.92	62.20	-0.80	0.64
04	74	59	57.96	-1.04	1.08	57.63	-1.37	1.87
05	80	60	60.00	0.00	0.00	59.92	-0.08	0.01
06	78	59	59.32	0.32	0.10	59.16	0.16	0.02
07	86	61	62.04	1.04	1.08	62.20	1.20	1.45
08	82	59	60.68	1.68	2.82	60.68	1.68	2.82
09	87	64	62.38	-1.62	2.62	62.58	-1.42	2.01
10	80	59	60.00	1.00	1.00	59.92	0.92	0.84
平均値	80.00	60.00	60.00	0.93	11.10	59.92	0.92	11.59
				最大の誤差			最大の誤差	
				1.68			1.68	
				最小の誤差			最大の誤差	
				-1.62			-1.42	

「推定誤差の最大値を最小にする」 ((6) 式)) を分析の目的にした場合の結果は、(14) 式と表 3-2-4 (GA 推定値、誤差、自乗) の通りである。この場合も、遺伝的アルゴリズムの“戦略”を前と同様なものとする、計算は 30 ~ 40 世代で収束した。

$$(14) \hat{y}_k = 0.343 x_k + 32.514$$

表 3-2-4 問題 1 の「誤差の最大値を最小にした場合」の結果

#	国語	英語	標準の 推定値	誤差	自乗	GAの 推定値	誤差	自乗
i	x_i	y_i	${}_1y_i$	${}_1e_i$	${}_1e_i^2$	${}_2y_i$	${}_2e_i$	${}_2e_i^2$
01	73	57	57.62	0.62	0.38	57.55	0.55	0.31
02	74	59	57.96	-1.04	1.08	57.90	-1.10	1.22
03	86	63	62.04	-0.96	0.92	62.01	-0.99	0.98
04	74	59	57.96	-1.04	1.08	57.90	-1.10	1.22
05	80	60	60.00	0.00	0.00	59.95	-0.05	0.00
06	78	59	59.32	0.32	0.10	59.27	0.27	0.07
07	86	61	62.04	1.04	1.08	62.01	1.01	1.02
08	82	59	60.68	1.68	2.82	60.64	1.64	2.69
09	87	64	62.38	-1.62	2.62	62.36	-1.65	2.71
10	80	59	60.00	1.00	1.00	59.95	0.95	0.91
平均値	80.00	60.00	60.00	0.00	11.10	59.95	-0.05	11.12
				最大の誤差			最大の誤差	
				1.68			1.64	
				最小の誤差			最小の誤差	
				-1.62			-1.65	

「誤差の2乗和を最小にする」((2)式)を目的に、「推定値が観測値を下回らない」((7)式)を条件にした結果は、(15)式と表3-2-5の $[{}_2y_i] [{}_2e_i] [{}_2e_i^2]$ の通りである。(ちなみに、 $[{}_1y_i] [{}_1e_i] [{}_1e_i^2]$ は、“解析的な”方法による解である。)この場合も、遺伝的アルゴリズムの“戦略”は前と同様にすると、計算は30～40世代で収束した。

表3-2-5 問題1の「推定値が観測値を下回らない場合」と
「推定値が観測値を上回らない場合」の結果

#	国語	英語	標準の			GA下			GA上		
			推定値	誤差	自乗	推定値	誤差	自乗	推定値	誤差	自乗
i	x_i	y_i	${}_1y_i$	${}_1e_i$	${}_1e_i^2$	${}_2y_i$	${}_2e_i$	${}_2e_i^2$	${}_3y_i$	${}_3e_i$	${}_3e_i^2$
01	73	57	57.62	0.62	0.38	58.61	1.61	2.59	56.69	-0.31	0.09
02	74	59	57.96	-1.04	1.08	59.01	0.01	0.00	56.95	-2.05	4.21
03	86	63	62.04	-0.96	0.92	63.81	0.81	0.65	60.02	-2.98	8.88
04	74	59	57.96	-1.04	1.08	59.01	0.01	0.00	56.95	-2.05	4.21
05	80	60	60.00	0.00	0.00	61.41	1.41	1.98	58.48	-1.52	2.30
06	78	59	59.32	0.32	0.10	60.61	1.61	2.59	57.97	-1.03	1.06
07	86	61	62.04	1.04	1.08	63.81	2.81	7.88	60.02	-0.98	0.96
08	82	59	60.68	1.68	2.82	62.21	3.21	10.29	59.00	-0.00	0.00
09	87	64	62.38	-1.62	2.62	64.21	0.21	0.04	60.28	-3.72	13.87
10	80	59	60.00	1.00	1.00	61.41	2.41	5.80	58.48	-0.52	0.27
平均値	80.00	60.00	60.00	0.00	11.10	61.41	1.41	31.82	58.48	-1.52	35.8
			最大の誤差	1.68		最大の誤差	3.21		最大の誤差	-0.00	
			最小の誤差	-1.62		最小の誤差	0.01		最小の誤差	-3.72	

$$(15) \quad \hat{y}_k = 0.4 x_k + 29.408$$

同様に、「誤差の2乗和を最小にする」((2)式)を目的に、「推定値が観測値を上回らない」((8)式)を条件にした結果は、(16)式と表3-2-5の $[{}_3y_i] [{}_3e_i] [{}_3e_i^2]$ の通りである。この場合は、遺伝的アルゴリズムの計算は20～30世代で収束した。

$$(16) \quad \hat{y}_k = 0.256 x_k + 38.004$$

3-2-4 項 「計算的な重回帰分析」の方法と計算の例示

次の例題は、国語の成績 $[x_i]$ と物理の成績 $[y_i]$ の2つを“独立変数”に、IQ $[z_i]$ を“従属変数”にした重回帰分析の問題である（以下、「問題2」という）。回帰式の形式を1次式とした場合の「解析的な方法」による結果を、(17)式と表3-2-6に示しておく。

$$(17) \hat{z}_k = 1.15 x_k + 1.5125 y_k - 67.625$$

表 3-2-6 問題2の「解析的な重回帰分析」の結果

#	国語	物理	I Q	標準推定値	誤差	誤差の二乗
i	x_i	y_i	z_i	$\hat{z}_i$	e_i	e_i^2
01	73	48	84	88.93	4.92	24.26
02	74	48	94	90.08	-3.93	15.41
03	86	52	116	109.93	-6.08	36.91
04	74	48	88	90.08	2.07	4.31
05	80	42	92	87.90	-4.10	16.81
06	78	58	112	109.80	-2.20	4.84
07	86	48	100	103.88	3.88	15.02
08	82	52	100	105.33	5.32	28.36
09	87	52	108	111.08	3.07	9.46
10	80	52	106	103.03	-2.98	8.85
平均値	80.00	50.00	100.00	100.00	0.00	164.20
						合計値
					最大の誤差	5.32
					最小の誤差	-6.08

前項の例題と同様に、遺伝的アルゴリズムの“戦略”として、1つの世代に20個の染色体を用意し、初期解つまり第1世代の染色体は、ランダムに生成した。倍率パラメーターが採りうる最大値 $[m_o]$ は、「解析的な方法」で得られた最適解を基に設定した。遺伝的操作として「優先生き残り」「2個のトーナメントをした後のランダム交叉」「突然変異」「ランダム生き残り」を行い、それぞれ2個、10個、6個、2個を生成した。ランダム交叉の交叉率は100%、突然変異の変異率は25%である。次の世代の染色体を生成したあと、それぞれ10%の感度分析によって解の改善をした。計算は40～50世代で収束した。この結果、「計算的な重回帰分析」で見つけ出した解は、「解析的な方法」による結果とほぼ同じものであった。

次に、「推定誤差の絶対値和を最小にする」((3)式)を目的して、「計算的な回帰分析」を行った。遺伝的アルゴリズムの計算は、上記と同じであるが、初期解つまり第1世代の染色

体は、「解析的な方法」で得られた最適解を含み、残りは“ランダムに”生成した。計算は40～50世代で収束した。結果は、(18)式と表3-2-7の通りである。

$$(18) \hat{z}_k = 1.004 x_k + 2.016 y_k - 83.110$$

表 3-2-7 問題 2 の「誤差の絶対値和を最小にした場合」の結果

# i	国語 x_i	物理 y_i	I Q z_i	標準			G A		
				推定値 $\hat{z}_i$	誤差 e_i	自乗 e_i^2	推定値 $\hat{z}_i$	誤差 e_i	自乗 e_i^2
01	73	48	84	88.93	4.92	24.26	86.95	2.95	8.70
02	74	48	94	90.08	-3.93	15.41	87.95	-6.05	36.55
03	86	52	116	109.93	-6.08	36.91	108.07	-7.93	62.95
04	74	48	88	90.08	2.07	4.31	87.95	-0.05	0.00
05	80	42	92	87.90	-4.10	16.81	81.88	-10.12	102.37
06	78	58	112	109.80	-2.20	4.84	112.13	0.13	0.02
07	86	48	100	103.88	3.88	15.02	100.00	0.00	0.00
08	82	52	100	105.33	5.32	28.36	104.05	4.05	16.40
09	87	52	108	111.08	3.07	9.46	109.07	1.07	1.14
10	80	52	106	103.03	-2.98	8.85	102.04	-3.96	15.67
平均値	80.00	50.00	100.00	100.00	0.00	164.20	98.01	-1.99	243.81
				最大の誤差 5.32			最大の誤差 4.05		
				最小の誤差 -6.08			最小の誤差 -10.12		

「推定誤差の最大値を最小にする」((6)式)を目的にした場合の結果は、(19)式と表3-2-8の通りである。この場合も、計算は40～50世代で収束した。

$$(19) \hat{z}_k = 1.332 x_k + 1.437 y_k - 78.650$$

表 3-2-8 問題 2 の「誤差の最大値を最小にした場合」の結果

# i	国語 x_i	物理 y_i	I Q z_i	標準			G A		
				推定値 $\hat{z}_i$	誤差 e_i	自乗 e_i^2	推定値 $\hat{z}_i$	誤差 e_i	自乗 e_i^2
01	73	48	84	88.93	4.92	24.26	87.56	3.56	12.69
02	74	48	94	90.08	-3.93	15.41	88.89	-5.11	26.07
03	86	52	116	109.93	-6.08	36.91	110.63	-5.37	28.88
04	74	48	88	90.08	2.07	4.31	88.89	0.89	0.80
05	80	42	92	87.90	-4.10	16.81	88.26	-3.74	13.96
06	78	58	112	109.80	-2.20	4.84	108.59	-3.41	11.61
07	86	48	100	103.88	3.88	15.02	104.88	4.88	23.79
08	82	52	100	105.33	5.32	28.36	105.30	5.30	28.07
09	87	52	108	111.08	3.07	9.46	111.96	3.96	15.67
10	80	52	106	103.03	-2.98	8.85	102.63	-3.37	11.33
平均値	80.00	50.00	100.00	100.00	0.00	164.20	99.76	-0.24	172.87
				最大の誤差 5.32			最大の誤差 5.30		
				最小の誤差 -6.08			最小の誤差 -5.37		

「誤差の2乗和を最小にする」((2)式)を目的に、「推定値が観測値を下回らない」((7)式)あるいは「推定値が観測値を上回らない」((8)式)を条件にした結果は、それぞれ(20)式あるいは(21)式と表3-2-9の通りである。2つの場合とも、遺伝的アルゴリズムの“戦略”は前と同様で、その計算は30～40世代あるいは10～20世代で収束した。

$$(20) \hat{z}_k = 1.286 x_k + 1.610 y_k - 78.316$$

$$(21) \hat{z}_k = 1.191 x_k + 1.322 y_k - 66.406$$

表3-2-9 問題2で「推定値が観測値を下回らない場合」と

「推定値が観測値を上回らない場合」の結果

#	国語	英語	IQ	標準の			GA下			GA上		
				推定値	誤差	自乗	推定値	誤差	自乗	推定値	誤差	自乗
i	X _i	Y _i	Z _i	$\hat{z}_i$	e_i	e_i^2	$\hat{z}_i$	e_i	e_i^2	$\hat{z}_i$	e_i	e_i^2
01	73	48	84	88.93	4.92	24.26	92.84	8.84	78.18	83.99	-0.01	0.00
02	74	48	94	90.08	-3.93	15.41	94.13	0.13	0.02	85.18	-8.82	77.72
03	86	52	116	109.93	-6.08	36.91	116.00	0.00	0.00	104.76	-11.24	126.25
04	74	48	88	90.08	2.07	4.31	94.13	6.13	37.55	85.18	-2.82	7.93
05	80	42	92	87.90	-4.10	16.81	92.18	0.18	0.03	84.40	-7.60	57.79
06	78	58	112	109.80	-2.20	4.84	115.37	3.37	11.37	103.17	-8.83	78.00
07	86	48	100	103.88	3.88	15.02	109.56	9.56	91.39	99.48	-0.52	0.27
08	82	52	100	105.33	5.32	28.36	110.86	10.86	117.85	100.00	-0.00	0.00
09	87	52	108	111.08	3.07	9.46	117.29	9.29	86.23	105.96	-2.05	4.18
10	80	52	106	103.03	-2.98	8.85	108.28	2.28	5.22	97.62	-8.38	70.26
平均値	80.0	50.0	100.0	100.00	0.00	164.20	105.06	5.06	427.85	94.97	-5.03	422.41
				最大の誤差	5.32		最大の誤差	10.86		最大の誤差	0.00	
				最小の誤差	-6.08		最小の誤差	0.00		最小の誤差	-11.24	

この節では、「解析的な(重)回帰分析」の問題点を検討し、その問題点を解決するために、実際に回帰分析を適用する問題が要求する目的に応じた解を“柔軟に”求めることができる、遺伝的アルゴリズムを利用した「計算的な回帰分析」の方法を説明した。この方法の開発によって、対象のカラー画像の特徴量などの“連続的な”量を独立(説明)変数に、その対象に対して回答者が申告する“連続的な”印象の程度(大きさ)を従属(被説明)変数にした「印象の推定モデル」を構築することができる方法論を手に入れることができた。

4章 「印象の推定」と「印象の推定モデル」の今後の研究

この論文の主題である「印象の推定」と「印象の推定モデル」、（「印象の推定モデル」の構築の方法である）「判別分析」に対する「遺伝的アルゴリズムを利用した計算的な方法」、そして、これらを含んだより大きな研究課題としての「『印象を扱う仕事』のための工学」である「印象の工学」の今後の研究についてまとめておく。（以下、「印象の工学」、「印象の推定」と「印象の推定モデル」、「遺伝的アルゴリズムを利用した計算的な方法」の順に述べる。）

（1）「印象の工学」は、「印象を“工学的に”取り扱う意図」を表す筆者の造語で、筆者は、これを、実際に存在する「印象を扱う仕事」のための工学と定義している。そして、その目的は、「印象を扱う仕事」の成果の品質とプロセスの効率を上げる（改善・向上する）ことと解釈し、この目的を「印象を扱う仕事」の方法とプロセスを明らかにし、それについてその合理性や妥当性を問う、という方法で達成しようとしている。

筆者は、この目的と方法にしたがって、まずは「印象を扱う仕事」のうち、その機能的なあるいはプロセス的な意味での“要素的な”仕事である「印象の計測」「印象の分析」「印象の推定」「印象の計画」「印象の設計」「印象の実現・構築」「印象の管理」などを切り出して、それぞれの「仕事」の成果の品質とプロセスの効率を上げるための検討に必要な“検討事項”を集めた「チェックリスト」の作成を進めている。実際の作業としては、筆者の身の回りやほかの研究者と実務家の仕事の観察・インタビュー調査・文献調査などを行い、「仕事」の成果の品質とプロセスの効率を上げるための問題点を抽出し検討し整理している。

「印象を扱う仕事」に取り組む人たちは、この「チェックリスト」を参照することによって、自分に取り組むあるいは取り組んでいる仕事の内容、より詳細に言えば目的、内容、前提条件、環境条件、方法、プロセス、成果、評価の仕方などについて、より深くより慎重に理解し検討することができ、その結果として、その仕事の成果の品質とプロセスの品質を上げることが期待できる。そういったものの完成を目指している。

1－3節でもその一部を例示した「なんのためにその仕事をするのか」「その成果は誰が使うのか」「すでに分かっていることはなにか」「まだ分からないことはなにか」「どんな方法があるのか」「その方法の仮説や前提条件はなにか」などの仕事の検討事項は、「印象を扱う仕事」の内容の“定義”に相当するものであったが、これを実際の問題で生じるいろいろな疑問点や問題点などを、できるだけきめ細かく用意しておこう、実際の問題をできるだけきめ細かく検討・整理しよう、できるようにしようというのが狙いである。これは「印象を扱う仕事」

を進めるための「ガイドライン」作りの仕事であるともいえる。

この仕事については、今後も、「心理学」を含むより多くの分野の研究者と実務家の協力を得て、より多くの「印象を扱う仕事」の“実務”を観察・検討し、“実務”の上でしばしば生じる疑問点や問題点の発見と整理を進めていこう、と考えている。また、これらの成果を「印象を扱う仕事」の“実務”に適用して、その有効性を確認しながら、その不具合を修正していこう、と考えている。

さらに、「印象を扱う仕事」には、実際の仕事の実行つまり具体化の段階で、技術的な選択が必要になる。たとえば「印象の計測」の仕事であれば、いま取り組んでいる「印象の計測」の仕事にはどの方法を選択すればよいか、被験者はどのように集めればよいか、刺激はどのように提示すればよいかなどの選択が必要になる。こういった“技術的な”疑問点や問題点に対しても、（とくに経験の少ない）利用者に対する支援が欲しいであろう。

「印象の計測」について少しく補足をすると、利用者の必要に応じて、計測方法の選択のしかた、それぞれの計測方法の仮説や前提条件、問題点や注意点、結果の限界、被験者（回答者）の集め方、ガイダンスのしかたなどの「知識」を取り出すことができれば、利用者への支援はより有効なものになるであろう。（ちなみに、現在は、この種の技術的な知識の多くは、書籍などの形式で提供されて、ノウハウは、人からの伝達という形式で提供されている。）そういった観点から、「印象を扱う仕事」のそれぞれの“要素的な”仕事について、技術的な知識の「知識ベース」化をはかれば、「チェックリスト」とともに、「印象を扱う仕事」に取り組む人たちを有効に支援するであろう。「チェックリスト」の作成と併せて、この課題にも取り組みたい、と考えている。

またさらに、「印象を扱う仕事」のプロセスの“効率”を上げること、つまり、仕事のプロセスにおいて、時間的な無駄はなかったか、コスト的な無駄はなかったか、その他の資源の無駄はなかったかななどを検討し、その効率を改善・向上をすることも、「印象の工学」の大きな目的の1つである。この課題に対しても、実際に発見や解決が求められている「印象を扱う仕事」の“実務”の観察と検討を通じて、また、「印象を扱う仕事」に“よく似た”「ソフトウェアを生産（開発）する仕事」の生産技術である「ソフトウェア工学」の方法と技術の技術移転を通じて検討を進めよう、と考えている。（ちなみに、「ソフトウェアを生産する仕事」も「印象を扱う仕事」も、基本的に人が作業をする、仕事の成果が目に見えない、利用者ニーズが不明確なことが多い、ニーズを記述するドキュメント技法がない、仕事の成果の評価がむずかしいなどといった共通点を持っている。）

具体的には、「仕事の要求仕様の定義」「成果の品質の評価」「プロセスの効率の評価」などの方法についての検討で、“プロセス的な”意味での要素的な仕事、たとえばなにかのモノ(製品)を対象にして、その「モノ作り」に関して、「印象の計画」「印象の設計」「印象の評価」などに対する「チェックリスト」の作成を検討したい、と考えている。

(2) 2章と3章で述べた「判別分析」「数量化理論2類」「回帰分析」などの「外的な基準を持った多変量解析」の方法を利用した、数学的な「印象の推定モデル」の構築は、おもに「印象の計測」によって得られた、回答者(被験者)の“平均的な”回答を推定するためのモデルであった。しかし、なにかの対象(つまり刺激)に対する「印象」には、人が違っててもほぼ同じものもあるが、その人のその対象に対する利害や興味・関心、あるいは、その人の持っている知識や経験や判断の基準などによって異なっているのが普通である。

そこで、個人のその対象に対する利害や興味・関心あるいは個人の知識や経験などを評価して、“個人の特性に合った”回答を出力つまり推定する、数学的な「印象の推定モデル」の構築が考えられる。こういったモデルが構築できれば、本当に“個人の好みに合った”あるいは“状況に合った”情報の検索や提供が可能になろう。ちなみに、1-4-1項で簡単に紹介した「感性情報処理」での「印象の推定モデル」の構築の試みは、こういった意図を持ったものなのであろう。しかし、これらの多くは、たとえば平均的なデータと個人のデータの“ずれ”や“違い”をなにかの知識や論理で修正する、といった程度のもので、「個人のその対象に対する利害や興味・関心あるいは知識や経験などの評価」に明に取り組んだものとはいえない。

(筆者も、1-4-2項で説明した「印象の工学」の研究のなかの「印象の推定モデル」の構築の方法の1つである「同時関係分析法」の検討で、「ベイズの定理」を利用して、「同時関係」の“平均的な”確かさの値を、観察された個人の特性によって修正して、“個人の特性に合わせた”モデルを構築する方法を提案しているが、「個人のその対象に対する利害や興味・関心あるいは知識や経験などの評価」への取り組みといったことからすると、そのゴールにはまだまだ程遠いものであろう。)

また、「印象」は、ときや場所や状況などによっても異なる。これらについても、ときや場所や状況などに合わせた回答を出力する、数学的な「印象の推定モデル」の構築という課題がある。こういったことから、人が感じる「印象」が、人、とき、場所、状況などで違うことを前提として、その「印象」を推定できる「印象の推定モデル」の構築という課題に取り組みたい、と考えている。多分、この課題は、「印象のメカニズム」の分析・解明と記述という課題と同じものであろう。

さて、「印象」は、基本的に、時間が経つと変化をする“ダイナミックな”(動的な)もの

である。「流行」「飽き」「新鮮さ」などは、“ダイナミックに”変化する「印象」の代表的な現象である。企業によるモノやサービスなどの製品の提供も、製品のライフサイクルにしたがって、利用者が企業が意図した「印象」を感じてくれるように、企業は「計画的な陳腐化」などといったことをするのである。しかし、概念的な取り組みを除けば、現在できている「印象」への取り組みは、特定の時間の断面で切った“スタティックな”（静的な）「印象」の作り込みや評価だけのようであり、これを改善するためには、「印象」をそもそも“ダイナミックな”ものとして捉えた「印象のダイナミクス」とでもいうべき取り組みが必要とされている。

こういった必要に対しては、社会システムの予測で使われている「システムダイナミクス」(systems dynamics, SD)の方法を参考に、「印象のダイナミクス」を作っていきたい、と考えている。ちなみに、システムダイナミクスでは、人間を含んだシステムである社会システムの特徴を、非線型、複雑、時間的な変化、情報の時間的な遅れ、歪み、増幅、原因と結果の時間的・空間的な隔たりなどと考え、“定量的な”データだけでなく、“定性的な”知識をも利用して、システムの構造を記述する。（ちなみに、人間を含んだシステムでは、一般に、数値化できる“定量的な”データよりも、言葉や概念で記述される“定性的な”知識の方が豊富である。）これらの特徴は、人の感じる「印象」を含む現象にも共通のものと考えられる。こういった認識を基に、システムダイナミクスの方法を参考に、“ダイナミックな”「印象」の構造を解明・記述し予測をすることができる「印象のダイナミクス」の構築にも取り組んでいきたい、と考えている。

(3) この論文で検討した遺伝的アルゴリズムを利用した「外的な基準を持った多変量解析」に対する「計算的な方法」は、その方法としての“柔軟性”から、「主成分分析」「因子分析」「クラスター分析」などの「外的な基準を持たない多変量解析」にも適用して効果を上げることができる可能性が考えられる。つまり、現在のそれぞれの多変量解析のモデルに対して開発され利用されている「解析的な方法」は、（この論文で取り上げた「判別分析」や「数量化理論2類」や「回帰分析」と同様に、）計算技術的な意味も含めて、いろいろな仮説や条件を前提として構築されている。しかし、これらの仮説や条件は、実際の問題が要求するものとは必ずしも一致していない。

そこで、これらの多変量解析のモデルに、遺伝的アルゴリズムを利用した「計算的な方法」を適用し、「解析的な方法」が前提（必要）としているいろいろな仮説や条件を前提とする必要がない、そして、実際の問題が要求するいろいろな条件を満たすことができる“有効な”解を見つけ出すことができる“柔軟な”方法の開発を模索したい、と考えている。

文 献

1. 松村明（編）：大辞林、1988年11月、三省堂
2. 大澤光：「印象の工学」の実現に向けて、大澤光（編著）：「印象の工学」とはなにか（人の「印象」を正しく分析・利用するために）、58-79ページ（第2章）、2000年1月、丸善プラネット
3. 堤正義（著）：逆問題の数学、2000年3月、共立出版
4. 長町三生（著）：感性工学のおはなし、77-124ページ、1995年7月、日本規格協会
5. 長町三生（著）：感性工学（感性をデザインに活かすテクノロジー）、1989年11月、海文堂出版
6. 長町三生（編著）：感性商品学（感性工学の基礎と応用）、1993年9月、海文堂出版
7. C. E. Osgood, G. J. Suci and P. H. Tannenbaum : The Measurement of Meaning, University of Illinois press, 1957
8. 岩下豊彦（著）：オスグッドの意味論とSD法、1979年4月、川島書店
9. 岩下豊彦（著）：SD法によるイメージの測定（その理解と実施の手引）、1983年1月、川島書店
10. 増山英太郎・小林茂雄（著）：センソリー・エバリュエーション（官能検査へのいざない）、1989年3月、垣内出版
11. 浅井邦二（編著）：こころの測定法（心理学における測定の方法と課題）、1994年3月、実務教育出版
12. 神宮英夫（著）：印象測定の心理学（感性を考える）、1996年4月、川島書店
13. 大澤光：印象はどんな特性を持っているのだろうか（「印象の計測」の問題点と「印象の特性」の評価の試み）、大澤光・西川泰夫（共編）：印象の工学（印象はどう測ればよいか）、現代のエスプリ 364号、37-65ページ、1997年11月、至文堂
14. 加藤俊一・栗田多喜男：画像の内容検索（電子美術館への応用）、情報処理、第33巻第5号、情報処理学会、1992年
15. 栗田多喜夫・加藤俊一・福田郁美・板倉あゆみ：印象語による絵画データベース検索、情報処理学会論文誌、第33巻第11号、1992年
16. 木本晴夫：感性語による画像検索とその精度評価、情報処理学会論文誌、第40巻第3号、1999年
17. 戸嶋朗・八村広三郎：絵画からの画面構成の抽出と検索への応用、情報処理学会論文誌、

第 40 巻第 3 号、1999 年

18. 長谷川隆・北原義典：マルチメディアに対する感性合成処理のコンセプトと実験システムの評価、情報処理学会論文誌、第 38 巻第 8 号、1997 年
19. 原田将治・伊東幸宏・中谷広正：感性語句を含む自然言語文による画像検索のための形状特徴空間の構築、情報処理学会論文誌、第 40 巻、1999 年
20. 尾畑貴信・萩原将文：感性を反映できるカラーポスター作成支援システム、情報処理学会論文誌、第 41 巻第 3 号、2000 年
21. 阿部友一・萩原将文：遺伝的アルゴリズムを用いた 2 次元画像からの物体認識、日本ファジィ学会誌、第 11 巻第 5 号、1999 年
22. 坂本崇・梶川嘉延・野村康雄：音楽感性空間における非線形判別分析を用いた曲印象別グループの分割、情報処理学会論文誌、第 40 巻第 4 号、1999 年
23. 白木渡ほか：ニューラルネットワークによるアーチ橋の景観評価システム、構造工学論文集、第 37 巻 A、687-697 ページ、1991 年
24. 大沢光：「感性工学」は“センスのよい”コンピューターを作れるだろうか、大沢光ほか（共著）：感性・人間・コンピューター（感性を分析する・理解する・取り入れる）、170-226 ページ（第 5 章）、1995 年 9 月、富士通経営研修所（富士通ブックス）
25. Mitsuru Osawa : “Method of Simultaneity Analysis” for Constructing “Model for Estimating Impression”, Kansei Engineering International, Vol.1 No.1, page 9-16, Nov.1999
26. 大澤光・善方日出夫：「同時関係分析法」の検証を考える！、第 3 回「印象の工学」ワークショップ（印象'98）予稿集、富士通株式会社・感性技術推進部、1998 年 6 月
27. 大澤光・王晋民：「同時関係分析法」による類義キーワードの抽出、情報処理学会第 51 回（1995 年後期）全国大会講演論文集、1995 年 9 月
28. 大澤光・王晋民：「同時関係分析法」による類義キーワードの抽出（その 2）、情報処理学会第 52 回（1996 年前期）全国大会講演論文集、1996 年 3 月
29. 大澤光：感性工学的な方法による「画像検索システム」の開発、日本心理学会第 59 回大会論文集、1995 年 10 月
30. Mitsuru Osawa : New Method of Discriminant Analysis by Utilizing Genetic Algorithm, Kansei Engineering International, Vol.1 No.1, page 1-8, Nov.1999
31. 杉山高一（著）：多変量データ解析入門、1983 年 4 月、朝倉書店
32. 田中豊・垂水共之（編）：（Windows 版）統計解析ハンドブック・多変量解析、1995 年 7 月、朝倉書店

33. J. H. Holland : Adaptation in Natural and Artificial Systems, University of Michigan Press, 1975
34. D. E. Goldberg : Genetic Algorithm In Search, Optimization and Machine Learning, Addison-Wesley, 1989
35. 北野宏明 (編) : 遺伝的アルゴリズム、1993 年 6 月、産業図書
36. S. Bagchi, S. Uckun, Y. Miyabe and K. Kawamura : Exploring Problem Specific Recombination Operators for Job Shop Scheduling, Proc. of the Fourth International Conference on Genetic Algorithm, 1991
37. T. Yamada and R. Nakano : A Genetic Algorithm Applicable to Large Scale Job-Shop Problems, Parallel Problem Solving form Nature 2, Elsevier Science Publishers, 1992
38. 山田武士・中野良平 : 遺伝的局所探索法によるジョブショップスケジューリング問題の解法、情報処理学会論文集、第 38 巻第 6 号、1997 年
39. 宇佐美芳明・鶴沼宗利・好永俊昭 : プラント建設のための仮設足場の配置計画手法、情報処理学会論文集、第 38 巻第 7 号、1997 年
40. 仙石浩明・吉原郁夫 : G A によるヒューリスティック探索の最適化、情報処理学会論文集、第 37 巻第 10 号、1996 年
41. J. P. Cophoon and W. D. Paris : Genetic Placement, Proc. of the IEEE International Conference on Computer -Aided Design, 1986
42. K. Shahookar and P. Mazumder : A genetic approach to standard cell placement using meta genetic parameter optimization, IEEE Transactions on Computetr-Aided Design, vol.9. 1990
43. B. A. Myers, B. V. Zanden and R. B. Dannenberg : Creating Graphical Interactive Application Objects by Deminstration, Proc. of the ACM SIGGRAPH Symposium on User Interface Software and Technology (UIST'89), ACM Press, 1989
44. D. R. Olsen and K. Allan : Creating Interactive techniques by Symbolically Solving Geometric Constraints, Proc. of the ACM SIGGRAPH Symposium on User Interface Software and Technology (UIST'90), ACM Press, 1990
45. B. T. V. Zanden : Constraint Grammers (A New Model For Specifying Graphical Applications), Proc. of the ACM Conference on Human Factors In Computing Systems (CHI'89), Addison-Wesley, 1989
46. Alastair Aitkin : Identification of Protein Consensus Sequences, Ellis Horwood Series In Biochemistry and Biotechnology, 1990
47. A. Konagaya and H. Kondo : Stochastic Motif Extraction using a Genetic Algorithm with the MDL

Principle, Hawaii International Symposium of Computer Systems, 1993

48. 戸谷智之・石川幹人：マルチ個体群の並列遺伝的アルゴリズムを用いたタンパク質の配列解析、情報処理学会論文集、第36巻第11号、1995年
49. 矢田哲士・石川幹人・田中俊・浅井潔：隠れマルコフモデルと遺伝的アルゴリズムによるDNA配列のシグナルパターン抽出、情報処理学会論文集、第37巻第06号、1996年
50. 武田和時：遺伝的アルゴリズムを用いた光アクセスネットワークの設計法、情報処理学会論文集、第38巻第12号、1997年
51. 棟朝雅晴・高井昌彰・佐藤義治：遺伝的アルゴリズムによる負荷分散機構を有する適応型ルーティング、情報処理学会論文集、第39巻第02号、1998年
52. 棟朝雅晴・高井昌彰・佐藤義治：確率学習による適応度評価を導入した遺伝的アルゴリズムに基づく動的負荷均衡、情報処理学会論文集、第36巻第04号、1995年
53. S. Geman and D. Geman : Stochastic relaxation (Gibbs distribution and the Bayesian Restriction of Images), IEEE Transaction, PAMI-6, 1984
54. K. Takatsu, S. Watanabe, H. Sawai and M. Yoneyama : A Proposal of Image Restoration Using Genetic Algorithms, Proc. of International Joint Conference on Neural Networks, IEEE & INNS, vli.1, pp.642-647, Nov. 1992
55. 金偉其・桂井浩：画素相関性を考慮した遺伝的アルゴリズムによる白黒濃淡画像の復元、情報処理学会論文集、第39巻第4号、1998年
56. 藤原義久・澤井秀文：進化的計算による3次元曲面の最適メッシュ生成、情報処理学会論文集、第40巻第5号、1999年
57. 吉本富士市・原田利宣・森山真光・吉本芳英：実数を遺伝子とした遺伝的アルゴリズムによるデータあてはめ、情報処理学会論文集、第41巻第1号、2000年
58. 吉本富士市・森山真光：スプライン関数を用いたデータあてはめ（遺伝的アルゴリズムによる節点の自動的な決定）、情報処理学会論文集、第39巻第9号、1997年
59. 山田雅之・伊藤裕一郎・世木博久・伊藤英則：編物デザインを支援するための編目模様生成システムの作成、情報処理学会論文集、第36巻第11号、1995年
60. J. David Shaffer : Multiple objective optimization with vector evaluated genetic algorithms, Proc. of the First International Conference on Genetic Algorithms, 1985
61. Carlos M. Fonseca and Peter J. Fleming : Genetic algorithms for multiobjective optimization (Formulation, discussion and generalization), Proc. of the Fifth International Conference on Genetic Algorithms, 1993

62. 石淵久生・村田忠彦：多目的フローショップ・スケジューリング問題のための遺伝的局所探索法、日本経営工学会誌、第48巻第6号、1998年
63. 村川宏・吉澤修治：近傍モデル遺伝的アルゴリズムによる多目的最適化、情報処理学会論文集、第40巻第04号、1999年
64. 越前谷博・荒木健治・桃内佳雄・柄内香次：実例に基づく帰納的学習による機械翻訳手法における遺伝的アルゴリズムの適用とその有効性、情報処理学会論文集、第37巻第08号、1996年
65. J. H. Holland and J. Reitman : Cognitive systems based on adaptive algorithms, D. Waterman and F. Hayes-Roth (eds.) : Pattern directed Inference systems, 313-329, Academic Press, 1978
66. L. B. Booker, D. E. Goldberg and J. H. Holland : Classifier systems and genetic algorithms, Artificial Intelligence, 40, 1989
67. S. F. Smith : A Learning System Based on Genetic Adaptive Algorithms, Ph.D. Thesis, University of Pittsburgh, 1980
68. 加藤常員・小沢一雅：遺伝的アルゴリズムを用いた非階層的クラスタリング、情報処理学会論文集、第37巻第11号、1996年
69. 川本芳久・松田秀雄・橋本昭洋：遺伝的アルゴリズムによる分子系統樹の作成、情報処理学会論文集、第37巻第06号、1996年
70. 丸山勉・小長谷明彦・小西弘一：遺伝的アルゴリズムを用いた並列グラフ分割アルゴリズム、情報処理学会論文集、第34巻第04号、1993年
71. R. Tanese : Distributed genetic Algorithm, Proc. of the Third International Conference on Genetic Algorithm, 1989
72. J. P. Cohoon, S. U. Hegde, W. N. Martin and D. Richards : Punctuated equilibria (A parallel genetic algorithm), Proc. of the Second International Conference on Genetic Algorithms, 1987
73. P. Spiessens and B. Manderick : A massively parallel genetic algorithm (Implementation and first analysis), Proc. of the Fourth International Conference on Genetic Algorithms, 1991
74. J. Stender (Ed.) : Parallel genetic algorithms, IOS Publishing at Amsterdam, 1993
75. D. Goldberg, B. Korb and K. Deb : Messy Genetic Algorithm (Motivation, Analysis and First Results), TCGA Report 89003, Department of Mechanics, The University of Alabama, 1989
76. D. Powell, S. Tong and M. Stolnick : EnGENEous (Domain independent, machine learning for design optimization), Proc. of the Third International Conference on Genetic Algorithm, 1989
77. Darrel Whitley : GENITOR (a different genetic algorithm), Proc. of the Rocky Mountain

Conference on Artificial Intelligence, 1988

78. H. Kitano : Designing Neural Networks using Genetic Algorithms with Graph Generation System, Complex System, vol.4 No.4, 1990
79. D. Ackley and M. Littman : A case for Lamarckian evolution, in C. Langton (Ed.) : Artificial life III, Addison-Wesley, 1994
80. J. J. Grefenstette : Lamarckian learning in multi-agent environments, Proc. of the Fourth International Conference on Genetic Algorithms, 1991
81. 吉井伸一郎・鈴木恵一・嘉数侑昇 : ラマルク進化を導入した L L T - G A の構築、情報処理学会論文集、第 36 巻第 8 号、1995 年
82. 中谷直司・金杉昭徳・近藤邦雄 : ウイルス進化論に基づく進化型アルゴリズム、情報処理学会論文集、第 40 巻第 05 号、1999 年
83. J. R. Koza : Genetic Programming (On the programming of computers by means of natural selection), MIT Press, 1992
84. J. R. Koza : Genetic Programming II (Automatic discovery of reusable programs), MIT Press, 1994
85. R.A.Fisher : The Use of Multiple Measurements in Taxonomic Problems, Annals of Eugenics 7, Part II, page 179-188, 1936
86. 平野広美 (著) : C でつくるニューラルネットワーク、1991 年 3 月、パーソナルメディア
87. N. Otsu, and T. Kurita : A new scheme for practical, flexible and intelligent vision systems, Proceeding of IAPR Workshop on Computer Vision, page 431-435, 1988.
88. B.Efron : The Jackknife, the Bootstrap and other Resampling Plans, CBMS 38, S.I.A.M., 1982
89. B.Efron and R.J.Tibshirani : An Introduction to the Bootstrap, Chapman & Hall, 1993
90. D. A. Freedman : A note on screening regression equations, The American Statistician. 37, 152-155, 1983
91. ケンドール (著) : 多変量解析の基礎、1972 年 10 月、サイエンス社
92. Mitsuru Osawa : Discriminant Analysis of Logical Relations, Proc. of the 2000 Spring Conference of KOSES & International Sensibility Ergonomics Symposium, page 157-162, June 2000, Korean Society for Emotion and Sensibility
93. 武藤真介 (著) : 初等多変量解析、2 ページ、1999 年 2 月、朝倉書店

謝 辞

筆者は、1991年に、モノ（工業製品や生活用品）のデザインなどに関連して「印象」を対象にした研究を開始し、これを「印象を“工学的に”取り扱おうとする意図」、より具体的には「『印象を扱う仕事』のための工学」の研究を、「印象の工学」という名称で提唱し、この意図と研究に興味と関心を寄せていただいた多くの研究者や実務家の方々にご協力やご支援やご指導をいただきながら、研究を進めてきました。

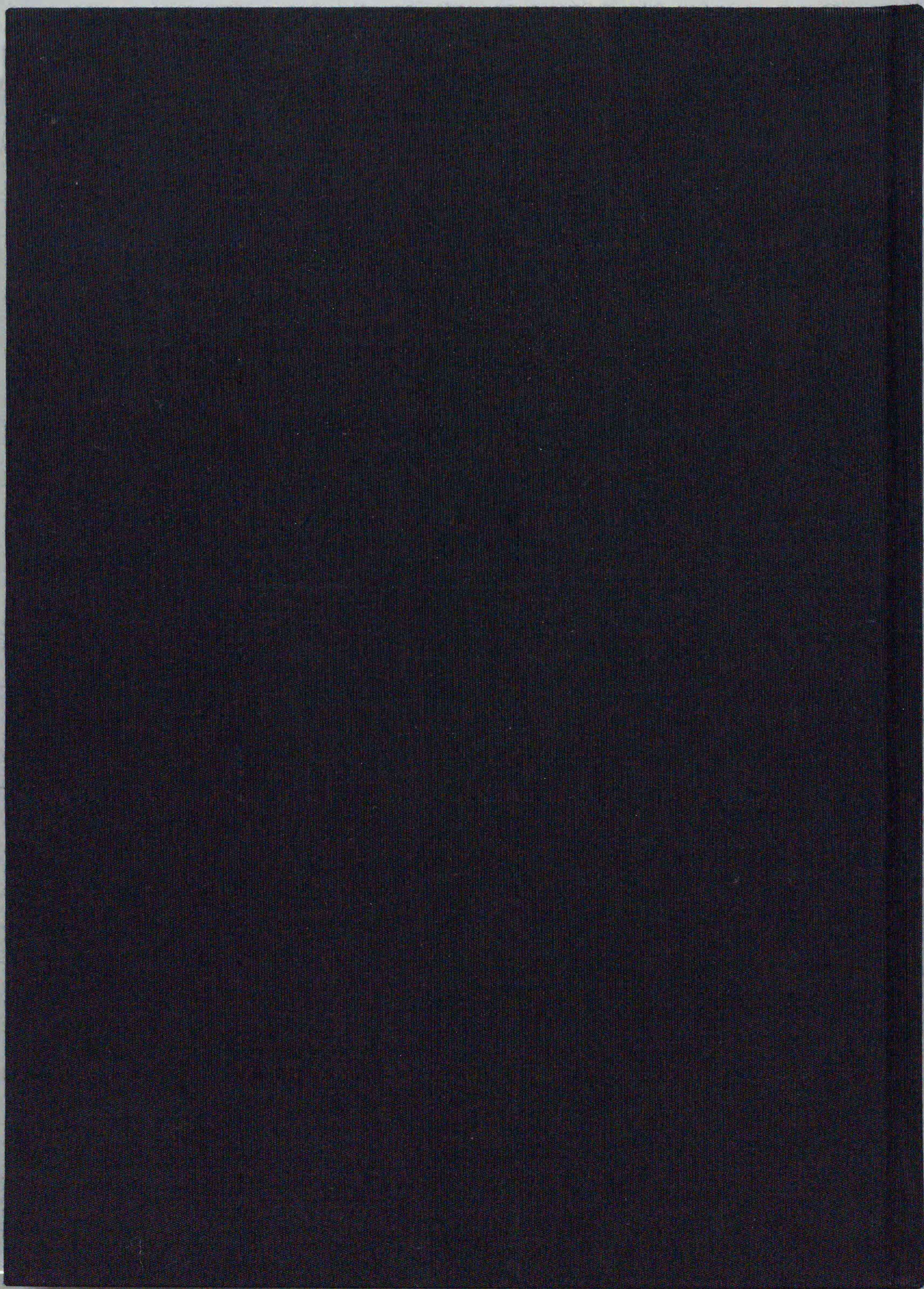
このなかでも、北海道大学大学院文学研究科心理システム科学講座・教授の西川泰夫先生には、「印象の工学」の研究全般にわたるご指導をいただき、また、度々にわたって有益なご示唆とご助言をいただき、いろいろな機会に「印象の工学」の研究成果の発表と議論の場を与えていただきました。この論文のまとめと内容の改善についても、怠惰な筆者に対して本当に辛抱強くまたきめ細かくご指導をいただきました。ここに記して、心からの感謝をいたします。

また、北海道大学大学院文学研究科心理システム科学講座・教授の阿部純一先生と行動システム科学講座・教授の天津起夫先生には、この論文の内容の改善についてのご指導をいただきました。ここに記して、心からの感謝をいたします。

千葉大学・名誉教授の鈴木邁先生と東京工芸大学大学院・教授の森典彦先生には、いろいろな機会に「印象の工学」の研究成果の発表と議論の場を与えていただき、また、度々にわたって貴重なご示唆をいただきました。さらに、1996年から1999年の間に、筆者が主催した4回の「印象の工学」ワークショップに“直接に”ご参加いただき、あるいは、「印象の工学」ワークショップのホームページとインターネットを通じて、また、筆者の「印象の工学」に関する著作や講演などを通じてご連絡をいただきました、200名を超える研究者と実務家のみなさんには、貴重なご意見と議論の機会をいただきました。これらのすべての方々に、心から感謝いたします。

筆者が勤務している富士通株式会社には、「印象の工学」という新しい課題の研究に、時間と予算および研究スタッフと場を提供していただきました。このご好意に対しまして、心からの感謝をいたします。

2001年1月 大 澤 光



inches 1 2 3 4 5 6 7 8
cm 1 2 3 4 5 6 7 8 9 10 11 12 13 14 15 16 17 18 19

Kodak Color Control Patches

© Kodak, 2007 TM: Kodak

Blue	Cyan	Green	Yellow	Red	Magenta	White	3/Color	Black

Kodak Gray Scale



© Kodak, 2007 TM: Kodak

A 1 2 3 4 5 6 M 8 9 10 11 12 13 14 15 B 17 18 19

