DISCRIMINANT ANALYSIS OF LOGICAL RELATIONS

Mitsuru Osawa

FUJITSU Limited

Abstract: Discriminant analysis is a method to relate whether the objects have a specific characteristic or not with their 'continuous' attribute values and, for given objects, to estimate whether they have a specific characteristic or not by their values of discriminant scores gotten from their attribute values. The author developed the new 'computational' method of discriminant analysis without specific hypotheses or assumptions and, by this new method, we can find 'feasible' solutions under the conditions required by our actual problems. In this paper, the author tried to apply this new method to the discrimination of logical relations. If this trial could be a success, we can apply this new method of discriminant analysis to the problems about relating the specific characteristic of the objects with their 'discrete' attribute values. The result was successful and the applicability of discriminant analysis could be expanded as a method for constructing the models for "estimating impressions".
Keywords: discriminant analysis, logical relation, genetic algorithm, estimation of impression

1. INTRODUCTION

Discriminant analysis is a method to relate whether the objects belong to a specific group with a certain characteristic or not with their 'continuous' attribute values. In other words, we can estimate whether the objects should belong to a specific group with a certain characteristic or not, by converting their attribute values to the single values through the discriminant function and by checking whether the single values are larger than a certain value i.e. the discriminant value. For the viewpoint of 'estimation of impressions', we can estimate whether we may feel a specific impression for the objects by their attribute values, by relating the impressions reported for the objects with their attribute values such as the characteristic values of their color images.

For discriminant analysis, the 'analytical' method has already been developed and widely been used. This 'analytical' method premises some kinds of hypotheses or assumptions, such as the objective functions of the analyses being differentiable. For the same problem, the author developed the new 'computational' method of discriminant analysis without specific hypotheses or assumptions by utilizing genetic algorithm[1] and, by this new method, we can get 'feasible' solutions under the conditions required by our actual problems to be solved. The author makes good use of this new method of discriminant analysis to construct the "models for estimating impressions".

In this paper, the author tried to apply this new method of discriminant analysis to the discrimination of logical relations. Logical variables cannot take 'continuous' values but 'discrete' values and it is characteristic of logical relations to contain complex interactions. If this trial could be a success, we can relate whether we may feel the specific impressions for given objects or not with their 'discrete' attribute values, such as their certain specifications or class categories.

In this connection, the discriminative errors of logical relations cannot be permissible and therefore the objective of discrimination analysis should not be maximization of 'correlation ratio' but clearing of 'discriminative errors' and maximization of difference between minimum discriminant score of 'true'(1) and maximum discriminant score of 'false'(0).

And logical relations are not linear, and therefore the discriminant function should take the forms of not only a linear connection of variables (*form (1)*) but also its absolute value (*form (2)*), which is easy to deal with non-linearity. For more complex logical relations, "two phase method of discriminant analysis", which is powerful to cope with interaction of variables, could be applied. [a],[b],[c],[d],[z],[u] and [v] represent logical variables, [f] and [g] are discriminant functions or scores, and [w_1], [w_2] and [w_3] are parameters of discriminant function or weights of variables to be determined in the following explanation.

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

$$(2)\ f = \pm\left|w_1 a + w_2 b + w_3\right| \quad (-1 \leq w_1, w_2, w_3 \leq +1)$$

2. DISCRIMINANT ANALYSIS OF FUNDAMENTAL LOGICAL RELATIONS

Firstly, "logical addition" i.e. "OR" was analyzed. The variable [u] is a 'logical sum' of two variable [a] and [b]. In this case the discriminant function [f] took the form of a linear connection and each parameter of the discriminant function i.e. each weight of the variables was regarded as 'gene' and the array of the parameters as 'chromosome'. This problem was analyzed by the new 'computational' method of discriminant analysis by utilizing genetic algorithm.

To explain in details, twenty chromosomes were prepared in a generation. Two 'best' chromosomes in a generation were survived to the next generation and, at the same time, two chromosomes randomly selected from the rest of chromosomes were also survived. Five pairs of chromosomes were randomly selected from all chromosomes and ten children chromosomes were generated by "random crossover". A hundred percents of 'genes' of parent chromosomes were randomly crossed. And six chromosomes were randomly selected from all chromosomes and one 'gene' selected randomly in each chromosome was mutated i.e. replaced by a random value. After all chromosomes in the next generation were generated, the chromosomes were improved by the method of "sensitivity analysis" with five- percent changes of values of parameters.

The calculation was repeated and converged until the tenth generation. The result is *relation (3)* and *[f] and [u] in Table 1*, and when the discriminant score of the combination of [a] and [b] is larger than *0.5*, the combinations is discriminated as 'true'(1), or else it is discriminated as 'false'(0).

$$(3)\ \text{if}\ f = a + b \geq 0.5 \quad \text{then}\ \hat{u} = 1\ \text{else}\ \hat{u} = 0$$

***Table 1**: Truth-values and discriminant-scores of "logical addition" and "logical multiplication" of two variables*

a	b	$f : f_0$	u	$g : g_0$	v
0	0	0<0.5	0	0<1.5	0
0	1	1>0.5	1	1<1.5	0
1	0	1>0.5	1	1<1.5	0
1	1	2>0.5	1	2>1.5	1

In the same way, the discrimination of "logical addition" of [a], [b] and [c] was analyzed and the result is *relation (4)* and *[f] and [u] in Table 2.*

$$(4)\ if\ \ f = a + b + c \geq 0.5 \quad then\ \ \hat{u} = 1\ \ else\ \ \hat{u} = 0$$

***Table 2**: Truth-values and discriminant-scores of "logical addition" and "logical multiplication" of three variables*

a	*b*	*c*	*f : f₀*	*u*	*g : g₀*	*v*
0	0	0	0<0.5	0	0<2.5	0
0	0	1	1>0.5	1	1<2.5	0
0	1	0	1>0.5	1	1<2.5	0
0	1	1	2>0.5	1	2<2.5	0
1	0	0	1>0.5	1	1<2.5	0
1	0	1	2>0.5	1	2<2.5	0
1	1	0	2>0.5	1	2<2.5	0
1	1	1	3>0.5	1	3>2.5	1

Secondly, "logical multiplication" i.e. "AND" was analyzed. The variable [v] is a 'logical product' of [a] and [b]. The discriminant function [g] also took the form of a linear connection and the new method of discriminant analysis was also applied. The method, the condition and the process of the calculation were almost as same as that of "logical addition". The result is *relation (5)* and *[g] and [v] in Table 1*.

$$(5)\ if\ \ g = a + b \geq 1.5 \quad then\ \ \hat{v} = 1\ \ else\ \ \hat{v} = 0$$

Similarly, the discrimination of "logical multiplication" of [a], [b] and [c] was analyzed and the result is *relation (4)* and *[g] and [v] in Table 2*.

$$(6)\ if\ \ g = a + b + c \geq 2.5 \quad then\ \ \hat{v} = 1\ \ else\ \ \hat{v} = 0$$

By the way, the result of the discrimination of "logical negation" i.e. "NOT" of variable [a] is *relation (7)* and *Table 3*.

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

***Table 3**: Truth-values and discriminant-scores of "logical negation"*

a	*f : f₀*	*z*
0	0>-0.5	1
1	-1<-0.5	0

3. DISCRIMINANT ANALYSIS OF "EXCLUSIVE LOGICAL ADDITION"

"Exclusive logical addition" i.e. "exclusive OR" of [a] and [b] is represented by means of *formula (8)*. And it contains complex interaction, and therefore it could not be discriminated when the discriminant function [f] took the form of a linear connection. But when the discriminant function [f] took the form of an absolute value of a linear connection, this relation was perfectly discriminated. The result is *relation (9)* and *[f] in Table 4*. The method and the condition of the calculation were as same as the cases of fundamental logical relations. The calculation was repeated and was convergent until the one hundredth generation. In the connection, *relation (10)* and *[g] in Table 4* is another solution of "exclusive logical addition".

$$(8)\ z = a \mathbin{\bar{\cup}} b = (\bar{a} \cap b) \cup (a \cap \bar{b})$$

$$(9)\ if\ \ f = -|a + b - 1| \geq -0.5 \quad then\ \ \hat{z} = 1\ \ else\ \ \hat{z} = 0$$

$$(10)\ if\ \ g = |a - b| \geq 0.5 \quad then\ \ \hat{z} = 1\ \ else\ \ \hat{z} = 0$$

***Table 4**: Truth-values and discriminant-scores of "exclusive logical addition" of two variables*

a	b	f : f0	g : g0	z
0	0	-1<-0.5	0<0.5	0
0	1	0>-0.5	1>0.5	1
1	0	0>-0.5	1>0.5	1
1	1	-1<-0.5	0<0.5	0

Incidentally the 'extended' definition of "exclusive logical addition" of [a],[b] and [c] was defined as the combination of [a],[b] and [c] is to be 'true'(1) when only one 'true'(1) variable exists in three. In this case, the result of the discrimination is *relation (11)* and *Table 5*.

$$(11)\ \ if\ \ f = -|a + b + c - 1| \geq -0.5\quad then\ \ \hat{z} = 1\ \ else\ \ \hat{z} = 0$$

***Table 5**: Truth-values and discriminant-scores of extended "exclusive logical addition" of three variables*

a	b	c	f : f0	z
0	0	0	-1<-0.5	0
0	0	1	0>-0.5	1
0	1	0	0>-0.5	1
0	1	1	-1<-0.5	0
1	0	0	0<-0.5	1
1	0	1	-1<-0.5	0
1	1	0	-1<-0.5	0
1	1	1	-1<-0.5	0

4. DISCRIMINANT ANALYSIS OF 'GIVEN' LOGICAL RELATION WITH THREE VARIABLES

The next problem is the discrimination of the logical relation [z] with *[a], [b] and [c], given by Table 6*. This relation is represented by *logical expression (12)*.

$$(12)\ \ z = (\bar{a} \cap \bar{b}) \cup (a \cap b) \cup (\bar{a} \cap b \cap \bar{c})$$

The method and the condition of the calculation were as same as the cases of fundamental logical relations. In this case the discriminant function took the form of an absolute value of a linear connection and *formula (13)* and *relation (14)* had gotten as the 'best' solution until the two hundredth generation. This result shows that when the discriminant score of the combination of variables is larger than *0.333*, the combination is discriminated as 'true'(1), or else it is discriminated as 'false'(0).

$$(13)\ \ f = |0.667a + 1.000b - 0.334c - 0.500|$$

$$(14)\ \ if\ \ f \geq 0.333\ \ then\ \ \hat{z} = 1\ \ else\ \ \hat{z} = 0$$

***Table 6**: Truth-values and discriminant-scores of 'given' logical relation with three variables*

a	b	c	f : f0	z
0	0	0	0.500>0.333	1
0	0	1	0.834>0.333	1
0	1	0	0.500>0.333	1
0	1	1	0.166<0.333	0
1	0	0	0.167<0.333	0
1	0	1	0.167<0.333	0
1	1	0	0.167>0.333	1
1	1	1	0.833>0.333	1

Needless to say, the 'better' solution might be gotten by the continuation of the calculation, but the author considers that the gotten solution is sufficiently feasible and practical.

5. DISCRIMINANT ANALYSIS OF 'GIVEN' LOGICAL RELATION WITH FOUR VARIABLES

The next problem to be solved is the discrimination of the logical relation [z] with *[a],[b],[c] and [d], given by Table 7*. This relation is represented by *logical expression (15)*. This problem was too complex to solve by the 'simple' method of discriminant analysis and therefore "two phases method of discriminant analysis", which is powerful to cope with complexity of the problems, was applied.

$$(15)\ z = (\bar{a} \cap c \cap d) \cup (a \cap b \cap c) \cup (\bar{a} \cap b \cap \bar{c} \cap \bar{d}) \cup (a \cap \bar{b} \cap \overline{(c \cap d)})$$

The method and the condition of the calculation were almost as same as the cases of fundamental logical relations. In the 'first' phase, the discriminant function [f] took the form of an absolute value of a linear connection of variables and this problem could not be perfectly discriminated until the three thousandth generation. The 'best' discriminant function was *formula (16),* and by this solution the combination of variables is discriminated as 'false'(0) when the discriminant score gotten by this function is smaller than *0.4160* i.e. the minimum value of scores of combinations in 'true'(1) and the combination is discriminated as 'true'(1) when the discriminant score is larger than *0.8342* i.e. the maximum value of scores of combinations in 'false'(0). Seven combinations of variables in sixteen could be discriminated but nine combinations could not discriminated. The discriminant errors might occur when their scores are between *0.4160* and *0.8342.*

$$(16)\ f = \left| 0.4176a - 0.4172b - 0.4182c - 0.8322d + 0.8337 \right|$$

$$(17)\ g = -0.4098a + 0.4098b + 0.8197c + 0.8107d$$

$$(18)\ \text{if } f \geq 0.8342 \text{ then } \hat{z} = 1 \text{ else if } f \leq 0.4160 \text{ then } \hat{z} = 0 \text{ else if } g \geq 0.2005 \text{ then } \hat{z} = 1 \text{ else } \hat{z} = 0$$

***Table** 7: Truth-values and discriminant-scores of 'given' logical relation with four variables*

a b c d	$f : f_0$	$g : g_0$	z
0 0 0 0	0.8337>0.4160	0.0000<0.2005	0
0 0 0 1	0.0015<0.4160	============	0
0 0 1 0	0.4156<0.4160	============	0
0 0 1 1	0.4166<0.8342	1.6304>0.2005	1
0 1 0 0	0.4166<0.8342	0.4098>0.2005	1
0 1 0 1	0.4157<0.4160	============	0
0 1 1 0	0.0016<0.4160	============	0
0 1 1 1	0.8338<0.8342	2.0402>0.2005	1
1 0 0 0	1.2513>0.8342	============	1
1 0 0 1	0.4191<0.8342	0.4009>0.2005	1
1 0 1 0	0.8332<0.8342	0.4098>0.2005	1
1 0 1 1	0.0010<0.4160	============	0
1 1 0 0	0.8342>0.4160	0.0000<0.2005	0
1 1 0 1	0.0020<0.4160	============	0
1 1 1 0	0.4160<0.8342	0.8197>0.2005	1
1 1 1 1	0.4162<0.8342	1.6304>0.2005	1

In the 'second' phase, nine combinations, which could not be perfectly discriminated in the 'first' phase, should be discriminated. In this phase the discriminant function [g] took the form of a linear connection. The calculation was convergent until the three thousandth generation and nine combinations were perfectly discriminated. The 'best' discriminant function was *formula (17)* and the combination is discriminated as 'true'(1) when its score gotten by the 'second' discriminant function is larger than *0.2005*, or else the combination is as 'false'(0).

Needless to say, the 'better' solution might be gotten by the continuation of the calculation, but the author considers that this solution is sufficiently feasible and practical.

6. CONCLUSION

In this paper the author illustrated that the new 'computational' method of discriminant analysis is applicable to the discrimination of logical relations and demonstrated the discrimination of the fundamental and complex logical relations.

Through this research, the author believes that the applicability of the new 'computational' method of discriminant analysis could be expanded as a method for constructing the "model for estimating impression", which relates the impressions reported for the objects with not only their 'continuous' attribute values but also their 'discrete' attribute values, such as certain specifications or classification categories.

References

[1] Mitsuru Osawa: NEW METHOD OF DISCRIMINANT ANALYSIS BY UTILIZING GENETIC ALGORITHM, KANSEI Engineering International, No.1 Vol.1, page 1-8, Nov. 1999