コレスポンデンス分析を用いた一対比較法および配偶法の解析に関する考察と提案

内 田　 治*

本研究は度数データを統計的に解析する方法論の一つであるコレスポンデンス分析に関するものである。コレスポンデンス分析はアンケート調査における度数データの解析に有用な方法であるが、この手法を官能検査によって得られたデータに適用することで、従来の官能検査手法よりも詳細な知見を得ることができることを数値例によって報告する。

本研究で取り上げた官能検査手法は、一対比較法と配偶法である。一対比較法はパネルの好みを調査する嗜好型官能検査の手法で、評価対象を１次元尺度上に順位づけすることを目的としている。本研究では、一対比較法によって得られたデータにコレスポンデンス分析を適用することで、１次元では表すことができない順位づけを多次元で表現できることを報告する。

一方、配偶法はパネルの識別力を調査する分析型官能検査で、従来はパネルに識別能力があるかないかの判定にのみ使われていた。本研究では、配偶法によって得られたデータから類似度行列を作成し、その行列をコレスポンデンス分析で解析することにより、評価対象間の識別難易度を視覚的に表現することが可能になることを報告する。

キーワード：コレスポンデンス分析，官能検査，一対比較法，配偶法，類似度行列

Consideration and Proposal Concerning the Analysis for Paired Comparison and Matching Test by Correspondence Analysis

Osamu UCHIDA

This research is concerned with “correspondence analysis” that is one of the methodologies that statistically analyze the frequency data. Correspondence analysis is a useful method for the analysis of the frequency data in the questionnaire survey. In this research, it was found that more detailed finding can be obtained by applying this technique to acquired data by the sensory test than a past sensory inspection technique with the numerical example.

The sensory test techniques which are taken up in this research are “paired comparison” and “matching test”. Paired comparison aims at the evaluation object in the technique of the preference sensory test that investigates the favor of the panel and it has aimed at order in one dimension. In this research, it was found that by the application of the correspondence analysis to acquired data by paired comparison, the order that cannot be shown by one dimension is multi-dimensional expressible.

On the other hand, matching test is the analytical sensory test that investigates the

*東京情報大学総合情報学部環境情報学科
Tokyo University of Information Sciences, Faculty of Informatics, Department of Environmental Information
2007年７月５日受理

panel discrimination, and is used only to judge whether the panel have or do not have the discrimination ability. In this research, it was found that applying correspondence analysis to the similarity matrix made from collected data by matching test can make a sight expression of the discrimination difficulty between the objects.

Keyword：correspondence analysis, sensory test, paired comparison, matching test, similarity matrix

1．緒言

コレスポンデンス分析は2元表の形に整理されたデータに対して適用可能な統計手法であり、特に、クロス集計表、01型データ表、アイテム・カテゴリ型データ表の解析に用いられている。一方、官能検査の分野における代表的な試験方法として、一対比較法と配偶法がある。この２つの試験法は、目的も試験方法もまったく異なるものであるが、どちらにも共通しているのが、識別可能性という問題である。一対比較法における回答の矛盾や一貫性の欠如、配偶法における誤回答は、評価者の識別力の低さに起因するものであるが、同時に、差が顕著でない、似た試料を提示されたという場合にも、回答の矛盾や誤回答は生じやすくなる。コレスポンデンス分析のねらいの１つは、似た対象（人や物）同士は近くに、似ていない対象同士は遠くに位置する布置図を作成することにある。したがって、一対比較法や配偶法によって得られたデータにコレスポンデンス分析を適用することで、回答の矛盾や誤回答の様子を試料間の近さという観点から考察することが可能になると考える。このことを具体的な数値例を使って示すことが本研究の目的である。

2．一対比較法における一意性とコレスポンデンス分析

2-1　一意性の係数

官能評価によって比較したい試料が3つ以上あるときに、試料を2個ずつ比較する方法を一対比較法という。一対比較法には、優劣だけを問題にする方法と、優劣および優劣の差を問題にする方法がある。優劣だけを問題にする一対比較法としては、サーストンの方法とブラッドレイの方法があり、優劣の差までも問題にする方法としては、シェッフェの方法がある。シェッフェの方法では、さらに、芳賀の変法、浦の変法、中屋の変法が提唱されている。本研究で取り上げる一対比較法は、優劣だけを問題にする場合である。

いま、A、B、Cという３つの試料があって、AとBの比較ではAのほうが良い、BとCの比較ではBのほうが良いと判定したのであれば、AとCの比較ではAのほうが良いと判定されると予測するであろう。なぜならば、その判定には矛盾がないからである。しかし、実際の官能評価の場面では、AとCの比較において、Cのほうが良いと判定されることがある。このような状態を一巡三角形（circular triad）になっているという言い方をする。[1]

一対比較法では、結果を矢線で表すことが行われ、矢線が出ている試料と矢線を受けている試料では、受けている試料が出ている試料よりも優れていることを意味するように矢線が描かれる。

図1-1の三角形は矛盾のない判定結果であることを示しており、1位がC、2位がB、3位がAと順位をつけることができる。一方、図1-2の

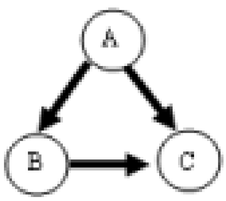


図1-1　一意性のある三角形

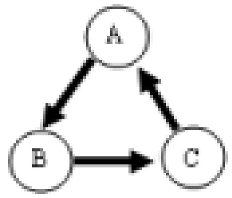


図1-2　一巡三角形

三角形では、順位を決めることができない。このような形の三角形が一巡三角形である。一巡三角形になってしまう原因には、評価者の識別能力のなさ、試料間の差異のなさ、評価の多次元構造が考えられる。したがって、一巡三角形が少なければ、評価者は識別する能力を持っている、試料間には差異がある、評価が一元的あると考えることができる。一巡三角形の数を用いて、識別能力の有無を見る指標として、一意性の係数（coefficient of consistency）がある。

一巡三角形の数dは、試料の数をk、多角形の各頂点から外に向かって出ている矢線の数をa_iとするとき、つぎの式で計算できる。

$$\mathrm{d}=\frac{1}{6}k(k-1)(k-2)-\frac{1}{2}\sum_{i=1}^{k}a_i(a_i-1)$$

このdによって、一意性の係数ζは、つぎの式で計算される。

kが偶数のとき $\zeta=1-\dfrac{24d}{k^3-4k}$ 、

kが奇数のとき $\zeta=1-\dfrac{24d}{k^3-k}$

一巡三角形が全くないとき、すなわち、d＝0のときには、$\zeta=1$となり、順位が一意に決まる。

2-2　データの形式

７つの試料（A、B、C、D、E、F、G）を一対比較することを想定すると、評価者はつぎに示すようなデータ表を作成することになる。

表1　一対比較法における原表

	A	B	C	D	E	F	G
A	×	1	0	1	1	1	1
B	0	×	1	1	1	1	1
C	1	0	×	1	1	1	1
D	0	0	0	×	1	1	1
E	0	0	0	0	×	1	1
F	0	0	0	0	0	×	1
G	0	0	0	0	0	0	×

行iと列jを比較したときに、iのほうが好ましいときには１、jのほうが好ましいときには０とつける。このようにすると行の合計は勝ち数を示しているので、この数の大きいものから順位をつければよい。

さて、このデータ表を分割表とみることにより、コレスポンデンス分析の適用を考える。一意性があれば、一次元構造を視覚的に摘出することができると考えるからである。

コレスポンデンス分析の適用に際しては、便宜的に対角線のところに任意の同じ数値を代入しておく必要がある。本研究では1と入力している。

2-3　コレスポンデンス分析の結果と一意性

データ表として、４つのパターンを想定して、コレスポンデンス分析を適用した。以下に結果を示す。なお、試料の数k＝7とする。

〈パターン1〉一巡三角形がない例

一巡三角形の数が0（一意性の係数ζ＝1）となるデータ表を以下に示す。

このデータ表にコレスポンデンス分析を適用する。なお、×のところには1を入力しておく。

コレスポンデンス分析によって得られる布置図（図1）を見ると、馬蹄形に試料が散布していることがわかる。1次元構造のときには、コレスポンデンス分析における布置図おいて、馬蹄形を示すことが知られており、馬蹄形状とな

表2　ζ＝1のデータ表

	A	B	C	D	E	F	G	勝ち数	順位
A	×	1	1	1	1	1	1	6	1
B	0	×	1	1	1	1	1	5	2
C	0	0	×	1	1	1	1	4	3
D	0	0	0	×	1	1	1	3	4
E	0	0	0	0	×	1	1	2	5
F	0	0	0	0	0	×	1	1	6
G	0	0	0	0	0	0	×	0	7

表3　ζ＝0.9286のデータ表

	A	B	C	D	E	F	G	勝ち数	順位
A	×	1	0	1	1	1	1	5	1
B	0	×	1	1	1	1	1	5	1
C	1	0	×	1	1	1	1	5	1
D	0	0	0	×	1	1	1	3	4
E	0	0	0	0	×	1	1	2	5
F	0	0	0	0	0	×	1	1	6
G	0	0	0	0	0	0	×	0	7

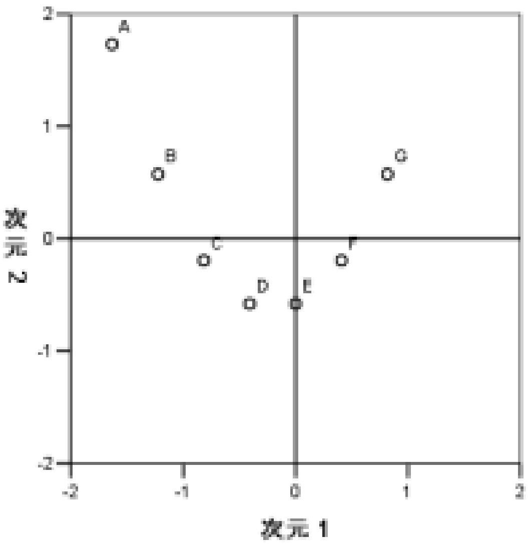


図2　コレスポンデンス分析による1次元と2次元の布置図（パターン1）

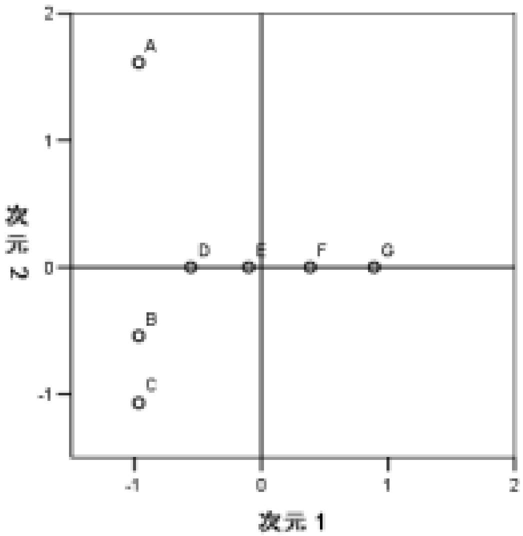


図3　コレスポンデンス分析による次元1と次元2の布置図（パターン2）

った場合には、一意性があると判断できる。[2]

〈パターン2〉一巡三角形が1つある例

一巡三角形の数が1（一意性の係数ζ＝0.9286）のときのデータ表を以下に示す。試料A、B、C間の評価が一巡三角形となっている。

コレスポンデンス分析によって得られる布置図（図3）を見ると、次元1においては、試料A、B、Cが重なっており、識別が難しいことが示されている。

〈パターン3〉一巡三角形が2つある例

一巡三角形の数が2（一意性の係数ζ＝0.8571）のときのデータ表を以下に示す。試料A、B、C間の評価が一巡三角形、D、E、F間の評価が一巡三角形となっている。

表4　ζ＝0.8571のデータ表

	A	B	C	D	E	F	G	勝ち数	順位
A	×	1	0	1	1	1	1	5	1
B	0	×	1	1	1	1	1	5	1
C	1	0	×	1	1	1	1	5	1
D	0	0	0	×	1	0	1	2	4
E	0	0	0	0	×	1	1	2	4
F	0	0	0	1	0	×	1	2	4
G	0	0	0	0	0	0	×	0	7

コレスポンデンス分析によって得られる布置図（図4）を見ると、次元1について、試料A、B、Cが重なっていて、さらに、次元1および次元2において、試料D、E、Fが重なっている。

このことから、3つずつの試料の識別が難しいことが示されている。

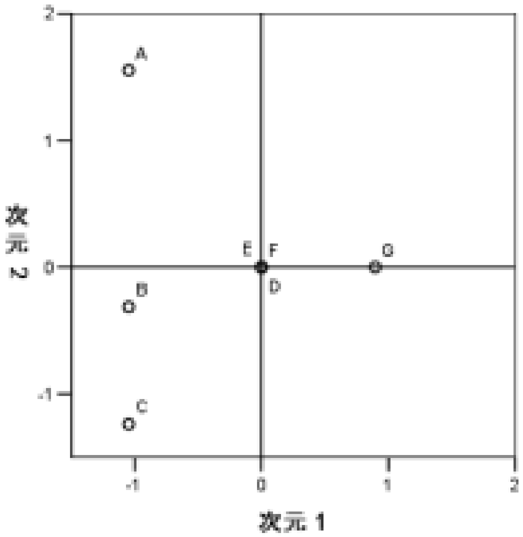


図4　コレスポンデンス分析による次元1と次元2の布置図（パターン3）

〈パターン4〉一意性が全くない例

一意性がなく（一意性の係数$\zeta=0$）、順位づけができないときのデータ表を以下に示す。

表5　$\zeta=0$のデータ表

	A	B	C	D	E	F	G	勝ち数
A	×	1	0	1	0	0	1	3
B	0	×	1	0	1	1	0	3
C	1	0	×	1	0	0	1	3
D	0	1	0	×	1	0	1	3
E	1	0	1	0	×	1	0	3
F	1	0	1	1	0	×	0	3
G	0	1	0	0	1	1	×	3

コレスポンデンス分析によって得られる布置図（図5）を見ると、円形に試料が散布していていることが見て取れる。全く識別能力がないときには、すべてが重なるのではなく、円形になるということに注意する必要がある。

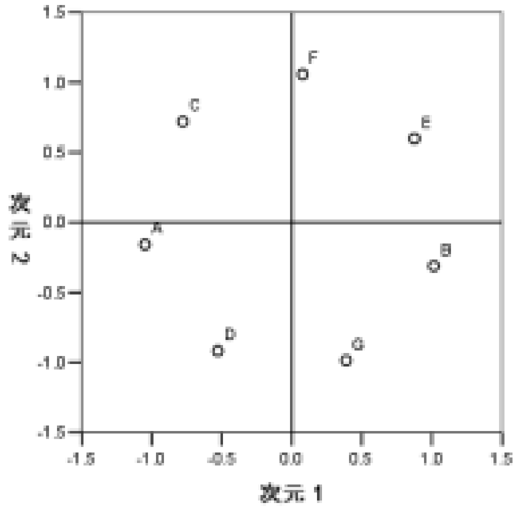


図5　コレスポンデンス分析による次元1と次元2の布置図（パターン4）

3．一対比較法における一致性とコレスポンデンス分析

3-1　一致性の係数

一対比較法における一意性の係数は1人の評価者に関する識別能力を考えるものであるが、n人が同じ一対比較の判断を行ったときに、n人の検査員全体の判定の一致性を評価するための指標が、一致性の係数（coefficient of agreement）である。[3]

一致性の係数uは以下の式により求めることができる。

$$u=2m/({}_nC_2\times{}_kC_2)-1 \quad \cdots\cdots(1)$$

ここに、mは2人ずつを組にしたときの判定の一致数、nは評価者の数、kは試料の数である。

一致性の係数uはn人の判定が完全に一致しているときには、1となり、n人の半分が一方を良いといい、残りの半分が他方を良いと判定したときには、$-1/(n-1)$か$-1/n$となる。

一致性があると判定されたときには、サーストンの一対比較法により試料を一次元の尺度上

に配置することができる。

3-2 データの形式

7つの試料（A、B、C、D、E、F、G）を40人の評価者が一対比較することを想定する。評価者一人一人の原データ表は表1のような形式となる。これを40人について集計した結果は以下のような形式となる。

表6 一対比較法における集計表

	A	B	C	D	E	F	G
A	×	30	35	20	35	30	35
B	10	×	25	20	30	35	30
C	5	15	×	15	25	30	30
D	20	20	25	×	30	35	25
E	5	10	15	10	×	25	30
F	10	5	10	5	15	×	15
G	5	10	10	15	10	25	×

サーストンの一対比較法はこの集計表が出発点となる。この表は、たとえば、AとBの比較において、40人中Aを良いとした者が30人、Bを良いとした者が10人いたことを意味している。

ここで、式（1）を用いて一致性の係数を計算するときに必要となるmの値は、以下のように求める。

$$\mathrm{m} = {}_{\mathrm{n}}\mathrm{C}_2 \times {}_{\mathrm{k}}\mathrm{C}_2 + \sum_{j>i} x_{ij}{}^2 - \mathrm{n}\sum_{j>i} x_{ij}$$

$\sum_{j>i}$ は表6の対角線の上半分の和をとることを意味する。

この集計表にコレスポンデンス分析を適用することを考える。[6][7][8]

ただし、コレスポンデンス分析を適用するには、このデータ表を、親近性（数値が大きいものほど親近性が近い）を示す表に変換する必要がある。このために以下のような変換方法を提案する。

① 優劣の人数差（絶対値）を求める。この値が小さいほど親近性が近く、大きいほど親近性が遠いと考えることにする。このような表は距離行列と呼ばれている。

表7 一対比較法における距離行列

	A	B	C	D	E	F	G
A	×	20	30	0	30	20	30
B	20	×	10	0	20	30	20
C	30	10	×	10	10	20	20
D	0	0	10	×	20	30	10
E	30	20	10	20	×	10	20
F	20	30	20	30	10	×	10
G	30	20	20	10	20	10	×

② 距離行列を数値が大きいほど親近性が近いことを意味する親近性行列に変換する。このためには、評価者の数40から各数値を引く。×のところは40とする。

表8 一対比較法における親近性行列

	A	B	C	D	E	F	G
A	40	20	10	40	10	20	10
B	20	40	30	40	20	10	20
C	10	30	40	30	30	20	20
D	40	40	30	40	20	10	30
E	10	20	30	20	40	30	20
F	20	10	20	10	30	40	30
G	10	20	20	30	20	30	40

このようにして変換した表8に対して、コレスポンデンス分析を適用する。

3-3 コレスポンデンス分析の結果と一致性

集計表として、2つのパターンを想定して、コレスポンデンス分析を適用した。

〈パターン1〉 1次元構造を示す例（一致性の係数u＝0.22）

コレスポンデンス分析の布置図とサーストンの一対比較法の布置図を以下に示す。

コレスポンデンス分析の布置図上では、各試料が馬蹄形状に散らばっており、1次元構造が示唆される。次元1における並び方は、サース

表9　一対比較法における集計表（パターン1）

	A	B	C	D	E	F	G
A	×	28	32	22	34	36	35
B	12	×	25	19	28	31	30
C	8	15	×	16	26	33	29
D	18	21	24	×	30	34	32
E	6	12	14	10	×	24	28
F	4	9	7	6	16	×	15
G	5	10	11	8	12	25	×

表10　一対比較法における集計表（パターン2）

	A	B	C	D	E	F	G
A	×	30	10	22	34	36	22
B	10	×	30	19	28	32	30
C	30	10	×	16	26	33	29
D	18	21	24	×	30	34	32
E	6	12	14	10	×	20	28
F	4	8	7	6	20	×	12
G	18	10	11	8	12	28	×

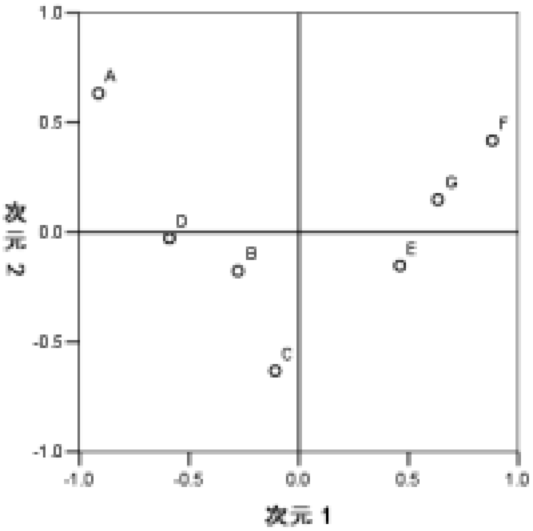


図6　コレスポンデンス分析による布置図（パターン1）

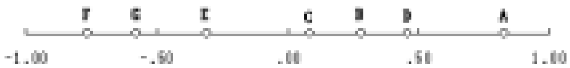


図7　サーストンの一対比較法による布置図（パターン1）

トンの一対比較法による布置図の並び方と正反対になっている。コレスポンデンス分析におけるスコアの符号は一意的なものではなく、相対的なものであるから、符号を逆にして見てもかまわない。したがって、コレスポンデンス分析の結果と、サーストンの一対比較法の結果は順序関係に関して一致していると考えられる。

〈パターン2〉 1次元構造を示さない例（一致性の係数u＝0.21）

コレスポンデンス分析の布置図上では、各試料が馬蹄形状に散らばっておらず、1次元構造は示唆されない。また、サーストンの一対比較法による布置図と比較しても符号の違いだけでなく、並び方に若干の差異が見られる。

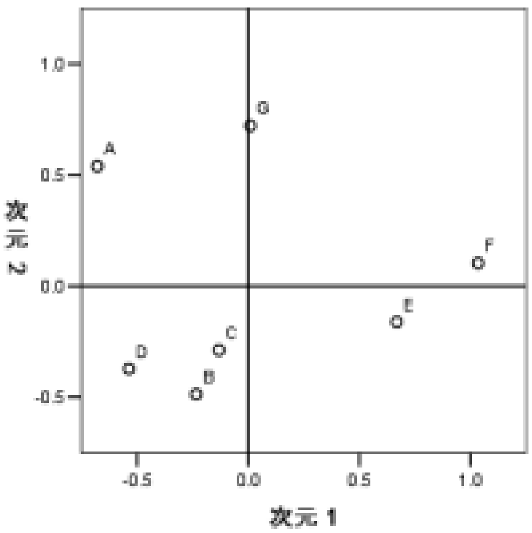


図8　コレスポンデンス分析による布置図（パターン2）

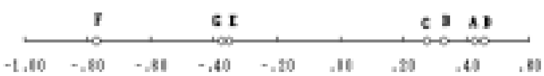


図9　サーストンの一対比較法による布置図（パターン2）

パターン1とパターン2を比較したとき、一致

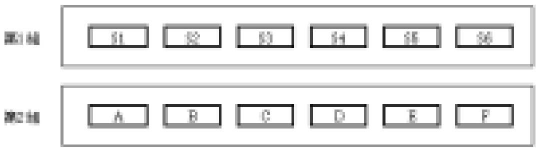
性の係数にほとんど差異が見られないことがわかる。このことから、一致性の係数の値によって、1次元尺度であることの妥当性を完全には判断できないということがわかる。この指標はあくまでも参考にとどめるべきであろう。

同時に、優劣だけを問題にするサーストン型の一対比較法において、コレスポンデンス分析も試料の差異を視覚化する方法として有効であることがわかる。

4. 配偶法とコレスポンデンス分析

4-1　配偶法のデータ

t種類の試料S1、S2、. . . 、Stの１個ずつから構成される組を2組作り、各組から1個ずつ取り出して、同じ試料同士のペアを作らせる試験方法を配偶法あるいはマッチングテストという。これは官能検査において、評価者の識別力を試験するときに用いられる方法である。



S1と同じ試料はAからFのどれか？

図10　配偶法のイメージ

tが4以上のときは、tに関係なく、正解数が4以上ならば、有意水準5%で識別力ありと判定される。

いま、6種類の試料S1、S2、S3、S4、S5、S6の１個ずつから構成される組を2組作り、20人の評価者に、各組から1個ずつ取り出して、同じ試料同士のペアを作らせる配偶法を実施すること考える。

2組の正しい対応は図11の通りとする。

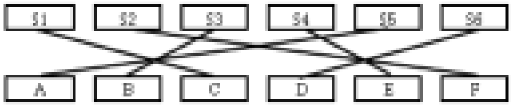


図11　正しい対応

さて、配偶法を20人に実施して、その回答結果が表11のようになったとしよう。

表11　配偶法の原データ

人	S1	S2	S3	S4	S5	S6
1	C	D	B	E	A	F
2	C	F	B	E	D	A
3	C	F	E	B	A	D
4	D	F	B	E	A	C
5	C	F	B	E	A	D
6	F	C	B	E	D	A
7	F	C	B	A	D	E
8	C	F	B	E	A	D
9	C	F	B	E	D	A
10	C	F	E	B	A	D
11	C	B	A	E	F	D
12	C	F	E	B	A	D
13	C	F	E	B	A	D
14	C	E	D	F	B	A
15	A	F	B	E	C	D
16	B	F	C	E	A	D
17	C	F	B	E	A	D
18	D	A	B	E	C	F
19	E	C	F	B	A	D
20	F	C	B	D	E	A

従来、配偶法の解析は誰が有意であるか（識別力があるか）ということだけを解析していた。筆者はこのデータに対して、コレスポンデンス分析を適用することで、より詳細な知見を得ることができるということを以下に示すことにする。

4-2　試料に着目した解析

まずは、表11の原データをもとに、表12に示す集計表を作成する。

S5は試料Aであり、後の解析で第1組と第2組の試料を区別するために、A’と’をつけて表現する。集計表からS5とAをペアにしたのが8

表12　配偶法の集計表

	S5	S3	S1	S6	S4	S2
	A’	B’	C’	D’	E’	F’
A	8	1	1	8	1	1
B	1	13	1	0	4	1
C	2	1	12	1	0	4
D	8	1	2	8	1	0
E	0	3	1	2	14	0
F	1	1	3	1	0	14

人いることがわかる。対角線上の人数は正しいペアを答えた人数（正解者数）を示している。集計表を見ると、AとDの区別が難しいという結果になっている。この表に対して、コレスポンデンス分析を適用すると、つぎに示す布置図が得られる。

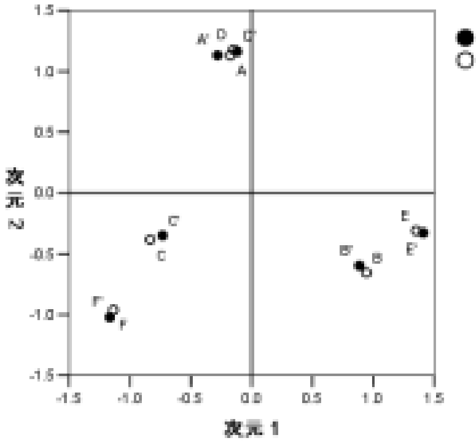


図12　配偶法における試料の布置図

AとDが近くに位置しており、この2つの試料の区別が難しいことがわかる。また、CとF、BとEが似ている試料であることが発見できる。

4-3　人（評価者）に着目した解析

表11の原データに対して、多重コレスポンデンス分析を適用する。この目的は、評価者同士の近さを視覚的に表現するためである。[9]

図13の布置図を見ると、11番と18番の評価者が離れたところに位置している。このような人の布置図は、性別や年齢といった評価者のプロフィールで層別すると新しい知見を得られる可能性がある。なお、全問正解した8番と17番は●としてある。この2人の位置から遠い人ほど、識別力が劣るという見方ができる。

ここで、図13上に正解数を記入したのが図14である。有意な人は原点の周辺に位置していることが発見できる。どのようなデータにおいても、常に有意な人が原点の周辺に位置するとは限らないが、多重コレスポンデンスにおいては、標準的な回答（多数派の回答）をした人が原点周辺に位置されることが多いという性質があるため、識別能力の全くない人ばかりを集めた場合や、識別不可能な資料を用意しない限り、正解数の多い人が原点の近くに位置することが予想される。[4] [5]

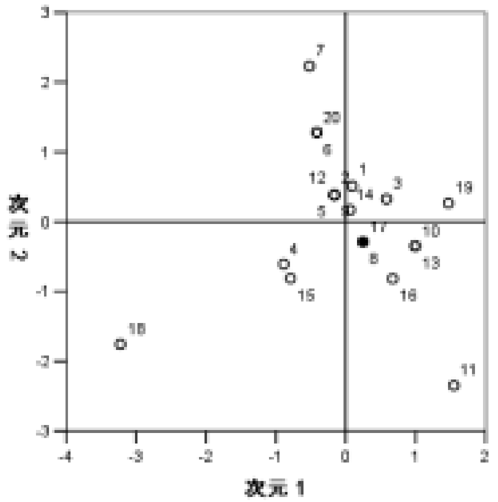


図13　評価者の布置図

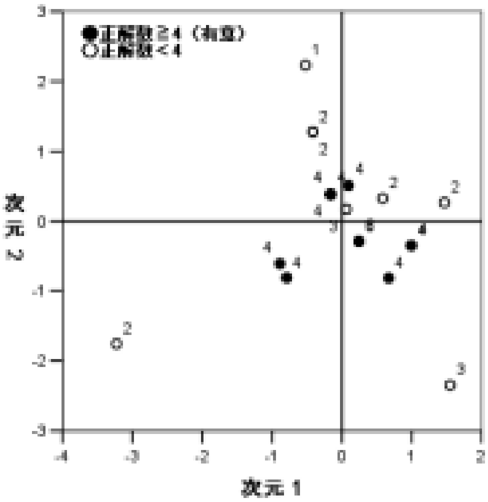


図14　正解数を記入した評価者の布置図

5．結語

コレスポンデンス分析は適用できるデータ表の範囲が広いため、アンケート調査のデータ解析には頻繁に用いられている。本研究では、コレスポンデンス分析を官能検査の分野の代表的な試験方法である一対比較法と配偶法によって得たデータの解析に活用することで、従来の有意差検定だけによる解析よりも詳細な解析結果を得ることができるということを示した。一対比較法においては、評価の一意性とコレスポンデンス分析の結果に対応があること、サーストンの方法と同様な結果が得られることが発見できた。同時に、従来から提唱されている一致性の係数の問題点も浮かびあがった。配偶法においては、従来は識別力があるかどうかだけの解析しか行われていなかったが、コレスポンデンス分析を適用することで、提示試料および評価者の詳細な解析が可能になるということを発見することができた。

参考文献

[1] 佐藤 信：官能検査入門, 日科技連出版社（1978）

[2] 大隅 昇, 馬場 康維, Alain Morineau, Ludovic Lebart, Kenneth M. Warwick：記述的多変量解析法, 日科技連出版社（1994）

[3] 日科技連官能検査委員会：新版　官能検査ハンドブック, 日科技連出版社（1978）

[4] 小林 龍一：数量化理論入門, 日科技連出版社（1981）

[5] 内田 治：すぐわかるSPSSによるアンケートのコレスポンデンス分析, 東京図書（2006）

[6] Michael Greenacre, Jorg Blasius：Correspondence Analysis in the Social Sciences, ACADEMIC PRESS（1994）

[7] Sten Erik Clausen：Applied Correspondence Analysis：An Introduction（Quantitative Applications in the Social Sciences）, SAGE Publications（1998）

[8] Susan C. Weller, A. Kimball Romney：Metric Scaling-Correspondence Analysis（Quantitative Applications in the Social Sciences）, SAGE Publications（1990）

[9] Jorg Blasius, Michael Greenacre：Visualization of Categorical Data, ACADEMIC PRESS（1998）