

数量化Ⅲ類のプロット・サブルーチンの手直し

社会学部 米 谷 淳

はじめに -いささかくどい多変量解析への招待

ファンシーという言葉がある。これは、最近の女性がよく買うキャラクター商品やギフトグッズの特徴であるそうだ。島村(1991)によればファンシーとは白くて小さくて丸くて柔らかい感じを与えるもので、マシュマロやイチゴミルクや人間や動物の赤ちゃんのようなものであるという。このようなイメージや印象をとらえるには単一の形容詞では不十分であり、それらは複数の形容詞のセットを用いてそれらのまとまりからかもしだされる、いわゆる、創発的な意味やコンセプトと考えられる。

最近、感性という言葉が心理学だけでなく工学や情報科学でもよく研究テーマに使われるようになってきている。感性をとらえるためには、物理量を測定するような方法だけにたよっているわけにはゆかず、様々な形容詞の対を対極に配したリッカート尺度のセットに表現されたプロフィールをもとにある事物のイメージや印象をとらえるsemantic differential法が多用されることになる。こうした手法では、質問紙の作成や調査結果の分析において、用いた質問項目を別個に扱うのではなく、多数の項目群をある基準に基づいて分類し、複数の項目または形容詞から構成されたグループになんらかの意味づけを行い、それらにラベルを付けて、分析対象を数個のグループに絞りこんでおいてから記述や他のグループとの関連性についての検討にうつることが有効であり、また要求されもする。

例えば、コンピュータに対するユーザーのイメージを調査することを考えてみよう。コンピュータに対して使える形容詞は「好き-嫌い」、「親しみやすい-親しみにくい」から、「賢い-鈍い」、さらに、「ドライ-ウェット」など、形容詞はいくらでもあるといつてよい。しかしながら、質問紙に使う項目はできるだけ少なくしたい。そこで、イメージ調査を妥当なものとするためには必要十分な数の適切な形容詞群を選定することから始めなければならない。そのためには、コンピュータ・イメージについておなじような意味として回答者に認知される形容詞の種類がどのようなものであり、また、コンピュータ・イメージがいかなる要因によって規定されているかについての知識が大いに役立つ。また、イメージの分析作業においても、それらの知識があるとないとでは得られる情報の質や有用性に雲泥の差を生じさせるものとなりかねない。もしも、アンケート調査結果の報告書が何十とある項目のひとつひとつについて、それも、似たような意味をもつ項目についての冗長な分析記述がだらだらと続くのであれば、読み手は途中で投げ出したくなるろうし、その中

のどれとどれがどのように関連しているか、また、情報を圧縮するとどういふことが言えるかについて直ちに知りたくなるだろう。このような時に、救いの手となるのが多変量解析の手法である。多数の項目について関連の高いものをどうしを群化させる客観的で、最近のコンピュータの発達によって容易になった因子分析や主成分分析はしばしば用いられている。ちなみに、筆者が因子分析により見出したコンピュータ・イメージを規定する主要な要因は「親近性」、「知性」、「評価」であり、なかでも「親近性」がコンピュータ学習の動機づけに少なくない影響を与えることが確かめられている。

しかしながら、因子分析や主成分分析で分析にかけることのできるデータは連続量であつて、「まったくそう思う」、「まあそう思う」、「どちらともいえない」、「あまりそう思わない」、「まったくそう思わない」などといった5段階の尺度から得られた離散的なデータを使うことに対してうしろめたさを感じないわけにはいかない。それらに順に1、2、3、4、5の得点を与えて計算する時には、1と2との差と2と3との差が概ね等しいことを前提としているが、実際の評価においてこう考えられる場合はまずない。「まったくそう思う」と「まあそう思う」との間の心理的間隔は、たぶん、「まあそう思う」と「どちらともいえない」との間の心理的間隔より、人によってその大小関係に違いこそあれ、かなり違っているように思われる。実はこのような、リッカート尺度上の1、2、3、4、5という数字は、上記の5つの区分にあたえられたラベルに過ぎず、せいぜいそれらの順序関係が保証されているだけにすぎない、いわゆる、質的な（あるいはカテゴリカルな）尺度の値である。こういった場合は、無思慮、無判別に因子分析や主成分分析を使うべきではなく、質的データに応じた、質的データの分析のために開発された手法、すなわち、数量化Ⅲ類を用いるべきなのだろう。

単にプログラムが手近になく、すぐに使えない（SASやSPSSには標準装備されていない）、他の人も気にせず使っているなどといった理由で因子分析で間に合わせてきた筆者の無反省な態度をここで改めるための懺悔をするかわりに、奈良大学情報処理センターでも数量化Ⅲ類を使えるようにするためにささやかな努力をしてみたので、ここにその作業の一部を報告する。

1. 数量化Ⅲ類のコンセプト

測定とは事物の属性を数量化することであり、AかBかといったカテゴリカルなデータを扱っていても、それらの生起頻度を問題とするので、科学的な計測のすべてが数量化の作業を行っているといえる。しかしながら、質的なデータを扱う場合、どのようなカテゴリのセットを用いるべきかは慎重に考慮すべき重要な問題となる。例えば、コンピュータについてのイメージを「好き」、「優しい」、「親しみのある」、「暖かい」、「重厚」、「力強

い」、「大胆」、「活動的」、「進歩的」という7つの形容詞からあてはまるものをいくつでも選ばせる形式で調べ、その結果、表1のようなデータを得たとする。これを、表2のように右上から左下に反応が並ぶように回答者の並べ替えを試みる。もしこのような表が得られれば、これから容易に、コンピュータのイメージが「好き」、「優しい」、「親しみのある」、「暖かい」という評価の要因と、「重厚」、「力強い」、「大胆」というパワーの要因と、「活動的」、「進歩的」という活動性の要因の3つの要因によって規定されていると考えることができよう。ちなみに、回答者の1は評価の要因のみを選択し、2は評価とパワー、3はパワーのみ、4はパワーと活動性、5は活動性の要因を選択している。このように、回答者の反応パターンの似たものどうしをまとめることにより、イメージの規定因を見つけ出すことができる。

表1 コンピュータ・イメージについてのアンケート結果（例）

回答者	好き	優しい	親しみのある	暖かい	重厚	力強い	大胆	活動的	進歩的
1					○	○	○		
2	○	○	○	○					
3					○	○	○	○	○
4	○	○	○	○	○	○	○		
5								○	○

表2 表1を並べ替えたもの

回答者	好き	優しい	親しみのある	暖かい	重厚	力強い	大胆	活動的	進歩的
2	○	○	○	○					
4	○	○	○	○	○	○	○		
1					○	○	○		
3					○	○	○	○	○
5								○	○

表1を表2のように、回答（○印）を回答者の列をX軸、回答項目の並びをY軸とみただたXY平面上に位置する点としたときに、XとYとの相関が最も強くなるように回答者の並べ替えを行うような作業を手作業でなく機械的、自動的にできると非常に便利である。数量化Ⅲ類は、そのような作業を数学的な手続きに基づいて行うものであり、複数のカテゴリカルな質問項目間の類似性を探るときに有効な道具となる。特に、最も相関を高くす

る並び方Aと、次に高い並び方Bのベスト2をとり、上下に並べ方Aによる並びをつくった上で、さらに、左右には並べ方Bに従った並びになるように、すなわち、平面上に並べてみる方法は、同じ位置に集まる項目をひとまとめにすれば似通った項目からなる群をつくるのに役立つ。このような類似性を視覚化するひとつの方法が数量化Ⅲ類といえるのではなかろうか。なお、林の数量化理論については、数式をたくさん盛り込んだ本格的な概説書（駒澤 1990；駒澤・橋口 1988；圓川 1988）や、平易にコンセプトを解説した啓蒙書の類（大村 1984；木下 1989）がでているので、詳しくはそれらをご覧ください。

2. 数量化Ⅲ類フォートラン・プログラムの改良

林の数量化のプログラムは大型計算機用のFORTRAN形式のもの（駒澤 1990）からBASIC形式のものまであり、さらに、BASICで書かれたものにはPC9801のような高級パーソナルコンピュータ用もの（駒澤・橋口 1988）だけでなく、記憶容量が小さいハンドヘルドコンピュータ用もの（小林 1984）まである。これ以外にも数量化を含んだパソコン用統計パッケージがいくつか市販されている。その中には、数値計算の精度が低く、逆行列計算の手順に問題があったりして桁落ちが激しく、実用に耐えないものも少なくなく、いまのところ、大量データを高速処理するには大型計算機を利用することが推奨される。

奈良大学においても大型計算機で林の数量化を使いたいという希望は前からあったが、昨年、統計数理研究所の御厚意により奈良大学情報処理センターに数量化FORTRANプログラムの提供を受けることができた。前年度までの所員であられた市ノ瀬氏により数量化Ⅰ類と数量化Ⅱ類が当センターのACOSシステムで使用できるようになった。しかし、数量化Ⅲ類については計算結果をグラフに表示するプロット・サブルーチンに問題があり、直ちに使用するには少々難点があることがわかった。計算過程にはまったく問題はないが、数量化Ⅲ類のプロット・サブルーチンのプログラム中のFORMAT文に旧式のシステム用の「1H+」という重ね打ちようのコマンドが使用されている。当センターのシステムではそういう場合、ページプリンタからの出力においては別々な行に打ち出すことになるので、同じ位置に複数の点がプロットされるようなグラフがいささか見にくいものになってしまう。そこで、筆者が若干の手直しを行い、重ね打ちがあっても、通常のグラフの形で出力できるようにした。以下に、その手直しのポイントをあげる。

- ・FORMAT (1H+, ...) をすべて削除した。すなわち、元のプログラム（駒澤 1990、p.276～p.277）のQ(3)7800、Q(3)7970の2行を削除した。

- ・元のプログラムにおいてもXY平面を配列AI (102) で仮想表現しているのを、

重ね打ちを、その配列に重複して書き込むことにし、打ち出しは最後に配列A Iを1回だけ打ち出すこととした。

- ・以上のために影響をうける部分を手直した。とくに、座標軸や外枠の書き込みを配列A Iの計算の終了後に行うようにした。

付録に手直したプロット・サブルーチンプログラムを載せる。こうして手直した数量化Ⅲ類プログラムにより駒澤(1990)の入力データ事例(P.233表7.1)を分析し、期待通りの結果(同書の出力例7-8 P.245~P.247)を得たことを報告する。なお、プログラムの手直しについては別な方法やもっとスマートな方法があると思われる。先輩諸兄のご指摘、ご批判をお願いしたい。

おわりに

現在、電子工学の専門家でも、もはやCPUの回路図を書いて、どこがどんな機能を受けもっているかなどを説明することは容易ではない。今日、われわれは、自動車の運転や、銀行のATMを、その機械の中身を知らなくてもなんとかうまくつかいこなしている。機械が故障したら、専門家に直させればよいし、使いにくいときにはメーカーに文句を言うか、別な使いやすいシステムを買えばよい。そのような通念が一般的になっている時代に、統計パッケージなどのコンピュータ・プログラムをつかうのに、いちいちその内部処理の細部について理解するひつようななどないのではないか、という見方が育ちつつあるのも不思議ではない。問題は、どういう機能が備わっており、どこを操作するとどのような結果が得られるかを体得することである。コンピュータの統計パッケージも、今や、そういったコンセプト、すなわち、エンドユーザーコンピューティングと呼ばれるコンセプトに基づいて開発されている。また、林の数量化は精密な数値計算とは縁がなく自分にはコンピュータなど必要ないと考えておられる諸先生こそお使いになるべき手法ではないかと考える。自分の勘や経験からの判断をデータによって裏付けたり、確かめたりするために、また、収集した調査結果をとりあえず分類するための目安を得るために数量化は実に便利で強力な道具となるのである。

文 献

- | | | |
|------|------|----------------------------|
| 圓川隆夫 | 1988 | 多変量のデータ解析 朝倉書店 |
| 大村 平 | 1984 | 評価と数量化のはなし 日科技連 |
| 木下栄蔵 | 1989 | わかりやすい数学モデルによる多変量解析入門 啓学出版 |
| 小林竜一 | 1984 | パソコンによる多変量解析 培風館 |

駒澤 勉	1990	数量化とデータ処理	朝倉書店
駒澤 勉・橋口捷久	1988	パソコン数量化分析	朝倉書店
島村麻里	1991	ファンシーの研究	ネスコ

付録 ー手直ししたプロット・サブルーチンプログラムのリスト

```

SUBROUTINE PLOT3(N,X,Y,ALFA)
DIMENSION X(1),Y(1),ALFA(1),NO(999)
DIMENSION MO(999),AI(102),IB(200),S1(999),S2(999)
REAL      MINUS(2)
DATA MINUS,BLNK,ONEI/'-','_',' ','/'
WRITE(6,2222)
2222 FORMAT(1H , 44X,'P L O T T I N G   G R A P H',/8X,'!',9X,'!',9X,
1'!',9X,'!',9X,'!',8X,' ',9X,'!',9X,'!',9X,'!',9X,'!',8X,' ! ')
DO 95 I=1,N
NO(I) =I
95 CONTINUE
NM1=N-1
DO 10 J=1,NM1
KI=0
IJ=N-J
DO 20 I=1,IJ
IF(X(I).LE.X(I+1)) GO TO 20
XI=X(I)
X(I)=X(I+1)
X(I+1)=XI
YI=Y(I)
Y(I)=Y(I+1)
Y(I+1)=YI
NOI=NO(I)
NO(I)=NO(I+1)
NO(I+1)=NOI
KI=1
20 CONTINUE
IF(KI.EQ.0) GO TO 30
10 CONTINUE
30 YMAX=ABS(Y(1))
DO 60 I=2,N
IF(YMAX.LE.ABS(Y(I))) YMAX=ABS(Y(I))
60 CONTINUE

```

```

XMAX=ABS(X(1))
IF(ABS(X(N)).GT.XMAX) XMAX=ABS(X(N))
DO 400 K=1,61
400 IB(K)=0
DO 450 K=1,N
IX=30*X(K)/XMAX+31.5
MO(K)=IX-31
450 IB(IX)=IB(IX)+1
IK=1
IKK=-31
I1=1
I2=31
DO 454 K=1,101
454 AI(K)=MINUS(2)
WRITE(6,600) AI
600 FORMAT(1H1,7X,102A1)
DO 444 L=1,2
DO 300 I=I1,I2
IKK=IKK+1
IF(I.LT.I2) GO TO 3
DO 455 K=2,100
455 AI(K)=MINUS(L)
GO TO 2
3 DO 460 K=2,102
460 AI(K)=BLNK
2 AI(1)=ONEI
AI(51)=ONEI
AI(101)=ONEI
IF(IB(I).EQ.0) GO TO 335
KK=IB(I)
DO 230 KKK=1,KK
IY=50*Y(IK)/YMAX+51.5
IALFA=NO(IK)
AI(IY)=ALFA(2*IALFA-1)
AI(IY+1)=ALFA(2*IALFA)
S1(IK)=AI(IY)
S2(IK)=AI(IY+1)
IK=IK+1
230 CONTINUE
335 IF(I.NE.31) GO TO 340
WRITE(6,610) IKK,AI
610 FORMAT(1H ,I5,2X,102A1,'(Y)')
GO TO 300

```

```

340 WRITE(6,620) IKK,AI
620 FORMAT(1H ,I5,2X,102A1)
300 CONTINUE
    IF(I.GE.61) GO TO 444
    I1=32
    I2=61
444 CONTINUE
    WRITE(6,1000)
1000 FORMAT(1H ,56X,'(X)')
    WRITE(6,7776)
7776 FORMAT(1H1,///1X,'***** CO-ORDINATE ( X,Y ) TABLE *****',///,
110X,'X''',2X,'NO',7X,'X',9X,'Y',12X,'X''',2X,'NO',7X,'X',9X,'Y',12
2X,'X''',2X,'NO',7X,'X',9X,'Y')
    WRITE(6,7777) (MO(K),S1(K),S2(K),X(K),Y(K),K=1,N)
7777 FORMAT(1H ,/3(I12,2X,2A1,2F10.2))
    WRITE(6,6000)
    N1=1
    N2=20
    N3=1
    N4=40
510 IF(N.LT.N2) GO TO 520
    WRITE(6,3000) (I,I=N1,N2)
    WRITE(6,4000) (ALFA(I),I=N3,N4)
    WRITE(6,5000)
    N1=N2+1
    N2=N2+20
    N3=N4+1
    N4=N4+40
    GO TO 510
520 CONTINUE
    IF(N.EQ.N1-1) GO TO 530
    WRITE(6,3000) (I,I=N1,N)
    N4=N*2
    WRITE(6,4000) (ALFA(I),I=N3,N4)
6000 FORMAT(1HO,///,1X,'***** TABLE OF PLOTTING MARKS FOR CATEGOR
1Y NO. *****',//)
3000 FORMAT(1H ,'CATEGORY NO.',6X,20I5)
4000 FORMAT(1H ,'PLOTTING MARK',5X,20(3X,2A1),/)
5000 FORMAT(1HO)
530 RETURN
    END

```