分割表項目数量化の荻野法に寄せて

水 谷 静 夫

0. 本稿の意図

本誌120集に載った荻野綱男の方法［文献１］は，属性的項目の分布を示す分割表における《最適尺度法》の一解法と言へる．荻野が出したやうな例に限らず，この方法が使へる問題は多い．原理的には，$a \times b$ 分割表（即ちデータが a 行 b 列の行列として表示されるもの）の形にまとめるのが適当な場合には，この方法が利用出来る；但し a や b が少なくとも３でなければ，使っても実際的価値は乏しい．

荻野法の計算自体は別にむづかしくはないが，逐次近似計算なので手間が掛かる．該論文左15頁に言ふ通り，計算機やプログラム電卓を使はなければやり切れまい．そこで本稿では，先づ，これを計算する Basic プログラムを提示する．第二に，荻野の例より相当に大きい或計量語彙論的問題に適用した結果を報告する．第三に，荻野法を使ふに当っての若干の注意を述べる．以上が本稿の意図である．理論的な事柄は，予定されてゐる岩坪秀一の論文に待つ．

1. 荻野の例と水谷のプログラム

荻野が120集に表１として掲げたデータに対し，筆者の書いたプログラム CRQ１を走らせた出力が，図１である．実用上の収束判定条件を，繰返し計算の前回との差がすべて 5×10^{-4} 未満とした；即ち小数第三位まで有効数字が得られるやうにした．ループ７回で６回目と有効数字が全く同じになり，従って数量化計算はこゝで打切られた．

図１の６回目の値を読者が120号の結果と比べる労を惜しまなければ，筆者のの方がすべて正確に１だけ小さい事に気づくであらう．その理由は後に述べる．又，図１の出力には，荻野がしなかった，結果の規準化と最大化された積率相関係数の計算とが，加へてある．念の為に言へば，規準化とは，変量 x，y それぞれに，平均が０で分散が１となるやうに変換を行ふ事である．この方法による結果は相対的なもの故，かゝる変換を施しても何ら本質的な違ひは生ぜず，かつ同類の他の結果と比べるのに便利なので，規準化する事はこの種の扱ひでは常套的な手段である．

さてこの計算をしたプログラム CRQ１を図２に掲げる．これは，ACOS 6（東芝）の対話型時分割システムの下で働く Basic で書いてある．やゝ冗漫な所もあるが，分りよさを旨として書いたつもりである．プログラムの構成は，

行0001–0040　制御用定数などの設定

行0100　データ読込み

図1　荻野の例の CRQ1 出力

```
        ﾌﾞﾝｶﾂ_ﾋｮｳ ｺｳﾓｸ ｽｳﾘｮｳｶ (ｶｸﾁｮｳ ｵｷﾞﾉ=ﾎｳ)
## 「ｺｸｺﾞｶﾞｸ」 120ｼｭｳ ﾉ ｹｲｺﾞ ｽｳﾘｮｳｶ ##

Xﾞﾉ ｼｭｳﾍﾝ ﾄﾞｽｳ: 1077   1146    436    290
Y ﾉ ｼｭｳﾍﾝ ﾄﾞｽｳ:   485    501    492    488    493    490
ﾃﾞｰﾀ ﾉ ｵｵｷｻ: 2949

X ﾉ ｼｮｷﾁ                  .000   1.000   2.000   3.000
===     1 ｶｲﾒ ﾉ Y                  .000    .536   1.882   2.842   4.372   5.000
  ----   ｵﾅｼﾞｸ X          .000   1.830   2.529   3.000
===     2 ｶｲﾒ ﾉ Y                  .000    .723   2.536   3.443   4.666   5.000
  ----   ｵﾅｼﾞｸ X          .000   1.987   2.605   3.000
===     3 ｶｲﾒ ﾉ Y                  .000    .752   2.637   3.534   4.709   5.000
  ----   ｵﾅｼﾞｸ X          .000   2.011   2.617   3.000
===     4 ｶｲﾒ ﾉ Y                  .000    .757   2.652   3.547   4.716   5.000
  ----   ｵﾅｼﾞｸ X          .000   2.015   2.619   3.000
===     5 ｶｲﾒ ﾉ Y                  .000    .757   2.654   3.549   4.717   5.000
  ----   ｵﾅｼﾞｸ X          .000   2.015   2.619   3.000
===     6 ｶｲﾒ ﾉ Y                  .000    .757   2.655   3.550   4.717   5.000
  ----   ｵﾅｼﾞｸ X          .000   2.015   2.619   3.000
                 ｽｳﾘｮｳｶ ﾉ ｹｲｻﾝ ｼﾞｶﾝ: .063 SEC

     *** ｹｯｶ ﾉ ｷｼﾞｭﾝｶ ***

X[I]:--
  -1.271    .477   1.000   1.331
Y[J]:--
  -1.482  -1.078   -.066    .412   1.034   1.185
ｿｳｶﾝ ｹｲｽｳ: .736

                 ｷｼﾞｭﾝｶ ﾉ ｹｲｻﾝ ｼﾞｶﾝ: .007 SEC
```

行1000-1310　データ要目の計算・提示（入力の正しさの確認用）

行2000-2910　数量化計算とループ各回の結果の提示

行3000-3600　規準化計算・提示（相関係数も）

行4000-4199　印刷の下請

行4300-4399　最小値・最大値決定の下請

配列名と主な変数名との使途は，

F：　分割表の各ますの値を入れる配列；

X，Y：　第g回目の計算結果を入れる配列；　　U，V：　作業用配列；

N1，N2：　行・列の周辺度数を記す配列；　　N：　データの大きさ；

L，S：　最大値・最小値；　　W，C，D：　作業用；

M1，M2，D1，D2：　x，yの平均値・標準偏差の算出用．

こゝで，図1の結果が何故1だけ荻野の示したものと違ふのかを述べよう．一言で述べれば計算時間を短縮する為である．荻野の論文の左15頁③の式に見る通り，（中間）数値の算出に当ってそれぞれ，比例させた値に1を足してゐるが，この足し算はたゞ値を1からaまで又はbまでに配る為にだけ役立ち，最終結果に対してはそれ以外に何の貢献もしない．もしこれを0から$a-1$又は$b-1$までに配るとすれば，全くいらなくなる足し算である．荻野法で得られる結果は相対的なもの故，この変更はこの方法にとって本質的な変更ではなく，殊に結果を規準化してしまへば全く同一の数値に達する．図1の例で言っ

ても，言はばこの無駄な計算に (6+4)×7=70回のステップを費す．後述の図4の計算ではそれが約一万六千回にも上る．幾ら計算機にさせるからとて，無駄は無駄である．従ってこの改修を施した．もし荻野式の値が欲しければ，最終値に1を加へればよい．

なほ荻野は「普通5〜7サイクルで値が収束する．」と言ふが，これは恐らく彼が手掛けた寸法の小さい分割表の場合の事である．（この点は後で再び取り上げる．）寸法がもっと大きい分割表データを処理するには，一々中間出力をする CRQ 1は不適当であり，又データの与へ方も行 5001-5004 の形式に限っては，使ひ勝手が頗る宜しくない．そこで別に CRQ 2が作ってあるが，紙幅を慮ってそのプログラムリストは割愛する．次の節に述べる例は，そちらで実行したものである．

2. 計量語彙論からの一例

水谷が文献2で扱った問題に荻野法を使ってみよう．その問題とは次のやうなものである．割り切って言ってしまへば飽かず別れた女を後日追ひ求める唄である『湯の町エレジー』と『上海帰りのリル』とではあるが，しかし，曲調の事を別にしても，この二つから受ける我々の感じは著しく違ふ．その違ひが用語面から客観的に捉へられないか．この二篇に周辺の十八篇を併せ，その用語情況を林知己夫の数量化理論 第IV類・第III類の方法によって分析し，成功と言へる結果を導いたのが，文献2である．

荻野法は或意味で第III類の方法に似た簡略版である．（双対(つゐ)的だからどちらをxと見てもよいが）キー語をx，作品をyとして両者を同時的に数量化しようとする．数量化第III類ではキー語出現の有無のパターンを手掛りとして作品の仕訳を企てる．そこではキー語の使用度数の多寡は捨象されてゐる．荻野式によれば使用度数の情報まで仕訳に参与させる事が出来る．だからと言ってどちらの方法が圧倒的に優れてゐるとは速断し難い．これは数理ばかりで決定出来る性質の問題ではない．計量語彙論の方法論の問題として大変に興味深くはあるが，本稿ではその議論を保留し，この類ひの分析にも荻野法の使途がある事だけを例示するにとゞめる．

図3はこの例題のデータであるが，見易さを考へて，CRQ 2で計算した図4に示す結果に基づいて排列してある．荻野法（に限らずこの類ひの分析法）は，要するに，素データから図3のやうな一層整った形のデータ排列を得，更にそこに隠れた秩序を読み取る為の方法とも言へよう．図3の右側にはキー語に，下側には作品に，この方法で与へられた値（即ち図4の規準化された x_i, y_i）を書き添へておいたが，その下側の値の組に見る通り，想定した作品群仕訳に成功してゐる．

こゝに挙げた20篇だけを考へてみても，およそ歌謡曲におのづからなる線型関係は期待しにくく，それらは多側面的な存在と見られるであらう．にも拘らずキー語をうまく選ぶ事によって，0.834といふ相当に高い積率相関係数を有つやうな，しかも我々の直観的な綜合判定による分類を裏付ける，一次元的排列が出来た事は，注目に価する．荻野法によるこの結果は，文献2での分析結果と大局的に合致する．相異なる三通りの方法でほゞ同じ結果が導けたから，これら歌謡曲の用語法にかなり安定した秩序・構造の存在を考へる事は，強ち無理でない．

図2 プログラムCRQ1

```
0001 PRINT "      ﾌﾞﾝｶﾂﾋｮｳ ｺｳﾓｸ ｽｳﾘｮｳｶ  (ｶｸﾁｮｳ ｵｷﾞﾉ=ﾎｳ)"
0002REM                              ﾐﾂﾞﾀﾆ ｼﾂﾞｦ  1980 04 19
0003   T$="「ｺｸｺﾞｶﾞｸ」 120ｼｭｳ ﾉ ｹｲｺﾞ ｽｳﾘｮｳｶ"
0004     PRINT "## ";T$;" ##"                                  ← 標題を与へる
0005     PRINT
0010DIM  F(4,6),N1(4),N2(6), X(1,4),Y(6,1), U(4,1),V(1,6)   分割表の大きさ
0020 LET  A=4  ¥  B=6  ¥  E=.0005                            に合はせて指定
0030     A0=A-1  ¥  B0=B-1
0040     G=1  ¥  N=Z0=0
0100 MAT READ X,F
0110      MAT Y=ZER
1000 PRINT "X ﾉ ｼｭｳﾍﾝ ﾄﾞｽｳ:";
1010 FOR  I=1 TO A
1020   W=0
1030   FOR  J=1 TO B
1040     W=F(I,J)+W
1050   NEXT J
1060   N1(I)=W  ¥  PRINT W;
1070   N=W+N
1080 NEXT I
1090 PRINT
1100 PRINT "Y ﾉ ｼｭｳﾍﾝ ﾄﾞｽｳ:";
1110 FOR  J=1 TO B
1120   W=0
1130   FOR  I=1 TO A
1140     W=F(I,J)+W
1150   NEXT I
1160   N2(J)=W  ¥  PRINT W;
1170 NEXT J
1180 PRINT
1190 PRINT "ﾃﾞｰﾀ ﾉ ｵｵｷｻ:";N
1200 PRINT
1300 PRINT "X ﾉ ｼｮｷﾁ";TAB(16);
1310      GOSUB 4110
1990                   Z=TIM(A)
2000 MAT V=X*F
2010   L=0  ¥  S=9999999
2020 FOR  J=1 TO B
2030   W=V(1,J)/N2(J)
2040    GOSUB 4300
2050   V(1,J)=W
2060 NEXT J
2100   K=B0/(L-S)
2110    D=1
2120 FOR  J=1 TO B
2130   W=(V(1,J)-S)*K
2140   IF ABS(Y(J,1)-W)<E THEN 2160
2150   D=0
2160 Y(J,1)=W
2170 NEXT J
2240                   Z0=TIM(B)-Z+Z0
2250  GOSUB 4000
2490                   Z=TIM(C)
2500 MAT U=F*Y
2510   L=0  ¥  S=9999999
2520 FOR  I=1 TO A
2530   W=U(I,1)/N1(I)
2540    GOSUB 4300
2550   U(I,1)=W
2560 NEXT I
2600   K=A0/(L-S)
2610    C=1
2620 FOR  I=1 TO A
2630   W=(U(I,1)-S)*K
2640   IF ABS(X(1,I)-W)<E THEN 2660
2650   C=0
2660   X(1,I)=W
2670 NEXT I
2740                   Z0=TIM(D)-Z+Z0
2750  GOSUB 4100
2900 IF C*D=1 THEN 2980
2910        G=G+1  ¥  GOTO 1990
2980 PRINT TAB(15);"ｽｳﾘｮｳｶ ﾉ ｹｲｻﾝ ｼﾞｶﾝ:";INT(1000*Z0)/1000;"SEC".
2990 PRINT
3000 PRINT "       *** ｹｯｶ ﾉ ｷｼﾞｭﾝｶ ***"
3010    M1=M2=D1=D2=0
3020                   Z=TIM(E)
3100 FOR  I=1 TO A
3110   W=N1(I)  ¥  C=X(1,I)
3120   W=W*C    ¥  M1=M1+W
3130   W=W*C    ¥  D1=D1+W
3140 NEXT I
3150 M1=M1/N
3160 D1=SQR(D1/N-M1*M1)
3200 FOR  J=1 TO B
3210   W=N2(J)  ¥  D=Y(J,1)
3220   W=W*D    ¥  M2=M2+W
3230   W=W*D    ¥  D2=D2+W
3240 NEXT J
3250 M2=M2/N
3260 D2=SQR(D2/N-M2*M2)
```

図2 続き

```
3300  W=0
3310 FOR  I=1 TO A
3320   K=X(1,I)
3330   FOR  J=1 TO B
3340     W=K*Y(J,1)*F(I,J)+W
3350   NEXT J
3360 NEXT I
3370 W=(W/N-M1*M2)/D1/D2
3380                 Z=TIM(F)-Z
3390 PRINT
3500 PRINT "X[I]:--"
3510 FOR  I=1 TO A0
3520   PRINT USING 3777, (X(1,I)-M1)/D1;
3530 NEXT I
3540   PRINT USING 3777, (X(1,A)-M1)/D1
3550 PRINT "Y[J]:--"
3560 FOR  J=1 TO B0
3570   PRINT USING 3777, (Y(J,1)-M2)/D2;
3580 NEXT J
3590   PRINT USING 3777, (Y(B,1)-M2)/D2
3600 PRINT "ｿｳｶﾝ ｹｲｽｳ:";INT(1000*W)/1000
3610 PRINT
3620 PRINT TAB(15);"ｷｼﾞｭﾝｶ ﾉ ｹｲｻﾝ ｼﾞｶﾝ:";INT(1000*Z)/1000;"SEC"
3630      STOP
3777: ###.###
4000 PRINT "===";G;"ｶｲﾒ ﾉ Y";TAB(24);
4010 FOR  J=1 TO B0
4020   PRINT USING 3777, Y(J,1);
4030 NEXT J
4040   PRINT USING 3777, Y(B,1)
4099     RETURN
4100 PRINT "  ----  ｵﾅｼﾞｸ X ";
4110 FOR  I=1 TO A0
4120   PRINT USING 3777, X(1,I);
4130 NEXT I
4140   PRINT USING 3777, X(1,A)
4199     RETURN
4300 IF W>S THEN 4320
4310      S=W
4320 IF W<L THEN 4399
4330      L=W
4399      RETURN
5000 DATA 0,1,2,3
5001 DATA 438,371,164, 90, 10,  4
5002 DATA  39,110,280,279,243,195
5003 DATA   7, 17, 39, 88,142,143
5004 DATA   1,  3,  9, 31, 98,148
9999 END
```

（5000〜5004）データはその都度与へる

序でに，話がやゝ脇道に入るが，**数量化第 III 類との関係**について一言しておく．荻野法のデータ行列を作り替へて，$f_{ij} \geqq 1$ の所は改めて 1，$f_{ij}=0$ の所はそのまゝ 0 にすれば，かくて得られるデータ行列はまさに第 III 類のそれである．かやうに作り替へたデータに荻野法を適用すればどうなるかを実験的に確かめたところ，次のやうになった．(1) 荻野法で得られる相関係数は，第 III 類の方法での第一軸値に関する相関係数と等しい．(2) x_i の値は，時に符号が逆になる場合がある事を別にすれば，即ち絶対値としては，第一軸値と一致する．(3) y_j の値は，両方法の算出法の違ひのせゐか，一致しない．この(3)に対する数理的究明をまだしてゐないが，荻野法は数量化第 III 類と緑が深い方法であると言へる．

図3 キー語の分布 附 荻野法による規準化した x_i 及び y_j

		昭和二六 上海帰りのリル	二一 上海リル	二七 霧の港のリル	一〇 上海リル	二七 私がリルよ	三四 夜霧に消えたチャコ	二七 ロンドンの街角で	一四 上海の花売娘	二一 東京の花売娘	一五 南京の花売娘	一五 広東の花売娘	一四 上海夜曲	一三 上海ブルース	一三 霧の四馬路	一三 上海の街角で	二六 湯の町物語	二八 別府エレジー	二四 湯の町夜曲	二七 湯の町よさようなら	三三 湯の町エレジー	x_i
リル	Lilu	21	8	18	4	10																−1.479
ヒトリボッチ		1		1		1																−1.421
キ・クル	来	2	3	1							1											−1.282
ドコ	何処	1		2		1				1												−1.068
オモイデ	思出	1			2	1		1														−1.034
ヒ	day				2	1	1															−.977
アカイ	赤		2	1			1	2	1			1										−.444
キリ	霧	1		3		1	1	2	1			1		1	1							−.383
ヒトミ	瞳	1	2		1			1	1	1	1		1	1								−.333
カナシイ	悲			1		3	2							1			1	1	1			.022
ユメ	夢	1			2		1		1	1	1	1	1	1		1	1				1	.216
ハナウリ	花売							3	6	3		1										.306
ハナカゴ	花籠							1	1	1	1											.321
ムスメ	娘							3	6	3	3	1										.327
メシ・ス	召							6		6		3										.339
ヌレル	濡							1			1	1										.377
ハナ	花		1	1		1	1			6	12	6	2	1	1	1	1		1			.380
カゲ	影・蔭							1	1	1			1									.381
スマロ	四馬路	1											1	1	2	1						.440
ワカレル	別	1				1							2		1	3		2		1		.734
チリ・ル	散					1					2		1		1	1		1			1	.763
ツキ	moon					1		1		1		1	1	3	1	1		1	1	1	1	.886
ムセビ・ブ	咽			1													1		1		1	.894
コヨイ	今宵			1										1	1	1		1		1	1	.999
キミ	君													1	1	1		1	3		3	1.597
ギタア	gitar											1							1	1	2	1.697
タビ	旅																1	1	1	1	1	1.772
ユ	湯																1	1	1	1	2	1.816
(ラ)レル	被																	1	1		1	1.869
カ	香																	1		1	1	1.922
y_j		−1.410	−1.262	−1.229	−1.072	−.918	−.199	.157	.236	.330	.347	.441	.549	.609	.779	.896	1.009	1.418	1.528	1.666	1.675	$\rho=$ .834

図4　キー語使用度数に基づく歌謡曲仕訳 (CRQ2の出力)

```
      ブンカツ_ヒョウ コウモク   スウリョウカ  (カクチョウ オキ゛ノ=ホウ)
## カヨウキョク:   ハナウリムスメ・・ リル ・ シャンハイ ・ ユ_ノ_マチ ##
                                          ケイサン ツキ・ヒ・ネン 05/30/80 / 17.334ジ
XCIJ: ムスメ      ハナウリ    ハナカゴ゛  メシ・ス    ヌレ=ル   カゲ゛
      リル        ヒ(DAY)  ヒトリボ゛ッチ キ・クル   オモイデ゛  ド゛コ
      ユ          タビ゛    カ(ODOR)  (ラ)レ=ル ギ゛タア  ムセビ゛・ブ゛
   アカ=イ     カナシ=イ    キミ   コヨイ    ツキ(MOON)     ハナ
   ユメ    キリ(FOG) ヒトミ    スマロ   ワカレ=ル  チリ・ル
YCJJ: 1-5 ハナウリムスメ_モノ   6-10 リル_モノ   11 ジュン_リル_モノ
      12-15 ショウワ_10ネン_ダ゛イ シャンハイ_モノ  16-20 ユ_ノ_マチ_モノ

X ノ シュウヘン ド゛スウ:   16     13     4    15     3     4    61     4     3
   7     5     5     6     5     3     3     5     4     8    10    10
   7    14    35    13    12    10     6    11     8
Y ノ シュウヘン ド゛スウ:   18     17     22     24     22     11     16     31     22
  30     7    10    11    10     9    15    11     6     7    11
デ゛ータ ノ オオキサ:   310
X ノ ショキチ:
  ( 1) 1.000  ( 2) 2.000  ( 3)  .000  ( 4) 6.000  ( 5) 3.000  ( 6) 4.000
  ( 7)10.000  ( 8) 9.000  ( 9)12.000  (10) 7.000  (11) 5.000  (12)11.000
  (13)21.000  (14)21.000  (15)24.000  (16)23.000  (17)19.000  (18)17.000
  (19) 8.000  (20)16.000  (21)20.000  (22)18.000  (23)15.000  (24)14.000
  (25)13.000  (26)10.000  (27)10.000  (28)10.000  (29)10.000  (30)10.000
===  316 カイメ デ゛ X ハ:
  ( 1)15.492  ( 2)15.317  ( 3)15.440  ( 4)15.599  ( 5)15.926  ( 6)15.960
  ( 7)  .000  ( 8) 4.308  ( 9)  .502  (10) 1.691  (11) 3.817  (12) 3.529
  (13)28.063  (14)27.693  (15)29.000  (16)28.528  (17)27.053  (18)20.359
  (19) 8.884  (20)12.879  (21)26.376  (22)21.229  (23)20.289  (24)15.950
  (25)14.540  (26) 9.404  (27) 9.833  (28)16.463  (29)18.989  (30)19.235
-------  オナジ゛ク Y ハ:
  ( 1)10.141  ( 2)11.399  ( 3)10.818  ( 4)10.718  ( 5) 9.649  ( 6) 2.081
  ( 7)  .910  ( 8)  .000  ( 9) 3.033  (10) 1.115  (11) 7.457  (12)12.064
  (13)12.434  (14)14.202  (15)13.481  (16)19.000  (17)18.096  (18)14.898
  (19)18.944  (20)17.418

                 スウリョウカ ノ ケイサン ジ゛カン:   12.698 SEC

      *** ケッカ ノ キジ゛ュンカ ***
XCIJ:--
  ( 1)   .327  ( 2)   .306  ( 3)   .321  ( 4)   .339  ( 5)   .377  ( 6)   .381
  ( 7)-1.479  ( 8) -.977  ( 9)-1.421  (10)-1.282  (11)-1.034  (12)-1.068
  (13) 1.792  (14) 1.749  (15) 1.901  (16) 1.846  (17) 1.674  (18)  .894
  (19) -.444  (20)  .022  (21) 1.595  (22)  .995  (23)  .886  (24)  .380
  (25)  .216  (26) -.383  (27) -.333  (28)  .440  (29)  .734  (30)  .763
YCJJ:--
  ( 1)  .237  ( 2)  .441  ( 3)  .347  ( 4)  .330  ( 5)  .157  ( 6)-1.072
  ( 7)-1.262  ( 8)-1.410  ( 9) -.918  (10)-1.229  (11) -.199  (12)  .549
  (13)  .609  (14)  .896  (15)  .779  (16) 1.675  (17) 1.528  (18) 1.009
  (19) 1.666  (20) 1.418
ソウカン ケイスウ: .834

                 キジ゛ュンカ ノ ケイサン ジ゛カン: .087 SEC
```

図4に結果を示した計算を始めるに当って，素データ行列を一見しただけでは各キー語にどんな初期値を与へるのが良ささうか見当がつけ難かったので，図4に見るやうな，かなりでたらめな値の組を与へてみた．これは最終値を近似するやうなものでなかったので (実験の為わざとさうしてみたのであるが)，実用上の収束に至るまでに三百回余りの繰返し計算を要した．文献1に，初期値がどうであれ「わずかにサイクル数が1～2回多くなるだけ」と言ふのは，荻野の勇み足である．因みに，図4の最終結果を知った上で，規準化した x_i の小数第二位までの値 (従って負数もある) を初期値として与へて験したとこ

ろ，今度は6回の繰返しで実用上の収束を見，この計算時間は0.2秒（図4のは12.7秒程）であった．

3. 荻野法における初期値の問題等

今述べた通り，荻野法を使ふ場合には初期値の与へ方に一考を要する．文献1に明記してないが，解法の形式的性質からして，

すべての初期値を同じ値にしてはならない．

その極端な場合として，すべてを0にすればどんな事が起るかを考へてみるがよい．かういふ初期値の組が無効である事は，すぐ分るであらう．又，{0, 1, 2, 1, 1, 3, 1, 1} のやうな

等しい初期値の多用は，それらに対応するデータ分布が似てゐない限り，計算能率を落す原因となり勝ちである．

分割表の寸法が小さい時には，等しい値が多くても（すべてではない限り）苦になる程の計算時間の開きが出ない．もともとごく短時間に処理し切れるからである．しかし前節で例に挙げた程度の規模になると，大きな開きが出る．最終結果（の少なくとも大小順）を近似する初期値であれば，当然，計算時間が短くて済む．しかしどんなものが近似値なのか，一般には，事前に分らない．だからこそ計算に訴へる必要が有る訳である．それにしても，理由もなく変な値の組を使ふのは避けた方がよい．特に，消費時間による打切り制を採ってゐる計算センタを利用する場合には，かうした注意が必要であらう．

本稿の計算はすべて筆者の大学の中型機で行ったが，十秒を越す計算といふのは筆者自身は他に余りしてゐない．この点からすると，荻野法は解き方の直観的な分りよさにも拘らず能率的な計算法ではないやうに思はれる．算法 (algorithm) の改良が期待される．

結果の解釈法は計算方式とは別の事であり切り離して論ずべき事柄ながら，一つだけ言及しておかう．図3の「リル」は，特定作品にだけ集中してしかも著しく多く使はれてゐる点で，特異なキーである．さういふ特異なキーを含むと，相関関係が大きい割にすっきりした結果が得られない事が有る．文献2の25キー語を本稿では五語ふやして30としたが，25キーによる荻野法の結果では相関係数が0.883と大きかったのに作品仕訳としては鋭い結果にならなかった；それで五つのキー語を追加して特異キーの影響を弱めようとしたのであった．一般には，相関係数の大小が数量化の成功の度合の指標になる．しかしさうでない場合が有る事を心得ておくのも無駄ではない．

4. 附言

計量国語学の分野では特に，具体的問題としては異なっても形式的には同じ処理法が反復して使はれるのが普通である．さういふものをプログラム化して蓄積して行けば，単体としては簡単なものであっても幾つかをつないで使ふ事により案外の仕事が出来，研究能率が上がる．荻野法もさうした単体の一つとしてプログラミングした．筆者は，この学年の演習題目に計量語彙論支援プログラム集の作成を取り上げたが，本稿のプログラムはその教材の一つである．この演習で，実用に耐へるプログラムをせめて十種は作らせたいと思ってゐる．ユーザ層の拡大により使用体験からのフィードバックでプログラムの質を高

め範囲を拡げる途も拓ける．因みに数量化理論第 III 類を語彙論に適用する西村恕彦作 LIKE 3 の移植により，それを使って卒業論文を書いた又はまとめつゝある学生が既に四名あり，文法論・表記論等を含め計算機利用の卒業論文は本年度予定のものまで併せて十五六篇になる．上記の演習題目はかうした動きに支へられてゐる．文科系だからプログラムが書けないといふ事は無い．況や，自動車の運転と同様，作れない者には使へないといふ事は全く無い．但し工科系の使ひ方を徒らにまねるのでなく，文科系には文科系の使ひ様が有る点を銘記すべきであらう．国語学界にさういふ気運が現れてもよい頃ではないか．本稿でプログラムを公開したのは，その呼び水になる事を欲してである．

なほ先に一寸触れた荻野法と数量化第 III 類との関係の考察には，線型代数学の知見がいる．旧世代人である筆者はこの種の教育を受けて来なかった．現代日本語に関する卒業論文の数が，理・工学部で文学部を上廻る勢ひである現況を見るにつけ，大学の国語学教育のカリキュラムが在来のまゝでいいかどうか，大いに気懸りである．蛇足御容赦．

文献

[1] 荻野綱男 (1980) 敬語における丁寧さの数量化——札幌における敬語調査から(2)——．『国語学』[120] 左13-24.

[2] 水谷静夫 (1980) 用語類似度による歌謡曲仕訳，『湯の町エレジー』『上海帰りのリル』及びその周辺．『計量国語学』12 [4] 145-161.

——東京女子大学教授——