㊺ 数量化の或る問題について

統計数理研究所　橋爪浅治

m 人に n ケの問題を呈出した時，問題 j（$j=1,2,\cdots,n$）に対し，i なる人（$i=1,2,\cdots,m$）が a_{ij} なる得点を取つたとする。今各問題間でそれぞれすべて相関が高いと仮定しておく。

未定常数 x_i，y_j（$i=1,2,\cdots m$；$j=1,2,\cdots,n$）を $\{a_{ij}\,x_i\,y_j\}$ なる cross tabilation に於て，全体の variance 一定のもとに，列間の variance が最小になる様に，きめることが出来た数とするならば，かゝる x_i，y_j は何を意味するであろうか。最初に各問題の間で相関が高いと仮定してあるから $\left\{\frac{1}{y_j}\right\}$ は問題 j が解答者に及ぼす寄興量と考へても良いと思われる。又 x_i はもし解答者 i が，各問題が相関があるにも拘らず無相関に答へているとすれば，列間の variance を最小にするためには，x_i の値は小さくなるであろう。

よつて x_i を解答者 i の正常係数と名付けることにする。又 y_j を問題係数と名付けてみる。

さて $\{a_{ij}\,x_i\,y_j\}$ に於て，平均，全体の variance，列間の variance．を夫々 V，W，S とおけば

$$(1)\qquad V=\frac{1}{mn}\sum_{i,j} a_{ij}\,x_i\,y_j$$

$$(2)\qquad W=\sum_{i,j}(a_{ij}\,x_i\,y_j)^2-mnV^2$$

$$(3)\qquad S=m\sum_{j=1}^{n} y_j^2\Big(\frac{1}{m}\sum_{i=1}^{m} a_{ij}\,x_i\Big)^2-mnV^2$$

又

(4)　$E^2 = S/W$

とおけば E^2 が最小なる様に $\{x_i\}$ $\{y_j\}$ をきめればよい。その爲には

$$\frac{\partial}{\partial x_i}\left(S/W\right) = \frac{\frac{\partial S}{\partial x_i}W - R\frac{\partial W}{\partial x_i}}{W^2} = 0$$

即ち

(5)　$$E^2\frac{\partial W}{\partial x_i} = \frac{\partial S}{\partial x_i}$$

同様に

(6)　$$E^2\frac{\partial W}{\partial y_j} = \frac{\partial S}{\partial y_j}$$

を解けばよい。　さて

(7)　$$\frac{\partial W}{\partial x_i} = 2\sum_{j=1}^{n} a_{ij}^2\, x_i y_j^2 - 2V\sum_{j=1}^{n} a_{ij} y_j$$

(8)　$$\frac{\partial S}{\partial x_i} = \frac{2}{m}\sum_{j=1}^{n} y_j^2\left(\sum_{i=1}^{m} a_{ij} x_i\right)a_{ij} - 2V\sum_{j=1}^{n} a_{ij} y_j$$

(9)　$$\frac{\partial W}{\partial y_j} = 2\sum_{i=1}^{m} a_{ij}^2\, y_j - 2V\sum_{i=1}^{n} a_{ij} x_i$$

(10)　$$\frac{\partial S}{\partial y_j} = \frac{2}{m}\left(\sum_{i=1}^{m} a_{ij} x_i\right)^2 y_j - 2V\sum_{i=1}^{n} a_{ij} x_i$$

(7), (8), (9), (10) を (5), (6) に代入し E^2, $\{x_i\}$, $\{y_j\}$ について解き，その中で E^2 の最小なる價をとるものが求むる解である。

ごく特別の場合を除いて上の聯立方程式を數値計算で解くことは難かしい。

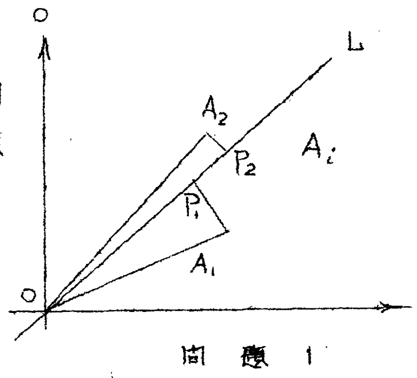
そこで別方法を考えてみる。

各列 j についての平均と偏差を m_j, σ_j とし a_{ij} を

$$(11) \qquad b_{ij} = \frac{a_{ij} - m_j}{\sigma_j}$$

なる一次変換を施して，各二ツの列間 $\{a_{ij}\}$, $\{a_{ij'}\}$ $(i = 1, \cdots\cdots, m)$ の相関に於て，それが高い i に大きな数 x_i を與える一方法を考える。

此処で相関を問題にしているから a_{ij} の代りに b_{ij} で置換えても差支ない。　今 $j = 1, 2$ として，$\{b_{i1}\}$ と $\{b_{i2}\}$ の相関図を画いてみると，その回帰直線は原点を通り方向係数 45° の直線Lに近似する（最初に問題間の相関が高いと假定しているから）



座標点 $A_1 : (a_{11}, a_{12})$，$A_2 : (a_{21}, a_{22})$ とし A_1, A_2 より L への垂線の足を P_1, P_2 とする，そして $\overline{OA_1}$, $\overline{OA_2}$ の大小関係の如何に係らず

$$A_1P_1 \lesseqgtr A_2P_2$$

なら

$$|x_1| \gtreqless |x_2| \quad \text{（不等号同順）}$$

なる如く與えられたとする。　しからば $\{x_i b_{i1}\}$, $\{x_i b_{i2}\}$ なる相関点列は L の廻りに密に集まり，したがってそれぞれの平均

$$\frac{1}{m} \sum_i x_i b_{i1} \ , \quad \frac{1}{m} \sum_i x_i b_{i2}$$

も直線 L 上の附近にくる。　よつて，この二つの平均点の差は小さい。

同様の考えを他の列について考察すれば $\{x_i b_{ij}\}$ なる crosstabilation に於て。列間の variance を最小にする様に x_i を與えれば求める一方法が得られるわけである。

以下，これの数学的方法を取扱つてみる。

先づ，$\{ x_i \; b_{ij} \}$ に於ける全体の mean と variance 列間の variance を夫々 V，W，S とすれば

$$(12) \qquad V = \frac{1}{mn} \sum_{i,j} b_{ij} x_i$$

$$(13) \qquad W = \sum_{i,j} (b_{ij} x_i)^2 - mn V^2$$

$$(14) \qquad S = m \sum_{j=1}^{n} \left(\frac{1}{m} \sum_{i=1}^{m} b_{ij} x_i \right)^2 - mn V^2$$

$$E^2 = S/W$$

とおき E^2 を最小ならしめる如く x_i を決めればよい。 即ち

$$\frac{\partial}{\partial x_i} \left(S/W \right) = \frac{\frac{\partial S}{\partial x_i} W - S \frac{\partial W}{\partial x_i}}{W^2}$$

$$= \frac{1}{W} \left\{ \frac{\partial S}{\partial x_i} - E^2 \frac{\partial W}{\partial x_i} \right\} = 0$$

さて $W = 0$ となる事はない。 何となれば $x_1 = x_2 = \cdots = x_n$ とおけば 列間の variance は 0 であるが，行間の variance は一般に 0 にならない。 又一般に $x_1 = x_2 = \cdots\cdots = x_n$ 以外に $S = 0$ とならないからである。 よつて，

$$(15) \qquad E^2 \frac{\partial W}{\partial x_i} = \frac{\partial S}{\partial x_i}$$

なる方程式をとけばよい。

$$(16) \qquad \frac{1}{2} \frac{\partial W}{\partial x_i} = \sum_{j=1}^{n} b_{ij}^2 x_i - \frac{1}{mn} \left(\sum_{i',j'} b_{i'j'} \right) \left(\sum_j b_{ij} \right)$$

$$(17) \qquad \frac{1}{2} \frac{\partial S}{\partial x_i} = m \sum_{j=1}^{n} \left(\frac{1}{m} \sum_{i'} b_{i'j} x_{i'} \right) \frac{b_{ij}}{m} - \frac{1}{mn} \left(\sum_{i'j'} b_{i'j'} \right) \left(\sum_j b_{ij} \right)$$

$$(18) \qquad \begin{cases} \sum_j b_{ij} = c_i \\ \sum_j b_{ij} b_{i'j} = d_{ii'} \end{cases}$$

とおけば（16）（17）（18）を（15）に代入して

$$(19) \qquad \begin{aligned} & E^2 \left\{ d_{ii} x_i - \frac{c_i}{mn} \sum_{i'} c_{i'} x_{i'} \right\} \\ & = \frac{1}{m} \sum_{i'} d_{ii'} x_i - \frac{c_i}{mn} \sum_{i'} c_{i'} x_{i'} \end{aligned}$$

さて，F，G を

$$(20) \qquad F = (b_{ii'}) = \left(\delta_{ii'} d_{ii'} - \frac{c_i c_{i'}}{mn} \right)$$

$$(21) \qquad G = (g_{ii'}) = \left(\frac{d_{ii'}}{m} - \frac{c_i c_{i'}}{mn} \right)$$

なる行列とすれば，（19）は

$$(22) \qquad E^2 (x_1, x_2, \cdots\cdots, x_m) F = (x_1, x_2, \cdots\cdots, x_m) G$$

と書き直せる，但し $\delta_{ii'}$ はクロネッカー の δ-記号である.

Fも Gも正値行列である事は明らかである。何となれば

$$(23) \qquad (x_1, \cdots\cdots x_m) F \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \equiv W > 0$$

$$(24) \qquad (x_1 \cdots\cdots x_m) G \begin{pmatrix} x_1 \\ \vdots \\ x_m \end{pmatrix} \equiv S \geqq 0$$

$$|F| \neq 0$$

適当なる直交行列 P によつて F を

$$P'FP = \begin{pmatrix} \lambda_1 & & 0 \\ & \ddots & \\ 0 & & \lambda_m \end{pmatrix} \qquad \lambda_i > 0$$

ならしめる事が出來る。

$$H = P \begin{pmatrix} \sqrt{\lambda_1} & & 0 \\ & \ddots & \\ 0 & & \sqrt{\lambda_m} \end{pmatrix} P'$$

とおけば明らかに

$$(25) \qquad F = H^2$$

となる。 又 G は (22) から容易に

$$(G) = (L'L + M'M)$$

なる形にあらはされる。 しからば (22) は

$$E^2(x_1 \cdots\cdots x_m)H = (x_1 \cdots\cdots x_m)GH^{-1}$$

$$= (x_1 \cdots\cdots x_m)HH'GH^{-1}$$

となり

$$H^{-1}GH^{-1} = H^{-1}(L'L + M'M)H^{-1}$$

$$= (LH^{-1})'LH^{-1} + (MH^{-1})'MH^{-1}$$

で対稱行列である。 この対稱行列の固有値が正である事は容易に分る。

前にも一寸述べた如く (22) は $E^2 = 0$

$$x_1 = x_2 = \cdots\cdots = x_n$$

なるとき解を有するから，これ以外の最小の固有値に対するベクトルが求むる解である。