

処理効率よりみた分類方式の比較*

水野幸雄 加藤隆二**

データ処理の中の重要な機能にデータの分類機能があり、この分類に関しては処理速度の面からいろいろの考察が加えられて来た。ここでは各分類方式についてその処理効率を考え、それらを比較することによって最も能率的な分類方法の条件を求めたものである。

1. 分類方式

現在用いられている分類方式には、桁分類法¹⁾、併合分類法²⁾、挿入分類法³⁾、交換分類法³⁾などがある。このうち最も広く用いられているのは桁分類法と併合分類法であり、大容量のデータに関してはこれらが最も多く用いられている。そこで、ここではその2方法について、その処理能力などを比較検討する。

1.1 桁分類法

この方法は出力側にポケットを分類数 (m 分類の場合は m 個) だけ置き、まず最小桁のキーについて分類する。ここでテープ分類ならば m 本の出テープができる。次にその上の桁のキーで前の m 本のテープを順に分類して m 本の出テープを作る。これを繰り返すことによって最後に集録したものはキーの大きさの順になっている。カード分類ではこの方式を用いているものがほとんどであり、テープ分類の際にも有用な手段の一つとされている。この方法による入出力回数、比較回数などは次のようになる。ただし分類すべきデータの量を N とし、桁数 r として m 分類を行う。キーの数は m^r である。

入出力回数 : $N_r = N \log_m k_1 N$

比較回数 : $N_r = N \log_m k_1 N$ (1.1.1)

磁気記録装置: $1+m$ 台

ここで $k_1 = m^r / N$ であり、分類すべき全ポケットの数とデータの数との比を表わす。

1.2 併合分類法

この方法は、主に磁気テープ分類に用いられる方法

である。入力側に m 本、出力側に n 本の磁気テープを持ち、各項目のキーを比較し、順次大きさの順に並ぶように分類して行くもので、例えば $m=n=2$ の場合には次のような規則に従う。

テープ A から読出されたキーを X, B から読んだものを Y, その前に出たテープ側へ移されたものを L とする。そのとき

X > Y > L ならば L と同じテープへ Y を移す。

Y > X > L ならば L と同じテープへ X を移す。

X > L > Y ならば L と同じテープへ X を移す。

Y > L > X ならば L と同じテープへ Y を移す。

L > X > Y ならば L と反対のテープへ Y を移す。

L > Y > X ならば L と反対のテープへ X を移す。

このとき下に示すような情報の移動があり、小さい順に並んだ一種のデータ (これは糸連と呼ばれる) の数は前の $1/2$ になり、これを繰返すことにより 1 本の糸連にまとめることができる。

Tape A 047, 482, 148, 126, ……

Tape B 315, 736, 945, 005, 342, 245, ……

Tape C 047, 315, 482, 736, 945, 126, ……

Tape D 005, 148, 342, ……

この糸連の分布に関しては飯島、木沢の理論^{4,5)}があり、それによれば無作為に並べられた数の作る糸連の数の平均値は $N/2$ であり、併合分類の途中で並べ換えによる糸連のつながる確率を無視すれば分類回数は

$$S = 2 \log_{(mn)} \frac{N}{2} \quad (1.2.1)$$

となる。

$m=n=2$ の場合には上に述べたような比較が行われるので、項目当りの比較回数は $8/3$ 回となる〔註1〕。故に桁分類の場合と同様に入出力回数、比較回数を出すと

(i) $m=n=2$ の場合

比較回数 $\frac{8}{3} N \log_2 \frac{N}{2}$

入出力回数 $N \log_2 \frac{N}{2}$

* Comparison between radix and merge sorting methods based on their processing efficiencies, by Yukio Mizuno and Ryuzi Kato (Nippon Electric Company, Ltd., Tokyo.)

** 日本電気株式会社

磁気記録機	4台
(ii) $m=n=3$ の場合	
比較回数	$3N \log_3 \frac{N}{2}$
入出力回数	$N \log_3 \frac{N}{2}$
磁気記録機	6台

現在用いられている高速分類機においては入出力回数を削減するために内部記憶装置を用い一旦高速記憶回路にデータを入れて、内部分類を行なった後にテープへ出すことが考えられているが、この方法を用いると記憶容量の半分だけのデータが分類できるとして、

$$\begin{aligned} \text{比較回数} & \quad \frac{8}{3} N \log_2 \frac{N}{2} \\ \text{入出力回数} & \quad N \left(\log_2 \frac{N}{M} + 1 \right) \\ \text{内部入出力回数} & \quad N \left\{ \log_2 \frac{M}{2} + 1 \right\} \end{aligned}$$

ただし M は記憶容量

となる。

ここで桁分類の場合と比較する前に、分類途上における糸連の分布およびそのつながる確率について考えてみる。

1.3 糸連の分布

ここではデータの項目 N が十分大きなものとして考えることとする。まず糸連の長さ l に着目して考えよう。

無作為に並べられた項目の作る糸連の長さの分布は次のように与えられている⁵⁾。

$$P_{(l)} = \frac{2(l+1-1)}{(l+2)!} \quad (\text{註2}) \quad (1.3.1)$$

この式は、重複を許さない順列、すなわち同じキーの値を複数個の項目が持たない場合についてのみ成り立つものであるが、キーの数が十分大きい場合には、(1.3.1)式の重複を許す順列に対しても近似的に用いることができる。実際問題として併合分類の分類速度(比較回数)に直接関係するのは分類前の糸連の数だけで、それから後の糸連を統合するための比較回数は、重複を許さない場合と同じことである。そして分類前の糸連の分布が重複を許さない場合と異なるのは同じキーの値を持つ項目が互に隣り合う場合だけで、これはキーの数が適当に大きく、(1.1.1)式の k_1 が0でない値をとる限りほとんど無視してもさしつかえない。

この(1.3.1)式で示される分布が、併合分類を行う場合にどのように変わるかは興味深い。それは糸連

のつながる確率とも関連し、分類の各段階で糸連のつながる確率がどのように変わるかを示唆するからである。

いま(1.2.1)式において $m=n$ とすると、分類回数 S は

$$S = \log_m \frac{N}{2} \quad (1.3.2)$$

となるが、分類途上における糸連のつながる確率が無視できず、ほぼ一定値 p であると見なされる場合には1回の分類によって糸連の数は $(1-p)/m$ 倍になるので、分類前の糸連の数を $N/2$ と置けば $N(1-p)/2m$ 個の糸連に統合されることになる。これを S 回繰り返すことによって1本の糸連に統合されると考えられるから、分類回数 S に関して次の式を得る。

$$\begin{aligned} \frac{N}{2} \cdot \left(\frac{1-p}{m} \right)^S &= 1 \\ \therefore S &= \frac{\log N/2}{\log m/1-p} \quad (1.3.3) \end{aligned}$$

この(1.3.3)式を用いると、併合分類の場合の入出力回数、比較回数は

(i) $m=n=2$ のとき

$$\begin{aligned} \text{比較回数} & \quad \frac{8}{3} N \frac{\log N/2}{\log 2/1-p} \\ \text{入出力回数} & \quad N \frac{\log N/2}{\log 2/1-p} \end{aligned}$$

となり、 $m=n=3$ のときも同様にして求まる。このように p によって比較回数、入出力回数の異なることは充分注意しなければならない。

第 n 回の併合分類を行うことによって得られる糸連の分布を $P_{n(l)}$ とすれば、そのとき長さ l の糸連は、長さ i の糸連と長さ $(l-i)$ の糸連を併合してできると考えられる。このときの同時確率は、 $P_{n-1(i)} \cdot P_{n-1}(l-i)$ となるから、次の漸化式が成立する。

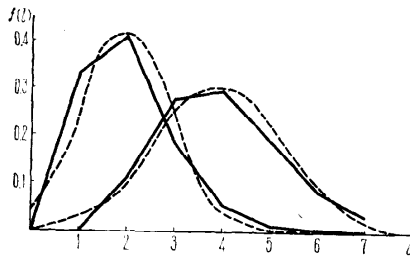
$$P_{n(l)} = \sum_{i=1}^{l-1} P_{n-1(i)} P_{n-1}(l-i) \quad (1.3.4)$$

たとえば第1回の併合分類を行った後の糸連の分布 $P_{1(l)}$ は次のように表わされる。

$$\begin{aligned} P_{1(l)} &= \sum_{i=1}^{l-1} P_{(i)} P_{(l-i)} \\ &= \sum_{i=1}^{l-1} \frac{4(i^2+i-1)\{(l-i)^2+(l-i)-1\}}{(i+2)!(l-i+2)!} \\ &= \frac{4(l+2)}{(l+4)!} \{2^l(l^3-13l+4) + l^3+8l^2+14l-7\} \quad (1.3.5) \end{aligned}$$

$P_{(l)}$ 、および $P_{1(l)}$ の分布を図に表わしたものが第1.1図である。

次に(1.3.4)式であたえられた分布において、



第 1・1 図

l および l^2 の期待値を $E_n(l)$, $E_n(l^2)$ で表わすと、

$$E_n(l) = \sum_{l=1}^{\infty} l P_n(l) \tag{1.3.6}$$

$$E_n(l^2) = \sum_{l=1}^{\infty} l^2 P_n(l)$$

となるから、(1.3.4) 式より次のようになる。

$$\begin{aligned} E_n(l) &= \sum_{l=1}^{\infty} l \sum_{i=1}^{l-1} P_{n-1}(i) P_{n-1}(l-i) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (i+j) P_{n-1}(i) P_{n-1}(j) \\ &= 2 \sum_{i=1}^{\infty} i P_{n-1}(i) = 2 E_{n-1}(l) \end{aligned}$$

$$\begin{aligned} E_n(l^2) &= \sum_{l=1}^{\infty} l^2 \sum_{i=1}^{l-1} P_{n-1}(i) P_{n-1}(l-i) \\ &= \sum_{i=1}^{\infty} \sum_{j=1}^{\infty} (i+j)^2 P_{n-1}(i) P_{n-1}(j) \\ &= 2 \sum_{i=1}^{\infty} i^2 P_{n-1}(i) + 2 \left\{ \sum_{i=2}^{\infty} i P_{n-1}(i) \right\}^2 \\ &= 2 E_{n-1}(l^2) + 2 E_{n-1}^2(l) \end{aligned}$$

$$\begin{aligned} \therefore E_n(l) &= 2 E_{n-1}(l) \\ E_n(l^2) &= 2 E_{n-1}(l^2) + 2 E_{n-1}^2(l) \end{aligned} \tag{1.3.7}$$

ただし、この場合には糸連のつながる確率を無視している。

一方無作為に並べられた糸連の長さの分布における $p(l)$ の一次、二次のモーメントは次のようになる。

$$\left. \begin{aligned} E(l) &= \sum_{l=1}^{\infty} \frac{2l(l^2+l-1)}{(l+2)!} = 2 \\ E(l^2) &= \sum_{l=1}^{\infty} \frac{2l^2(l^2+l-1)}{(l+2)!} = 4e-6 \end{aligned} \right\} \tag{1.3.8}$$

そこで (1.3.8) 式を初期条件として (1.3.7) 式を解くと、次式を得る。

$$\begin{aligned} E_n(l) &= 2^{n+1} \\ E_n(l^2) &= 2^n(4e-10) + 2^{2(n+1)} \end{aligned} \tag{1.3.9}$$

中心極限定理によれば独立な確率変数の多数個の和の分布は近似的に正規分布をなすことが知られている。この定理により分類途上における糸連の分布 $P_n(l)$ は n が大きくなるに従って正規分布に近づく。

すなわち近似的に次の式が成立する。

$$P_n(l) \doteq P_{Nn}(l) = \frac{1}{\sqrt{2\pi\sigma_n^2}} e^{-\frac{(l-m_n)^2}{\sigma_n^2}}$$

(1.3.9) より

$$P_{Nn}(l) = \frac{1}{\sqrt{2^{n+1}\pi(4e-10)}} e^{-\frac{(l-2^{n+1})^2}{2^n(4e-10)}}$$

$$\tag{1.3.10}$$

第 1・1 図には点線で正規分布に近似した場合の確率密度関数が書かれてある。

1.4 糸連のつながる確率

前節の結果によって我々は分類途上における糸連の分布が (1.3.10) 式で近似されることを知った。そこで次に糸連のつながる確率がどのようになるかを考えることによって糸連の併合される状態、また、その回数などを確めることができる。

糸連の分布を $P(l)$ とすれば、無作為に並べ替えたときの糸連のつながる確率は、次の式で表わされる。

$$q = \sum_{m=1}^{\infty} \sum_{l=1}^{\infty} \frac{l! m!}{(m+l)!} P(l) P(m) \tag{1.4.1}$$

n 段階の分類を終った糸連につながる確率を q_n とし、 n を適当に大きなものとすれば、次のようになる。

$$\begin{aligned} q_n &= \sum_{l=1}^{\infty} \sum_{m=1}^{\infty} \frac{l! m!}{(m+l)!} \\ &= \frac{1}{2^{n+1}\pi(4e-10)} e^{-\frac{(l-2^{n+1})^2 + (m-2^{n+1})^2}{2^n(4e-10)}} \end{aligned} \tag{1.4.2}$$

いま、ここで適当に大きな数 M を導入し、(1.3.10) 式において $l=M$ の点で Tschebyscheff の不等式を用いれば、

$$Pr\{l \leq M\} \leq \frac{1}{2} \frac{2^n(4e-10)}{(2^{n+1}-M)^2}$$

$$\lim_{n \rightarrow \infty} \frac{1}{2} \frac{2^n(4e-10)}{(2^{n+1}-M)^2} = 0$$

$$\therefore \lim_{n \rightarrow \infty} Pr\{l \leq M\} = 0 \tag{1.4.3}$$

すなわち、如何に大きく M をとっても $P_{Nn}(l)$ において、 $l \leq M$ となる確率は n を大きくするに従って 0 に近づく。(1.4.2) 式において $l! m! / (m+l)!$ は m, l が大きくなるに従って 0 に近づくから、(1.4.3) 式の条件から次の式が得られる。

$$\lim_{n \rightarrow \infty} q_n \leq \frac{\{M!\}^2}{\{2M!\}} \tag{1.4.4}$$

この式の右辺で $M \rightarrow \infty$ とすると 0 となるから、 q_n は 0 に近づくことがわかる。

この近づく度合を求めるために、一次の近似として糸連の長さをその平均値で置き換えることとする。その時には

$$q_n \doteq \frac{\{E_n(2I)\}^2}{E_n(I)!} = \frac{\{2^{n+1}\}^2}{2^{n+2}} \left. \vphantom{q_n} \right\} (1 \cdot 4 \cdot 5)$$

$$q \doteq \frac{1}{6}, \quad q_1 \doteq \frac{1}{70}, \quad q_2 \doteq \frac{1}{12870}$$

このうち q の値については飯島氏により $1/6$ であることが証明されている⁹⁾。

(1・4・5) からわかるように糸連のつながる確率は急速に減少する。そして実際分類を行う場合には最初に並べ換えてから分類を行うことはないから、最初の分類段階で q_1 の確率でつながるのが最大のつながる確率である。それ故、次のことが言える。すなわち分類途上における糸連のつながる確率は分類が進むにつれて急速に減少して 0 に収斂する。そのため分類途上において糸連のつながる確率は実用上無視してさしつかえない。

故に次章ではこの考えに基づき、(1・3・3) 式で $P=1$ とした式、すなわち (1・3・2) 式を用いて論を進めることとする。

2. 分類方式の比較

前節で代表的な 2 種の分類方式について説明し、その処理回数を算出した。これら分類法はそれぞれ一長一短があり、これを定量的に比較するためには何らかの尺度が必要になる。それは処理速度、装置の大きさなどを含んだものになろう。これらの尺度についてはいろいろの考え方ができるが、最も経済的な分類という点から見れば、処理時間と装置の大きさ（費用または素子の数）の積を最小にすることが考えられる。そこでこの積を W で表し、処理効率を表わす尺度とする。

$$W = (N_t t_t + N_c t_c)(am + b) \quad (2 \cdot 1)$$

ただし N_t ; 入出力回数 t_t ; 入出力時間
 N_c ; 比較回数 t_c ; 比較時間
 m ; 磁気記録機台数
 a ; 磁気記録機の大きさを表わす定数
 b ; 計算機本体の大きさを表わす定数

この値を第 1 章の結果から各方式について求めると

(i) $m=2, n=2$ の場合の併合分類

$$W_{m2} = \{8/3 t_c + t_t\} N \log_2 \frac{N}{2} \cdot (4a + b)$$

(ii) $m=3, n=3$ の場合の併合分類

$$W_{m3} = (3 t_c + t_t) N \log_3 \frac{N}{2} \cdot (6a + b)$$

(iii) 内部記憶装置を用いる場合

$$W_{i2} = \left\{ 3 t_c \log_2 \frac{N}{2} + t_t \left(\log_2 \frac{N}{M} + 1 \right) \right.$$

$$\left. + t_c' \left(1 + \log_2 \frac{M}{2} \right) \right\} (4a + b)$$

ただし t_c' ; 内部入出力時間

(iv) 桁分類 (m 分類)

$$Wdm = (t_c + t_t) N \log_m k_1 N \cdot \{(m+1)a + b\}$$

ここで計算を簡単にするために次のような係数を導入する。

$$k_2 = b/(a+b), \quad k_3 = t_c/t_t$$

すなわち k_2 は計算機本体と磁気記録機とのうち、計算機本体の大きさの割合であり、処理速度のみを問題にする場合には $k_2=1$ 、磁気記録機の台数を主に考える場合には $k_2=0$ となる。また k_3 は比較時間と入出力時間の比であり、入出力時間だけを問題にする場合には $k_3=0$ となる。

このとき、各方式の W を求めると、

$$W_{m2} = t_t N (a+b) (4 - 3k_2) \left(1 + \frac{8}{3} k_3 \right) \log_2 \frac{N}{2}$$

$$W_{m3} = t_t N (a+b) (6 - 5k_2) (3k_3 + 1) \log_3 \frac{N}{2} \quad (2 \cdot 2)$$

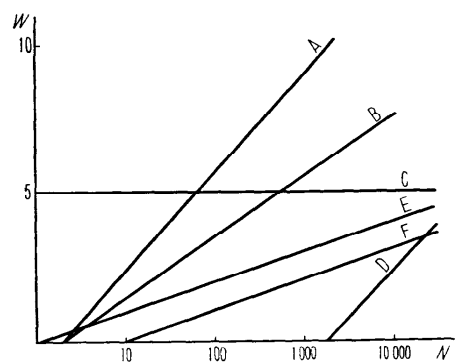
$$Wdm = t_t N (a+b) (1 + m - mk_2) (k_3 + 1) \log_m k_1 N$$

$$W_{i2} = t_t N (a+b) (4 - 3k_2) \left\{ 3k_3 \log_2 \frac{N}{2} + k_3' \log_2 M + \log_2 \frac{2N}{M} \right\}$$

ただし $k_3' = t_c'/t_t$

(2・2) 式によって k_1, k_2, k_3, N 等をあたえることにより W を最小にする分類方式が求まる。

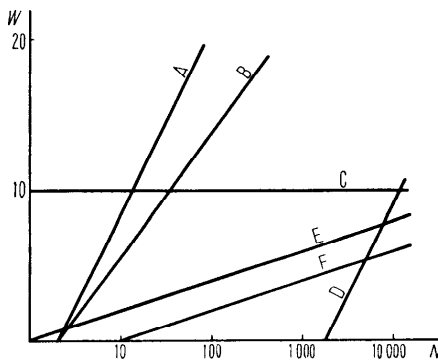
まず $k_2=1, k_3=0$ の場合について考えると、そのときの各方式についての W の値を図示したものが第 2・1 図である。この中で処理時間は一項目当りの処理時間で表わしてある。



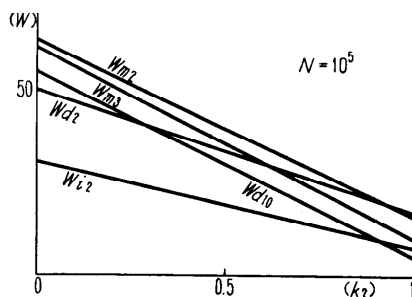
第 2・1 図

- A; $m=n=2$ の場合の併合分類
 - B; $m=n=3$ " "
 - C; $r=\text{constant}$ の場合の桁分類 ($r=10^5$)
 - D; 内部記憶装置の容量を, 1,000 語とした場合の内部分類
 - E; $k_1=1$ の場合の桁分類 (10 分類)
 - F; $k_1=10$ の場合の桁分類 (10 分類)
- すなわち処理速度だけを考えると
- a. 10分類した場合の速度は併合分類の場合より著るしく速い。(特に分類ポケットの小さい場合)
 - b. しかし内部分類法を用いると母数の少ない場合には威力を発揮する。
 - c. 併合分類と桁分類の平衡点は $10^5 \sim 10^6$ の所にある。

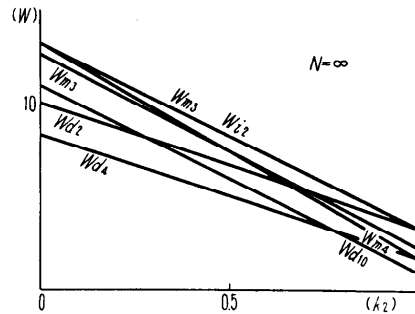
次に入出力時間に比して比較時間が無視できない場合を考える。このとき k_3 は 0 でない値を持つ。そして現在用いられている情報処理装置では $k_3 < 1$ と考えてよいため、今極端な場合として $k_3 = 1$ とし、その時の W の値を図に示す。その結果は第 2.2 図のようになり、桁分類の優位性がさらに増すことがわかる。



第 2.2 図



第 2.3 図



第 2.4 図

最後に項目数が第 2.1 図の場合の平衡点にあるとき、すなわち $N=10^5$ のときと、 $N \rightarrow \infty$ と考えたときのそれぞれについて、計算機本体と磁気記録機の大きさの割合を示す定数 k_2 によって W の値がどのように変わるかを図示すると第 2.3 図、第 2.4 図を得る。これによれば

- (a) $N=10^5$ 程度までは内部分類が非常に速い。
- (b) 桁分類において、 k_2 が小さいときには $N=10^5$ の場合は $m=2$ 、すなわち二分類の最適だが N を限りなく大きくすると $m=4$ のときが最適になる。
- (c) 併合分類において、 $N=10^5$ の場合 $m=3$ のときが良いのは $k_2 > 0.63$ の範囲でそれを過ぎると $m=4$ の方が良くなる。特に N が限りなく大きいときには k_2 の値に応じて最適方式が $m=4$ から $m=\infty$ にまで変わる。

ここで桁分類および併合分類において k_2 の変化に応じてどのように最適処理方式が変わるかを見るために、(2.2) 式より m を連続変数と考えて、

$$Wd(x) = t_c N(a+b)(1+x-xk_2) \log_x k_1 N \quad (2.3)$$

$$Wm(x) = t_c N(a+b) \{2x - (2x-1)k_2\} \log_x \frac{N}{2} \quad (2.4)$$

(2.3) 式を x で微分して 0 と置くと、

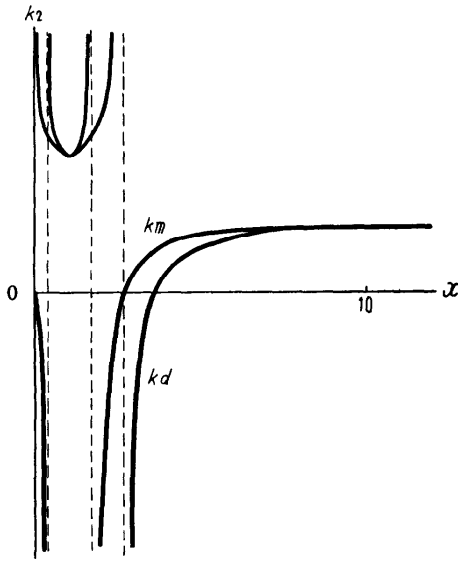
$$(1-k_2) \log x - (1+x-xk_2) \frac{1}{x} = 0$$

$$\therefore k_2 = 1 - \frac{1}{x(\log x - 1)} \quad (2.5)$$

(2.5) 式は k_2 に対する分類数の最適値 $m=x$ を表わすものである。

同様に (2.4) 式より

$$2(1-k_2) \log x - \frac{1}{x} \{2x - (2x-1)k_2\} = 0$$



第 2・5 図

$$k_2 = 1 + \frac{1}{2x(1 - \log x) - 1} \quad (2.6)$$

(2.5), (2.6) 式は、 $x > 1$ で不連続な増加関数なので、 k_2 の値によって最適方式は第 2・5 図に示すように徐々に変わる。第 2・5 図において k_d は桁分類の場合の k_2 の曲線を示し、 k_m は併合分類の場合を示す。特に $k_2 = 0$ のときには桁分類方式では $m = 4$ 、併合分類方式では $m = 3$ が最適となり、 m の値がこれより小さい場合には、最適分類方式になり得ないことを示している。ただし併合分類方式では m が大きくなるにつれて比較回数が増すため、実際には k_3 の値によって最適条件が決まる。(2.2) 式より

$$W_{m2} < W_{m3}$$

の条件を求めると $k_3 > 0.31$ となり、比較時間が入力時間の 0.31 倍より大きければ入力テープを 2 本ずつ用いた併合分類が最適方式となる。

4. 結 論

前節に述べた結果により、与えられた分類システムに対し、項目数 N , k_1 , および情報処理装置の能力を表わす係数 k_2, k_3 を知ることによって最も効率の高い分類方式を決定することができた。しかし、これには N が十分大きいという制約があり、また最適とはいうものの、他と比較した上での最適であり、最上のものであるとはいえない、などの問題を残している。

この最上のものに近づけるためには、各方式の分類途上でテープの持つ情報量を分析し、情報の利得の度合から最も損失の少ない分類方法を定義することが望ましく、この点が今後の課題となろう。

(昭和 36 年 7 月 18 日受 付)
(昭和 36 年 10 月 30 日再受付)

参 考 文 献

- 1) 相磯秀夫: SORTING について, 電子計算と制御, (1959 年 9 月)
- 2) A.S. Douglas: Techniques for Recording of, and Reference to Data in a computer, Computer Journal Apr. (1959)
- 3) P. Hildebrandt and Harold Isbitz: Radix Exchange—an internal sorting method for Digital Computer, JACM, Apr. (1959)
- 4) 木沢誠: テープによる並べ替えの理論 I, 電試集報, (1959 年 11 月)
- 5) 飯島泰蔵: テープによる並べ替えの理論, 電試集報, (1960 年 5 月)

[註 1] この値は次のようにして求まる。すなわち L, X, Y の 3 種を比較する場合、2 回の比較で済むときと 3 回の比較を必要とする場合がある。たとえば、 $L < X, X < Y$ ならばもはや L と Y とを比較する必要はない。このような場合の起る確率は $1/3$ だから、平均比較回数は $8/3$ となる。

[註 2] この値は任意形の連続分布関数の考えからも求めることができる。

糸連の長さが l であるということは次のような状態にあることを意味する。ただし $X_1, X_2, X_3, \dots, X_l, Y_1, Y_2$ は各項目のキーの大きさとする。

$$Y_1 > X_1 < X_2 < X_3 < \dots < X_l > Y_2 \quad (1)$$

または

$$Y_1 < X_1 > X_2 > X_3 > \dots > X_l < Y_2$$

ここに

$$X_i \neq X_j \neq Y_k \quad (i, j = 1, 2, \dots, l; k = 1, 2,)$$

すなわち重複を許さない場合についてのみ考えることとする。

これらの X, Y は、適当な分布を持つ母集団から抽出した要素であると考えられるから、この母集団の確率密度関数を $f(x)$ とおくことができる。すなわち

$$P\{x = x_j\} = f(x_j) dx$$

$$\int_{-\infty}^a f(x) dx = F(a) \leq 1$$

このとき (1.3.1) 式を満たす確率は次式によってあたえられる。

$$P(\omega) = \frac{1}{A} \int_{-\infty}^{\infty} \int_Y^{\infty} \int_{-\infty}^{X_1} \int_{-\infty}^{X_2} \dots \int_{-\infty}^{X_{l-1}} \int_{X_l}^{\infty} f(x_1) f(x_2) \dots f(x_l) f(y_1) f(y_2) \dots dY_1 dX_1 dX_2 \dots dX_l dY_2 \quad (2)$$

ここに

$$A = \int_{-\infty}^{\infty} \int_Y^{\infty} f_{(Y_1)} f_{(X_1)} dY_1 dX_1$$

これは次のように変形される。

$$A = \int_{-\infty}^{\infty} f_{(Y_1)} [F_{(\infty)} - F_{(Y_1)}] dY_1 \\ = \int_{-\infty}^{\infty} f_{(Y_1)} dY_1 - \int_{-\infty}^{\infty} f_{(Y_1)} F_{(Y_1)} dY_1 = -\frac{1}{2}$$

$$P_{(l)} = 2 \int_{-\infty}^{\infty} \int_Y^{\infty} \int_{-\infty}^{X_1} \dots \int_{-\infty}^{X_{l-1}} f_{(Y_1)} f_{(X_1)} \dots \\ \dots f_{(X_l)} dY_1 dX_1 \dots dX_{l-1} - 2 \int_{-\infty}^{\infty} \int_Y^{\infty} \dots \\ \dots \int_{-\infty}^{X_l} f_{(Y_1)} f_{(X_1)} \dots f_{(X_l)} f_{(Y_2)} dY_1 dX_1 \dots \\ \dots dX_l dY_2$$

このうち右辺第1項を $q_{(l)}$ とおけば、第2項は $q_{(l+1)}$ となり、次式を得る。

$$p_{(l)} = q_{(l)} - q_{(l+1)} \quad (3)$$

ここで $q_{(l)}$ を求めると、

$$q_{(l)} = 2 \int_{-\infty}^{\infty} \int_Y^{\infty} \int_{-\infty}^{X_1} \dots \int_{-\infty}^{X_{l+2}} f_{(Y_1)} f_{(X_1)} \dots \\ \dots f_{(X_{l-1})} F_{(X_{l-1})} dY_1 dX_1 \dots dX_{l-1}$$

$$= 2 \int_{-\infty}^{\infty} \int_{Y_1}^{\infty} \int_{-\infty}^{X_1} \dots \int_{-\infty}^{X_{l-i+1}} f_{(Y_1)} f_{(X_1)} \dots \\ \dots f_{(X_{l-i})} \left\{ \frac{1}{i!} - F_{(X_{l-i})} \right\} dY_1 dX_1 \dots dX_{l-i} \\ = 2 \int_{-\infty}^{\infty} f_{(Y_1)} \frac{1}{l!} [F_{(\infty)}^l - F_{(Y_1)}^l] dY_1 \\ = \frac{2l}{(l+1)!}$$

(1・3・3) 式より

$$P_{(l)} = \frac{2(l^2 + l - 1)}{(l+2)!} \quad (4)$$

故に糸連の長さの分布は確率密度関数の形に無関係に(4)式で表わされる。これは飯島氏の得た結果⁴⁾と全く一致している。

【註3】厳密にいえば、糸連のつながりをここで考えなければならぬが、後述するように糸連のつながる確率は非常に小さいので、ここでは無視して考える。

電力潮流問題のデジタル計算*

泥堂多積** 茂木一雄** 乗松立木***
藤木久男**** 小林徳也****

1. 緒言

計数型電子計算機の進歩と普及に伴い、電気事業における電子計算機の利用度は急速に進展し、その応用分野は技術的な面に限っても極めて広く、従来からの最も有力な武器であった交流計算盤にとって代りつつあるのが世界的情勢である。

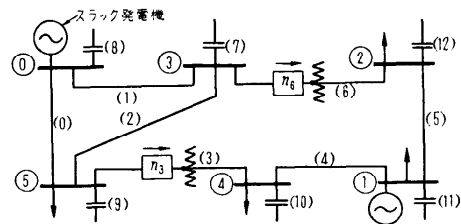
ところで、交流計算盤が最も多く利用されるのは電力潮流問題(使用時間の約85%を占めるといわれる)であるので、著者等は国産電子計算機(NEAC-2203)を用いて電力潮流問題をデジタル計算することを試み、そのための実用的なプログラムを開発し例題に適用したところ、充分実用になる成果を得たのでその概

要を紹介する。

2. 電力潮流問題のデジタル計算法¹⁾

2.1 電力潮流問題の性格

電力回路網の簡単な例(発電機端子2, 負荷端子4)



第1図 例題の回路図

①: 端子番号 (①~②: 発電機端子, ③~⑤: 負荷端子)
(j): 枝番号

n_j : off-nominal 巻線比(第2図参照)

* Digital computer solution of power-flow problems, by Tatsumi Deido, Kazuo Mogi Tatsuki Norimatsu, (Electrotechnical Laboratory), Hisao Fujiki, Tokuya Kobayashi (Yokogawa Electric Works Co., Ltd.)

** 電気試験所, *** 元電気試験所, **** 横河電機