

確率のある組合せ分割問題

(数学モデルとその解法)

A Division Problem of Combinations with Probabilities
(Mathematical Model and it's Solution)

福 永 純 三*
平 田 隆 教**

To ask many optimization problems, we need to calculate the optimization value in the combinations. We have asked the optimized parts kind number in the combinations conducted by an integer linear programming. In this report correlating such the optimization problem, we explain an optimization problem to divide the combinations into the several groups. The combinations are resembled to the objects of a class with several attributes which kinds have the occurring probabilities and when the number of objects with same attribute kind is divided by the total objects number, the quotient is equal to the occurring probability of the attribute kind. Satisfying such conditions, we aim to ask the optimized combinations division groups evaluated by a function. In the ideal explained problems, it happens hardly to divide into the combinations groups which combinations numbers are integers. So it's necessary to ask the combinations groups which combinations numbers are reals and we approximate the real numbers into the integer numbers. As samples of the functions, we state the function averaging the combinations number belonging to the each divided group and the general function. We present a simple example where the number of the each divided combinations group is approached to the mean calculated by the combinations number and the groups number.

Key Words: Optimization Problem, Combinations Division, Probability

* 九州情報大学経営情報学部

** (株) アライドエンジニアリング

1 はじめに

最適化問題を考える場合、多くの組合せの中に最適値が存在する組合せを求める例は多い。整数線形計画問題に帰着される最適化問題はそのような場合に相当すると言える。

筆者らは、今までに部品種類数の最適化問題に取り組んできた^{[1],[2]}。これは、製品を構成する部品種類数を、製品を製造・販売をする上で、最適な部品種類数に設定する最適化問題である。その最適部品種類数問題の一環で、同じ要素区分から構成される組合せの生起確率を保つことを条件の一つにして、ある評価のもとで最適な組合せ群を求める問題（組合せの最適分割と言える。）を検討してきた。このレポートでは、その問題を若干一般化した記述を取り入れながら説明をする。

すなわち、今ここに数種の要素があり、各要素には数種の区分がある場合、その区分同士から構成される組合せ全体に対して、ある評価軸からの最適値を与える組合せの分割群を求める問題である。各要素にある区分の種類にはそれぞれ生起確率があり、種類毎の生起確率を確保しつつ最適値を評価をする問題である。

身近な事に例えて説明をする。オブジェクト指向におけるクラスとその属性、各属性の種類が実現されることにより生じるオブジェクトが一つ一つの組合せに当てはまると考える。あるいは、集約クラスやコンポジションとそのオブジェクトに置き換えて考えてもよい。組合せをオブジェクトと解釈することにする。人をクラスとして取り上げるなら、性別や血液型、出身県等が属性となり、性別や血液型、出身県等の各属性の違いが組合わさって組合せ（人、オブ

ジェクト）が実現する。

ところで、いま人の集団で、各属性の種類には予測があり、性別に関して男女の生起確率はそれぞれ0.5、血液型のそれはO型0.3、A型0.4、B型0.3、AB型0となる100人の集団が居たとする。その集団は性別と血液型の違いから6グループに分割されてグループ全体として何らかの評価をされるがグループ全体でみれば、性別と血液型の割合は保たれているが各グループの人数は不確定とする。その時に、グループに対するある評価に最適な組合せのグループ分けはどのようなになるか、あるいは、男性でO型のグループは10人であることが既に確定しているとき、評価に最適な組合せのグループ分けを求めようとするものである。そのような問題を解くことをテーマに考える。オブジェクト指向における振る舞いは例えようとするものではないが、強いて言えば、各属性の同じ種類を持つオブジェクトの集団が生起確率を満たすことと、同じ属性の種類からなるオブジェクトが集団として評価されることが、振る舞いとも考えられる。

どのような評価をするかは問題によるものであるが、分割されたそれぞれの組合せの数が平均化される場合と一般的な評価をする場合の解法について説明をする。

2 組合せの表現

オブジェクト指向における属性は、情報処理では普通に使用されており、理解され易いと考え、組合せの仕組みを説明するために今後も属性を使用することとする。

また、この章で説明する内容は、今迄のレポー

ト^[2]の説明と重複する部分もあるが、後述の内容に必要なことを追加していること、この度、“属性”の言葉を使用することから、改めて説明することとする。

2.1 組合せ

今、あるもの(クラス)が、ある数の属性の組合せによって決められ、各属性にはいくつかの種類があることによりいくつかの組合せ(オブジェクト)ができるものとする。今、それらの属性を

$$A^{(i)}, i = 1, \dots, a$$

と表す。属性 $A^{(i)}$ には $t^{(i)}$ 個の種類があるとし、それらを

$$A_j^{(i)}, j = 1, \dots, t^{(i)}$$

と表す。 a 個の属性がもつ、それぞれの種類の組合せを全て(数学で言ういわゆる直積にあたる。)を考えれば、その属性とそれぞれの種類からなる組合せは図1のようにまとめられる。

属性の数およびそれらの種類数は、既に検討してきた最適化問題^[1]について言えば、概ね

$$a \leq 200, t^{(i)} \leq 40, i = 1, \dots, a$$

であった。しかし、本レポートで述べる確率のある組合せの分割問題のみに限れば、これらの

	属性	
	$A^{(1)} \dots A^{(s)}$	
	$A_1^{(1)} \dots A_1^{(a)}$	$\vdots \dots \vdots$
組合 せ	$\vdots \dots \vdots$	$\vdots \dots \vdots$
	$A_t^{(1)} \dots A_t^{(a)}$	

図1 属性の種類組合せ

数より小さい数(現実的な最適化問題が考えられるであろうとする意味で。)と言える。

上のような、 a 個の属性の種類組合せで構成される組合せの数は $t^{(1)} \times \dots \times t^{(s)}$ 個である。これらの組合せの任意の一つを

$$(A_{i_1}^{(1)}, \dots, A_{i_a}^{(a)})$$

と表すことができ、従って、属性の種類組合せ全体は集合

$$A = \{(A_{i_1}^{(1)}, \dots, A_{i_a}^{(a)}) \mid 1 \leq i_j \leq t^{(j)}, j = 1, \dots, a\}$$

と表わせる。さらに表現を簡略にするため、個々の組合せをその属性の種類の下付きの添字の組と同一視することもできる。すなわち、対応

$$A_i = (A_{i_1}^{(1)}, \dots, A_{i_a}^{(a)}) \leftrightarrow \mathbf{i} = (i_1, \dots, i_a)$$

を考えることにより、属性の種類組合せ全体を次のような a 個の自然数の組についての、ある一定な集合として表せる。

$$I_A = \{\mathbf{i} = (i_1, \dots, i_a) \mid 1 \leq i_j \leq t^{(j)}, j = 1, \dots, a\}$$

2.2 禁止と限定の組合せ

本レポートで述べる組合せ問題の実際問題への適用を考慮するとき、考えられる組合せ全ては対象にならない。全体の中のある組合せは禁止や限定という考え方で除外される。現実問題では、この組合せは販売の対象としないことや、この組合せに限って販売する等に当たる。そのことを説明する。

各属性の種類から構成される組合せのうち、属性の内のある数の属性と、それら属性がもつ種類の内のある種類についての組合せは、考慮

の対象として除かれることがあるとする。すなわち、禁止の組合せがあるものとする。禁止の組合せは h 個あるものとし、それぞれについて禁止を規定する属性の種類の集合を

$$H_\alpha, \alpha = 1, \dots, h$$

とする。ここに、集合 H_α は属性の全体 $\{1, \dots, a\} = \{j | j = 1, \dots, a\}$ の部分集合であって、また、各 $j (j = 1, \dots, a)$ について属性との対応 $j \leftrightarrow A^{(j)}$ が定められているとする。禁止の対象となる組合せの種類の規定するためには、 H_α に属する各属性について、属性の種類を特定する必要がある。そこで、 $j \in H_\alpha$ である各 j に対して、 $A_{ij}^{(j)} = A_{ij(a)}^{(j)}$ となる属性のある種類が禁止の対象になるとする。このとき、禁止属性集合 H_α に基づく禁止の対象となるものの組合せ (単に禁止ともいう。) の全体は

$$I_{H_\alpha} = \{\mathbf{i} = (i_1, \dots, i_a) | \mathbf{i} \in I_A, \\ j \in H_\alpha \text{ である各 } j \text{ について } A_{ij}^{(j)} = A_{ij(a)}^{(j)}\}$$

と表せる。また、禁止の全体は

$$I_H = I_{H_1} \cup \dots \cup I_{H_h}$$

である。

属性の種類から組合せ可能な組合せ全体のうち、ある属性とその種類については、ある特定の種類の組合せを含むものだけに限定して考える場合がある。これらは、禁止と反対の考えに立つものであるが、禁止と同様に、以下のように表せる。限定の対象となる属性、すなわち、限定される属性の種類の組合せが d 個あるものとし、それぞれについて限定を規定する属性の集合を

$$D_\beta, \beta = 1, \dots, d$$

とする。ここに、 D_β は属性全体の集合 $\{1, \dots, a\}$ の部分集合である。また、 D_β に含まれる各 j について、 $A^{(j)} = A_{ij(\beta)}^{(j)}$ となる属性のある種類が限定の対象となるとする。このとき、限定される属性の集合 D_β に基づく限定の対象となる属性の種類の組合せ (単に限定ともいう。) の全体は

$$I_{D_\beta} = \{\mathbf{i} = (i_1, \dots, i_a) | \mathbf{i} \in I_A, \\ j \in D_\beta \text{ である各 } j \text{ について } A_{ij}^{(j)} = A_{ij(\beta)}^{(j)}\}$$

と表せる。さらに、限定の全体は

$$I_D = I_{D_1} \cap \dots \cap I_{D_d}$$

となる。

属性の種類の可能な組合せ全体は集合 I_A として表現された。最適化問題での対象となる組合せの種類の全て I_U は、 I_A の部分集合であり、

$$I_U = I_D \cap I_H^c$$

となる。ここに、 I_H^c は集合 I_H の I_A における補集合である。

考える問題にもよるが、各種のデータの多くの場合に対して、繰り返し最適値を求める必要もあり、全体の組合せの中で、如何に計算の時間を短く、禁止の組合せを除き限定となる組合せに絞るかが大切になる。

2.3 組合せの問題

このように I_U として表現できる組合せの中で、最適化問題が考えられることになる。たとえば、属性の種類の組合せに値を持たせて最適化問題が考えられる。その値の持ち方は、組合せを構成する属性の種類の値の総和から組合せの値を考えること等である。実際問題への応用では、たとえば、値はその属性を実現する費用

を意味することになる。

属性種類

$$A_{ij}^{(i)}, i = 1, \dots, a, 1 \leq i_j \leq t^{(i)}$$

の値が

$$c_{ij}^{(i)}, i = 1, \dots, a, 1 \leq i_j \leq t^{(i)}$$

であるとすれば、組合せを構成する属性の種類が

$$(A_{ij}^{(1)}, \dots, A_{ij}^{(a)}), 1 \leq i_j \leq t^{(i)}$$

であれば、その組合せのもつ値は

$$C_i = \sum_{j=1}^a c_{ij}^{(i)} + (\text{一定値}) \cdots (1)$$

とする。このような値を持つとしたとき、組合せ間に順序を設けて組合せ間の代替問題の解決策を述べたものが既に報告しているレポート^[1]である。値の持ち方もいろいろ考えられ、非線形の式で表現される最適化問題であれば計算時間に何らかの工夫を要することになる。

2.4 確率のある組合せ

組合せに確率を考えることについて述べる。

考慮の対象としているものの属性の種類についてその生起確率が決まっている、あるいは予測できるものとする。実際問題で考えるなら、それは、過去からの商品の販売実績であったり、計画している商品の販売予測であったりする。すなわち、属性 $A^{(i)}$, $i = 1, \dots, s$ には、属性の種類によって生起確率があり、次のように表せるとする。

種類	$A_1^{(i)}$	$A_2^{(i)}$	\dots	$A_{t^{(i)}}^{(i)}$	計
確率	$q_1^{(i)}$	$q_2^{(i)}$	\dots	$q_{t^{(i)}}^{(i)}$	1

先に述べたように、ある属性の組合せには禁止や限定がある。組合せは集合 $I_{A:U}$ の要素として表されているが、組合せができる条件に各属性の種類の生起確率を考えることにより、新たな最適化問題が考えられる。全ての組合せが実現可能で、各属性が独立に生起確率を持つとすれば、ある数からなる組合せ全体の分割は、全体を各組合せを構成する属性種類の生起確率の積に比例配分することで決めることが極自然である。しかし、属性の組合せの禁止や限定によって全ての属性の全ての種類の組合せにならないことに加え、独立以外の各属性間の関係を考慮しなければならないある目的をもった組合せの分割問題が考えられることになる。

3 確率のある組合せ分割問題

3.1 問題の説明

今迄に述べた属性の種類の組合せ全体とそれぞれの組合せ数を決める変数を次の表1に示す。

すなわち、

$$m = \prod_{i=1}^a t^{(i)}$$

であり、 x_i , $i = 1, \dots, m$ は、それぞれの属性の種類の組合せ数とする。組合せの全体数を

表1 組合せ全体とその数

No.	$A^{(1)}$	\dots	$A^{(a)}$	変数
1	$i_{11} = 1$	\dots	$i_{a1} = 1$	x_1
2	i_{12}	\dots	i_{a2}	x_2
\dots	\dots	\dots	\dots	\dots
j	i_{1j_1}	\dots	i_{aj_a}	x_j
\dots	\dots	\dots	\dots	\dots
m	$i_{1m} = t^{(1)}$	\dots	$i_{am} = t^{(a)}$	x_m

$$v = \sum_{i=1}^m x_i$$

としたときに、予め v は与えられるとする。

それぞれの No. に対応する組合せの数の集合を

$$X = \{x_j | j = 1, \dots, m\}$$

とする。可能な組合せ全てを列挙したが、禁止や限定を考えるため、

$$X_h = \{x_j = 0 | (i_{1j_1}, \dots, i_{aj_a}) \in I_U^c\}$$

とする。また、予め与えられた値を持つ組合せ、

$$X_g = \{x_j = x'_j | x'_j \text{ は既知}, 1 \leq j \leq m\}$$

があるとする。これは、何かの理由で、強制的に決めたい組合せの数を指示することにあたる。今、

$$\hat{x}_{ij}, i = 1, \dots, a, j = 1, \dots, t^{(i)}$$

を属性 $A^{(i)}$ の、 j 番目の属性種類をもつ組合せの全体数とする。

このレポートで考える組合せの分割問題は、ある評価のもとに、

$$q_j^{(i)} = \frac{\hat{x}_{ij}}{v}, i = 1, \dots, a, j = 1, \dots, t^{(i)}$$

を満たして、

$$X_s = X \cap X_h^c \cap X_g^c$$

を求める問題である。

3.2 分割最適化問題とその解法

組合せ分割の評価として、

- 1) 既に述べた式(1)の値の評価
- 2) 分割された組合せ群の数を平均化する
- 3) 他の評価指標を設ける

が考えられる。

1) に関して言えば、属性実現に要する値を式(1)のみで表すなら、属性の種類の生起確率を組合せ全体で保証できる限りにおいては評価にはならない。生起確率と値との何らかの関係の評価することになれば新たな問題も考えられる。

2) の各群の組合せ数を平均化することの意味は、言葉通りに、それぞれの分割された組合せ群がその数において出来るだけ平均化されることであり、式で表せば次のようになる。

$$z = \sum_{i=1}^m (x_i - \bar{x})^2 + \lambda \left(\sum_{i=1}^m x_i - v \right) + \sum_{i=1}^a \sum_{j=1}^{t_i-1} \mu_{ij} (\hat{x}_{ij} - v q_j^{(i)}) \dots (2)$$

$$\text{ただし、} \bar{x} = \frac{\sum_{i=1}^m x_i}{m}$$

$$\lambda, \mu_{ij}, i = 1, \dots, a, j = 1, \dots, t_i - 1$$

ラグランジェ定数

禁止や既知の組合せにより、必ずしも一意には解けない。また、必ずしも整数の組合せ数で決まらないので、実数の組合せ数から実数に近い整数の組合せ数を求めることになる。

3) に関して言えば、何かの評価の式があり、しかも各属性の種類の生起確率を保ちながら最適解を計算するには、組合せ全体に対する直接の計算で分割を求めることも一つの方法である。組合せ全体への直接の計算を考える場合、生起確率から与えられる条件式の数 j_0 は、禁止等を考慮しなければ、

$$j_0 = \sum_{i=1}^a t^{(i)} - a + 1$$

であり、各組合せの数 $x_j, j = 1, \dots, m$ に関して、組合せを構成する属性の種類の生起確率が、

$$\{q_{j_1}^{(1)}, \dots, q_{j_a}^{(a)}\}$$

であるとすれば、

$$0 \leq x_j \leq vq_{j_i}^{(i)}, i = 1, \dots, a, j = 1, \dots, m$$

が言える。この条件と、

$$t^{(i)} \geq 2, i = 1, \dots, a$$

$$m = t^{(1)} \times \dots \times t^{(a)} \geq (t^{(1)} + \dots + t^{(a)})$$

であることから、全体数と生起確率に関するの条件式からなる連立方程式、

$$x_1 + \dots + x_m = v$$

$$\hat{x}_{ij} = vq_j^{(i)}, j = 1, \dots, t_i - 1, i = 1, \dots, a$$

により、消去計算をすることより、 $(m - j_0 + 1)$

の数の変数からなる一次方程式

$$\sum_{i=j_0}^m c_i x_i = b \quad \dots (3)$$

を満たす、整数、実数、あるいは（丸めた）整数からなる $\{x_i | i = j_0, \dots, m\}$ の組合せをもとめながら、生起確率条件を満たす最適な組合せを求めることになる。しかし、禁止や限定、既知の値の組合せを考慮に入れるなら、より少ない変数や条件式での計算になる。

ところで、一組の $x_j, j = 1, \dots, m$ を挙げたが、同時に k 個の組、すなわち、

$$x_{jl}, j = 1, \dots, m, l = 1, \dots, k$$

を考えることもできる。しかし、組全体として、また各組としての生起確率に対する考えを取り入れた条件式を立てなければならない。

3.3 計算例

素直で簡単な例ではあるが、実際の計算結果を示す。

表2 属性とその生起確率

属性 1		属性 2		属性 3	
種類	確率	種類	確率	種類	確率
黒	0.1	赤	0.4	青	0.5
白	0.9	茶	0.6	紫	0.5

表3 属性とその組合せ

No.	属性 1	属性 2	属性 3
1	黒	赤	青
2	黒	赤	紫
3	黒	茶	青
4	黒	茶	紫
5	白	赤	青
6	白	赤	紫
7	白	茶	青
8	白	茶	紫

すなわち、あるクラスは3個の属性からなり、第1の属性は {黒、白} からなり、黒の生起確率は0.1、白のそれは0.9であるとする。同様に、第2の属性は {赤、茶} からなり、それぞれ、0.4、0.6である。第3の属性は、{青、紫} からなり、それぞれ、0.5、0.5であるとする。この属性と生起確率をもとに、100の組合せを分割する計算を試みる。

表4 表3に基づく組合せ分割

No.	解 1	解 2	解 3	解 4	解 5	解 6
1	2	0	0	0	2	2
2	2	0	(8)	(8)	2	2
3	3	5	2	(8)	4	2
4	3	5	0	0	2	28
5	18	20	18	16	(0)	(0)
6	18	20	15	16	36	39
7	27	25	30	26	(0)	(0)
8	27	25	27	26	54	27
計	100	100	100	100	100	100

解 1 の組は確率の積で全体の組合せ数100を分割した結果である。解 2 の組は各属性の生起確率をそれぞれ保ちながら各組合せ数を平均に近くなるように分割した結果であり、解 3 の組は No.2 の組合せ数を予め 8 に設定した時に平均化した結果である。また、解 4 の組は、No.2, 3 の組み合わせ数を予め 8 に設定した結果であり、解 5 の組は、No.5, 7 の組み合わせを禁止 (0) に設定した結果である。

計算は、既に述べた平均化の式(2)を直接に解くことによって求めた。解 4 では、属性 1 の条件は既知の値から既に成立していない。しかし、属性 1 の条件に最も良い条件で近づけられている。解 5 では、属性 1、2 の生起確率を優先させた結果になっている。一方、生起確率を満足することは不可能なので、各属性の生起確率に対する誤差を最小化することも考えられる。すなわち、次の式を解くことを考える。

$$z = \sum_{i=1}^a \sum_{j=1}^{t_i-1} (\hat{x}_{ij} - v q_j^{(i)})^2 + \lambda (\sum_{i=1}^m x_i - v)$$

そのようにして計算した結果が解 6 である。

4 おわりに

確率を持つ属性種類から構成される組合せ群 (オブジェクト群) がある。ただし、ある特定の属性種類については、その組合せに禁止も存在する。そのような組合せの群に対して、各属性種類の出現確率を満たすことに加えて、他の関数で表現される条件をも充たして、組合せ群を数種類の群に分割する問題を考えた。その際、各属性種類の出現確率を充たしながら整数個の組合せの群を分割することは、多くの場合、また組合せを構成する属性とその種類の数が増す

ほどに、整数個の組合せからなる数個の群に分割することは困難である。しかし、数個の組合せ群を評価する関心の持てる指標を見出すならば、改めて考察を試みたい。

< 参考文献 >

- [1] 福永純三, 平田隆教, 永村和照; 製品の構成部品種類数最適化問題 (数学モデルとその近似解法), 広島大学工学研究科研究報告, 2002, Vol.51, NO.1, 17-22.
- [2] 二神かほる, 福永純三, 平田隆教, 新中裕, 藤越康祝; 商品の多様化対応における最適化問題 - 自動車商品に関する最適部品種類の自動算出 -, 広島大学経済学論叢, 19(1996)-3・4, 43-68.