（3）「実験計画とその周辺における数理構造の解明とその応用」に関する研究報告

Conectedness and Resistance in Designs – an Overview

Satyabrata Pal
Department of Agricultural Statistics, Faculty of Agriculture
Bidhan Chandra Krishi Viswavidyalaya, Mohanpur, Nadia, West Bengal, Pin-741252, India

The talk presented some important results in the areas of connectedness, resistance, and optimality with special reference to the important contributions (of Prof. S. Pal) in these areas published in different journals. Though in the areas of resistance and optimality, the set-up was block design, in the area of connectedness, the intended set-up was multiway elimination of heterogeneity designs.

At the very outset, the author presented the form of the coefficient matrix (C^{vxv} matrix) in case of a multiway elimination of heterogeneity design published in Pal and Katyal (JISA, 1988). The explicit expression of the C-matrix of the multifactor design (multiway elimination of heterogeneity design) was not available in the literature till then.

The discourse then dwelt on the concepts of connectedness with respect to two factor designs ($f_i \sim f_j$ connected designs) and total connectedness with respect to a single factor. While ($f_i \sim f_j$) connected designs refers to a connected (w.r.t. the factors, i and j) two-way design, total connectedness refers to a factor, say, 't' (or f_t connected), when the parameter effects of the factor 't' are estimated from the whole design after eliminating the effects of all other factors.

In the context of multifactor designs, the estimability of the effects w.r.t. the factor 't' can be obtained if we look at the (t-1) two factor designs of the type (t, i), i = 1,2,.., t-1. The following important theorem in this regard was proved in the lecture (Pal and Katyal, JISA- 1988).

Theorem: In a general multi-way (t-1 way) elimination of heterogeneity set-up given under the t-way model with t-th factor being considered as the treatment factor (also $n_{ij} = r_i r_j^T / n$ is satisfied), if, $n_{ij} = n_{it} r_t^{-ä} n_{tj}$ for i j å (1,2,..,v), the symbols having usual significance, the design is f_t connected, if it is ($f_t \sim f_l$) connected for all l = 1,2,…,t-1.

Example 1:

Column Row	1	2	3	4
1	A	A	C	D
2	A	A	D	C
3	C	D	B	B
4	D	C	B	B

Here t = 3, 3rd factor being treatment factor, 1st factor being row factor, 2nd factor being column factor. The above design satisfies $n_{ij} = r_i r_j^T / n$, and it is also ($f_t \sim f_l$) connected

for all l = 1,2, t = 3, but n_{ij} $n_{it} r_t^{-ä} n_{tj}$, for all (i, j), i j å (1,2), and hence the design is not f_t connected (t =3).

Example 2:

Column Row	1	2	3	4
1	A	A	B	C
2	B	C	A	A
3	C	B	A	A

This design satisfies $n_{ij} = r_i r_j^T / n$, for all i j å (1,2) and also $n_{ij} = n_{it} r_t^{-ä} n_{tj}$, i =j å (1,2), t =3 [and also ($f_t \sim f_l$) connected or all l å (1,2)], so the design is f_t connected (t=3).

Thus from the above theorem total connectedness with respect to a factor can be checked by checking the connectedness of two factor designs which can be built up from the whole design in addition to some conditions (which again can be built up from two factor designs) to be satisfied.

Some results on connectedness for nested R/C designs are obtained (with S. Kageyama and V. Katyal) and the related communication will appear soon.

Next, the author spoke on universal optimality of non-proper designs, the results were obtained in their paper appeared in Communications in Statistics- Theory and Methods (1988). This was the first paper on optimality results in the area of non-proper designs. Several constructions on universally optimum non-proper designs were presented.

The final part of the talk dealt with the concept of resistance in block designs and constructions of resistant and susceptible non-proper variance balance designs were presented. Before arriving at the results, two kinds of variance balanced designs were proposed (Pal and Pal, 1991,JISAS) and their statuses (resistant or susceptible) with respect to the loss of one treatment (all observations on that treatment in the design) were obtained.

References

Pal, S. and V. Katyal (1988): "On Multiway Elimination of Heterogeneity Designs", Journal of Indian Statistical Association, Vol.26, pp 51-58.

Pal, S. and Pal, S.N. (1988): "Non-proper Variance Balanced Designs and Optimality" Communications in Statistics- Theory and Methods, 17 1685-1695.

Pal, S.N. and Pal, S. (1991): "A note on Resistancy of Non-proper Variance Balanced Designs" Journal of Indian Statistical Association, Vol.28, pp 13-18.

Balanced (C_4, C_5)-$2t$-Foil System

Kinki University Kazuhiko Ushio

1. Introduction

Let K_n denote *the complete graph* of n vertices. Let C_k be *the k-cycle*. *The (C_4, C_5)-$2t$-foil* is a graph of t edge-disjoint 4-cycles and t edge-disjoint 5-cycles with a common vertex and the common vertex is called *the center of the (C_4, C_5)-$2t$-foil*. In particular, the (C_4, C_5)-2-foil is called *the (C_4, C_5)-bowtie*. When K_n is decomposed into edge-disjoint sum of (C_4, C_5)-$2t$-foils, we say that K_n *has a (C_4, C_5)-$2t$-foil decomposition*. Moreover, when every vertex of K_n appears in the same number of (C_4, C_5)-$2t$-foils, we say that K_n *has a balanced (C_4, C_5)-$2t$-foil decomposition* and this number is called *the replication number*.

It is a well-known result that K_n has a C_3 decomposition if and only if $n \equiv 1$ or 3 (mod 6). This decomposition is known as *a Steiner triple system*. See Colbourn and Rosa[1] and Wallis[3]. Horák and Rosa[2] proved that K_n has a C_3-bowtie decomposition if and only if $n \equiv 1$ or 9 (mod 12). This decomposition is known as *a bowtie system*.
In this sense, our balanced (C_4, C_5)-$2t$-foil decomposition of K_n is to be known as *a balanced (C_4, C_5)-$2t$-foil system*.

2. Balanced (C_4, C_5)-$2t$-foil decomposition of K_n

Notation. We denote a (C_4, C_5)-$2t$-foil passing through
$v_1 - v_2 - v_3 - v_4 - v_1 - v_5 - v_6 - v_7 - v_8 - v_1$,
$v_1 - v_9 - v_{10} - v_{11} - v_1 - v_{12} - v_{13} - v_{14} - v_{15} - v_1$,
$v_1 - v_{16} - v_{17} - v_{18} - v_1 - v_{19} - v_{20} - v_{21} - v_{22} - v_1$,
...,
$v_1 - v_{7t-5} - v_{7t-4} - v_{7t-3} - v_1 - v_{7t-2} - v_{7t-1} - v_{7t} - v_{7t+1} - v_1$
by
$\{(v_1, v_2, v_3, v_4), (v_1, v_5, v_6, v_7, v_8)\} \cup \{(v_1, v_9, v_{10}, v_{11}), (v_1, v_{12}, v_{13}, v_{14}, v_{15})\}$
$\cup \{(v_1, v_{16}, v_{17}, v_{18}), (v_1, v_{19}, v_{20}, v_{21}, v_{22})\} \cup \ldots$
$\cup \{(v_1, v_{7t-5}, v_{7t-4}, v_{7t-3}), (v_1, v_{7t-2}, v_{7t-1}, v_{7t}, v_{7t+1})\}$.

Theorem. K_n has a balanced (C_4, C_5)-$2t$-foil decomposition if and only if $n \equiv 1$ (mod $18t$).

Proof. **(Necessity)** Suppose that K_n has a balanced (C_4, C_5)-$2t$-foil decomposition. Let b be the number of (C_4, C_5)-$2t$-foils and r be the replication number. Then $b = n(n-1)/18t$ and $r = (7t+1)(n-1)/18t$. Among r (C_4, C_5)-$2t$-foils having a vertex v of K_n, let r_1 and r_2 be the numbers of (C_4, C_5)-$2t$-foils in which v is the center and v is not the center, respectively. Then $r_1 + r_2 = r$. Counting the number of vertices adjacent to v, $4tr_1 + 2r_2 = n - 1$. From these relations, $r_1 = (n-1)/18t$ and $r_2 = 7(n-1)/18$. Therefore, $n \equiv 1$ (mod $18t$) is necessary.

(Sufficiency) Put $n = 18st + 1$, $T = st$. Then $n = 18T + 1$. Construct n (C_4, C_5)-$2T$-foils as follows:

Department of Informatics, Faculty of Science and Technology, Kinki University, Osaka 577-8502, JAPAN. E-mail:ushio@is.kindai.ac.jp Tel:+81-6-6721-2332 (ext. 4615) Fax:+81-6-6730-1320

$B_i = \{(i, i+1, i+3T+2, i+T+1), (i, i+6T+1, i+10T+2, i+15T+3, i+7T+1)\}$
$\cup \{(i, i+2, i+3T+4, i+T+2), (i, i+6T+2, i+10T+4, i+15T+6, i+7T+2)\}$
$\cup \{(i, i+3, i+3T+6, i+T+3), (i, i+6T+3, i+10T+6, i+15T+9, i+7T+3)\}$
$\cup$ ...
$\cup \{(i, i+T, i+5T, i+2T), (i, i+7T, i+12T, i+18T, i+8T)\}$ $(i = 1, 2, ..., n)$.
Decompose each (C_4, C_5)-$2T$-foil into s (C_4, C_5)-$2t$-foils. Then they comprise a balanced (C_4, C_5)-$2t$-foil decomposition of K_n.

Note. We consider the vertex set V of K_n as $V = \{1, 2, ..., n\}$.
The additions $i + x$ are taken modulo n with residues $1, 2, ..., n$.

Example 1. A balanced (C_4, C_5)-2-foil decomposition of K_{19}.
$B_i = \{(i, i+1, i+5, i+2), (i, i+7, i+12, i+18, i+8)\}$ $(i = 1, 2, ..., 19)$.

Example 2. A balanced (C_4, C_5)-4-foil decomposition of K_{37}.
$B_i = \{(i, i+1, i+8, i+3), (i, i+13, i+22, i+33, i+15)\}$
$\cup \{(i, i+2, i+10, i+4), (i, i+14, i+24, i+36, i+16)\}$ $(i = 1, 2, ..., 37)$.

Example 3. A balanced (C_4, C_5)-6-foil decomposition of K_{55}.
$B_i = \{(i, i+1, i+11, i+4), (i, i+19, i+32, i+48, i+22)\}$
$\cup \{(i, i+2, i+13, i+5), (i, i+20, i+34, i+51, i+23)\}$
$\cup \{(i, i+3, i+15, i+6), (i, i+21, i+36, i+54, i+24)\}$ $(i = 1, 2, ..., 55)$.

Example 4. A balanced (C_4, C_5)-8-foil decomposition of K_{73}.
$B_i = \{(i, i+1, i+14, i+5), (i, i+25, i+42, i+63, i+29)\}$
$\cup \{(i, i+2, i+16, i+6), (i, i+26, i+44, i+66, i+30)\}$
$\cup \{(i, i+3, i+18, i+7), (i, i+27, i+46, i+69, i+31)\}$
$\cup \{(i, i+4, i+20, i+8), (i, i+28, i+48, i+72, i+32)\}$ $(i = 1, 2, ..., 73)$.

Example 5. A balanced (C_4, C_5)-10-foil decomposition of K_{91}.
$B_i = \{(i, i+1, i+17, i+6), (i, i+31, i+52, i+78, i+36)\}$
$\cup \{(i, i+2, i+19, i+7), (i, i+32, i+54, i+81, i+37)\}$
$\cup \{(i, i+3, i+21, i+8), (i, i+33, i+56, i+84, i+38)\}$
$\cup \{(i, i+4, i+23, i+9), (i, i+34, i+58, i+87, i+39)\}$
$\cup \{(i, i+5, i+25, i+10), (i, i+35, i+60, i+90, i+40)\}$ $(i = 1, 2, ..., 91)$.

Example 6. A balanced (C_4, C_5)-12-foil decomposition of K_{109}.
$B_i = \{(i, i+1, i+20, i+7), (i, i+37, i+62, i+93, i+43)\}$
$\cup \{(i, i+2, i+22, i+8), (i, i+38, i+64, i+96, i+44)\}$
$\cup \{(i, i+3, i+24, i+9), (i, i+39, i+66, i+99, i+45)\}$
$\cup \{(i, i+4, i+26, i+10), (i, i+40, i+68, i+102, i+46)\}$
$\cup \{(i, i+5, i+28, i+11), (i, i+41, i+70, i+105, i+47)\}$
$\cup \{(i, i+6, i+30, i+12), (i, i+42, i+72, i+108, i+48)\}$ $(i = 1, 2, ..., 109)$.

References

[1] C. J. Colbourn and A. Rosa, Triple Systems. Clarendom Press, Oxford (1999).
[2] P. Horák and A. Rosa, Decomposing Steiner triple systems into small configurations, *Ars Combinatoria* 26 (1988), pp. 91–105.
[3] W. D. Wallis, Combinatorial Designs. Marcel Dekker, New York and Basel (1988).

Dudeney の円卓問題の展望

静岡県立大学　小林みどり
半導体研究所　喜安善市
東海大学　中村義作

1. Dudeney の円卓問題とその歴史

Dudeney の円卓問題とは，次のような問題である.「n 人の人が $(n-1)(n-2)/2$ 回円卓を囲む．その際，どの人についても，自分以外の任意の 2 人が自分の両隣りにくるようにせよ.」

この問題は 1899 年 Judson が Amer.Math.Monthly に提出し，その後，1917 年 Dudeney が彼の著書 Amusements in Math. で紹介して，広く知られるようになった.

上の問題を，グラフ理論の言葉で定式化すると次のようになる.「K_n を n 個の頂点をもつ完全グラフとする．K_n の任意の 2-path (長さ 2 の path) をちょうど 1 回ずつ含む Hamilton cycle の集合を求めよ.」このような集合を K_n の Dudeney 集合と呼ぶ．すなわち，Dudeney の円卓問題は，Dudeney 集合を構成する問題である.

この問題は上記のように，今から 100 年も前に提出された問題であるが，現在でもなお未解決である．この問題は，初め Safford (1904)，Dickson(1905) らが研究した．Berghot らが $n = p+1$ (p は素数) のときに解決し，1975 年には $n = 2p$ (p は素数) の場合が解かれ (Anderson, Nakamura)，1980 年には $n = p^e + 1$ (p は素数，e は自然数) の場合が解かれた (Nakamura, Kiyasu, Ikeno)．その後，$n = p+2$ (p は奇素数，2 が p の原始根)，$n = pq+1$ (p. q は相異なる奇素数)，$n = p^e+1$ (前述の $n = p^e+1$ とは構成法が異なる)，$n = p^e q^f + 1$ (p, q は素数で，$p \geq 5, q \geq 11$) の場合が解決され，そしてついに，n がすべての偶数の場合が解決された [4].

その後, $n = p+2$ (p は奇素数，-2 が p の原始根)，$n = p+2$ (p は奇素数，2 が p の原始根の 2 乗, $p \equiv 3 \pmod 4$)，$n = p+2$ (p は奇素数，2 が p の原始根の 2 乗, $p \equiv 1 \pmod 4$), 3 が mod p の平方非剰余) の場合が解決された．n が一般の奇数のとき, Dudeney 集合の構成は未解決であるが, 二重 Dudeney 集合の構成については, 最近解決された [6].

2. Dudeney 集合のいろいろな構成法

Dudeney 集合の構成法としては種々のものが考えられている．K_n の 完全 1 因子分解から作る方法，有限体の原始根を利用する方法，枝交換法，直積法，帰納的構成法などがある．いずれも Dudeney 集合の自己同型群を利用するもので，固定点が高々 2 個であることから，自己同型群が n 次 cyclic group を含むもの，$n-1$ 次 cyclic group を含むもの，$n-2$ 次 cyclic group を含むものが考えられる．それらに対して，Dudeney 集合の構成法が考えられている．鼓型 starter から構成する方法もある.

3. 黒色 1 因子と open problem

前述のように, 偶数次 Dudeney 集合の構成は既に解決済みであるが, 奇数次 Dudeney 集合については, $p+2$ 次（p は奇素数）の一部の場合しか解決されていない．$p+2$ 次（p は奇素数）Dudeney 集合は, 次の命題のように, 黒色 1 因子から構成することができる.

命題1 K_{p+1} の黒色1因子が存在すれば, K_{p+2} の Dudeney 集合が存在する.

しかし, 黒色1因子は, すべての素数に対して存在するわけではなく, $p = 5, 11, 13, 19$ などには存在しない. 次の命題が証明できる.

命題2 K_{p+1} の黒色1因子が存在すれば, 2 は mod p の平方剰余である.

予想 2 が mod p の平方剰余ならば, K_{p+1} の黒色1因子が存在する.

上の予想が言えれば, 2 が mod p の平方剰余のとき, K_{p+2} に Dudeney 集合が存在することが言える.

4. Dudeney の円卓問題の関連問題とその応用

Dudeney の円卓問題は,「完全グラフ K_n の 2-path の Hamilton cycle による 完全被覆」を求める問題である. それに関連する問題として, 完全グラフでなく完全二部グラフや完全有向グラフの場合, また, Hamilton cycle でなく Hamilton path, k-cycle, k-path, k-circuit の場合などが考えられる. 2-path でなく 1-path (edge) の完全被覆を求める問題は, 古くから, Hamilton cycle decomposition 問題としてよく知られており簡単な解がある. k-circuit decomposition 問題は, 1977 年に Hwang, Lin によって解決されており, k-cycle decomposition 問題も, Alspach らにより最近解決された. これらの問題はすべて, 隣りとの関係においてバランスのとれた配置を求める問題である. このような問題は, 総称して, 隣接デザインの問題と呼ばれている.

隣接問題は, たとえば血清学の実験で,「k 個のウィルスを抗血清の回りに並べる. その際, どの 2 つのウィルスも隣り合うようにする必要がある. 実験回数を最小にしたいが, どのように並べたらよいか.」という問題に応用されている.

参考文献

[1] H. E. Dudeney, "Amusements in Mathematics," Thomas Nelson and Sons, London, 1917, Dover Reprint, New York, 1970.

[2] K. Heinrich, M. Kobayashi and G. Nakamura, Dudeney's Round Table Problem, *Annals of Discrete Math.* **92** (1991) 107-125.

[3] K. Heinrich, D. Langdeau and H. Verrall, Covering 2-paths uniformly, J. Combin. Des. **8** (2000) 100-121.

[4] M.Kobayashi, Kiyasu-Z. and G.Nakamura, A solution of Dudeney's round table problem for an even number of people, *Journal of Combinatorial Theory*, Ser.A **62** (1993) 26-42.

[5] M. Kobayashi, N. Mutoh, Kiyasu-Z. and G. Nakamura, New Series of Dudeney Sets for $p+2$ Vertices, *Ars Combinatoria*, to appear.

[6] M. Kobayashi, N. Mutoh, Kiyasu-Z. and G. Nakamura, Double Coverings of 2-paths by Hamilton Cycles, *J. Combinatorial Designs* **10** (2002) 195-206.

Constructions of Weighted Graph Designs

Keio University, Yukiyasu Mutoh

1 Introduction

Let $G = (V, E)$ denote a graph with a vertex set V and an edge set E and $k := |V|$. A weight function w of G is a function $w : E \mapsto \mathbb{Z} \setminus \{0\}$, where the integer $w(e)$ is called the *weight* of the edge $e \in E$. Then a triple (V, E, w) is called a *weighted graph*, which is denoted by (G, w). If w is clear from the context, we just write G instead of (G, w). If $w(e)$ is constant $(= \lambda)$ for each e, then (G, w) is denoted by λG.

Let $G = (V, E, w)$ and $G' = (V', E', w')$ be weighted graphs. The *union of G and G'* is the graph $G \cup G' := (V \cup V', E \cup E', w + w')$. Here, $w + w'$ is defined as follows:

$$(w + w')(e) = \begin{cases} w(e) + w'(e) & \text{if } e \in E \cap E', \\ w(e) & \text{else if } e \in E \setminus E', \\ w'(e) & \text{else } e \in E' \setminus E. \end{cases}$$

Let K_v be a complete graph with v vertices and let X be the point set of K_v. Let $\varphi : G \to K_v$ be an injective map, then $\varphi(G)$ is called a *G-block* or a *block*. We define Φ as a set of such maps φ. A pair (X, Φ) is called a *G-decomposition of λK_v* or a *G-design*, denoted by $WGD(G, \lambda; v)$, if $\bigcup_{\varphi \in \Phi} \varphi(G) = \lambda K_v$ holds.

Let $\Omega = \{w_1, w_2, \ldots, w_s\}$ be the set of distinct weights in (G, w). We define $\lambda_t(i, j)$ as the number of G-blocks in which two distinct points i and j of X occur in an edge having weight $w_t \in \Omega$. If a pair (X, Φ) is a $WGD(G, \lambda; v)$, then $\sum_{t=1}^{s} w_t \lambda_t(i, j) = \lambda$ holds for any pair of distinct points i and j. A pair (X, ϕ) is called a *completely balanced* $WGD(G, \lambda; v)$, if $\lambda_t(i, j)$ is constant $(= \lambda_t)$ not depending on the distinct pair i and j for $t = 1, 2, \ldots, s$.

Let E_x be the set of edges containing $x \in V$ and $W_x = \sum_{e \in E_x} w(e)$ be the sum of the weights in E_x. We define d as the greatest common divisor of the set $\{W_x | x \in V\}$ and $W = \sum_{e \in E} w(e)$ as the sum of all weights. Then the following conditions are necessary for the existence of a $WGD(G, \lambda; v)$: (i) $\lambda v(v - 1) \equiv 0 \pmod{2W}$ and (ii) $\lambda(v - 1) \equiv 0 \pmod{d}$.

If λ is an integer such that there exists a completely balanced $WGD(G, \lambda; v)$, then the following theorem holds.

Theorem 1.1 (Lamken and Wilson [1]) *Given a weighted graph G and a certain integer λ, there exists a completely balanced $WGD(G, \lambda; v)$ for all sufficiently large integers v satisfying the necessary conditions (i) and (ii).*

In this talk, we will consider the case where $WGD(G, \lambda; v)$ is not assumed to be completely balanced.

2 Existence of a weighted graph design

A *pairwise balanced design* (PBD) on v points with block sizes from a set $L(\subset \mathbb{N})$ consists of a set X of v points together with a family of subsets (called *blocks*) $B_1, B_2, \ldots, B_b$ such that any two distinct points occur together in a unique block B_i and $|B_i| \in L$ for each i. Let $B[L]$ be the set of integers v such that there exists a PBD with v points and block sizes from L. L is called a *closed set*, if $B[L] = L$ holds. For a set $M(\subset \mathbb{N})$, M is called *eventually periodic with period* n_M if the following condition satisfies. For each $m \in M$, $m + tn_M$ belongs to M for all sufficiently large integer t.

Theorem 2.1 (Wilson [2]) *Every closed set L is eventually periodic with period $\beta(L)$, where $\beta(L) = \gcd\{l(l-1) | l \in L\}$.*

Now, let $WGD[G, \lambda]$ be the set of integers v such that there exists a $WGD(G, \lambda; v)$. Then it is easy to show that $WGD[G, \lambda]$ is a closed set. By Theorem 2.1, the following proposition holds.

Proposition 2.2 *Let G be a weighted graph and let λ be an integer. If there exists a $WGD(G, \lambda; v_0)$, then there exists a $WGD(G, \lambda; v)$ for all sufficiently large integers $v \equiv v_0 \pmod{\beta(WGD[G, \lambda])}$.*

If the following conjecture is valid then for any λ there exists a $WGD(G, \lambda; v)$ for all sufficiently large integers v satisfying the necessary conditions.

Conjecture *For every integer u satisfying the necessary conditions, there is an integer $v_0 \equiv u \pmod{\beta(WGD[G, \lambda])}$, such that there exists a $WGD(G, \lambda; v_0)$.*

Moreover, we will show an existence theorem for a $WGD[G, \lambda]$ with $v \equiv 1 \pmod{2W}$ satisfying some linear constraints for λ.

References

[1] E.R. Lamken and R.M. Wilson. Decompositions of edge-colored complete graphs. *J. Combin. Theory Ser. A* **89** (2000) 149-200.

[2] R. M. Wilson. An existence theory for pairwise balanced designs II. *J. Combin. Theory Ser. A* **13** (1972) 246-273.

完全二部グラフのcluttered orderingとRAIDへの応用

慶應大・理工　足立智子

1. RAID

ハードディスク（以下，ディスクと呼ぶ）の読み込み・書き込みをスピードアップするために，一般的には，複数のディスクに並列に記憶する方法をとる．このようにしてディスクの数が多くなってくると，ディスクの破損の可能性が増大するという問題点が生じる．そこで，ディスクの破損個所の発見・修復のためにcheck diskを用いる．このように，記憶すべきデータと障害回復のための冗長データを複数のディスクに分散して格納することにより，アクセス性能と対障害性を同時に確保するための技法がRAID (redundant arrays of independent disks) である．安全性を高めるためにcheck diskを多くすると，追加のコストが増えてしまう．そこで，安全性と追加のコストのバランスを考えることが重要になってくる．

まず，information diskには保存したいデータを分割して格納し，check diskにはinformation disk内のデータが破損した場合に復旧するための冗長データを格納するとする．今，k個のinformation diskとc個のcheck diskがあるとし，これらの関係を0, 1を成分にもつ$c \times (k+c)$行列$H = [P|I]$で表す．Hはパリティ検査行列と呼ばれる([2]参照)．ただし，Iは単位行列であり，Pはc行k列の$\{0,1\}$-行列である．Hの最初のk列はinformation diskに対応し，後半のc列はcheck diskに対応しているHの1つの行に現れるinformation diskの内容の排他的論理和が計算され，その行に対応するcheck diskに書き込まれている．そして，1つのディスクが壊れても復旧できるようにHの列はmod 2で線形独立になっている(図1参照)．本稿で扱う二次元のRAIDでは，check diskを縦横の二次元に配列する．図1のように，check diskを頂点，information diskを辺とみなすことで，RAIDを二部グラフで表現することができる．

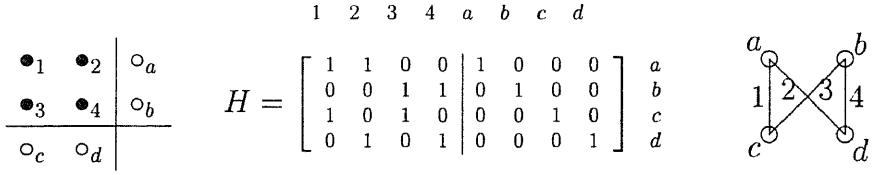


図 1: 2次元のRAID，パリティ検査行列H, 対応する完全二部グラフ$K_{2,2}$.

2. Cluttered Ordering

あるグラフ$G = (V, E)$について$n = |V|$, $E = \{e_0, e_1, \cdots, e_{m-1}\}$とする．ある正の整数$d \leq m$を考え，*window*と呼ぶ．$\{0, 1, \cdots, m-1\}$上の置換$\pi$に対して$V_i^{\pi,d}$を$\{e_{\pi(i)}, e_{\pi(i+1)}, \cdots, e_{\pi(i+d-1)}\}$の各辺に含まれる点の集合とする．インデックスはmod mで計算し，$0 \leq i \leq m-1$である．d本の辺を持つ部分グラフのアクセスコストをその部分グラフの頂点数で測るのだが，上限(d-最大アクセスコスト)は$\max_i |V_i^{\pi,d}|$で与

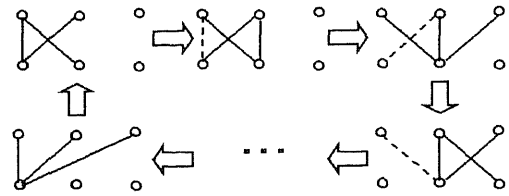


図 2: $K_{3,3}$ の $(3,4)-$ cluttered ordering.

えられる．d-最大アクセスコストが f となる辺の順序付けを (d,f)-*cluttered ordering* と呼ぶ (図 2 参照)．

3. 完全二部グラフの Cluttered Ordering

完全グラフの cluttered ordering の構成法は Cohen 等 [3] によって与えられた．本稿では，2次元の RAID に自然に対応するように，完全二部グラフの cluttered ordering の構成法について考察する．そのために，*wrapped ρ-labelling* と (d,f)-*movement* という 2 つの概念を導入する．

二部グラフ $H=(U,E)$ について $U=V\cup W$, $d=|E|$ とする．写像 $\rho: U\to \mathbb{Z}_d\times\mathbb{Z}_2$ が $\rho(V)\subset \mathbb{Z}_d\times\{0\}$, $\rho(W)\subset \mathbb{Z}_d\times\{1\}$ を満たし，$\mathbb{Z}_d$ の各要素が $\{\rho(v)-\rho(w)|v\in V, w\in W, (v,w)\in E\}$ に一つずつ存在するとき，ρ を H の *ρ-labelling* と呼ぶ．さらに，U の部分集合 X, Y に対して，$\mathbb{Z}_d\times\mathbb{Z}_2$ において $\rho(Y)=\rho(X)+(\kappa,0)$, $(\kappa,d)=1$ を満たす整数 κ が存在するとき，ρ を H の *wrapped ρ-labelling* と呼ぶ．次に，(d,f)-*movement* について述べる．同形な二つの二部グラフ $H=(U,E)$, $H'=(U',E')$ について $U=V\cup W$, $U'=V'\cup W'$, $|V|=|V'|$, $|W|=|W'|$, $E=\{e_0,e_1,\cdots,e_{d-1}\}$, $E'=\{e'_0,e'_1,\cdots,e'_{d-1}\}$ とする．$\{0,1,\ldots,d-1\}$ 上の置換 π を用いて，完全二部グラフ G を $H_0:=H$, $H_i:=(U_i,E_i), 1\le i\le d$ と $d+1$ 個の部分グラフに分割する．但し，$E_i:=(E_{i-1}\setminus\{e_{\pi(i-1)}\})\cup\{e'_{\pi(d+i-1)}\}$, U_i は E_i の各辺に含まれる頂点の集合とする．このとき，$H_d=H'$ となり，$\max_{0\le i\le \ell}|U_i|=f$ ならば π を H から H' への (d,f)-*movement* と呼ぶ．ここで，次の定理が得られる (証明は [1] を参照)．

定理 同形な二部グラフ H,H' に対し，wrapped ρ-labelling と (d,f)-movement が存在するならば，完全二部グラフ $K_{d,d}$ において (d,f)-cluttered ordering は存在する．

参考文献

[1] T. Adachi, M. Jimbo, M. Müller, Constructions of a cluttered ordering for the complete bipartite graph, in preparation.

[2] M. Cohen and C. Colbourn (2001), Ordering disks for double erasure codes, *ACM Symposium on Parallel Algorithms and Architectures*, Theory Comput. Syst. **34**, Springer-Verlag, 229-236.

[3] M. Cohen, C. Colbourn and D. Froncek (2001), Cluttered orderings for the complete graph, *COCOON 2001*, Lect. Notes Comp. Sci. **2108**, Springer-Verlag, 420-431.

A Practical Algorithm for Searching Consistent Sets of Shares in a Threshold Scheme

筑波大学システム情　左　瑞麟

1 Introduction

Any set of k shares in a (k,n) Shamir threshold scheme can be used to reconstruct the secret. However, if one or more of the n shares are faulty, then the secret may not be reconstructed correctly. Supposing that at most t of the n shares are faulty, Rees et al. (1999) described two algorithms to determine consistent sets of shares so that the secret can be reconstructed from the k legitimate shares in any of these consistent sets. In this paper, we propose a modified algorithm for this problem. Its efficiency is compared with those of the two algorithms mentioned above, and its expected number of shares used to reconstruct the secret is compared with those of McElice and Sarwate (1981) and Rees et al. (1999), in some special cases.

Suppose that the (k,n) Shamir threshold scheme is implemented in $GF(q)$. Let

$$S = \{(x_i, y_i) : 1 \le i \le n\} \subseteq (GF(q) \setminus \{0\}) \times GF(q)$$

be the set of n shares, and assume that at most t of the n shares are faulty. That is, there exists a polynomial $P_0(x) \in GF(q)[x]$ of degree at most $k-1$ such that $y_i = P_0(x_i)$ for at least $n-t$ of the n shares. The secret, which can be reconstructed from any k non-faulty shares, is the value $P_0(0)$.

Denote the subset of *good shares* by $G = \{i : y_i = P_0(x_i), 1 \le i \le n\}$, and the subset of *bad shares* by $B = \{1,2,\ldots,n\} \setminus G$. Then $|G| \ge n-t$ and $|B| \le t$.

For any $T \subseteq \{1,2,\ldots,n\}$ such that $|T| = k$, there is a unique polynomial P_T of degree at most $k-1$ such that $P_T(x_i) = y_i$ for all $i \in T$, which can easily be computed by Lagrange interpolation

$$P_T(x) = \sum_{i \in T} y_i \prod_{j \in T \setminus \{i\}} \frac{x - x_j}{x_i - x_j}.$$

Define $C_T = \{i : P_T(x_i) = y_i, 1 \le i \le n\}$. Then $|C_T| \ge n-t$ if $T \subseteq G$, and $|C_T| \le k+t-1$ if $T \cap B \ne \emptyset$. If $n-t \le k+t-1$, then there could exist a polynomial $P_T \ne P_0$ of degree at most $k-1$ such that at least $n-t$ shares lie on P_T. Therefore, Rees et al. required that the inequlity $n \ge k+2t$ always holds.

Let v, k, λ and t be positive integers such that $v \ge k \ge t$. A t-(v,k,λ) *covering* is a pair $(\mathcal{V}, \mathcal{B})$, where $\mathcal{V}$ is a v-set of elements (usually called *points*) and $\mathcal{B}$ is a collection of k-subsets (usually called *blocks*) of $\mathcal{V}$, such that every t-subset of points occurs in at least λ blocks in $\mathcal{B}$. The *covering number* $C_\lambda(v,k,t)$ is the minimum number of blocks in any t-(v,k,λ) covering. A t-(v,k,λ) covering $(\mathcal{V}, \mathcal{B})$ is *optimal* if $|\mathcal{B}| = C_\lambda(v,k,t)$.

Let $\mathcal{T}$ be a set of k-subsets of $\{1,2,\ldots,n\}$ such that its complement $\{\{1,2,\ldots,n\} \setminus T : T \in \mathcal{T}\}$ is the collection of blocks of a t-$(n, n-k, 1)$ covering. Then Rees et al. claimed two algorithms to compute the polynomial P_0. One is a randomized algorithm (Algoeithm 1) which used randomized k-subset chosen from the set $\{1,\cdots,n\}$, and the other (Algorithm 2)used a k-subset $\mathcal{T} \subseteq \binom{n}{k}$ with its complement form a t-$(n, n-k, 1)$ covering to compute P_0

$C_1(n, n-k, t)$ provides an upper bound on the number of iterations required by Algorithm 2. Unfortunately, the construction for optimal t-$(v,k,1)$ coverings, especially for $t \ge 3$, is not easy at

all. Algorithm 2 is a Las Vegas type algorithm, and it will take, generally speaking, a rather long time to find the correct polynomial P_0.

2 The new practical algorithm

Let $\mathcal{R}$ be a collection of k-subsets of $\{1,2,\ldots,n\}$ such that its complement $\{\{1,2,\ldots,n\} \setminus R : R \in \mathcal{R}\}$ forms the collection of blocks of a t-$(n, n-k, \lambda)$ covering. Then the following deterministic algorithm can compute the polynomial P_0, since for any $B \subseteq \{1,2,\ldots,n\}$ with $|B| \leq t$, there exist at least λ k-subsets $T \in \mathcal{R}$ such that $B \cap T = \emptyset$. Algorithm 3 will succeed when the first such subset T appears.

Algorithm 3

Input $\mathcal{R}, S, n, k, t$.

For each $T \in \mathcal{R}$, perform the following steps:

1. compute P_T
2. compute C_T
3. **if** $|C_T| \geq n - t$ **then** $P_0 = P_T$ and QUIT

3 A few 3-$(n, n-k, \lambda)$ Coverings with $\lambda > 1$ and $6 \leq n-k \leq 3k$

In order for Algorithm 3 to succeed efficiently, compared with Algorithm 2, we need to construct (optimal) t-$(n, n-k, \lambda)$ coverings with $2t \leq n-k \leq kt$. The first inequality was required for the uniqueness of the polynomial of degree at most $k-1$ such that at least $n-t$ shares lie on it. The second one is because that $C_1(n, n-k, t) = t+1$ for $n-k \geq kt$ had been already proved by Mills, so that the application of Algorithm 1 is sufficient for efficiently searching consistent sets of shares.

Theorem 3.1 *Let* $\lambda' = \mu\lceil\lambda\frac{(n-1)(n-2)}{(k-1)(k-2)}\rceil + 3\mu\frac{l(n-1)}{l-2}\lceil\lambda\frac{n-2}{k-2}\rceil + \lambda\mu(\frac{(nl-1)(nl-2)}{(l-1)(l-2)} - 1 - 3\frac{l(n-1)}{l-2})$, *where* $\lceil x\rceil$ *denotes the smallest integer greater than or equal to* x. *If there exists a resolvable* 3-(nl, l, μ) *design, then*

$$C_{\lambda'}(nl, kl, 3) \leq \mu\frac{(nl-1)(nl-2)}{(l-1)(l-2)}C_\lambda(n, k, 3).$$

Theorem 3.2 *A resolvable* 3-$(4n, 4, 1)$ *design exists for every* $n \equiv 1, 2 \pmod 3$ *except possibly for* $n \in P = \{55, 59, 73, 91, 115, 149, 169, 181, 269, 275, 313, 329, 455, 559, 577, 581, 595, 635, 685, 703, 905, 955, 1589\}$.

参考文献

[1] E. F. Brickel and D. R. Stinson, *The detection of cheaters in threshold schemes*, SIAM J. Disc. Math. 4 (1991), 502–510.

[2] R. S. Rees, D. R. Stinson, R. Wei and G. H. J. van Rees, *An application of covering designs: determining the maximum consitent set of shares in a threshold scheme*, Ars Combin. **53** (1999), 225–237.

Mutually M-intersecting Varieties and Optical Orthogonal Codes

東京理科大学　理工学部　　宮本　暢子[1]
明星大学　情報学部　　篠原　聡[2]

1　はじめに

f を斉次多項式とする。$PG(n,q)$ 上で $f(\boldsymbol{x})=0$ を満たすすべての点集合を *variety* と呼び、$V(f)$ で表わす。次に以下の 3 つの条件を満たす variety の集合 $V(f_1),V(f_2),\ldots,V(f_s)$ を考える。

(i) M は非負整数の集合とする。

(ii) $1\leq i\leq s$ に対して、$|V(f_i)|=\rho$ を満たす。

(iii) $1\leq i,j\leq s$, $i\neq j$ に対して $|V(f_i)\cap V(f_j)|\in M$ を満たす。

このような集合を *mutually M-intersecting varieties* と定義し、$\mathcal{V}(\rho,M)$ と書く。mutually M-intersecting varieties は、直交配列や均斉配列の構成に用いることができるが、ここでは光ファイバを用いた符号分割多元接続 (Fiber Optic Code Division Multiple Access: FO-CDMA) を実現するために利用される optical orthogonal code の構成への応用を考える。*Optical orthogonal code* C とは以下の 2 つの条件を満たす長さが n で重みが w である $(0,1)$-sequence の集まりで、$(n,w,\lambda_a,\lambda_c)$-OOC と書く。

- (*auto-correlation property*) 任意の $(c_0,c_1,\ldots,c_{n-1})\in C$ と任意の $1\leq t\leq n-1$ なる整数 t に対し、$\sum_{i=0}^{n-1}c_ic_{i+t}\leq\lambda_a$
- (*cross-correlation property*) 任意の異なる $(c_0,\ldots,c_{n-1}),(c'_0,\ldots,c'_{n-1})\in C$ と任意の $0\leq t\leq n-1$ なる整数 t に対し、$\sum_{i=0}^{n-1}c_ic'_{i+t}\leq\lambda_c$

ただし c,c' の添字は n で剰余をとる。また $\lambda_a=\lambda_c=\lambda$ のときは、(n,w,λ)-OOC と書く。

C の各元を**符号語**と呼び、より多くの符号語がある事が望ましいとされている。与えられたパラメータ (n,w,λ) に対して、符号語の数が最大であるような OOC を *optimal* であるという。constant weight code に対する Johnson bound より導かれる、OOC の符号語数についての上限式

$$\Phi(n,w,\lambda)\leq\lfloor\frac{1}{w}\lfloor\frac{n-1}{w-1}\lfloor\cdots\lfloor\frac{n-\lambda}{w-\lambda}\rfloor\cdots\rfloor\rfloor\rfloor$$

が存在し、この上限を達成する事で Optimal な OOC であると保証できる。本報告では、PG$(2,q)$ 上の conic を用いた mutually M-intersecting varieties を紹介する。さらに q が偶数のときの幾何的な結果を述べ、これらを用いた $\lambda\geq 2$ の場合の OOC の構成法を提案する。

2　構成法

PG$(2,q)$ 上の conic を用いた結果を挙げる。p を PG$(2,q^2)$ 上の点であって、PG$(2,q)$ 上にはない点とする。このとき、$\mathcal{C}$ を p を通るような PG$(2,q)$ 上の conic の集合とする。

[1]miyamotois.noda.tus.ac.jp
[2]sshinohami.meisei-u.ac.jp

定理 1　$\mathcal{C}$ は、q^3-q^2 個の conic から成る mutually M-intersecting varieties $\mathcal{V}(q+1,\{0,1,2\})$ である。

PG$(3,q)$ において平面 π 上の conic を C,C' とし、ϕ を PG$(3,q)$ の Singer cycle とするとき、PG$(3,q)$ の点に対して、conic 上の点は 1, それ以外の点は 0 とおくことにより、conic と OOC の符号語を対応させると、

auto-correlation: for $i=1,\ldots,n-1$

$$C\cap\phi^i(C)=(C\cap\pi)\cap(\phi^i(C)\cap\phi^i(\pi))\subseteq C\cap(\pi\cap\phi^i(\pi))$$

cross-correlation: for $j=1,\ldots,n-1$

$$C\cap\phi^j(C')=(C\cap\pi)\cap(\phi^j(C')\cap\phi^j(\pi))\subseteq C\cap(\pi\cap\phi^j(\pi)).$$

のように考えられる。また $j=0$ の場合には、定理 1 で得られる conic を用いれば、異なる conic の交点は高々2 点であることが保証され、次の定理が言える。

定理 2　$\mathcal{C}$ を用いて、符号語数が $q^3-q^2+\lfloor\frac{q^3-1}{q^2-1}\rfloor$ の $(q^3+q^2+q+1,q+1,2)$-OOC が構成できる。

q が偶数のとき、conic C にその nucleus $\mathcal{N}(C)$ を追加した点集合を考え、

$$\mathcal{C}'=\{C_i\cup\{\mathcal{N}(C_i)\}:C_i\in\mathcal{C}\}$$

とする。$C_1',C_2'\in\mathcal{C}'$ に対して、$\mathcal{N}(C_i)=P_{N_i}$ とするとき、次の結果が得られる。

補題 1　$\mathcal{C}$ に含まれる異なる 2 つの conic が共有点を持たないならば、$|(C_1\cup\{P_{N_1}\})\cap(C_2\cup\{P_{N_2}\})|\leq 2$ である。

補題 2　$\mathcal{C}$ に含まれる異なる 2 つの conic の交点が 1 点であるならば、2 つの nucleus は互いの conic 上の点ではない。すなわち、$|(C_1\cup\{P_{N_1}\})\cap(C_2\cup\{P_{N_2}\})|\leq 2$ である。

補題 3　$\mathcal{C}$ に含まれる異なる 2 つの conic の交点が 2 点であるならば、2 つの nucleus は異なる点である。すなわち、$P_{N_1}\neq P_{N_2}$ である.

予想　$\mathcal{C}$ に含まれる異なる 2 つの conic の交点が 2 点であるならば、2 つの nucleus は同時に互いの conic 上の点であることはない。すなわち、"$P_{N_1}\in C_2$ かつ $P_{N_2}\in C_1$" とはならない。

以上の補題と予想により、

$$|(C_1\cup\{P_{N_1}\})\cap(C_2\cup\{P_{N_2}\})|=|(C_1\cap C_2)\cup(C_1\cap\{P_{N_2}\})\cup(C_2\cap\{P_{N_1}\})\cup(\{P_{N_1}\}\cap\{P_{N_2}\})|\leq 3.$$

またこの結果を OOC へ応用すると、

予想　$\mathcal{C}'$ を用いて符号語数が $q^3-q^2+\lfloor\frac{q^3-1}{q^2-1}\rfloor$ の $(q^3+q^2+q+1,q+2,2,3)$-OOC が構成できる。

より高次元に一般化することによって、以下の定理が得られる。

定理 3　$\mathcal{V}(\rho,M)$ を PG(d,q) の mutually M-intersecting varieties とする。このとき $(\frac{q^{d+2}-1}{q-1},\rho,\lambda)$-OOC が構成できる。　ただし、$\lambda=\max_{m_i\in M}m_i$ とする。

Bounds for Robust Metering Schemes and Their Relationship with A^2-code

東京工業大学・理財セ　尾形わかは
茨城大学・工　黒澤　馨

A (k, n)-metering scheme allows a correct counting on the number of hits that a Web site received during a certain period. That is, a Web server $\mathcal{S}$ can compute a *proof* if and only if k or more clients visited $\mathcal{S}$ during a certain period. Naor and Pinkas proposed the first cryptographically secure (k, n)-metering scheme [1]. Ogata and Kurosawa showed that their scheme is not as secure as they claimed and presented a more secure scheme [2].

More specifically, there exist four kinds of participants, a Web server $\mathcal{S}$, n clients $\mathcal{C}_1, \ldots, \mathcal{C}_n$, an audit agency $\mathcal{A}$ and an outside enemy $\mathcal{E}$ in this model. (We consider that n clients are monitors and the outside enemy is not.) We then require the following three kinds of security.

Security against servers A malicious Web server $\mathcal{S}$ tries to forge a *proof* from only $k-1$ or less shares (authenticators) of clients and to cheat $\mathcal{A}$. Hence $\mathcal{S}$ should not be able to inflate her hit counts. (There appears to be no way to detect whether $\mathcal{S}$ is deflating her hit counts.)

Security against clients Malicious clients try to forge an illegal share which would be accepted by $\mathcal{S}$, but would not allow $\mathcal{S}$ to compute the correct *proof*. Hence $\mathcal{S}$ must be able to detect illegal shares forged by clients.

Security against outside enemy An outside enemy $\mathcal{E}$ tries to forge a (legal or illegal) share which would be accepted by $\mathcal{S}$. If it is legal, it causes a counting error because he is not a monitor. If it is illegal, it does not allow $\mathcal{S}$ to compute the correct *proof*. Hence $\mathcal{S}$ must be able to detect a share forged by $\mathcal{E}$.

We say that a (k, n)-metering scheme is

- *robust* if it satisfies all the three security requirements.
- *non-robust* if it satisfies only the security against servers.

We further say that a (k, n)-metering scheme is perfect if $\mathcal{S}$ gains no information on *proof* from any $k-1$ or less shares. (It is interesting that the metering schemes proposed so far are all perfect.)

For *non*-robust and perfect metering schemes, a lower bound on the communication complexity $|V_i|$ $(i = 1, \ldots, n)$ was shown by Masucci and Stinson [3], where V_i is a set of possible values v_i which is sent by client $\mathcal{C}_i$ to $\mathcal{S}$ when $\mathcal{C}_i$ has access to $\mathcal{S}$. (They considered a more general model than ours such that there are multiple Web servers and there exists a ramp structure among clients.)

However, *non*-robust metering schemes are not practical. We cannot assume that clients are all honest. We cannot assume that there is no outside enemy, either.

In this paper, we derive lower bounds on the communication complexity $|V_i|$ $(i = 1, \ldots, n)$ and the size of server's secrets $|E_s|$ for *robust* (k, n)-metering schemes.

We first derive lower bounds on $|V_i|$ and $|E_s|$ for "perfect and robust" (k, n)-metering schemes by using counting arguments. We also present a slightly modified version of the Ogata-Kurosawa scheme [2] and prove that it satisfies all the equalities of our bounds. This means that our bounds are all tight.

We next show an almost equivalence between robust (k, n)-metering schemes and k-multiple-use A^2-codes such that we can always construct a k-multiple-use A^2-code from a (k, n)-metering scheme, and in some cases, we can do the reverse. By using this equivalence, we derive lower bounds on $|V_i|$ and $|E_s|$ for robust (but not necessarily perfect) (k, n)-metering schemes. This equivalence is of independent interest because no relationship has been known between them so far.

	Lower bound on $\lvert V_i \rvert$	Lower bound on $\lvert E_s \rvert$
Non-robust and perfect	[3]	Meaningless*
Robust and perfect	This paper	This paper
Robust	This paper	This paper

References

[1] M. Naor and B. Pinkas, "Secure and Efficient Metering," Eurocrypt '98, LNCS 1403, pp.576–589 (1998)

[2] W. Ogata and K. Kurosawa, "Provably Secure Metering Scheme," Asiacrypt 2000, LNCS 1976, pp.388–398 (2000)

[3] B. Masucci, D. R. Stinson, "Efficient Metering Schemes with Pricing," IEEE Trans. on IT, Vol.47, pp.2835–2844 (2001)

暗号の乱数性について

名古屋工業大学・知能情報システム学科　萩田 真理子

インターネットを用いて電子商取引を行う時には、第３者に知られることなく安全に通信を行うために、データは暗号化して伝送されている。その時に利用される暗号技術として、現在広く使用されているのは、AESやDESに代表される共有鍵暗号方式である。これは、通信を行う二人だけが知っている共通暗号鍵をもとに、送信するデータを第３者に分からないデータに変換する方法で、この共通鍵を秘密に共有するための手段として、公開鍵暗号方式と呼ばれる、すべての操作を見られても秘密のデータを共有できる方法が使われている。

ここでは、共有鍵暗号方式の変換を、鍵・文書を入力すると暗号化データを出力する写像として捉えたときの散らばり具合を安全性の指標として提案する。

＜公開鍵暗号方式＞

公開鍵暗号には、x を与えられたときに y=f(x) を計算するのは簡単だが、y=f(x) を与えられたときに x を求めるのは難しい数学の問題が使われている：

素因数分解問題　与えられた素数 p, q について、n=pq を計算するのは簡単だが、n(=pq) を与えられて因数分解して p, q を求めるのは難しい。

離散対数問題　与えられた p, g, x について、$y=g^x \pmod p$ を求めるのは簡単だが、p, g, y を与えられて、$y=g^x \pmod p$ なる x を求めるのは難しい。

公開鍵暗号方式の例として、SSHにも使われているRSAを紹介する：

＜RSA(1978)＞　p,q:大きな素数について、n=pq、M=(p-1)(q-1)、　e：M と素な整数とし、ed=1 mod なる d を用意する。(MはZ/pZ×Z/qZ の乗法群 Z/(p-1)Z×Z/(q-1)Z の位数。) 公開鍵を(n,e)、秘密鍵を(d,p,q,M)とし、送り手は公開鍵を使い、メッセージ m (0<m<n)を、C = me mod n と変換して C を送る。受け手は秘密鍵を用いて、m = Cd mod n と変換して元に戻す。

鍵交換に最も良く使われているのは、Diffie-Hellmanの鍵交換方式である：

＜Diffie-Hellman の鍵交換方式(1976)＞　公開された大きな素数ｐと、位数ｐの有限体GF(ｐ)の乗法群の生成元ｇを用いて、ＡとＢがそれぞれ乱数ａ，ｂを発生させる。ＡはＢに($g^b \bmod p$)を送り、ＢはＡに ($g^b \bmod p$) を送る。それぞれ、$k=(g^b)^a \bmod p=(g^a)^b \bmod p$ を計算し、秘密鍵として共有する。

公開鍵暗号方式は安全だが、計算に時間がかかるため、秘密の共有鍵を交換するときだけこの方式を使い、大量の文書のやり取りにはその鍵を用いて共有鍵暗号方式で暗号化することが多い。DES、AESと呼ばれる共有鍵暗号方式が特に有名である。

＜共有鍵暗号方式＞

DES(Data Encryption Standard)は、これまでアメリカ合衆国政府の国立標準研究所(NIST)が、データ暗号の標準として推奨してきた共有鍵暗号方式である。これまで広く使われてきたが、最近の技術の進歩で解読される危険が出てきたため、NISTはDESに代わるデータ暗号の標準としてAES(Advanced Encryption Standard)を公募し、数学者J.Daemen, V.Rijmenによって開発されたRijndaelと呼ばれる暗号化方式が採択された。どちらも共有鍵を利用して、固定サイズのブロックに区切られたデータを高速に暗号化する共有鍵暗号方式で、ブロックサイファーと呼ばれている。AESは、128（又は192,256）ビットの共有鍵を用いて、128（又は192,256）ビットに区切られた情報列をかき混ぜて、鍵を知らなければ元に戻せなくする暗号化方式で、変換は次のように行われている：

＜AESのしくみ＞

共有鍵のサイズも、一度に変換する情報列のサイズも128ビットの場合を紹介する。この１６バイトに区切られた情報はブロックと呼ばれ、４×４に並べて扱う。鍵は１１個の４×４バイトのラウンドキー０から１０に拡張しておく。AESの具体的な変換は、次の４つの操作からなる：

1．Byte Sub：バイトの中での入れ替え（近いものを遠くにずらし、特に全て0、全て1のビット列の1の数が変わるように入れ替える操作。具体的にはバイトを有限体GF（28）の元と見て逆元をとり（ただし0は0に移す）、行列の積と和をとって変換する。） 2. Shift Row：行ごとのいれかえ（バイト単位で、i行目（i=0,1,2,3）を左にiバイトシフトする。） 3. Mix Column：列ごとにかきまぜる（各列をバイト係数の多項式とみなし、環上の多項式の積で変換する。） 4. Add Round Key：鍵を知らないとできない唯一の変換（ビットごとにラウンドキーをEXORで加える。） この４つの変換を、4,（1,2,3,4），...，（1,2,3,4），1,2,4の順に行う。ただし、中間部のラウンド(1,2,3,4)は9回繰り返す。また4で加えるラウンドキーは、ラウンドキー０から１０を前から順に使う。

＜暗号の安全性の指標としての乱数性＞

暗号は、与えられた鍵・文書に対して、暗号化データを出力する関数である。上記のAESのように、鍵・文書・暗号化データが128ビットならば、$\varphi: F_2^{128} \times F_2^{128} \to F_2^{128}$という関数になっているととらえることができる。鍵・文書の選び方と、出力データの相関ができるだけ小さい方が、解読し難いと予想される。つまり、鍵データや文書をいろいろな方向に少しだけ（例えば距離1だけ）変えたときに、暗号文がランダムに散らばっている、乱数性の高い暗号が好ましいと考えられる。具体的には、暗号の安全性の指標として、原点から距離１のベクトル全体がどのように散らばるか、距離１のベクトル全体がどんなベクトルの集合を与えるか、ｔ次元部分空間がどこに散らばるか、などを統計的検定によりラウンド数をどこまで減らすと偏りが見られるか調べることで、簡単な変換を繰り返すタイプの暗号を評価できる。特にDESやAESのように、共有鍵暗号方式で簡単な変換（ラウンド）を繰り返して作られる暗号化方式について、変換の乱数性を統計的、代数的に検定して評価してみたいと考えている。

グループ鍵を用いた暗号化ファイル共有システムの構築について

慶應大・理工　大原 幸多
(株)デンソークリエイト　陳 志松

1　はじめに

近年，ソフトウェアのバグやメールに添付されたウィルスなどにより，フォルダへ不正進入されるという現象が多発している．通信を暗号化し，ファイルのアクセス権を管理するよりもファイルを暗号化することによって，このような攻撃を防ぐことができ，またアクセス権を意識しないファイルの共有を実現する．ファイルを暗号化するソフトウェアは既に市場に存在するが，その多くは複数ユーザーによるネットワークを介したファイル共有を実現するものではない．本報告では，Microsoft Windowsプラットホーム上で複数ユーザーによるファイル共有に対応した暗号化ファイルシステムを紹介する．

2　ファイル暗号化

本システムの暗号化アルゴリズムには2000年10月にNIST[1]によって決定されたAES標準のRijndaelを採用した．ファイルを暗号化する際には，ECBモードを使用し，最終ブロック処理は，RSA社が開発した規格であるPKCS♯5パディング[2]を適用している．PKCS♯5パディングの適用によって，暗号化後のファイルサイズは最小で1バイト，最大で16バイト増大する．

3　グループ鍵配送のためのWaterfallアルゴリズム

本システムではファイルの暗号化にAESによる共通鍵暗号方式を採用したため，複数ユーザーでファイルを共有するためには，各ユーザー共通の鍵を生成する必要がある．鍵生成には，Waterfallアルゴリズム[3]を用いている．

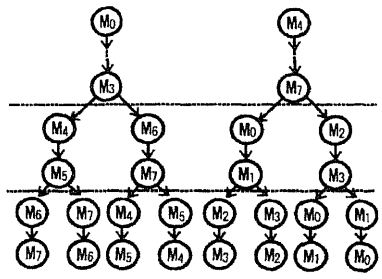


図 1: メンバー数8のWaterfall

グループ鍵の作成要求を発行するメンバーM_0は，$p' = 2p + 1$が素数となる素数pと$GF(p')$の原始元αを計算する．Waterfallアルゴリズムでは，グループは最初に二つの均等なサブグループに分けられ，順番

が決められている．そして，各サブグループのメンバーM_iは，前のユーザーから送られてきたパラメータと自分の秘密鍵r_iを用いてべき乗し，次のメンバーに送る．最後のメンバーまで同様の操作を行った後，各グループをさらに二つのサブグループに分け，各サブグループの最初のメンバーは前のグループにおける自分が所属していなかった方の最後のメンバーからパラメータを受け取る．同様の操作を，サブグループのメンバー数が1になるまで繰り返す．最終的に，各メンバーM_iは$\alpha^{r_0 \cdots r_{i-1} r_{i+1} \cdots r_{n-1}}$を受け取っているため，自分の秘密鍵を用いて共通鍵$\alpha^{r_0 \cdots r_{n-1}}$を生成することができる（図1）．

4 システム構築

本システムでは，鍵配送はP2P型に，ファイルの交換はサーバー・クライアント型として実現している（図2）．

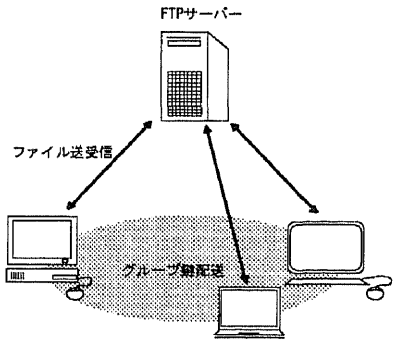


図 2: アーキテクチャー

各メンバー間ではRSA-デジタル署名を用いることによりメンバー認証を行っている．また，共有するファイルはFTPサーバー上に保存する．ファイルのダウンロード・アップロードは，Windows Explorerをインターフェイスとした，マウス操作によってファイルの暗号化/復号化とともに行うことが可能である．Microsoft社はWindowsシェルを拡張するためのインターフェイスの幾つかを公開しており[4]，本システムのユーザーインターフェイスはMicrosoft COM（Component Object Model）技術にそのインターフェイスを組み合わせることにより実現されている．なお，素数生成やべき乗計算にはNTL[5]を使用し，コーディングはすべてVisual C++6.0で行った．

参考文献

[1] National Institute of Standards and Technology, http://www.nist.gov/

[2] PKCS♯5(RSA Security Inc.),
http://www.rsasecurity.com/rsalabs/pkcs/pkcs-5/index.html

[3] 陳, 山本, 西野, 神保, “ピアグループにおける鍵共有”, FIT2002, 2002

[4] MSDN, http://msdn.microsoft.com/

[5] NTL:A Library for doing Number Theory, http://shoup.net/ntl/

汎直交配列による実験計画

東海大学　秋山 仁
徳島文理大学　近藤 衛
東海大学　中村義作

たとえば，つぎの構造模型

$$\begin{aligned} y_1 &= \mu + \alpha_1 + \beta_1 \quad + \varepsilon_1 \\ y_2 &= \mu + \alpha_1 + \beta_2 \quad + \varepsilon_2 \\ y_3 &= \mu + \alpha_2 + \beta_1 + \gamma_1 + \varepsilon_3 \\ y_4 &= \mu + \alpha_2 + \beta_2 + \gamma_1 + \varepsilon_4 \\ y_5 &= \mu + \alpha_2 + \beta_1 + \gamma_2 + \varepsilon_5 \\ y_6 &= \mu + \alpha_2 + \beta_2 + \gamma_2 + \varepsilon_6 \end{aligned} \quad \cdots\cdots\cdots\cdots (1)$$

を考えると，ふつうの意味での直交計画にならないが，本稿で述べる汎直交計画(汎直交配列による直交計画)にはなっている．ここに，μは母平均，(α_1, α_2),(β_1, β_2),(γ_1, γ_2)はそれぞれ因子A，B，Cの水準ごとの母数，$\varepsilon_1 \sim \varepsilon_6$は互いに独立な誤差項である．以下では，式(1)の構造模型を例として，汎直交計画の粗い考え方を説明する．

データ，母数，誤差を担うベクトルをそれぞれ

$$\mathbf{y} = \begin{vmatrix} y_1 \\ y_2 \\ y_3 \\ y_4 \\ y_5 \\ y_6 \end{vmatrix}, \quad \mathbf{g}(M) = \begin{vmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{vmatrix}, \quad \mathbf{g}(A_1) = \begin{vmatrix} 1 \\ 1 \\ 0 \\ 0 \\ 0 \\ 0 \end{vmatrix}, \quad \mathbf{g}(A_2) = \begin{vmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 1 \\ 1 \end{vmatrix}, \quad \mathbf{g}(B_1) = \begin{vmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \\ 0 \end{vmatrix},$$

$$\mathbf{g}(B_2) = \begin{vmatrix} 0 \\ 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{vmatrix}, \quad \mathbf{g}(C_1) = \begin{vmatrix} 0 \\ 0 \\ 1 \\ 1 \\ 0 \\ 0 \end{vmatrix}, \quad \mathbf{g}(C_2) = \begin{vmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \\ 1 \end{vmatrix}, \quad \boldsymbol{\varepsilon} = \begin{vmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \varepsilon_3 \\ \varepsilon_4 \\ \varepsilon_5 \\ \varepsilon_6 \end{vmatrix} \quad \cdots\cdots\cdots\cdots (2)$$

とすると，式(1)の構造模型は

$$\begin{aligned} \mathbf{y} = \mu\, \mathbf{g}(M) &+ \alpha_1 \mathbf{g}(A_1) + \alpha_2 \mathbf{g}(A_2) + \beta_1 \mathbf{g}(B_1) \\ &+ \beta_2 \mathbf{g}(B_2) + \gamma_1 \mathbf{g}(C_1) + \gamma_2 \mathbf{g}(C_2) + \boldsymbol{\varepsilon} \quad \cdots\cdots\cdots\cdots (3) \end{aligned}$$

となる．いま，データを担う因子ベクトルにたいし，つぎの正方行列(対称行列)

$$\mathbf{M} = \frac{\mathbf{g}(M)\mathbf{g}(M)'}{\mathbf{g}(M)'\mathbf{g}(M)}, \quad \mathbf{A}_i = \frac{\mathbf{g}(A_i)\mathbf{g}(A_i)'}{\mathbf{g}(A_i)'\mathbf{g}(A_i)}, \quad \mathbf{B}_i = \frac{\mathbf{g}(B_i)\mathbf{g}(B_i)'}{\mathbf{g}(B_i)'\mathbf{g}(B_i)}, \quad \mathbf{C}_i = \frac{\mathbf{g}(C_i)\mathbf{g}(C_i)'}{\mathbf{g}(C_i)'\mathbf{g}(C_i)}$$

を定義する．ここに，分子はベクトルのダイアド，分母はベクトルの内積である．すると，$\mathbf{A}_i$にたいしては

$$(\mathbf{A}_i)^2 = \frac{\mathbf{g}(A_i)\mathbf{g}(A_i)'}{\mathbf{g}(A_i)'\mathbf{g}(A_i)} \cdot \frac{\mathbf{g}(A_i)\mathbf{g}(A_i)'}{\mathbf{g}(A_i)'\mathbf{g}(A_i)} = \frac{\mathbf{g}(A_i)\{\mathbf{g}(A_i)'\mathbf{g}(A_i)\}\mathbf{g}(A_i)'}{\{\mathbf{g}(A_i)'\mathbf{g}(A_i)\}\{\mathbf{g}(A_i)'\mathbf{g}(A_i)\}}$$

$$= \frac{\mathbf{g}(A_i)\mathbf{g}(A_i)'}{\mathbf{g}(A_i)'\mathbf{g}(A_i)} = \mathbf{A}_i$$

となるように，$\mathbf{M}$, $\mathbf{A}_i$, $\mathbf{B}_i$, $\mathbf{C}_i$($i = 1, 2$)にたいして

$$\mathbf{M}^2 = \mathbf{M}, \quad \mathbf{A}_i^2 = \mathbf{A}_i, \quad \mathbf{B}_i^2 = \mathbf{B}_i, \quad \mathbf{C}_i^2 = \mathbf{C}_i$$

が成り立つ．これは$\mathbf{My}$, $\mathbf{A}_i\mathbf{y}$, $\mathbf{B}_i\mathbf{y}$, $\mathbf{C}_i\mathbf{y}$がそれぞれ$\mathbf{y}$のある方向への射影であることを示す．

いま，正方行列形A，B，Cを

$$A=A_1+A_2-M,\ B=B_1+B_2-M,\ C=C_1+C_2-A_2 \quad \cdots\cdots\cdots\cdots (4)$$

で定義すると，M，A，B，Cにたいして

$$A^2=A,\ B^2=B,\ C^2=C,\ MA=MB=MC=AB=AC=BC=0 \quad \cdots\cdots\cdots\cdots (5)$$

が成り立つ．そこで，単位行列Iを使って，Eを

$$E=I-(M+A+B+C) \quad \cdots\cdots\cdots\cdots (6)$$

で定義すれば，データの直交成分の分解式は

$$y=Iy=(M+A+B+C+E)y=My+Ay+By+Cy+Ey \quad \cdots\cdots\cdots\cdots (7)$$

となり，2乗和の分解式は

$$y'y=y'Iy=y'My+y'Ay+y'By+y'Cy+y'Ey \quad \cdots\cdots\cdots\cdots (8)$$

となる．分散分析は式(8)に基づいてなされるが，このときの自由度は，行列のトレース(tr)を使うと，

実験回数$=tr(I)$，$f_M=tr(M)$，$f_A=tr(A)$，$f_B=tr(B)$，$f_C=tr(C)$，$f_e=tr(E)$

となる．

以上の結果を式(1)の構造模型にたいして具体的に計算すると，単位行列の分解はつぎのようになる．

$$\begin{pmatrix} 1&0&0&0&0&0\\ 0&1&0&0&0&0\\ 0&0&1&0&0&0\\ 0&0&0&1&0&0\\ 0&0&0&0&1&0\\ 0&0&0&0&0&1 \end{pmatrix} = \frac{1}{6}\begin{pmatrix} 1&1&1&1&1&1\\ 1&1&1&1&1&1\\ 1&1&1&1&1&1\\ 1&1&1&1&1&1\\ 1&1&1&1&1&1\\ 1&1&1&1&1&1 \end{pmatrix} + \frac{1}{12}\begin{pmatrix} 4&4&-2&-2&-2&-2\\ 4&4&-2&-2&-2&-2\\ -2&-2&1&1&1&1\\ -2&-2&1&1&1&1\\ -2&-2&1&1&1&1\\ -2&-2&1&1&1&1 \end{pmatrix}$$

$$+\frac{1}{6}\begin{pmatrix} 1&-1&1&-1&1&-1\\ -1&1&-1&1&-1&1\\ 1&-1&1&-1&1&-1\\ -1&1&-1&1&-1&1\\ 1&-1&1&-1&1&-1\\ -1&1&-1&1&-1&1 \end{pmatrix} + \frac{1}{4}\begin{pmatrix} 0&0&0&0&0&0\\ 0&0&0&0&0&0\\ 0&0&1&1&-1&-1\\ 0&0&1&1&-1&-1\\ 0&0&-1&-1&1&1\\ 0&0&-1&-1&1&1 \end{pmatrix} + \frac{1}{6}\begin{pmatrix} 2&-2&-1&1&-1&1\\ -2&2&1&-1&1&-1\\ -1&1&2&-2&-1&1\\ 1&-1&-2&2&1&-1\\ -1&1&-1&1&2&-2\\ 1&-1&1&-1&-2&2 \end{pmatrix}$$

以上の考察を踏まえて，汎直交計画の条件を示すと，以下の3項になる．

（1）各要因(因子と交互作用)の水準ごとの反復数は等しくなくてもよい．ただし，母数にたいする付帯条件は，母数とその反復数の積を水準ごとに加えた和が0になるように修正する．

（2）任意の2因子A，Bが直交するとは，iとjに独立な定数kが存在して，

$$kN(A_i,B_j)=N(A_i)N(B_j)$$

となることを指す．ここに，$N(A_i)$，$N(B_j)$はそれぞれA_i，B_jを含む実験回数，$N(A_i,B_j)$はA_i，B_jを同時に含む実験回数を表す．

（3）因子Cが因子Aの特定の水準だけに現れるときは，因子Cを因子Aの隣接高次因子といい，これでも因子Aと因子Cは直交すると解釈する．このため，平均Mはすべての要因にたいして低次である．高次の要因同士の直交条件は，低次の因子の特定の水準にたいするものだけで決める．

［参考書］ 近藤衛，中村義作：『工科系の実験計画法』，工学図書(株)，1981年7月

直交実験と Discrepancy

筑波大学　社会工学系　藤原　良

1．はじめに

まずはじめに，つぎのような小さなモデルを考えてみよう：　いま鉢に植えたゼラニュームの成長速度を観たいのであるが，要因として次の３つを考える．

要因１：　ゼラニュームの品種（赤，ピンク，白）

要因２：　１日あたり与える水の量

要因３：　一月当たり与える肥料の量

このとき要因１は３つの水準しかないので，直交実験では他の要因も３水準に合わせる．要因２と３はもともとは連続的な量なので，例えば要因２では，

第１水準＝100cc/日，　第２水準＝500cc/日，　第３水準=1000cc/日

といったように適当な値を割り当てることになる．統計モデルを作れば次のようになる．

$$y_{ijk} = \mu + \alpha_i + \beta_j + \gamma_k + \varepsilon_{ijk}\ ,\ \ i,j,k = 0,1,2$$

$$\sum_i \alpha_i = \sum_j \beta_j = \sum_k \gamma_k = 0$$

直交配列を作ってみると次のような配列ができる．

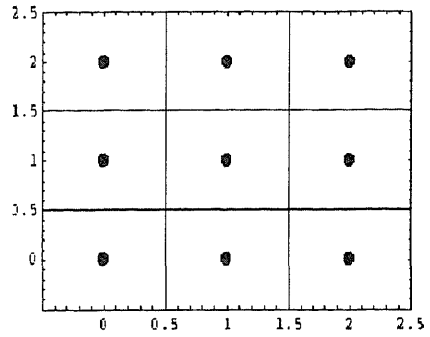
0	0	0	1	1	1	2	2	2
0	1	2	0	1	2	0	1	2
0	1	2	1	2	0	2	0	1

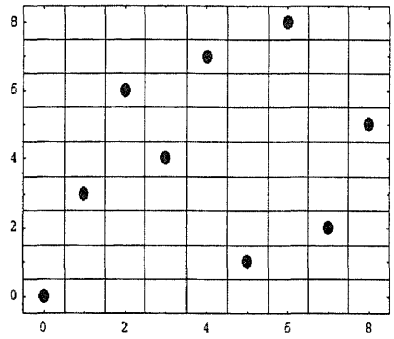
実験は９回であるので，各要因において，同じ水準を３回ずつ実験していることになる．

つぎに，第２，第３要因は連続量なので，実験回数は増やさないで，おっと多い水準とすることも可能である．そこで次のような計画を作ってみた，第２，３要因は９水準となっている．もちろん直交はしていない，また均整にもなっていない．

0	0	0	1	1	1	2	2	2
0	3	6	1	4	7	2	5	8
0	3	6	4	7	1	8	2	5

これら二つの計画を，第２，３要因のみを平面上に図示すると次のようになる．これらの点の散らばり具合が観測値や推定値にどう影響するかが，興味ある問題である．





2．Discrepancy

一方，数値積分，特にモンテカルロ法の分野で最近話題に上っている概念に Discrepancy というのがある．これはある点集合の一様性を評価する一つの基準である．Discrepancy は

$X=\{x_1,x_2,...,x_N\}$ ， $x_i\in[0,1)^s$　　$J:[0,1)^s$の任意の部分区間

$N(X,J)$：区間 J の中に入っている X の点の数

L_∞-discrepancy：　$D_N^{(s)}=sup_J|N(X,J)/N-J|$

と定義されている．何種類か定義があるがこの L_∞

定理（Koksma-Hawka）

$f(x):[0,1]s$　　$X=\{x_1,x_2,...,x_N\}$ ， $x_i\in[0,1)^s$

$V(f)$　: bounded variation

$$\left|\frac{1}{N}\sum f(x_i)-\int f(u)du\right|\le V(f)D_N^{(s)}$$

3．直交配列と一様性

Discrepancy の計算はそれほど容易ではなく，また低い Discrepancy をもつ点集合を作る一般的な方法はまだない．しかし，(t,m,s)-net と呼ばれる，Discrepancy が保証された点集合を作る方法がある．これは順序直交配列と呼ばれる直交配列の特殊形と同値である．直交配列が格子点上の点を全て含むような条件になっているのに対し，(t,m,s)-net などの Low-Discrepancy な点集合は部分区間（立体）上に点がバランスするようになっている．Discrepancy の基準から見ると格子点というのはけっして良くはない．

4．課題

問題1．同じ水準数，実験回数の均整配列と Low-Discrepancy 点集合はどちらが良いか，またどんな基準で比較すべきか．

問題2．実験回数は同じで，水準数を増やして実験した場合のメリットとその評価方法．

5．参考文献

[1]　H. Niederreiter, Point Sets and Sequences with Small Discrepancy, Monatsh. Math.Vol.104 (1984)pp.273-337

Group Testing and Positive Detectable Matrices

慶應大・理工　神保雅一, Meinard Müller

Let $C = \{c_1, \ldots, c_n\}$ a set of *items* and $\sigma : C \to \{0, 1\}$ a map indicating the *state* of each item. An item c_i is said to be *positive* if $\sigma(c_i) = 1$, otherwise *negative*. In applications such as DNA library screening (in this case, the items are *clones*) one has the goal to determine the set of all positive items in C, where a method is given to *test* the state of each item (e.g., by some chemical analysis). To reduce the number of tests, one chooses a subset $P \subset C$, also denoted as *group* or *pool*, and tests all items of P in one stroke. The state of a pool is *positive* if it contains at least one positive item, otherwise *negative*. This strategy is known as *group testing* which can be defined as the process of selecting pools and testing them to determine exactly which items are postive [1]. A group testing procedure is called *nonadaptive* if all pools are specified a priori without knowing the state of other pools. In this case, the *complexity* of the group testing algorithm is given by the number of its pools. Note, that it must be ensured by the group testing procedure that every possible set of positive items is distinguished.

For an overview of different group testing methods and some of their applications we refer to [2]. Colbourn [1] considers the setting were the set C is equipped with a linear order $c_i \prec c_{i+1}$, $1 \leq i < n$, and has the *d-consecutive positive property*, i.e., the set of positive items is a consecutive set with respect to the ordering $\prec$ and contains at most d items. His main result can be summarized as follows.

Theorem 1 *The complexity of nonadaptive group testing for a set C of n items having the d-consecutive positive property is $\Theta(d + \log_2 n)$.*

To proof the upper bound Colbourn designs a group testing algorithm which proceeds in two steps. Let $d \geq 2$. In the first step, the n items of C are partitioned into $\lceil n/(d-1) \rceil$ linearly ordered subpools of $(d-1)$ consecutive items respectively (except of the last subpool having possibly a smaller size). By assumption, at most two of these pools, which are then consecutive, are positive. The items of these positive pools can be tested individually in $O(d)$. Treating these subpools as items, the general case can in this way be reduced to the case $d = 2$ which is dealt with in the second step. To this means, Colbourn constructs an $m \times n$-matrix $H = (h_{j,i})$ over GF(2) by adding three suitable rows to an incidence matrix of some Gray code and possibly deleting some columns. From this matrix H he gets a group test with $m = \lceil \log_2 n \rceil + 3$ pools which accomplishes the task for the case $d = 2$. Here, the columns of H correspond with the items, the rows of H correspond with the pools, and $h_{ji} = 1$ means that the jth pool contains the ith item c_i, $1 \leq j \leq m$, $1 \leq i \leq n$.

In this talk we give a construction improving the group testing algorithm of Colbourn [1] described above. The main idea of our construction is that in

the case $d = 2$ one can distinguish up to any two consecutive positive items if all columns of H as well as any column arising as bitwise OR-sum of two consecutive columns of H are pairwise different. This observation leads to the following definition.

Definition 2 *Let $H = [x_1, x_2, \ldots, x_n]$ be an $m \times n$-matrix over* GF(2) *with column vectors x_i, $1 \leq i \leq n$. Define $y_i := x_i \vee x_{i+1}$ (bitwise OR-sum), $1 \leq i \leq n-1$. Then H is called a* 2-consecutive positive detectable matrix *or, for short, a* 2CPD-matrix *iff the list*

$$x_1, x_2, \ldots, x_n, y_1, y_2, \ldots, y_{n-1}$$

consists of non-zero, pairwise distinct vectors.

Since there are 2^m vectors in $\mathrm{GF}(2)^m$ one obviously has the bound $n \leq 2^{m-1}$. A 2CPD-matrix H is called *full* iff $n = 2^{m-1}$. For example, the following matrix is a full 2CPD-matrix of column size $m = 4$:

$$H = \begin{bmatrix} 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 1 & 0 & 1 & 0 & 0 & 0 \end{bmatrix}$$

Note that in a full 2CPD-matrix of column size m all non-zero vectors of $\mathrm{GF}(2)^m$ appear either as some x_i or some y_i. By a recursive construction we prove the following theorem.

Theorem 3 *Let $m > 2$, then if there exists a full* 2CPD*-matrix of column size m, then there also exists a full* 2CPD*-matrix of column size $m + 2$.*

One can easily show that there is no full 2CPD-matrix of column size $m = 3$. Giving explicite full 2CPD-matrices for the cases $m = 1$, $m = 2$, $m = 4$ and $m = 5$, one gets the following result.

Corollary 4 *There exists a full* 2CPD*-matrix of any column size $m \in \mathbb{N}$ except for the case $m = 3$.*

We finally note that from a full 2-consecutive positive detectable matrix one gets a group testing algorithm for the case $d = 2$ with $m = \lceil \log_2 n \rceil + 1$ pools. In terms of the maximal number of items which can be tested by using a prescribed number of pools, this improves Colbourn's construction by a factor of four.

References

[1] Colbourn, C. J.: Group testing for consecutive positives. Annals of Combinatorics **3** (1999), 37–41.

[2] Du, D.-Z., Hwang, F. K.: Combinatorial group testing and its applications. World Scientific, Singapore, 1993.

ゲノム情報および発現情報の解析

国立がんセンター研究所　疾病ゲノムセンター　水島　洋

昨年(2001年)2月にヒトのゲノム解析に関する論文が国際共同チームとCelera社によって、Nature と Science 誌に発表されるという歴史的イベントがあったばかりであるのに、すでにPerlegen社は、DNAチップ技術を使うことによって50人のゲノム情報を調べ、ハプロタイプ解析を行っていると報告している。このように、ゲノムが明らかになったことを前提としたポストゲノム解析時代になり、これまで解析手法の研究や大量解析に主眼を置いていたゲノムプロジェクトも、ゲノム配列の意味するところの情報解析や、疾患などとの関連性の研究へと移行してきている。また、ゲノム全体が明らかになった生物種も多くなってきていることから、違う生物種のゲノムを比較することによって、単に進化学的な解析にとどまらず、様々な解析が可能になってきている。

そこで重要になってくるのがこれらの大量情報を解析する学問、「バイオインフォマティクス」である。これは生物学と情報科学をむすびつけた学問であり、新しい分野だけあってまだ専門とする研究者は少ないものの、得られたデータからいかにして情報解析して新しい知識を抽出するかがこれからの焦点となってくることは間違いない。当初、Celera社のプロジェクトが採用したWhole Genome Shotgun法について、500塩基ずつに分断された個々の塩基配列情報から30億文字からなるパズルを完成させる処理には計算が膨大すぎて無理であると言われてきていた。しかし、この問題を解決したのがバイオインフォマティクスである。計算能力の向上と新しいアルゴリズムの組み合わせによってみごとに短時間でインフルエンザ菌、ハエ、ヒトゲノム、さらにマウスまでもゲノム構造を明らかにしてきている。Celera社には14テラバイト（14,000,000,000,000文字）の記憶容量をもつ高速なコンピュータを保有しており、米国のスーパーコンピュータセンターと肩を並べる性能を有している。Celera社の研究者の大半はコンピューターサイエンティストといわれており,これからは情報科学的な解析能力の差が大切であることがわかる。

また、データ量が大きくなっていることに加え、さまざまな種類のデータベースが存在することも注目される.インターネット上には500を越える生物系のデータベースがあるといわれ、データベースのデータベース（http://www.infobiogen.fr/services/dbcat/ ）も存在するほどである。フラットファイルで構築されていることが多いものの、これらは相互に関連していることが多く、データベース間連携を考えることも非常に重要になってきている。ドイツLION社ではSRSというソフトウェアで、フラットファイルのデータベースの相互連携を自動的に管理するシステムを開発しており、利用者にとって大変使いやすいものとなっている。

日本政府は平成11年12月にミレニアムゲノムプロジェクトの開始を決定した。このプロジェクトの目標は、同じような環境にあって人によってがんや痴呆、糖尿病、高血圧などの病気になったりならなかったりするのはなぜか、また同じ病気でもどうして人によって症状や予後が違うのか、さらに同じ薬でもどうして人によって効き方や副作用の出方が違うのかを、遺伝子のレベルで解明することである。その結果、個人個人に最適な治療法

や予防法を選択できるようにする、いわゆるオーダーメイド医療を進めるとともに、新しい薬などの画期的な治療法の開発を進めることを目指している。厚生省のミレニアムゲノムプロジェクトでは、国立がんセンター、国立循環器病センター、国立精神神経センター、国立国際医療センター、国立小児病院、国立医薬品食品衛生研究所の 6 箇所において、がん、高血圧などの循環器疾患、ぜんそくなどの小児アレルギー疾患、痴呆、糖尿病、薬剤反応性などに関連したゲノム解析を推進するものである。

筆者は以前より転写制御機構の解析を行っており、そのために必要な転写制御因子のデータベースの構築を行ってきている。これまで DNA に結合する部位の配列情報を集め、それを用いて転写制御未知な遺伝子の転写調節部位の予測を行ってきたが、結合配列が短いために多くの結合部位が予測され、なかなか本来の転写因子の推定がむずかしかった。そこで、ヒトとマウスのゲノム情報を比較し、タンパク質をコードしている Exon 部分以外でもさまざまな領域が保存されていることを利用し、これらの部分に保存されている転写制御因子の結合配列を検索表示するシステムを開発した。マウスとヒトのオーソロガスな遺伝子のゲノム配列（上流と下流を 10KB ずつ程度あると望ましい）を入力することによって広域において遺伝子配列が保存されている部分を表示するとともに、そこに保存されている転写制御因子結合部位を表示するものである。フグのゲノムも明らかになっているが、いくつかの遺伝子でヒトとフグでオーソロガスな遺伝子の比較をおこなったところ、マウスの場合とは違って、Exon 部分しかホモロジーのある部分がみられないことが多かった。

近年の技術開発によって、生物のすべての遺伝情報を解析してしまう網羅的な遺伝子解析のみならず、マイクロアレーによる網羅的な転写産物の解析、2 次元電気泳動による網羅的な蛋白質の解析、また、Two-Hybrid 法による網羅的な蛋白質間相互作用の研究など、加速度的にデータは増大している。新しい技術のためには新しい理論による解析が必要になってくる。ゲノム情報が明らかになることによって、これまでにない情報が得られるようになったが、その解析のためには生物学的な意味合いを考慮した実験計画をすることが重要であり、それにもとづいた関連ソフトウェアの開発も必要であろう。

参考：

Mizushima H., Ichikawa H., Ohki M.; Analysys Tool for finding Transcription Regulatory elements, Using Transcription Factor Data Base (TFDB). Proceedings of "Bioinformatics in Genome Regulation and Structure 2002 (Novosibirsk)".

参考文献：

Science Vol.291 No.5507 (2001)
Nature Vol.409 No.6822 (2001)
Science Vol.294 p1605-1776 (2001)
実践バイオインフォマティクス（オライリージャパン 2002）

凹性仮説検定のための最適実験計画

明星大学理工学部　広津 千尋

1 序論

順序制約下にある母数に対する推定，検定については既に多くの論文があるが，一方，最適実験計画を扱った論文は意外に少ない．その理由としては，順序制約問題で想定されるのは圧倒的に単調制約が多いこと，そして単調制約の場合，少なくとも検出力に関する限り両端に総実験数の半分ずつを割り振る自明な実験計画が最適になってしまうためと思われる．著者等は以前に1元配置の設定で単調仮説の検定問題を扱ったが，そこでは自明な最適計画の least favorable case（LFC）に対する検出力は確保したまま，出来る限り標本を内側に配分する maxmin 検定（最小検出力最大化検定）を導いた（Hirotsu and Herzberg, 1987; Hirotsu, 2002 参照）．新薬臨床試験の分野では，とくにICH（International Conference on Harmonization）以来，薬効のあること自体の証明として用量反応関係を示すことが求められると同時に，臨床至適用量の推測が求められる．そのような場合に，検定の検出力を確保したまま，やや内側の水準に可能な限りの例数を割り振るという考え方が正当化される．本論では凹性仮説に関して線形検定を用いる場合の maxmin 計画を導く．とくに水準数に応じてLFCが推移することと，それに応じて最適計画のパターンが変化するという興味ある結果が得られる．なお，凹（凸）性仮説は用量反応曲線の線形外挿等において本質的な形状制約である．

2 定式化

1元配置のモデル

$$y_{ij} = \mu_i + \varepsilon_{ij}, \quad i = 1, \ldots, a, \quad j = 1, \ldots, n_i$$

において，凹性仮説

$$\mathrm{H} : \mu_{i+2} - 2\mu_{i+1} + \mu_i \leq 0, \quad i = 1, \ldots, a-2$$

に対する線形 maxmin 検定を考える．すなわち，総実験数 n を与えられたものとして n_i の最適配分を考える．仮説H は凸錐をなし，その $a-2$ 個のコーナーベクトル $\boldsymbol{m}_k$ $(k = 1, \ldots, a-2)$ の第 j 要素は C_k を規準化定数として

$$\begin{cases} C_k\{2a - k - 1 + jv_k + (k - j + 1)u_k\}, & j \leq k; \\ C_k\{2a - k - 1 + jv_k\}, & j \geq k + 1; \end{cases}$$

で与えられる．ただし，

$$u_k = -\frac{a(a-1)}{k(k+1)}, \quad v_k = -\frac{3a - 2k - 1}{a+1}$$

である．すなわち，$\boldsymbol{m}_k$ は k 点を境とするスロープ変化モデルを形成するという特徴がある (Hirotsu and Marumo, 2002)．ここで，Hirotsu and Herzberg (1987) に従い，線形検定を

$$\left(\sum_k \lambda_k \boldsymbol{m}_k\right)' \boldsymbol{N}\boldsymbol{Q}\bar{\boldsymbol{y}}$$

と表すと，凹性仮説 H に対する maxmin 検定は，コーナーベクトル $\boldsymbol{m}_k$ $(k=1,\ldots,a-2)$ に対する線形 maxmin 検定に帰着する．ただし，$\boldsymbol{N}=\mathrm{diag}(\sqrt{n_i})$，$\boldsymbol{Q}'$ は $Var(\boldsymbol{NQ}\bar{\boldsymbol{y}}) \propto I$ となるような $(a-2)\times a$ 直交行列，$\bar{\boldsymbol{y}}$ は観測値平均ベクトル，そして λ_k は Lagrange 未定係数である．このとき，両端に $n/4$，中央に $n/2$ を配分するのが最適となるが，その最小検出力を最大に保ったまま，その他の水準にどのくらいの実験数を割り振れるかが問題である．その解は，最小検出力を与える LFC が a の値によって推移することから a の値によって異なるが，このたび一般の a について解が得られたので報告する．

3 主要な結果

3.1 Least Favorable Case

対立仮説 $\boldsymbol{m}_k$ に対する線形 maxmin 検定の非心度を γ_k とすると，以下に述べるように LFC（下線）が推移する．a 小の間は不規則だが，$a \geq 24$ では規則性が成り立つ．

(1) $a \leq 16$: $\underline{\gamma_1} < \gamma_2 < \gamma_3 < \gamma_4 < \gamma_5 < \cdots$
(2) $a = 17$: $\gamma_1 > \underline{\gamma_2} < \gamma_3 < \gamma_4 < \gamma_5 < \cdots$
(3) $18 \leq a \leq 20$: $\gamma_1 > \gamma_2 > \underline{\gamma_3} < \gamma_4 < \gamma_5 < \cdots$
(4) $21 \leq a \leq 23$: $\gamma_1 > \gamma_2 > \gamma_3 > \underline{\gamma_4} < \gamma_5 < \cdots$
(5) $a \geq 24$ で非心度は $k = a/4 - 1$ で最小値をとり，その両側で単調増大

3.2 maxmin 計画

maxmin 計画は a の値にかかわらず両端に $n/4$ のサンプルを配分する．γ_k $(k=1,2,\ldots,[(a-1)/2])$ は次式で与えられる:

$$\gamma_k = \frac{1}{2}\sqrt{n}C_k|u_k|\left[k - \frac{2}{n}\{kn_1 + (k-1)n_2 + \cdots + n_k\}\right],\ n_1 = n/4.$$

ここで残り $n/2$ のサンプルは，γ_k が LFC の γ を下回らない範囲で対称に配分される．なお，$k=1$ 以外の l が LFC のな場合は $2 \leq k \leq l$ の水準にサンプルは配分されない，つまり $n_k = 0$ である．これにより，実際にどの程度中間の水準にサンプルを割り振れるかの具体例は当日示す．

References

[1] Hirotsu, C. (2002). On an optimal design for an isotonic inference. *J. Statist. Planning and Inference*, **102**, 205-213.

[2] Hirotsu, C. and Herzberg, A. M. (1987). Optimal allocation of observations for inference on k ordered normal population means. *Australian J. Statist.*, **29**, 151-165.

[3] Hirotsu, C. and Marumo, K. (2002). Changepoint analysis as a method for isotonic inference. *Scandinavian J. Statist.*, **29**, 125-138.

ASYMPTOTIC EXPANSION FORMULA FOR THE OC-FUNCTION OF THE MODIFIED Λ-TEST IN THE MULTIVARIATE ANALYSIS OF VARIANCE

- PRACTICAL PRECISION AND EFECTIVE SAMPLE SIZE -

Minoru SIOTANI (Science University of Tokyo)
Toshiya IWASHITA (Meisei University)
Takashi SEO (Science University of Tokyo)

1. The Modified Λ-Test and its OC-Function

In the analysis of linear model or regression model, the modified Λ-test is based on the criterion

$$W = -\{n - \tfrac{1}{2}(p-m+1)\}\log\Lambda, \qquad \Lambda = |V_o| / |V_h + V_o|$$

where V_h is p×p dispersion matrix due to hypothesis, $W \sim W_p(m,\Sigma,\Omega)$ which is noncentral Wishart distribution with m d.f., noncentrality matrix Ω and covariance matrix Σ; V_e is p×p dispersion matrix due to error, which is distributed independently of V_h according to the central Wishart distribution $W_p(n,\Sigma)$. Hypothesis is $H: \Omega = 0$.

The operating characteristic (OC) function is defined by

$$Q_E \equiv Q_E(p,m,n;\alpha,\Omega) = \Pr\{W \leqq W(\alpha) \mid \Omega\} = 1 - \text{Power function}$$

where $W(\alpha)$ is the upper 100α% point of W. No exact formula of Q_E for executing numerical work is available.

The following asymptotic expansion formula for Q_E can be obtained from the formula for the nonnull distribution of W given by Sugiura & Fujikoshi [Ann.Math. Statist. Vol. 40, 942-952, 1969] :

$$Q_A \equiv Q_A(p,m,n;\alpha,\Omega) = G^*_{mp} + \frac{1}{2\nu}\sum_{i=1}^{3} a^*_i G^*_{mp+2i} + \frac{1}{48\nu^2}\sum_{i=0}^{6} b^*_i G^*_{mp+2i} + O(\nu^{-3})$$

where $\nu = n - \frac{1}{2}(p-m+1)$, $G^*_f \equiv G^*_f(W(\alpha);\omega^2)$ is the value of noncentral χ^2-distribution with f d.f. and noncentrality parameter $\omega^2 = \operatorname{tr}\Omega$ evaluated at $\chi^2 = W(\alpha)$ and a^*_i, b^*_i are given in Sugiura & Fujikoshi (1969), which are dependent on p,m and

$S_j = \operatorname{tr}\Omega^j = \theta_1^j + \theta_2^j + \ldots + \theta_p^j$, $\theta_1 \geqq \theta_2 \geqq \ldots \geqq \theta_p$ are latent roots of Ω. The OC-functions Q_E and Q_A depend on the noncentrality matrix Ω through S_j.

2. Aim of the Problem

(1) To set up a Reference Domain D of Parameters (p,m,n,ω), over which Q_A serves as the OC-function of the test; (2) To construct an approximate upper bound $U \equiv U(p,m,n;\alpha,\Omega)$ on the absolute error of Q_A: $Y \equiv Y(p,m,n;\alpha,\Omega) = |Q_A - Q_E|$ such that, for given α

$$Y \lesssim U \quad \text{for} \quad (p,m,n;\Omega) \in D$$

where symbol "$A \lesssim B$" is the relaxiation of "$A \leqq B$" and read it 'A is less than or approximately equal to B'. (3) To give the users a useful but simple formula by

which they can determine an effective sample size satisfying a requirement on the precision of Q_A.

3. Results and Comments for their Derivation

Main works were done numerically and experimentally along our previous papers [Siotani, Iwashita & Seo: Amer. J. Math. Manag. Sci., 15, 215-237, 1995; J. Japan Statist. Soc., 28, 135-152, 1998] Readers should refer them.

3.1 Reference Domain D of Parameters (p,m,M,ω): $M=n-p+1$ is used instead of n.

(a) $2 \leqq p \leqq m \leqq 12$,

(b) $0 \leqq \omega \leqq 6$,

(c) Minimum values M_o of M:
See the Table in the right side.

Table Minimum Values M_o at (p,m)

p \ m	2	3	4	5	6	7	8	9	10	11	12
12											80
11										70	75
10									60	64	70
9								50	54	59	64
8							42	46	50	54	59
7						35	39	43	46	51	56
6					29	32	36	39	43	46	50
5				22	26	29	32	35	38	40	43
4			17	20	23	25	27	29	32	34	36
3		13	15	17	19	21	23	25	27	29	31
2	9	10	12	14	16	17	19	20	22	24	26

Comments: Since the effect on the absolute error Y in the case of rankΩ=1 is worse than other cases, we have only to treat safely the case of rank $\Omega = 1$.

3.2 Experimental upper bound on Y (α=0.05)

$$U_{EXP} = 0.83107\, p^{0.75239} m^{1.22437} M^{-1.93142} \times \exp\{-0.28126(\omega - 3.83526)^2\};$$

$$U^*_{EXP} = \log U_{EXP} = -0.18504 + 0.75239 \log p + 1.22437 \log m - 1.93142 \log M - 0.28126 \times (\omega - 3.83526)^2, \quad (p,m,M,\omega) \in D.$$

Comments: Steps for the construction of U_{EXP} or U^*_{EXP} are as follows: (1) To set a funcform of U, (2) To estimate the constants involved in the functional form by the method of least squares based on the data set $\{Y_r : p_r, m_r, M_r, \omega_r : r=1,2,..,132\}$; Y_r were computed by carrying out the Monte Carlo simulation of large scale for Q_E. For the choice of parameter points $(p_r, m_r, M_r, \omega_r)$, refer to Siotani, Iwashita & Seo (1995, 1998). (3) To check the effectiveness of the constructed upper bound based on the data used plus additional data not used in the construction.

3.3 Relation between Effective Sample Size and Precision of Q_A:

Suppose we wish to know an effective sample size n or $M=n-p+1$ satisfying $Y \lesssim U_{EXP} \leqq \gamma$ or $Y^* \lesssim U^*_{EXP} \leqq \log\gamma$ for a given small positive γ and given p, m. Using the formulae in 3,2, we obtain ($\alpha = 0.05$)

$$\log M \geqq L_{p,m,\gamma} \quad \text{or} \quad M = n-p+1 \geqq M_L = \exp(L_{p,m,\gamma}) \text{ for all values of } \omega,$$

where

$$L_{p,m,\gamma} = -0.09572 - \frac{\log\gamma}{1.93142} + 0.38955 \log p + 0.63392 \log\omega.$$

If a value of ω is fixed, we have

$$M \geqq M_L(\omega) = \exp\{L_{p,m,\gamma}(\omega)\}, \text{ where } L_{p,m,\gamma}(\omega)\} = L_{p,m,\gamma} - 0.14562(\omega - 3.83526)^2.$$

4. An Application. If the users wish to design the modified Λ-test by determining a sample size satisfying the requirement; α=0.05, whenever $\omega \geqq \omega_o$, H: $\Omega = 0$ should be rejected with a probability greater than or equal to a given value P_o; is easily and safely excuted with Q_A function or Q_A-curve with the domain and precision given above.

超高分子ポリエチレン多孔質体の焼結条件の最適化

日東電工（株）金子興道

1．　はじめに、

従来から、超高分子ポリエチレン多孔質体は樹脂粉末の焼結により製造されている。この多孔質体は焼結工程を経て通気性透湿性などの特徴をいかした多くの加工品として生産販売されている。しかしながら、製造開始し以来、約１０年にわたり成形体の密度バラツキに起因する、歩留低下、外観などの潜在的不具合に悩まされ続けてきた。

今回、基本機能を用いた技術開発という視点から焼結工程の工程条件を再設計する事により、長年の潜在的な悩みを一挙に解決することを目的として本研究に着手した。

2．製造方法と基本機能

2．1 製造方法

本焼結工程で使用される焼結炉は一般的な装置であるため省略する。一定の容器に粉末状の超高分子ポリエチレンをいれ容器の上部に重りを置くことで加圧しながら焼結する。

一般に焼結に関して、金属粒子に関して種々の理論や実験式が提案されている。これらを要約すると焼結は２段階のステップを経て行われる。まず第一段階では急速に収縮が起こる。この段階では連続孔が存在している。続いてゆっくりと空隙が周囲からの収縮圧によりつぶされていき、細孔化していく。粉末状の超高分子ポリエチレンにおいても同様なプロセスが予想される。

2．2 基本機能

焼結工程の目的は、望む比重にし、しかもその比重をブロック全域にわたり均一にすることである。「比重が均一」と言うことは、ブロックのどの部分からサンプリングしても、その重量と体積は比例するということである。即ち、（１）式が成り立つことが基本機能である。

体積＝β×重量　　（１）　　但し、βは比例定数。

今回は、比重が大きい方が望ましいため、各々の逆数をとり、（２）式を基本機能とした。

（１／体積）＝β×（１／重量）　　（２）

3．要因の選定と割付

3．1 要因の選定

（１）信号因子

入力信号は、ブロック重量の逆数とした。仕込量を　500,1000,2000g　３水準変化して焼結した。焼結後５０℃で２４時間乾燥後重量を測定した。実際は焼結後バリなどがひどく、それを取り除いて重量を測定したためを仕込量より軽くなっている。出力信号であるブロック体積の逆数は水置換法で求めた。

（２）誤差因子

誤差因子として原料樹脂を取り上げた。具体的には、　Ｐ１単独（以下P1と表す）とＰ１,Ｐ２の混合物（特定製品用の配合であり、以下TOKと表す）とした。

（３）制御因子と割付

制御因子は影響が大きいと予想されている、H2速度（昇温速度）、粉末にかける圧力（荷重）、昇温過程の途中に降温過程を入れる事等を考慮して、下記の制御因子と水準を検討した。制御因子は直交表Ｌ１８に割り付けた。

4．実験方法と実験結果

4．1 実験方法

実験には内径約150㎜の円筒形の焼結用金型を使用した。樹脂を充填後金属製の重りを乗せ、焼結炉に入れ，焼結パターンに従って焼結した。仕込量５００，１０００，２０００ｇ　の３種、樹脂の配合２種類（TOK,P1）合計６種類の金型を同時に焼結した。

4．2 実験結果：省略

5．ＳＮ比の計算　前述したように、基本機能を

（１／体積）＝β×（１／重量）　とし、データ変換し下記の計算を行った。

5．1 計算方法

$y = \beta \times M$ において、y：1/体積　　M：1/重量　　β：比例定数　とする。

SN比（η）は、各々次式で示される。

$$\eta = 10\log \frac{\frac{1}{R_{TOK}+R_{P1}}\left(S_\beta - V_e\right)}{V_N}$$

５．２ 計算結果：ＳＮ比と感度の計算結果を省略する。

６．要因効果図と最適条件

昇温速度を小さくすればＳＮ比が大きくなる。さらに実験範囲より遅くすることで、よりSN比を大きく（均一化）出来ることも示唆している。感度の要因効果図からは、比重を大きくするには、焼結時間を長く、焼結温度を高くするのが有効である。意外な結果だが、検討範囲内では荷重を変化しても比重にはあまり影響ないことが判った。供熱方法の差もあまり大きくはなかった。成形容器の熱容量が大きく、自然冷却では思ったほど温度が低下しないためと思われる。

選択した最適条件を ◎ 印で示す。断熱材の効果は余り差がないので、作業性を考慮して断熱材なしを選択した。焼結時間は、感度（比重を大きくしたい）と後述する焼結度を優先して１時間を採用した。

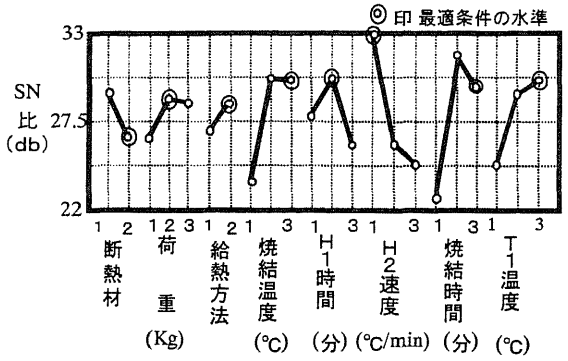


図3 SN比の要因効果図

表3 推定及び確認実験結果 (db)

	SN 比		感 度	
	推 定	確認結果	推 定	確認結果
最適条件	41.0	38.5	-2.84	-3.30
現行条件	-	24.8	-	-3.63
利 得	-	13.7	-	0.3

７．確認実験

図３の最適条件で、前述した方法と同様に成形し、重量と体積を測定した。推定値と確認実験結果を比較するとＳＮ比、感度ともに再現性がとれている。表３

８．製品での確認

以上にして得られた最適条件により、各種製品の製造に適用した結果、いずれもバラツキが小さくなった。

９、まとめ

今回の品質工学に適用により、長期間に渡る潜在的な品質問題を一挙に解決することが出来た。

（尚、今回の研究集会においてご指摘頂いた、信号の直交性の考察を行なった上で品質工学会誌の投稿予定である。）

Efficiency Factor of Split-Plot Designs

岐阜県立看護大学　小澤　和弘
大阪府立大・工　栗木　進二

2つの要因 A と C があり，各要因は v_1, v_2 個の処理（水準）をもつものとする．また，b 個のブロックがあり，各ブロックは k_1 個の wholeplot に分割され，各 wholeplot は k_2 個の subplot に分割されていて，要因 A の処理が wholeplot に対して施され，要因 C の処理が subplot に対して施されるものとする．ただし，1つのブロックにおいては，要因 C の同じ k_2 個の処理が各 wholeplot の subplot に施される．このような design を *split-plot design* と呼び，ここでは，$k_1 < v_1$，$k_2 < v_2$ である incomplete split-plot design (ISPD) を考え，各要因の処理の組合せ A_wC_j $(w = 1, 2, \ldots, v_1; j = 1, 2, \ldots, v_2)$ を含むブロック数は一定 (r) とする．

ISPD の線形モデルとして，処理効果が母数であり，ブロック効果，wholeplot 効果，subplot 効果が確率変数である混合モデルを考える．また，3段階の無作為化，(1) ブロックの無作為化，(2) 各ブロックの wholeplot の無作為化，(3) 各ブロックにおける各 wholeplot の subplot の無作為化，を考える．i 番目の処理効果 τ_i を

$$\tau_i = \mu + \alpha_w + \gamma_j + (\alpha\gamma)_{wj}, \quad i = (w-1)v_2 + j,$$

$(w = 1, 2, \ldots, v_1; j = 1, 2, \ldots, v_2)$ とする．ただし，μ は一般平均，α_w は A_w の主効果，γ_j は C_j の主効果，$(\alpha\gamma)_{wj}$ は A_w と C_j の交互作用効果である．Multistratum analysis において，これら処理効果の対比を推定するとき，ISPD では，(I) inter-block stratum, (II) inter-wholeplot stratum, (III) intra-unit stratum の3つの strata が重要となる．各 stratum の情報行列 A_1，A_2，A_3 は

$$A_1 = \frac{1}{k_1k_2}N_1N_1' - \frac{r^2}{bk_1k_2}\mathrm{J}_{v_1v_2}, \quad A_2 = \frac{1}{k_2}N_2N_2' - \frac{1}{k_1k_2}N_1N_1',$$
$$A_3 = r\mathrm{I}_{v_1v_2} - \frac{1}{k_2}N_2N_2'$$

によって与えられ，A_f/r $(f = 1, 2, 3)$ の固有値を ISPD の stratum efficiency factor (cf. Houtman and Speed (1983)) と呼ぶ．ここで，N_1，N_2 は，処理の組合せとブロック，処理の組合せと wholeplot の接合行列であり，I_n は n 次の単位行列，J_n は成分がすべて1の $n \times n$ の行列である．

本報告では，ISPD を構成するために，square lattice design と rectangular lattice design を用いる．Rectangular lattice design を wholeplot の処理計画，square lattice design を subplot の処理計画に対応させ，それらの semi-Kronecker 積 (cf. Mejza, Kuriki and Mejza (2001)) によって構成される ISPD $\mathcal{D}$ を考える．このとき，

$$v_1 = s_1(s_1-1),\ v_2 = s_2^2,\ k_1 = s_1 - 1,\ k_2 = s_2,$$

$$b = s_1 s_2 r,\ r \leq \min\{s_1, s_2+1\}$$

であり，次の定理が得られる．

定理 1. N_1N_1' は固有値 $r(s_1-1)s_2$, s_1s_2, $(s_1-1)s_2$, $(s_1-r)s_2$, 0 をもち，その重複度は 1, $(r-1)(s_1-1)$, $rs_1(s_2-1)$, s_1-1, $s_1s_2^2(s_1-1) - \{r(s_1-1)(s_2-1)+1\}$ である．

定理 2. N_2N_2' は固有値 rs_2, s_2, 0 をもち，その重複度は $s_1(s_1-1)$, $rs_1(s_1-1)(s_2-1)$, $s_1(s_1-1)(s_2-1)(s_2-r+1)$ である．

定理 3. $\mathcal{D}$ は *generally balanced* である．

これらの定理から表 1 の stratum efficiency factor が求められる．

表 1: stratum efficiency factor

Contrast	重複度	I	II	III
A	$(r-1)(s_1-1)$	$\frac{s_1}{(s_1-1)r}$	$1-\frac{s_1}{(s_1-1)r}$	-
	s_1-1	$\frac{s_1-r}{(s_1-1)r}$	$1-\frac{s_1-r}{(s_1-1)r}$	-
	$(s_1-1)(s_1-r)-1$	-	1	-
C	$r(s_2-1)$	$\frac{1}{r}$	-	$1-\frac{1}{r}$
	$s_2^2-r(s_2-1)-1$	-	-	1
$A\times C$	$r(s_1-1)(s_2-1)$	-	$\frac{1}{r}$	$1-\frac{1}{r}$
	$rs_1(s_1-2)(s_2-1)$	-	$\frac{1}{r}$	$1-\frac{1}{r}$
	$\{s_1(s_1-1)-1\}\{s_2^2-r(s_2-1)-1\}$	-	-	1

ある 2 つの計画により構成される ISBD の stratum efficiency factor が与えられているならば，同様の 2 つの計画により構成される ISPD の stratum efficiency factor を求めることができる．しかし，逆に ISPD から ISBD の stratum efficiency factor を求められるとは限らず，ISPD でのみ stratum efficiency factor を得ることができる ISPD は本報告が初めてである．

GD-OPTIMAL BALANCED FRACTIONAL 2^m FACTORIAL DESIGNS OF RESOLUTION $R^*(\{1\}|3)$

Masahide KUWADA (*Hiroshima University*)
Yoshifumi HYODO (*Okayama University of Science*)
Dong HAN (*Hiroshima University*)

We consider a fractional 2^m factorial design T with N assemblies, where the four-factor and higher-order interactions are assumed to be negligible. The vector of the ν_3 factorial effects not to be negligible is then given by $\Theta'=(\theta_0';\theta_1';\theta_2';\theta_3')$, where $\nu_3=1+m+m(m-1)/2+m(m-1)(m-2)/6$ and $\theta_u'=\{\theta_{t_1\ldots t_u} \mid 1\le t_1<\ldots<t_u\le m\}$ $(0\le u\le 3)$. Under these situations, the ordinary linear model is given by

$$y(T)=E_T\Theta+e_T,$$

where $y(T)$, E_T and e_T are a vector of N observations, the design matrix of size $N\times\nu_3$ and an error vector with mean 0_N and covariance $\sigma^2 I_N$, respectively. The normal equations for estimating Θ are given by

$$M_T\hat{\Theta}=E_T'y(T),$$

where $M_T(=E_T'E_T)$ is the information matrix of order ν_3.

Definition 1. When the $(\ell+1)$-factor and higher-order interactions are assumed to be negligible,

(I) if **all** the factorial effects from the p-factor interactions to the $(p+f)$-factor ones are estimable, then a design is said to be of **resolution** $R(\{p,p+1,\ldots,p+f\}|\ell)$, where $0\le p\le p+f\le\ell$. Especially when $p=0$ and $p+f=\ell$, it is of resolution $2\ell+1$, and when $p=0$ or 1 and $p+f=\ell-1$, it is of resolution 2ℓ,

(II) if **at least** the factorial effects from the q-factor interactions to the $(q+g)$-factor ones are estimable, then a design is said to be of **resolution** $R^*(\{q,q+1,\ldots,q+g\}|\ell)$, where $0\le q\le q+g\le\ell$.

Let T be a 2^m-BFF design derived from an SA$(m;\{\lambda_i\})$. Then the information matrix M_T is given by

$$M_T=\sum_{u=0}^{3}\sum_{v=0}^{3}\sum_{\alpha=0}^{\min(u,v)}\gamma_{|v-u|+2\alpha}D_\alpha^{(u,v)}=\sum_{u=0}^{3}\sum_{v=0}^{3}\sum_{\beta=0}^{\min(u,v)}\kappa_\beta^{u-\beta,v-\beta}D_\beta^{\#(u,v)},$$

where

$$\gamma_i=\sum_{j=0}^{m}\sum_{p=0}^{i}(-1)^p\binom{i}{p}\binom{m-i}{j-i+p}\lambda_j \quad \text{for } 0\le i\le 6,$$

$$\kappa_\beta^{u,v}(=\kappa_\beta^{v,u})=\sum_{\alpha=0}^{\beta+u}\gamma_{v-u+2\alpha}z_{\beta\alpha}^{(\beta+u,\beta+v)} \quad \text{for } 0\le u\le v\le 3-\beta \text{ and } 0\le\beta\le 3),$$

$$z_{\beta\alpha}^{(u,v)}=\sum_{b=0}^{\alpha}(-1)^{\alpha-b}\binom{u-\beta}{b}\binom{u-b}{u-\alpha}\binom{m-u-\beta+b}{b}\sqrt{\binom{m-u-\beta}{v-u}\binom{v-\beta}{v-u}}\Big/\binom{v-u+b}{b} \quad \text{for } u\le v.$$

By using the algebraic structure of the TMDPB association scheme, the M_T is isomorphic to $\|\kappa_\beta^{u,v}\|(=K_\beta$, say) of order $(4-\beta)$ for $0\le\beta\le 3$.

Proposition. *Let T be an* SA$(m;\{\lambda_i\})$. *Then* $\mathrm{rank}\{K_\beta\}=r_\beta$ $(0\le\beta\le 3)$ *if and only if just r_β of the indices λ_i $(\beta\le i\le m-\beta)$ are nonzero and λ_j $(j\ne i\ ;\ \beta\le j\le m-\beta)$ are all zero, where $r_\beta<4-\beta$. And if* $\mathrm{rank}\{K_\beta\}=r_\beta$ $(\le 4-\beta)$, *then the first r_β rows and columns of K_β are linearly independent.*

A parametric function $C\Theta$ of Θ is estimable for a known matrix C of order ν_3 if and only if there exists a matrix X of order ν_3 such that $XM_T=C$. If $C\Theta$ is estimable, then we have $\mathrm{Var}[C\hat{\Theta}]=\sigma^2XM_TX'$.

Theorem. *Let T be an* SA$(m;\{\lambda_i\})$ *with $N<\nu_3$ and $m\ge 6$. Then*

(I) *T is of resolution* $R(\{0,1,2\}|3)$, *i.e., resolution* VI, *if and only if one of the following holds*:

(i) $\lambda_0+\lambda_{2n}\neq 0$, $\lambda_i\neq 0$ (i=1,n,2n−1) *and* $\lambda_j=0$ ($j\neq i$; $2\leq j\leq 2n-2$), *where* $m=2n$,

(ii) $\lambda_0+\lambda_m\neq 0$, $\lambda_i\neq 0$ (i=1,2,m−2) *and* $\lambda_j=0$ ($j\neq i$; $3\leq j\leq m-1$), *or its complement*,

(iii) $\lambda_0\geq 0$, $\lambda_i\neq 0$ (i=1,2,m−2,m−1), $\lambda_m\geq 0$ *and* $\lambda_j=0$ ($3\leq j\leq m-3$), *where* $m\geq 7$,

(iv) $\lambda_i\neq 0$ (i=1,3,5) *and* $\lambda_j=0$ ($j\neq i$; $0\leq j\leq 6$), *where* m=6,

(v) $\lambda_i\neq 0$ (i=0,2,4,6) *and* $\lambda_j=0$ ($j\neq i$; $1\leq j\leq 5$), *where* m=6,

(II) *T is of resolution* R({0,1}| 3) *if and only if one of the following holds*:

(i) $\lambda_0+\lambda_m\neq 0$, $\lambda_i\neq 0$ (i=1,p,m−1; $3\leq p(\neq m/2)\leq m-3$) *and* $\lambda_j=0$ ($j\neq i$; $2\leq j\leq m-2$), *where* $m\geq 7$,

(ii) $\lambda_i\neq 0$ (i=0,1,n+1,2n+1) *and* $\lambda_j=0$ ($j\neq i$; $2\leq j\leq 2n$), *where* $m=2n+1$, *or its complement*,

(iii) $\lambda_0+\lambda_m\neq 0$, $\lambda_i\neq 0$ (i=1,2,m−1) *and* $\lambda_j=0$ ($3\leq j\leq m-2$), *or its complement*,

(III) *there does not exist a design of resolution* R({1,2}| 3), *i.e., resolution* VI,

(IV) *T is of resolution* R({1}| 3) *if and only if one of the following holds*:

(i) $\lambda_i\neq 0$ (i=1,2,7) *and* $\lambda_j=0$ ($j\neq i$; $0\leq j\leq 9$), *where* m=9, *or its complement*,

(ii) $\lambda_i\neq 0$ (i=1,2,5) *and* $\lambda_j=0$ ($j\neq i$; $0\leq j\leq 6$), *where* m=6, *or its complement*.

Let $K_0^*=PK_0P'$ and $K_\gamma^*=K_\gamma$ ($1\leq\gamma\leq 3$), where P=diag[$\begin{pmatrix}0&1\\1&0\end{pmatrix}$; I_2], and further let

$$\bar{\chi}_0^*(\alpha)=I_4 \quad \text{if } \det(K_0^*)\neq 0, \qquad \bar{\chi}_1^*(\alpha)=I_3 \quad \text{if } \det(K_1^*)\neq 0,$$

$$\text{diag}[1; g_0^{00}(\alpha); g_0^{22}(\alpha)] \quad \text{if rank}\{K_0^*\}=3, \qquad \text{diag}[1; g_1^{1,1}(\alpha)] \quad \text{if rank}\{K_1^*\}=2,$$

$$\bar{\chi}_2^*(\alpha)=I_2 \quad \text{if } \det(K_2^*)\neq 0, \qquad \bar{\chi}_3^*(\alpha)=1 \quad \text{if } K_3^*\neq 0,$$

$$g_2^{00}(\alpha) \quad \text{if rank}\{K_2^*\}=1, \qquad \text{vanish} \quad \text{if } K_3^*=0,$$

where if rank{ K_0^*}=3,

$$g_0^{00}(\alpha)=1 \quad \text{if } \alpha=0, \qquad g_0^{22}(\alpha)=1 \quad \text{if } \alpha=0,$$

$$1/(1+|w_0|) \quad \text{if } \alpha=1, \qquad 1/(1+|w_0^*|) \quad \text{if } \alpha=1,$$

$$1/\sqrt{1+(w_0)^2} \quad \text{if } \alpha=2, \qquad 1/\sqrt{1+(w_0^*)^2} \quad \text{if } \alpha=2,$$

and if rank{ K_γ^* }=3−γ ($1\leq\gamma\leq 2$),

$$g_\gamma^{2-\gamma,2-\gamma}(\alpha)=1 \quad \text{if } \alpha=0,$$

$$1/(1+|w_\gamma|) \quad \text{if } \alpha=1,$$

$$1/\sqrt{1+(w_\gamma)^2} \quad \text{if } \alpha=2.$$

Let $\bar{K}_\beta^*$ be the matrices given by the first r_β rows and columns of K_β^*, where rank{ K_β^* }=$r_\beta\geq 1$ for 0 $\leq\beta\leq 3$. Then the covariance matrix of the estimators of the linearly independent parametric functions in $C\Theta$ is isomorphic to $\sigma^2\bar{\chi}_\beta^*(\alpha)\bar{K}_\beta^{*-1}\bar{\chi}_\beta^*(\alpha)$. Thus we define $V_T(\alpha)$ as follows:

$$V_T(\alpha)=\Pi_\beta[\det\{\bar{\chi}_\beta^*(\alpha)\bar{K}_\beta^{*-1}\bar{\chi}_\beta^*(\alpha)'\}]^{\phi_\beta}.$$

Definition 2. Let T be a 2^m-BFF design of resolution R({1}| 3) (or R*({1}| 3)) with N assemblies derived from an SA(m;{ λ_i }). If $V_T(\alpha)\leq V_{T^*}(\alpha)$ for any T^*, which is a 2^m-BFF design of resolution R({1}| 3) (or R*({1}| 3)) with N assemblies derived from an SA(m;{ λ_i^* }), then T is said to be **GD$_\alpha$-optimal** ($0\leq\alpha\leq 2$).

Since $\bar{\chi}_\beta^*(\alpha)$ ($0\leq\beta\leq 3$; $0\leq\alpha\leq 2$) are diagonal and nonsingular, we define $V_T^*(\alpha)$ as follows:

$$V_T^*(\alpha)=\Pi_\beta[\det\{\{\bar{\chi}_\beta^*(\alpha)\bar{K}_\beta^{*-1}\bar{\chi}_\beta^*(\alpha)'\}^{-1}\}]^{\phi_\beta}=\Pi_\beta[\det\{\bar{K}_\beta^*\}/\{\det\{\bar{\chi}_\beta^*(\alpha)\}\}^2]^{\phi_\beta}.$$

Then we have $V_T^*(\alpha)=1/V_T(\alpha)$. And hence $V_{T_1}(\alpha)\leq V_{T_2}(\alpha)$ if and only if $V_{T_1}^*(\alpha)\geq V_{T_2}^*(\alpha)$, where T_1 and T_2 are 2^m-BFF designs of resolution R({1}| 3) (or R*({1}| 3)) with same number of assemblies derived from an SA(m;{ λ_i }) and an SA(m;{ λ_i^* }), respectively. Using $V_T^*(\alpha)$, we can obtain GD$_\alpha$-optimal 2^m-BFF designs of resolution R*({1}| 3).

Further results on partially balanced fractional $2^{m_1+m_2}$ factorial designs of resolution IV

Subir Ghosh (Univ. California)
葉田 正秀 (広島大・総合科学)
兵頭 義史 (岡山理大・理, 国際自然研)

単純部分均斉配列 SPBA$(m_1+m_2;\{\lambda_{i_1i_2}\})$ (i.e. PBA$(N,\ m_1+m_2,\ 2,\ m_1+m_2;\ \{\lambda_{i_1i_2}\})$) から得られる $2^{m_1+m_2}$- PBFF 計画 T を考える. ただし, 3 因子以上の高次交互作用は無視可能とし, $m_1, m_2 \geq 2$ とする. このとき, T に基づく線形模型は, $\varepsilon[\boldsymbol{y}(T)] = E_T\Theta$, $\mathrm{Var}[\boldsymbol{y}(T)] = \sigma^2 I_N$ で与えられる. ただし, $\boldsymbol{y}(T) : N \times 1$ 観測値ベクトル, $E_T : N \times \nu(m_1, m_2)$ 計画行列, $\Theta = (\Theta'_{00}; \Theta'_{10}; \Theta'_{01}; \Theta'_{20}; \Theta'_{02}; \Theta'_{11})'$: 2 - 因子交互作用までの $\nu(m_1, m_2) \times 1$ 要因効果ベクトル, $\nu(m_1, m_2) = 1 + m_1 + m_2 + \binom{m_1+m_2}{2}$ である.

ETMDPB アソシエーション代数 $\mathfrak{A} = [D_{\beta_1\beta_2}^{\#(a_1a_2,b_1b_2)}]$ の性質を用いて, 情報行列 $M_T(= E_T'E_T)$ は,

$$M_T = \sum_{\beta_1\beta_2}\sum_{a_1a_2}\sum_{b_1b_2} \kappa_{\beta_1\beta_2}^{a_1a_2,b_1b_2} D_{\beta_1\beta_2}^{\#(a_1+\beta_1a_2+\beta_2,b_1+\beta_1b_2+\beta_2)}$$

で与えられ, 次のブロック対角行列に相似となる :

$$\mathrm{diag}(K_{00}; \overbrace{K_{10}, \cdots, K_{10}}^{\phi_{10}}; \overbrace{K_{01}, \cdots, K_{01}}^{\phi_{01}}; \overbrace{K_{20}, \cdots, K_{20}}^{\phi_{20}}; \overbrace{K_{02}, \cdots, K_{02}}^{\phi_{02}}; \overbrace{K_{11}, \cdots, K_{11}}^{\phi_{11}})$$

ただし, $\phi_{\beta_1\beta_2} = \left\{\binom{m_1}{\beta_1} - \binom{m_1}{\beta_1-1}\right\}\left\{\binom{m_2}{\beta_2} - \binom{m_2}{\beta_2-1}\right\}$ で, 指標 $\lambda_{i_1i_2}$ のある線形式を要素とする実対称行列 $K_{\beta_1\beta_2}(= [\kappa_{\beta_1\beta_2}^{a_1a_2,b_1b_2}])$ $(\beta_1\beta_2 = 00, 10, 01, 20(m_1 \geq 4), 02(m_2 \geq 4), 11)$ は,

$$K_{\beta_1\beta_2} = \begin{cases} (D_{\beta_1\beta_2}F_{\beta_1\beta_2}D^*_{\beta_1\beta_2})(D_{\beta_1\beta_2}F_{\beta_1\beta_2}D^*_{\beta_1\beta_2})' & (H(\beta_1\beta_2) \neq \phi) \\ O & (H(\beta_1\beta_2) = \phi) \end{cases}$$

で与えられる. ここに $H(\beta_1\beta_2) \equiv \{(i_1, i_2) \mid \lambda_{i_1i_2} \neq 0,\ \beta_k \leq i_k \leq m_k - \beta_k (k = 1, 2)\}$, $D_{\beta_1\beta_2} : (m_1, m_2)$ のある関数を対角要素にもつ対角行列, $F_{\beta_1\beta_2}$: 非零指標 $\lambda_{a^i_{\beta_1\beta_2}x^i_{\beta_1\beta_2}}$ の添字 $(a^i_{\beta_1\beta_2}, x^i_{\beta_1\beta_2})$ のある関数を要素とする行列, $D^*_{\beta_1\beta_2}$: 非零指標 $\lambda_{a^i_{\beta_1\beta_2}x^i_{\beta_1\beta_2}}$ のある正値関数を対角要素にもつ対角行列である.

実対称行列 $L_{\beta_1\beta_2} = [\ell_{\beta_1\beta_2}^{a_1a_2,b_1b_2}]$ $(\beta_1\beta_2 = 00, 10, 01, 20(m_1 \geq 4), 02(m_2 \geq 4), 11)$ を次式で定義する :

$$L_{\beta_1\beta_2} = \begin{cases} F_{\beta_1\beta_2}F'_{\beta_1\beta_2} & (H(\beta_1\beta_2) \neq \phi) \\ O & (H(\beta_1\beta_2) = \phi) \end{cases}$$

このとき, $\mathrm{rank}(K_{\beta_1\beta_2}(i_1, \cdots, i_r)) = \mathrm{rank}(L_{\beta_1\beta_2}(i_1, \cdots, i_r))$ が成り立つ. ただし,
$X(i_1, \cdots, i_r)$ $(i_1 < \cdots < i_r)$: X の $i_1, \cdots, i_r$ 行, $i_1, \cdots, i_r$ 列からなる $r \times r$ 部分行列である.

本報告では, ETMDPB アソシエーション代数の性質および行列方程式を用いて, 次の (A)~(E) における要因効果ベクトルが推定可能となる分解能 IV の $2^{m_1+m_2}$- PBFF 計画の行列 $L_{\beta_1\beta_2}$ による特徴付けおよびそれらの例示を行った :

(A) $\Theta^{\mathrm{A}} \equiv (\Theta'_{00}; \Theta'_{10}; \Theta'_{01}; \Theta'_{20}; \Theta'_{02})'$
(B) $\Theta^{\mathrm{B}} \equiv (\Theta'_{00}; \Theta'_{10}; \Theta'_{01}; \Theta'_{20}; \Theta'_{11})'$ (or $(\Theta'_{00}; \Theta'_{10}; \Theta'_{01}; \Theta'_{02}; \Theta'_{11})'$)
(C) $\Theta^{\mathrm{C}} \equiv (\Theta'_{00}; \Theta'_{10}; \Theta'_{01}; \Theta'_{20})'$ (or $(\Theta'_{00}; \Theta'_{10}; \Theta'_{01}; \Theta'_{02})'$)
(D) $\Theta^{\mathrm{D}} \equiv (\Theta'_{00}; \Theta'_{10}; \Theta'_{01}; \Theta'_{11})'$
(E) $\Theta^{\mathrm{E}} \equiv (\Theta'_{00}; \Theta'_{10}; \Theta'_{01})'$

(A) の場合

定理 1 $\Theta^{\mathrm{A}} \equiv (\Theta'_{00}; \Theta'_{10}; \Theta'_{01}; \Theta'_{20}; \Theta'_{02})'$: 推定可能 $\iff L_{20}(= \ell_{20}^{00,00}) \neq 0\,(m_1 \geq 4)$, $L_{02}(= \ell_{02}^{00,00}) \neq 0\,(m_2 \geq 4)$ および次の条件 (I) または (II) が成り立つ :

(I) $\det(L_{00}) \neq 0$ の場合
(i) $\det(L_{10}) \neq 0$ のとき
(a) $\det(L_{01}) \neq 0$ のとき $L_{11}(= \ell_{11}^{00,00}) = 0$
(b) $\det(L_{01}) = 0$ のとき $m_1 = 2n_1\,(n_1 \geq 1)$, $\det(L_{01}^{\mathrm{A}}(11)) \neq 0$, $\ell_{01}^{10,10} = 0$

(ii) $\det(L_{10})=0$ のとき $m_2=2n_2\,(n_2\geq 1)$, $\det(L_{10}^{\mathrm{A}}(11))\neq 0$, $\ell_{10}^{01,01}=0$, $\det(L_{01})\neq 0$

(II) $\det(L_{00})=0$ の場合

$m_1=2n_1\,(n_1\geq 1)$, $m_2=2n_2\,(n_2\geq 1)$, $\det(L_{00}^{\mathrm{A}}(11))\neq 0$, $\ell_{00}^{11,11}=0$, $\det(L_{10})\neq 0$, $\det(L_{01})\neq 0$

ただし $L_{00}^{\mathrm{A}}(11)$: L_{00} の最初の 5×5 部分行列 ; $L_{10}^{\mathrm{A}}(11)$: L_{10} の最初の $2\times 2\,(m_1\geq 3)$, $1\times 1\,(m_1=2)$; $L_{01}^{\mathrm{A}}(11)$: L_{01} の最初の $2\times 2\,(m_2\geq 3)$, $1\times 1\,(m_2=2)$

[注意] ある $\eta_1\eta_2\,(=00,10,01,11)$ に対して, $\det(L_{\eta_1\eta_2})\neq 0 \Longrightarrow A_{\eta_1\eta_2}^{\#(11,11)}\boldsymbol{\Theta}_{11}$ も推定可能となる. ただし $A_{\beta_1\beta_2}^{\#(a_1a_2,b_1b_2)}$: $D_{\beta_1\beta_2}^{\#(a_1a_2,b_1b_2)}$ の $n(a_1a_2)\times n(b_1b_2)$ 部分行列 $(n(a_1a_2)=\binom{m_1}{a_1}\binom{m_2}{a_2})$

例 1 T を SPBA$(2+2;\{\lambda_{00}=\lambda_{01}=\lambda_{02}=\lambda_{11}=\lambda_{20}=\lambda_{22}=1,\ \lambda_{10}=\lambda_{12}=\lambda_{21}=0\})$ とする. このとき,

$$L_{00}=\begin{bmatrix} 6 & 2 & 0 & 8 & 4 & 0\\ 2 & 20 & 0 & 4 & -4 & 0\\ 0 & 0 & 16 & 0 & 0 & 0\\ 8 & 4 & 0 & 24 & 16 & 0\\ 4 & -4 & 0 & 16 & 24 & 0\\ 0 & 0 & 0 & 0 & 0 & 16 \end{bmatrix}$$

$\det(L_{00})=16777216$

$$L_{10}=\begin{bmatrix} 1 & 0\\ 0 & 0 \end{bmatrix}$$

$\det(L_{10})=0$

$L_{11}(=\ell_{11}^{00,00})=1$

$$L_{01}=\begin{bmatrix} 2 & 2\\ 2 & 4 \end{bmatrix}$$

$\det(L_{01})=4$

$\Longrightarrow \det(L_{00})\neq 0$, $\det(L_{10})=0$, $L_{10}^{\mathrm{A}}(11)\ (=\ell_{10}^{00,00})\neq 0$, $\ell_{10}^{01,01}=0$, $\det(L_{01})\neq 0$ であるから $\boldsymbol{\Theta}^{\mathrm{A}}$ は推定可能である. さらに $\det(L_{00})\neq 0$, $\det(L_{01})\neq 0$, $L_{11}(=\ell_{11}^{00,00})\neq 0$ より $A_{00}^{\#(11,11)}\boldsymbol{\Theta}_{11}$, $A_{01}^{\#(11,11)}\boldsymbol{\Theta}_{11}$, $A_{11}^{\#(11,11)}\boldsymbol{\Theta}_{11}$ も推定可能となる.

(B) の場合

定理 2 $\boldsymbol{\Theta}^{\mathrm{B}}\equiv(\boldsymbol{\Theta}'_{00};\boldsymbol{\Theta}'_{10};\boldsymbol{\Theta}'_{01};\boldsymbol{\Theta}'_{20};\boldsymbol{\Theta}'_{11})'$: 推定可能 $\Longleftrightarrow \det(L_{10}^{\mathrm{B}})\neq 0$, $L_{20}^{\mathrm{B}}(=\ell_{20}^{00,00})\neq 0\,(m_1\geq 4)$, $L_{11}^{\mathrm{B}}(=\ell_{11}^{00,00})\neq 0$ および次の条件 (I) または (II) が成り立つ :

(I) $\det(L_{00}^{\mathrm{B}})\neq 0$ の場合

(i) $\det(L_{01}^{\mathrm{B}})\neq 0$ のとき $m_2\geq 4$, $L_{02}^{\mathrm{B}}(=\ell_{02}^{00,00})=0$

(ii) $\det(L_{01}^{\mathrm{B}})=0$ のとき $m_2=2n_2\,(n_2\geq 2)$, $\det(L_{01}^{\mathrm{B}}(11))\neq 0$, $\ell_{01}^{01,01}=0$

(II) $\det(L_{00}^{\mathrm{B}})=0$ の場合

$m_2=k_2^2\,(k_2\geq 2)$, $\det(L_{00}^{\mathrm{B}}(11))\neq 0$, $\ell_{00}^{02,02}=0$, $\det(L_{01}^{\mathrm{B}})\neq 0$

ただし $L_{00}^{\mathrm{B}}(11)$: L_{00}^{B} の最初の 5×5 部分行列 ; $L_{01}^{\mathrm{B}}(11)$: L_{01}^{B} の最初の $2\times 2\,(m_2\geq 3)$

例 2 T を SPBA$(2+4;\{\lambda_{03}=\lambda_{11}=\lambda_{13}=\lambda_{21}=\lambda_{23}=1,\ \lambda_{00}=\lambda_{01}=\lambda_{02}=\lambda_{04}=\lambda_{10}=\lambda_{12}=\lambda_{14}=\lambda_{20}=\lambda_{22}=\lambda_{24}=0\})$ とする. このとき,

$$L_{00}^{\mathrm{B}}=\begin{bmatrix} 5 & -2 & -2 & 2 & -4 & 0\\ -2 & 12 & -4 & -4 & -8 & 0\\ -2 & -4 & 20 & -4 & -8 & 0\\ 2 & -4 & -4 & 20 & -8 & 0\\ -4 & -8 & -8 & -8 & 48 & 0\\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

$\det(L_{00}^{\mathrm{B}})=0$, $\det(L_{00}^{\mathrm{B}}(11))=262144$

$$L_{10}^{\mathrm{B}}=\begin{bmatrix} 2 & 0\\ 0 & 8 \end{bmatrix}$$

$\det(L_{10}^{\mathrm{B}})=16$

$$L_{01}^{\mathrm{B}}=\begin{bmatrix} 5 & -2 & -2\\ -2 & 12 & -4\\ -2 & -4 & 20 \end{bmatrix}$$

$\det(L_{01}^{\mathrm{B}})=960$

$L_{02}^{\mathrm{B}}(=\ell_{02}^{00,00})=0$, $L_{11}^{\mathrm{B}}(=\ell_{11}^{00,00})=2$

$\Longrightarrow \det(L_{00}^{\mathrm{B}})=0$, $\det(L_{00}^{\mathrm{B}}(11))\neq 0$, $\ell_{00}^{02,02}=0, \det(L_{10}^{\mathrm{B}})\neq 0$, $\det(L_{01}^{\mathrm{B}})neq0$, $L_{11}^{\mathrm{B}}(=\ell_{11}^{00,00})\neq 0$ であるから $\boldsymbol{\Theta}^{\mathrm{B}}$ は推定可能である. さらに $\det(L_{01}^{\mathrm{B}})\neq 0$ より $A_{01}^{\#(02,02)}\boldsymbol{\Theta}_{02}$ も推定可能となる.

(C)~(E) の場合に関する定理および例については省略する.

A-最適一部実施要因計画のシミュレーションによる検証

神戸市立工業高専　末次武明
神戸大学発達科学部　白倉暉弘

1．はじめに

次のような線形モデルを考える．

$$\mathbf{y} = E\theta + \mathbf{e}\,, \quad V(\mathbf{e}) = \sigma^2 I_N$$

ここで、$\nu = 1 + m + \binom{m}{2}$ とすると、$\mathbf{y}(N \times 1)$ は観測値ベクトル、$\theta(\nu \times 1)$ は2因子交互作用までの未知母数ベクトル、$E(N \times \nu)$ はそれに対する計画行列、$\mathbf{e}(N \times 1)$ は誤差ベクトル、σ^2 は誤差分散、I_N は大きさ N の単位行列である．

$\hat{\theta}$ を θ の最良線形不偏推定量、$M = E'E$ で M は正則であるとすると、

$$|\hat{\theta} - \theta|^2 = \mathbf{e}' E M^{-2} E' \mathbf{e}$$

$\psi = \mathbf{e}' E M^{-2} E' \mathbf{e}$ とするとき、各計画で、誤差を考えてシミュレーションによって ψ の値を比べることと、trace(M^{-1})（A-optimality）を比較してみる。

2．方法

ψ の計算で主要な部分である M^{-2} を計算するとき、BA(OA) のときには、計算の精度と速度を高めるために、次のような $\mathcal{K}_\alpha$ (Shirakura,T. (1993) 参照) の性質を利用している。

M の成分（$\gamma_0 \sim \gamma_4$）は、指数 $\mu_0 \sim \mu_4$ から導かれ、さらに、M は、対称行列 $\mathcal{K}_\beta$ $(\beta = 0, 1, 2)$ で表現される。

定理1 強さ4の BA から導かれる分解能5の計画に対して、M^{-2} は $\mathcal{K}_\beta^{-2}$ $(\beta = 0, 1, 2)$ で表現される。

したがって、M^{-2} の要素 $v_\alpha^{(i,j)}$ は、$\mathcal{K}_\beta^{-2}$ の要素 $\xi_{k,l}^{\beta}$ を用いて、表される。

但し、BA でない計画の場合は、E から、直接 M^{-2} を出して、計算している。

2つの計画の比較をするには、それぞれに $N(0,1)$ に従う違った誤差を与えて、ψ の値を比較し、値の小さい方がよいとして、その回数を数えている。

以下の結果では、$N = 32$, $m = 6$ の場合だけを挙げている。計算は基本的に 100 万回の繰り返しを行っている。

3．主な結果

3.1 OA どうしの比較

例1）次の OA(32,6,4;1) (trace $(M^{-1}) = 0.6875$) を考える。但し、$\Omega(6,k)$ は weight(1の個数) が k である処理組合せの集合とする。

$$\mathrm{OA1} : \Omega(6,1) + \Omega(6,3) + \Omega(6,5)$$

$$\mathrm{OA2} : \Omega(6,0) + \Omega(6,2) + \Omega(6,4) + \Omega(6,6)$$

$$\mathrm{OA3} : \begin{array}{c|c} OA(16,5,4;1) & \begin{matrix}1\\ \vdots \\ 1\end{matrix} \\ \hline OA(16,5,4;1) & \begin{matrix}0\\ \vdots \\ 0\end{matrix} \end{array}$$

・比較結果（計画1の方が良かった割合）

計画1－計画2	割合
$OA1 - OA2$	0.499608
$OA3 - OA2$	0.500254
$OA1 - OA3$	0.499859

［結論］OA どうしの比較では、ψ の値には、差がない。

3.2 OA と BA, BA どうしの比較

例２）次のような $N=32$, $m=6$, 強さ４の BA を考える。

BA1 : BA(32,6,4,{3,2,2,2,1}) (trace $(M^{-1})=0.7250$)
BA2 : BA(32,6,4,{4,2,2,1,4}) (trace $(M^{-1})=0.9179$)
BA3 : BA(32,6,4,{5,2,1,2,5}) (trace $(M^{-1})=0.9495$)
BA4 : BA(32,6,4,{4,2,1,2,6}) (trace $(M^{-1})=0.9510$)

・OA1 との比較、BA どうしの比較結果

但し、以下の「trace の差」では、符号を逆にしている。

計画１－計画２	割合	trace の差
$OA1-BA1$	0.543744	0.0375
$OA1-BA2$	0.725632	0.2304
$OA1-BA3$	0.757273	0.2620
$BA1-BA2$	0.719888	0.1929
$BA2-BA3$	0.532147	0.0316
$BA2-BA4$	0.534347	0.0331
$BA3-BA4$	0.501592	0.0015

［結論］OA と BA、あるいは BA どうしでは、trace (M^{-1}) に見合った差があるようである。

3.3 BA 以外の計画との比較

例４）次のような BA でない計画と、BA,OA を比較する。

DES1 : BA(30,6,4,{3,2,2,1,3})+110000+000011 (trace $(M^{-1})=0.8966$)
DES2 : BA(22,6,4,{1,2,2,2,1})+10 行 (trace $(M^{-1})=0.7965$)
DES3 : 64 行から１つおきに選んだ (trace $(M^{-1})=1.3254$)

計画１－計画２	割合	trace の差
$OA1-DES1$	0.707833	0.2091
$OA1-DES2$	0.62274	0.1090
$OA1-DES3$	0.87314	0.6421
$BA1-DES1$	0.668055	0.1716
$BA2-DES1$	0.479637	−0.0213

［結論］BA でない計画と比べても、trace (M^{-1}) に見合った差があるようである。

4．おわりに

simulation によって ψ の値を比べることと、A-optimality に用いられる trace (M^{-1}) の違いがよく対応していることが分かった。

特に、trace (M^{-1}) の差が小さい（0.03 程度）ときは、ほとんど差が出てこないことも分かった。

参考文献

Shirakua,T.(1993). Fractional factorial designs of two and three levels. Discrete Mathematics 116, 99-135,North-Holland.

Nishii,R. and Shirakura,T.(1986). More precise tables of Srivastava-Chopra balanced optimal 2^m fractional factorial designs of resolution V, $m \leq 6$, J. Statist. Plann. Inference 13,111-116.

Srivastava,J.N. and Chopra,D.V.(1971). Balanced optimal 2^m fractional factorial designs of resolution V, $m \leq 6$, Technometrics 13, 257-269.