数理解析研究所講究録
第 820 巻 1993 年 119-124

A series of identities for the coefficients of inverse matrices on a Hamming scheme

Hamming scheme 上の逆行列に関するある種の恒等式

岐阜大・工　三嶋美和子 (Miwako Mishima, Gifu Univ.)
岐阜大・工　神保　雅一 (Masakazu Jimbo, Gifu Univ.)

Abstract

In this paper, a series of identities on Hammimg schemes concerned with inverse matrices of linear combinations of association matrices is given, which is useful in the statistical design of experiments.

1. Statistical background

In time series analysis or spatial data analysis, correlated data are often treated. Similarly, in factorial experiments the closeness of level combinations (assemblies) of factors may cause the correlation between the observations. Here, we consider factorial experiments with p factors of q levels each.

Let $A_1, \cdots, A_p$ be factors with q levels. We assume that there are no interaction effects between these factors, and assume the following model :

$$y_i = \mu + \alpha^1_{\gamma_{i1}} + \alpha^2_{\gamma_{i2}} + \cdots + \alpha^p_{\gamma_{ip}} + \varepsilon_i \quad \text{for } i = 1, 2, \cdots, N, \tag{1}$$

where y_i and ε_i are the i-th observation and the error ; μ is the general mean ; γ_{ij} is the level of the j-th factor in the i-th experiments ; and α^j_k is the main effect for the k-th level of the j-th factor. Without loss of generality, we assume that

$$\alpha^j_0 + \alpha^j_1 + \cdots + \alpha^j_{q-1} = 0$$

holds for any j. Let $\Gamma = (\gamma_{ij})$ be an array of level combination (assembly). The model (1) can be also represented by

$$y = \mu\mathbf{1} + X\alpha + \varepsilon, \qquad \mathrm{cov}(\varepsilon) = \Sigma,$$

where $y = (y_1, \cdots, y_N)'$, $\varepsilon = (\varepsilon_1, \cdots, \varepsilon_N)'$, $\mathbf{1}$ is the all-one column vector, $\alpha = (\alpha^1_0, \cdots, \alpha^1_{q-1}; \cdots; \alpha^p_0, \cdots, \alpha^p_{q-1})'$, $X = (X_1, \cdots, X_p)$ is an $N \times pq$ design matrix and $X_j = (x^j_{ik})$ is an $N \times q$ matrix such that

$$x^j_{ik} = \begin{cases} 1 & \text{if } \gamma_{ij} = k, \\ 0 & \text{if } \gamma_{ij} \neq k. \end{cases}$$

Let $\gamma_i = (\gamma_{i1}, \cdots, \gamma_{ip})$ and $\gamma_k = (\gamma_{k1}, \cdots, \gamma_{kp})$ be the i-th row and the k-th row of Γ, respectively. The number of j such that $\gamma_{ij} \neq \gamma_{kj}$ is called the *Hamming distance* between γ_i and γ_k, denoted by $d(\gamma_i, \gamma_k)$. We assume that the correlation of errors ε_i and ε_k of the i-th and the k-th experiments depends only on the Hamming distance $d(\gamma_i, \gamma_k)$ and that two kinds of covariance structure are given as follows :

$$\mathrm{cov}(\varepsilon_i, \varepsilon_k) = \begin{cases} \sigma^2 & \text{if } i = k, \\ \sigma'^2 & \text{if } \gamma_i = \gamma_k \text{ and } i \neq k, \\ \sigma^2 \rho^{d(\gamma_i, \gamma_k)} & \text{if } i \neq k \text{ and } d(\gamma_i, \gamma_k) > 0. \end{cases} \tag{2}$$

$$\mathrm{cov}(\varepsilon_i, \varepsilon_k) = \begin{cases} \sigma^2 & \text{if } i = k, \\ \sigma'^2 & \text{if } \gamma_i = \gamma_k \text{ and } i \neq k, \\ \sigma^2\rho & \text{if } d(\gamma_i, \gamma_k) = 1, \\ 0 & \text{if } d(\gamma_i, \gamma_k) > 1. \end{cases} \tag{3}$$

We assume that $\rho \geq 0$, which may be natural in usual cases. If $\rho = 0$ (that is, for uncorrelated case), the "optimality" of orthogonal array is well-known (see, for example, Kiefer (1975), Kiefer and Wynn (1981)). In a practical use, "linear" orthogonal array is often utilized. Our aim is to find designs which are "optimal" not only for $\rho = 0$ but also for any $\rho \geq 0$. Thus we restrict ourselves to the class of linear orthogonal arrays. When we use the *generalized least square estimate* (*GLSE*), we have to calculate the C-matrix $C(X) = X'\Sigma^{-1}X$ in order to evaluate the "optimality" of the design. In the midst of the calculation, we need to obtain the values

$$\sum_{i=0}^{n} b_i \binom{n}{i} (q-1)^i, \qquad \sum_{i=1}^{n} b_i \binom{n-1}{i-1} (q-1)^{i-1}, \tag{4}$$

where n is an integer ($\leq N$) and b_i's are real numbers corresponding to Σ^{-1} which will be defined in the next section. These formulas are evaluated in the theorems of the next section. For the details in this section, we refer the reader to Mishima and Jimbo (1992).

2. Identities

Let $U = \{u_1, \cdots, u_v\}$ be a finite set. We assume that $n+1$ binary relations $R_0, \cdots, R_n$ are defined on the set U. Let $D_i = (d_{kl}^{(i)})$ be a $v \times v$ $(0,1)$-matrix such that

$$d_{kl}^{(i)} = \begin{cases} 1 & \text{if } (u_k, u_l) \in R_i, \\ 0 & \text{otherwise.} \end{cases}$$

Definition. The $(n+1)$-tuple $< D_0, \cdots, D_n >$ is called an *association scheme* on a finite set U of v points if D_i satisfies the following conditions :

(i) D_i is symmetric for $i = 0, 1, \cdots, n$,

(ii) $\sum_{i=0}^{n} D_i = J_v$, where J_v is the $v \times v$ all-one matrix,

(iii) $D_0 = I_v$, where I_v is the $v \times v$ identitiy matrix,

(iv) $D_i D_j = \sum_{k=0}^{n} c_{ijk} D_k = D_j D_i$ for $i, j = 0, 1, \cdots, n$, where c_{ijk} is a constant depending on i, j and k.

It is well-known that the number of 1's contained in a row or a column of D_i is a constant ($=v_i$) not depending on the particular choice of a row or a column. And the vector space consisting of all matrices $\sum_{i=0}^{n} a_i D_i$ is a ring (see, for example, MacWilliams and Sloane (1977)). It is obvious that if an element $\sum_{i=0}^{n} a_i D_i$ has the inverse, then $(\sum_{i=0}^{n} a_i D_i)^{-1}$ can be written by a linear combination of $D_0, \cdots, D_n$, if it exists.

The vector space consisting of all matrices $\sum_{i=0}^{n} a_i D_i$ has the unique basis of primitive idempotents $E_0, \cdots, E_n$. Let

$$D_i E_j = p_i(j) E_j,$$

where $p_i(j), (j = 0, \cdots, n)$, are the eigenvalues of D_i. Let

$$P = \begin{bmatrix} p_0(0) & p_1(0) & \cdots & p_n(0) \\ p_0(1) & p_1(1) & \cdots & p_n(1) \\ \vdots & \vdots & \ddots & \vdots \\ p_0(n) & p_1(n) & \cdots & p_n(n) \end{bmatrix}$$

be the first eigenmatrix and let

$$Q = vP^{-1} = \begin{bmatrix} q_0(0) & q_1(0) & \cdots & q_n(0) \\ q_0(1) & q_1(1) & \cdots & q_n(1) \\ \vdots & \vdots & \ddots & \vdots \\ q_0(n) & q_1(n) & \cdots & q_n(n) \end{bmatrix}.$$

be the second eigenmatrix.

It is known that $p_0(j) = q_0(j) = 1$ and $p_i(0) = v_i$. And the following proposition is obtained.

Proposition. For an association scheme $< D_0, \cdots, D_n >$ on U, let $(\sum_{k=0}^n a_k D_k)^{-1} = \sum_{i=0}^n b_i D_i$, then

$$b_i = \frac{1}{v} \sum_{j=0}^n \frac{q_j(i)}{\sum_{k=0}^n a_k p_k(j)} \tag{5}$$

holds and we have

$$\sum_{i=0}^n b_i p_i(l) = \frac{1}{\sum_{k=0}^n a_k p_k(l)} \qquad \text{for } l = 0, 1, \cdots, n\ . \tag{6}$$

Proof. By noting the relations $D_i = \sum_{j=0}^n p_i(j) E_j$ and $E_i = \frac{1}{v} \sum_{j=0}^n q_i(j)\, D_j$, we have

$$\begin{aligned} \sum_{i=0}^n b_i D_i &= (\sum_{k=0}^n a_k D_k)^{-1} = (\sum_{k=0}^n \sum_{j=0}^n a_k p_k(j)\, E_j)^{-1} \\ &= \sum_{j=0}^n (\sum_{k=0}^n a_k p_k(j))^{-1} E_j = \frac{1}{v} \sum_{i=0}^n \sum_{j=0}^n (\sum_{k=0}^n a_k p_k(j))^{-1} q_j(i)\, D_i. \end{aligned}$$

Thus we obtain (5). Furthermore, by using $\sum_{i=0}^n q_j(i) p_i(l) = \delta_{jl}$, we have

$$\begin{aligned} \sum_{i=0}^n b_i p_i(l) &= \frac{1}{v} \sum_{i=0}^n \sum_{j=0}^n (\sum_{k=0}^n a_k p_k(j))^{-1} q_j(i) p_i(l) \\ &= \sum_{j=0}^n (\sum_{k=0}^n a_k p_k(j))^{-1} \cdot \delta_{jl} = (\sum_{k=0}^n a_k p_k(l))^{-1}. \end{aligned}$$

Thus the proposition is proved. □

Now, let $U = F^n$, where $F = \{0, 1, \cdots, q-1\}$. And define the relation R_i as $(x, y) \in R_i$ if $d(x, y) = i$, then $< D_0, \cdots, D_n >$ is an association scheme, which is called a *Hamming scheme.* In this case, $p_i(0) = v_i = \binom{n}{i}(q-1)^i$ holds.

Theorem 1. For the Hamming scheme $< D_0, \cdots, D_n >$ on F^n, let $(\sum_{k=0}^{n} \rho^k D_k)^{-1} = \sum_{i=0}^{n} b_i D_i$, then

$$b_i = \frac{(-\rho)^i \{1 + \rho(q-2)\}^{n-i}}{(1-\rho)^n \{1 + \rho(q-1)\}^n} \tag{7}$$

holds for $i = 0, 1, \cdots, n$. Furthermore

$$\sum_{i=m}^{n} b_i \binom{n-m}{i-m} (q-1)^{i-m} = \frac{(-\rho)^m}{(1-\rho)^m \{1 + \rho(q-1)\}^n} \tag{8}$$

holds for $m = 0, 1, \cdots, n$.

Remark. Especially, let $m = 0$ and 1 in (8), then we can obtain the value of (4) for the covariance structure (2) in the previous section.

Proof. In the case of a Hamming scheme, it is known that

$$p_k(j) = q_k(j) = P_k(j; n) = \sum_{i=0}^{k} (-q)^i (q-1)^{k-i} \binom{j}{i} \binom{n-i}{k-i}, \tag{9}$$

where $P_k(j; n)$ is called a *Krawtchouk polynomial* (for the properties of the Krawtchouk polynomial, see MacWilliams and Sloane (1977)).

By using (9), we have

$$\begin{aligned}
\sum_{k=0}^{n} \rho^k p_k(j) &= \sum_{k=0}^{n} \sum_{i=0}^{k} (-q)^i (q-1)^{k-i} \binom{j}{i} \binom{n-i}{k-i} \rho^k \\
&= \sum_{i=0}^{j} \sum_{k'=0}^{n-i} (-q)^i (q-1)^{k'} \binom{j}{i} \binom{n-i}{k'} \rho^{k'+i} \\
&= \{1 + \rho(q-1)\}^{n-j} (1-\rho)^j.
\end{aligned}$$

Then by (5) and (9),

$$\begin{aligned}
b_i &= \frac{1}{q^n} \sum_{j=0}^{n} \frac{1}{(1-\rho)^j \{1 + \rho(q-1)\}^{n-j}} \sum_{l=0}^{i} (-q)^l (q-1)^{j-l} \binom{i}{l} \binom{n-l}{j-l} \\
&= \frac{1}{q^n} \frac{1}{\{1 + \rho(q-1)\}^n} \sum_{l=0}^{i} \sum_{j'=0}^{n-l} (-q)^l (q-1)^{j'} \binom{i}{l} \binom{n-l}{j'} \left\{ \frac{1 + \rho(q-1)}{1-\rho} \right\}^{j'+l}
\end{aligned}$$

holds. After an straightforward but somewhat tedious calculation, we obtain (7). And it is easy to show (8). □

Theorem 2. For a Hamming scheme $< D_0, \cdots, D_n >$ on F^n, let $(D_0 + \rho D_1)^{-1} = \sum_{i=0}^{n} b_i D_i$, then

$$\sum_{i=m}^{n} b_i \binom{n-m}{i-m} (q-1)^{i-m} = \frac{m!(-\rho)^m}{\prod_{k=0}^{m} \{1 + n\rho(q-1) - k\rho q\}} \tag{10}$$

holds for $m = 0, 1, \cdots, n$.

Remark. Let $m = 0$ and 1, then we can obtain the value of (4) for the covariance structure (3) in the previous section. We may also prove (10) by using the properties Krawtchouk polynomial in a similar manner to Theorem 1. But the following proof may be simpler and the explicit formula of b_i is more complicated than that of Theorem 1.

Proof. Let $g(m) = \sum_{i=m}^{n} b_i \binom{n-m}{i-m}(q-1)^{i-m}$. When $m = 0$, let $l = 0$, $a_0 = 1$, $a_1 = \rho$ and $a_2 = \cdots = a_n = 0$ in (6), then it is easy to show that

$$g(0) = \sum_{i=0}^{n} b_i \binom{n}{i} (q-1)^i = \frac{1}{1+n\rho(q-1)}, \tag{11}$$

since $p_i(0) = v_i = \binom{n}{i}(q-1)^i$ holds.

Now, let $< D_0^{(m)}, \cdots, D_{n-m}^{(m)} >$ be a Hamming scheme on F^{n-m} and let

$$Z^{(m)} = \left(D_0^{(0)} + \rho D_1^{(0)}\right)^{-1} \cdot \left\{\left(\bigotimes_{i=1}^{m} \mathbf{e}\right) \otimes \mathbf{1}_{q^{n-m}}\right\} = \left(\sum_{i=0}^{n} b_i D_i^{(0)}\right) \cdot \left\{\left(\bigotimes_{i=0}^{m} \mathbf{e}\right) \otimes \mathbf{1}_{q^{n-m}}\right\},$$

where $\mathbf{e} = (0, \cdots, 0, 1)'$ is a q-dimensional vector, $\mathbf{1}_{q^{n-m}}$ is the q^{n-m}-dimensional all-one column vector, $\otimes$ indicates the direct product and $\bigotimes_{i=1}^{m} \mathbf{e} = \underbrace{\mathbf{e} \otimes \cdots \otimes \mathbf{e}}_{m}$.

For any two matrices A and B, let $\mathbf{S}$ be the set of the all ordered $(l+h)$-tuples consist of l A's and h B's and define the following function :

$$f(A, B; l, h) = \sum_{(S_1, \cdots, S_{l+h}) \in \mathbf{S}} \bigotimes_{i=1}^{l+h} S_i.$$

Since $D_i^{(j-1)} = (J_q - I_q) \otimes D_{i-1}^{(j)} + I_q \otimes D_i^{(j)}$ holds for $0 \leq i \leq n-j+1$, where $D_{-1}^{(j)} = D_{n-j+1}^{(j)} = 0$, we obtain

$$\begin{aligned} \sum_{i=0}^{n} b_i D_i^{(0)} &= \sum_{i=1}^{n} b_i (J_q - I_q) \otimes D_{i-1}^{(1)} + \sum_{i=0}^{n-1} b_i I_q \otimes D_i^{(1)} = \cdots \\ &= \sum_{k=0}^{m} \left\{ f(J_q - I_q, I_q; m-k, k) \otimes \sum_{i=m-k}^{n-k} b_i D_{i-m+k}^{(m)}, \right\}, \end{aligned}$$

where I_q is the $q \times q$ identity matrix and J_q is the $q \times q$ all-one matrix.

Furthermore, since $(A_1 \otimes \cdots \otimes A_m) \cdot (B_1 \otimes \cdots \otimes B_m) = A_1 B_1 \otimes \cdots \otimes A_m B_m$ holds for any matrices $\{A_i\}$ and $\{B_i\}$,

$$f(J_q - I_q, I_q; m-k, k) \cdot \left(\bigotimes_{i=0}^{m} \mathbf{e}\right) = f(\mathbf{j} - \mathbf{e}, \mathbf{e}; m-k, k)$$

holds, where $\mathbf{j}$ is the q-dimensional all-one column vector. Thus by noting that $D_i^{(m)} \mathbf{1}_{q^{n-m}} = \binom{n-m}{i}(q-1)^i \mathbf{1}_{q^{n-m}}$ holds, we can rewrite $Z^{(m)}$ as follows :

$$Z^{(m)} = \sum_{k=0}^{m} \left\{ f(\mathbf{j} - \mathbf{e}, \mathbf{e}; m-k, k) \otimes \sum_{i=m-k}^{n-k} b_i \binom{n-m}{i-m+k} (q-1)^{i-m+k} \mathbf{1}_{q^{n-m}} \right\}.$$

On the other hand, let $\bar{\mathbf{e}} = (1, 0, \cdots, 0)'$, then similarly, we obtain

$$\begin{aligned}
0 &= \left\{ \left(\bigotimes_{i=1}^{m} \bar{\mathbf{e}}' \right) \otimes \mathbf{1}'_{q^{n-m}} \right\} \cdot \left\{ \left(\bigotimes_{i=1}^{m} \mathbf{e} \right) \otimes \mathbf{1}_{q^{n-m}} \right\} \\
&= \left\{ \left(\bigotimes_{i=1}^{m} \bar{\mathbf{e}}' \right) \otimes \mathbf{1}'_{q^{n-m}} \right\} \cdot \left(D_0^{(0)} + \rho D_1^{(0)} \right) Z^{(m)} \\
&= \left\{ \left(\bigotimes_{i=1}^{m} \bar{\mathbf{e}}' \right) \otimes \mathbf{1}'_{q^{n-m}} \right\} \cdot \left\{ \rho f(J_q - I_q, I_q; 1, m-1) \otimes D_0^{(m)} + \left(\bigotimes_{i=1}^{m} I_q \right) \otimes \left(D_0^{(m)} + \rho D_1^{(m)} \right) \right\} Z^{(m)} \\
&= \rho \left(f(\mathbf{j}' - \bar{\mathbf{e}}', \bar{\mathbf{e}}'; 1, m-1) \otimes \mathbf{1}'_{q^{n-m}} \right) Z^{(m)} + \{1 + \rho(n-m)(q-1)\} \left\{ \left(\bigotimes_{i=1}^{m} \bar{\mathbf{e}}' \right) \otimes \mathbf{1}'_{q^{n-m}} \right\} Z^{(m)}.
\end{aligned} \tag{12}$$

It is obvious that

$$f(\mathbf{j} - \mathbf{e}, \mathbf{e}; m-k, k) \cdot f(\mathbf{j}' - \bar{\mathbf{e}}', \bar{\mathbf{e}}'; 1, m-1) = \begin{cases} m(q-2) & \text{if } k = 0, \\ m & \text{if } k = 1, \\ 0 & \text{otherwise} \end{cases} \tag{13}$$

holds. Thus by using (13) and $\binom{n-m+1}{i-m+1} = \binom{n-m}{i-m+1} + \binom{n-m}{i-m}$, the first term of the right hand side of (12) is

$$\left(f(\mathbf{j}' - \bar{\mathbf{e}}', \bar{\mathbf{e}}'; 1, m-1) \otimes \mathbf{1}'_{q^{n-m}} \right) Z^{(m)} = q^{n-m} \left\{ m(q-2) \cdot g(m) + m \cdot g(m-1) \right\},$$

and similarly the second term is

$$\left\{ \left(\bigotimes_{i=1}^{m} \bar{\mathbf{e}}' \right) \otimes \mathbf{1}'_{q^{n-m}} \right\} Z^{(m)} = q^{n-m} \cdot g(m).$$

Therfore by (12) we have

$$\{1 + \rho(n-m)(q-1)\} \cdot g(m) = -m\rho \cdot g(m-1),$$

which proves the theorem together with (11). □

References

[1] J. Kiefer (1975), Construction and optimality of generalized Youden designs II. In *A Survey of Statistical Designs and Linear Models* (J. N. Srivastava, ed.) 333-353, North-Holland, Amsterdam.

[2] J. Kiefer and H. P. Wynn (1981), Optimum balanced incomplete block and Latin square designs for correlated observations. *Ann. Statist.* 737-757.

[3] F. J. MacWilliams and N. J. A. Sloane (1977), *The theory of error-correcting codes.* North-Holland, New York.

[4] M. Mishima and M. Jimbo (1992), Optimum orthogonal designs with multilevel-factors for correlated errors (in preparation).