The Korean Communications in Statistics
Vol. 13 No. 3, 2006, pp.777-786

On the Geometric Equivalence of Asymmetric Factorial Designs[1)]

DongKwon Park[2)] and EunHye Park[3)]

Abstract

Two factorial designs with quantitative factors are called geometrically equivalent if the design matrix of one can be transformed into the design matrix of the other by row and column permutations, and reversal of symbol order in one or more columns. Clark and Dean (2001) gave a sufficient and necessary condition (which we call the "*gCD* condition") for two symmetric factorial designs with quantitative factors to be geometrically equivalent. This condition is based on the absolute value of the Euclidean(or Hamming) distance between pairs of design points. In this paper we extend the *gCD* condition to asymmetric designs. In addition, a modified algorithm is applied for checking the equivalence of two designs.

Keywords : Asymmetric; geometrically equivalent; quantitative.

1. 서 론

요인실험은 대개 산업체나 과학적인 연구에서 활용되는데 이 경우 각 요인수준을 고정하고 몇 개의 요인조합을 선택하게 된다. 실험계획법에서 실험요인은 질적(qualitative) 혹은 양적(quantitative) 요인으로 나뉜다. 질적 요인의 경우 수준 간 순위는 없다. 따라서 질적 요인의 수준은 수적인 의미를 가지지 않는다. 반면에 양적 요인의 수준은 요인의 실제적인 숫자 값으로 표현된다. 질적 요인의 실험의 목적은 처리 평균 간의 비교와 차이가 있는 경우 그 수준들을 찾는 따위의 분석이 가능하여 ANOVA나 다중비교 방법들이 이용 가능하다. 반면에 반응표면 분석 등은 수준간의 순위에 관심을 두고 양적 요인에서 실험하게 된다. 주요 연구 목적도 어떻게 잘 반응값과 요인간의 관계를 다항식 등을 통해 잘 해석하느냐에 있다.

요인 수준의 양적, 질적 형태에 따라 두 종류의 동등성(equivalence 혹은 isomorphism)으로 나누어 생각 할 수 있다. Cheng과 Ye (2004)의 용어를 사용하여

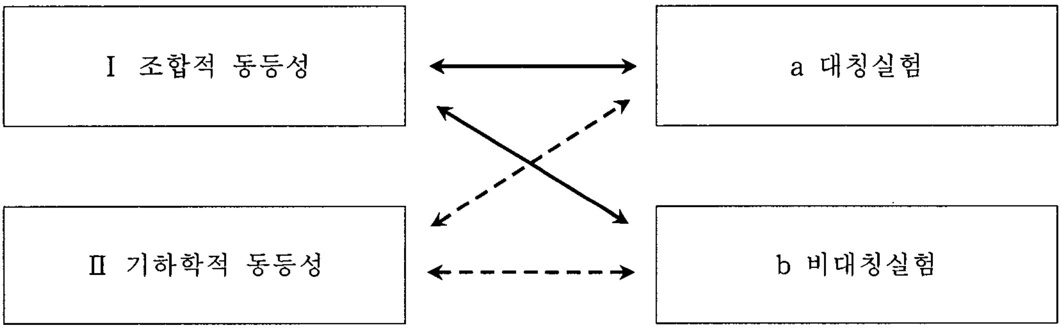
---

1) Park's work was supported by Yonsei University Research Fund of 2004.
2) Professor, Department of Informations and Statistics, Yonsei University, Wonju, Kangwon -Do, 220-710, Korea.
Correspondence : statpdk@yonsei.ac.kr
3) Graduate Student, Department of Applied Statistics, Yonsei University, Wonju, Kangwon -Do, 220-710, Korea.

“기하학적(geometric) 동등성“과 “조합적(combinatorial) 동등성”으로 각각 분류하기로 하고 아래와 같이 정의한다.

양적 요인실험 d_1과 d_2는 한 실험이 다른 실험의 요인과 처리 조합을 재배열(reordering)하고, 하나 혹은 그 이상의 요인 수준을 *역으로(reversing)* 배열함으로써 얻어질 수 있다면 기하학적 동등성을 갖는다고 한다. 반면에 조합적 동등성은 하나 혹은 그 이상의 요인 수준의 임의의 주어진 한 배열에 대해서 가능한 경우의 동등성이다. Cheng과 Wu (2001)에서 보였듯이 예로서 실험 3^{4-1}에서 양적 수준의 단순한 조합적 변환은 모형의 효율성에 영향을 줄 수 있다. 따라서, 양적 자료의 경우 효율성을 유지하기 위해서는 기하학적 동등성을 유지하는 것이 필요하다.

요인실험의 경우 모든 요인 수준이 같은 실험을 대칭(symmetric) 실험, 그렇지 않은 실험을 비대칭(asymmetric) 혹은 혼합(mixed) 실험이라 한다. 따라서, 요인실험간 동등성 비교는 다음과 같은 네 경우로 분리하여 생각 할 수 있다.



(I-a)의 경우, 특히 이요인 실험의 동등성을 비교하기 위한 방법들이 약 50여 년 동안 관심이 되어 왔고 실험계획의 분류와 순위 매김 등에 사용되어 왔다(참조 Xu (2005)). Draper와 Mitchell (1967)은 두 실험의 동등성에 대한 필요조건을 두 실험이 같은 단어길이패턴(Word Length Pattern)을 갖는 것이라고 정의하였으나 이후 동일한 단어길이패턴을 갖지만 글자패턴(Letter Pattern)이 달라 동등성을 만족하지 못하는 두 실험의 예를 보였다 (1968). 그 후 Draper와 Mitchel l(1970)은 두 실험 d_1과 d_2에 대해 A_d는 글자패턴이고 R을 행에 대한 어떠한 순서변환 행렬이라고 할 때, $A_{d_1} = RA_{d_2}$를 만족하면 두 실험은 동등성을 가지며 이는 실험의 동등성에 대한 필요조건이라고 생각하였다. 하지만 Chen과 Lin (1991)은 선명도 Ⅶ를 갖는 2^{31-16}실험 d_1과 d_2는 같은 단어길이패턴과 글자패턴은 4차 상호작용이 고차의 상호작용과 교락되는 것을 발견하였으며 같은 글자패턴을 갖지만 실험에서 모든 4차 상호작용이 3차와 4차의 상호작용과 교락되어 2^{k-p} 실험의 동등성에 대한 충분조건이 아니라고 결론을 내렸다. Clark과 Dean (2001)은 처음으로 대칭 실험의 조합적 동등성을 갖기 위한 필요충분조건을 제시하였는데 그 조건은 두 실험의 처리행렬 T_{d1}의 행과 T_{d2}의 행 사이의 해밍거리(Hamming distance)에 기초로 한다.

(II-a)의 경우, 즉 대칭 실험의 기하학적 동등성을 보기 위해 최근에는 대수기하학

(algebraic geometry)을 활용한 지수함수(Indicator function)를 이용하는 방법이 소개되고 있다.(Cheng and Ye (2004)). 이 방법은 기하학적 교락 구조를 이해하는데 매우 효과적인 방법이라 할 수 있다. 그러나, 비대칭요인 실험의 경우는 지수함수의 형태의 복잡함으로 인해 아직까지 활용이 어렵다는 문제점이 있다.

본 논문에서는 (*II-a*) 실험을 포함하는 (*II-b*)의 경우, 즉 비대칭 실험의 기하학적 동등성에 대해 Clark and Dean (2001)이 제안한 해밍 거리를 수정한 Euclidean 거리를 이용한 필요충분조건을 수정 제시한다. 이를 근거로 비대칭 양적 요인에 대한 알고리즘을 수정해 적용한다.

2. 해밍 거리와 Euclidean 거리

처리행렬 T_d를 요인실험 d의 i번째 실행에서 관찰된 k번째 요인에서의 수준인 (i,k)번째 요소로 구성된 $n\times p(i=1,2,\cdots,n;\ k=1,2,\cdots,p)$ 행렬이라고 하자. T_d의 행은 실험 d에서 관찰된 n개의 처리조합으로 나타난다. 실험에서의 조합적 동등성은 처리행렬의 동등성에 의해 다음과 같이 표현될 수 있다. 만약 T_{d_1}의 행과 열의 순서를 바꾸고, 하나 혹은 그 이상의 열에서 기호를 다르게 표시함으로써 T_{d_2}가 얻어진다면 두 실험 d_1과 d_2는 동등하다고 한다. 예를 들어 처리행렬 T_{d_1}과 T_{d_2}가 다음과 같다고 하자.

$$T_{d_1}=\begin{bmatrix}1&1\\2&1\\3&2\\3&2\\3&3\end{bmatrix}\qquad T_{d_2}=\begin{bmatrix}1&3\\2&3\\3&1\\3&2\\1&3\end{bmatrix}$$

처리행렬 T_{d_1}과 T_{d_2}에 해당하는 실험 d_1과 d_2는 T_{d_1}의 첫 번째 열에서 1을 2로, 2를 3으로, 3을 1로 각각 다르게 표시하면 T_{d_2}는 다음과 같다.

$$T_{d_2}=\begin{bmatrix}2&3\\3&3\\1&1\\1&2\\2&3\end{bmatrix}$$

그 다음 T_{d_2}의 첫 번째 열과 두 번째 열을 바꾸면 T_{d_2}는 다음과 같게 된다.

$$T_{d_2}=\begin{bmatrix}3&2\\3&3\\1&1\\2&1\\3&2\end{bmatrix}$$

위에서 얻어진 T_{d_2}의 처리조합의 순서를 재배열하게 되면 T_{d_1}과 동일함을 알 수 있다.

이와 같이 두 부분요인실험 중 한 실험의 요인의 이름을 다르게 표시하고(factor relabeling), 요인 수준의 이름을 다르게 표시하고(factor level relabeling), 처리조합의 순서를 재배열(reordering)함으로써 다른 실험이 얻어지면 이 두 실험은 조합적 동등

성을 갖는다고 정의한다.

Clark과 Dean (2001)은 두 이요인 대칭실험 d_1과 d_2가 조합적 동등성을 갖기 위한 필요충분조건을 제시하였는데 그 조건은 아래에서 정의되는 두 처리행렬간의 해밍거리(Hamming distance)에 기초로 한다. T_d의 k번째 열에서 i번째와 j번째의 행에서 요인이 다르다면 $\delta[T_d]_{i,j}^k=1$이라고 하고, 같다면 $\delta[T_d]_{i,j}^k=0$ 이라고 하자. 만일 p-차원 공간에서 n개의 점들로 실험을 표현할 수 있다면 $\sum_{k=0}^{p}\delta[T_d]_{i,j}^k$는 일치하지 않는 i번째와 j번째 점에서 차원의 수를 계산한다. 즉, 두 행 사이에서 차이가 나는 요인의 개수가 해밍거리이다. 따라서 (i,j)번째 요소를 갖는 실험 d의 거리행렬 H_d를 다음과 같이 정의할 수 있다.

$$[H_d]_{i,j}=\begin{cases}\sum_{k=1}^{p}\delta[T_d]_{i,j}^k\ , & i\neq j\\ 0 \qquad , & i=j\end{cases} \tag{2.1}$$

거리행렬 H_d는 열의 순서변환과 T_d의 열에서의 요인의 이름이 바뀌는 것에 영향을 받지 않는다. 다시 말해 실험 d에서 요인의 이름과 요인 수준의 이름을 다르게 표시해도 거리행렬은 변하지 않는다. 그러나 처리조합의 순서를 바꾼 경우 즉, 행의 순서가 변하는 경우 거리행렬을 변하게 된다. 따라서 두 실험 d_1과 d_2의 동등성에 대한 필요조건은 다음과 같은 변환행렬 R이 존재하는 것이다.

$$H_{d_1}=RH_{d_2}R' \tag{2.2}$$

이번에는 기하학적 동등성을 살펴보자. 요인 수준이 둘 인 경우는 물론 두 동등성이 같다. 각각의 양적 요인의 수준을 선형적으로 변환시킨다고 가정하면 각각의 범위는 [1, -1]이 된다. $[l,u]$의 정확한 범위를 갖는 요인에 대해 이 변환은 요인수준의 실제값 x를 $[\frac{(2x-u-l)}{(u-l)}]$로 연결시킴으로써 이루어진다. 예를 들어 네 번의 실험횟수를 가진 실험의 한 요인 수준이 [100, 200, 200, 300]으로 실행된다면 각각에 해당하는 변환된 값은 [-1, 0, 0, 1]이다. 네 번의 실험횟수를 가진 어떤 동등한 실험에서든 요인의 중간 수준에서 두 개의 실험 실행을 포함하게 된다. 예를 들어 요인 수준이 [100, 100, 200, 300]으로 실행되어 [-1 , -1, 0, 1]로 변환된 값을 갖는다면 이 실험은 기하학적으로 동등하지 않다. 요인수준 이름의 변화에 있어 유일하게 가능한 방법은 요인수준의 변환된 값에 -1을 곱함으로써 얻어지는 것이다. [-1, 0, 0, 1]의 경우는 -1을 곱해도 변화가 없지만 [-1, -1, 0, 1]의 경우에는 -1을 곱하면 변환된 요인수준 값이 바뀌므로 동등성을 만족하지 못한다.

이러한 이유로 변환에 따른 일치 여부에 근거한 해밍 거리를 확대한 Euclidean 거리행렬 E_d를 다음과 같이 정의하자.

$$[E_d]_{i,j}=\sum_{k=1}^{p}|[T_d]_{ij}-[T_d]_{jk}| \tag{2.3}$$

이때 $i,j\in 1,2,\cdots,n$ 이다. Euclidean 거리는 해밍거리의 같고 다른 점의 변환에 실질 거리를 추가함으로써 기하학적 동등성으로 국한시키는 역할을 한다.

3. 비대칭 실험의 기하학적 동등성

3.1. 비대칭 실험의 기하학적 동등성의 필요충분조건

기존의 부분요인실험의 동등성에 관한 연구는 대칭형 모형을 대상으로 이루어졌다. 먼저 요인의 수가 둘인 경우의 동등성을 살펴보자. 요인의 수준이 2인 p개의 요인과 실험횟수가 n번인 두 실험은 행렬로 표현될 수 있는데 실험에서 실행횟수로 표현되는 행은 +1 과 -1로 구성되며 이때 +1과 -1은 요인의 두 수준을 나타내고, 열은 요인을 나타낸다. 행렬 R, C, L을 다음과 같이 정의하자. R은 $n\times n$ 행 순서변환(permutation) 행렬이다. 즉, R은 행들의 순서변환에 의한 $n\times n$ 단위행렬 I_n으로부터 얻어진 행렬이다. C는 $p\times p$ 열 순서변환 행렬이다. L은 ± 1로 이루어진 $p\times p$ 대각행렬이라 하자. 두 실험 T_{d_1}과 T_{d_2}에 대해서 $d_1 = Rd_2CL$을 만족하는 R, C 그리고 L이 존재하면 두 실험 T_{d_1}과 T_{d_2}는 동등하다.

두 실험 T_{d_1}과 T_{d_2}에 대해 R, C 그리고 L의 존재 여부를 결정하는 한 가지 방법은 실험을 기하학적으로 살펴보는 것이다. 만일 두 실험 T_{d_1}과 T_{d_2}에 대한 n개의 점들을 p-공간 차원에 그린다고 할 때, $T_{d_1}\equiv T_{d_2}$라면 두 그림은 유사하게 그려질 것이다. 예를 들어 두 실험 T_{d_1}과 T_{d_2}가 다음과 같다고 하자.

$$T_{d_1} = \begin{bmatrix} -1 & -1 \\ 1 & -1 \\ 1 & 1 \end{bmatrix} \qquad T_{d_2} = \begin{bmatrix} 1 & 1 \\ -1 & -1 \\ -1 & 1 \end{bmatrix}$$

이 두 실험은 R, C 그리고 L이 다음과 같을 때, $T_{d_1} = RT_{d_2}CL$을 만족하며 두 실험은 동등성을 갖는다.

$$R = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} \qquad L = \begin{bmatrix} -1 & 0 \\ 0 & -1 \end{bmatrix}$$

또는 다음과 같은 R, C 그리고 L에 대해서도 두 실험은 동등성을 만족한다.

$$R = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{bmatrix} \qquad C = \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad L = \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix}$$

위의 정의로부터 두 실험 T_{d_1}과 T_{d_2}가 동등성을 갖는다면, $T_{d_1} = RT_{d_2}CL$를 만족하는 R, C 그리고 L에 대해 다음과 같이 쓸 수 있다.

$$\begin{aligned} T_{d_1}T_{d_1}' &= RT_{d_2}CLL'C'T_{d_2}'R' \\ &= RT_{d_2}T_{d_2}'R' \end{aligned}$$

L은 ± 1로 이루어진 대각행렬이므로 $LL' = I_p$이며, C는 정규직교행렬이므로 $CC' = I_p$이다. 뒤에 R을 곱한 것은 $T_{d_2}T_{d_2}'$의 열들을 재배열한 것이며 앞에 R을 곱한

것은 $T_{d_2}T_{d_2}^{'}R^{'}$의 행들을 재배열한 것이다. 그 결과 $T_{d_1}T_{d_1}^{'}$의 각각의 행들의 원소들 사이와 $T_{d_2}T_{d_2}^{'}$의 각각의 행들의 원소들 사이에 일대일 대응이 존재한다.

그러나, $T_{d_1}T_{d_1}^{'}=RT_{d_2}T_{d_2}^{'}R^{'}$을 만족하는 R의 존재는 실험의 동등성에 있어 충분조건이 아니다. Draper와 Mitchell (1968)은 12개의 요인을 가지고 512번 실험한 동일한 단어길이 유형을 가졌으나 두 실험이 동등성을 만족하지 못함을 보였다. $A\subseteq\{1,2,\cdots,p\}$에 대해 E_A는 집합 A를 삭제함으로써 나타나는 열들과 함께 $p\times p$ 단위행렬로부터 이루어진 행렬이다. E_A는 $p\times(p-|A|)$행렬로 이때$|A|$는 집합 A의 원소의 개수를 의미한다. $S_A=E_AE_A^{'}$로 정의되는데 S_A는 $i\in A$이면 $[i,i]$ 원소는 0의 값을 가지며 그 밖의 경우에는 1의 값을 갖는 대각행렬이다. 만약 $A=\varnothing$이면 $S_A=I_p$이다.

실험의 동등성에 대한 필요충분조건은 다음과 같이 연차적으로 성립해야 한다.

$$
\begin{aligned}
T_{d_1}T_{d_1}^{'} &= RT_{d_2}T_{d_2}^{'}R^{'} \\
T_{d_1}S_pT_{d_1}{'} &= RT_{d_2}S_{i_p}T_{d_2}{'}R' \\
T_{d_1}S_{p-1,p}T_{d_1}{'} &= RT_{d_2}S_{i_{p-1},i_p}T_{d_2}{'}R' \\
\vdots \quad & \quad \vdots \\
T_{d_1}S_{2,3,\cdots,p}T_{d_1}{'} &= RT_{d_2}S_{i_2,i_3,\cdots,i_p}T_{d_2}{'}R'
\end{aligned}
$$

즉, $T_{d_1}\equiv T_{d_2}$에 대한 필요충분조건은 위의 조건을 보듯, 모든 열들의 조합에서 모두 만족하는 $n\times n$ 행 변환행렬 R과 $\{1,2,\cdots,p\}$의 순서변환 $\{c_1,c_2,\cdots,c_p$가 존재하는 것이다. R을 알지 못하더라도 위의 정의는 실험의 동등성을 확인하는데 사용할 수 있다. 첫 번째로 $T_{d_1}T_{d_1}^{'}$의 행들이 변환될 수 있다면 $T_{d_1}T_{d_1}^{'}$의 각각의 행에서 원소들의 집합은 $T_{d_2}T_{d_2}^{'}$의 각각의 행에서 원소들의 집합과 대응시킬 수 있으므로 $T_{d_1}T_{d_1}^{'}$과 $T_{d_2}T_{d_2}^{'}$를 비교할 수 있다. 두 번째로 i_p는 $T_{d_1}S_pT_{d_1}^{'}$과 $T_{d_2}S_{i_p}T_{d_2}{'}$의 행들이 짝지어지는 것을 찾을 수 있다. 이 과정은 $T_{d_1}S_{2,3,\cdots,p}T_{d_1}{'}$과 $T_{d_2}S_{i_2,i_3,\cdots,i_p}T_{d_2}{'}$의 행들이 짝지어질 때까지 계속된다. 만약 모든 과정에서 수행된다면 두 실험은 동등성을 갖게 되며, 만일 어느 과정에서든 성공하지 못한다면 두 실험은 동등성을 갖지 못한다.

위와 같은 과정을 정리하여 Clark과 Dean (2001)은 두 실험 d_1과 d_2에 $n\times n$ 변환행렬 R이 존재하며, $q=1,2,\ldots,p$에 대해 $H_{d_1}^{1,2,\ldots,q}=R(H_{d_2}^{c_1,c_2,\ldots,c_q})R'$를 만족하는 $1,2,\ldots,p$의 순서변환 $c_1,c_2,\ldots,c_p$이 존재하면 두 실험 d_1과 d_2는 기하학적 동등성을 갖는다는 것을 보였다.

정리 3.1 (gCD 조건; Clark과 Dean, 2001) 두 대칭 요인실험 d_1과 d_2에 $n\times n$ 변환행렬 R이 존재하며, $q=1,2,\ldots,p$에 대해 $H_{d_1}^{1,2,\ldots,q}=R(H_{d_2}^{c_1,c_2,\ldots,c_q})R'$을 만족하는 $\{1,2,\ldots,p\}$의 순서변환 $\{c_1,c_2,\ldots,c_p\}$이 존재하면 두 실험 d_1과 d_2는 기하학적 동등

성을 갖는다.

기존의 부분요인실험의 동등성에 관한 연구는 대칭형 모형을 대상으로 이루어졌다. 본 논문에서는 대칭형 모형을 대상으로 한 동등성에 대한 연구를 비대칭형 모형으로 확대시켜 살펴보고자 한다. 아래의 정리는 Clark과 Dean (2001)이 제시한 방법을 해밍거리를 Euclidean 거리로 대치함으로써 비대칭 실험으로 확대할 수 있음을 보여준다. 증명 과정도 매우 유사하다.

정리 3.2 두 비대칭 요인실험 d_1과 d_2에 $n\times n$ 변환행렬 R이 존재하며, $q=1,2,\dots,p$에 대해 $E_{d_1}^{1,2,\dots,q}=R(E_{d_2}^{c_1,c_2,\dots,c_q})R'$를 만족하는 $\{1,2,\dots,p\}$의 순서변환 $\{c_1,c_2,\dots,c_p\}$이 존재하면 두 실험 d_1과 d_2는 기하학적 동등성을 갖는다.

증명 : 두 실험 d_1과 d_2가 기하학적 동등성을 갖는다고 가정하자. 거리행렬 E_{d_2}는 T_{d_2}의 어떤 열에서든지 역으로 재배열하여도 변하지 않는다. 그 결과 d_1과 d_2에서 요인들이 같은 수준으로 표시된다고 가정할 수 있다. 그러면 $T_{d_1}T_{d_1}'=RT_{d_2}C$ 라고 쓸 수 있다. 여기서 C는 실험 d_2에서 d_1의 요인의 이름을 바꾼 순서변환 $\{c_1,c_2,\dots,c_p\}$에 해당하는 변환행렬이며 R은 실험 d_2를 d_1에서와 같은 순서로 처리조합의 순서를 바꾸는(reorder) 변환행렬이다. 따라서 $1\le k\le p$에 대해 $[E_{d_1}^k]_{i,j}=\delta[RT_{d_2}C]_{i,j}^k=\delta[RT_{d_2}]_{i,j}^{c_k}=[E_{d_2}^{c_k}]_{r_i,r_j}$이다. 필요성은 $q=1,2,\dots,p$에 대해 $E_{d_1}^{1,2,\dots,p}=\sum_{k=1}^{q}[E_{d_1}^k]_{i,j}=\sum_{k=1}^{q}[E_{d_2}^{c_k}]_{r_i,r_j}=\sum_{k=1}^{q}R(E_{d_2}^{c_k})R'=R(E_{d_2}^{c_1,c_2,\dots,c_q})R'$ 사실에서 따른다.

충분성은 "$q=1,2,\dots,p$에 대해 $\{1,2,\dots,p\}$의 어떤 순서변환인 $\{c_1,c_2,\dots,c_p\}$와 변환행렬 $R_{n\times n}$의 행렬 $R(E_D^{c_1,c_2,\dots,c_q})R'$의 순서는 유일하게 동등성을 계획하는 처리행렬 T_d까지 결정한다" 라고 보인 Clark과 Dean (2001)의 [Lemma 2.1.]에서 $k=1,2,\dots,n+1$에 대해 $[T_{d^*}]_{1,q}=1$로 넣으면 모든 과정이 동일하여 성립된다. □

3.2. 알고리즘

이번 절에서는 Clark과 Dean (2001)이 제시한 실험의 동등성에 대한 알고리즘을 살펴보기로 한다. 쉽게 이해하기 위해 먼저 모든 요인들이 -1과 +1 두 수준을 갖는다고 가정하자. 그러면 두 처리행렬 T_{d_1}과 T_{d_2}는 변환행렬 R과 C가 존재하고 $T_{d_1}=RT_{d_2}CL$을 만족하는 $L^2=I$인 대각행렬 L이 존재한다면 두 실험은 동등성을 갖는다. 따라서 동등성을 확립하기 위해서 R, C와 L을 결정해야 한다. R과 C가 알려지면 L은 결정되며 L의 대각항에서 음수의 원소들은 T_{d_1}의 첫 번째 행에 상응하는 요소들로부터 부호가 다른 $RT_{d_2}C$의 첫 번째 행의 요소들에 해당한다. 변환행렬 R과 C

는 $H_{d_1}^q = R(H_{d_2}^{c_q})R'$를 이용해 결정한다.

두 실험의 동등성을 확인하기 위한 포트란 프로그램은 두 부분으로 나뉜다. *Deseq1* 이라 불리는 첫 번째 알고리즘은 동등성에 대한 첫 조사를 하며 *Deseq2* 는 동등성이 존재하는 경우 한 실험이 다른 실험으로 변환하는 순서의 변경을 찾는다.

특히 *Deseq1*은 $q = p, p-1, 1, p-2, 2, \cdots, p-[p/2], [p/2]$에 대해 $\{c_1, c_2, \ldots, c_p\}$의 어떤 부분집합 $\{c_1, c_2, \ldots, c_p\}$가 있어 $H_{d_1}^{1,2,\cdots,q}$와 $H_{d_2}^{c_1,c_2,\cdots,c_q}$가 동일한 개수의 어떤 집합을 포함하는지 확인한다. 알고리즘에서 q의 각각 값에서의 조사를 "stage"라고 하며 stage의 수는 $p-q$ 이다. 만약 조사들 중 하나가 실패하면 알고리즘은 멈추고 그 stage에서 실패한 상태를 나타낸다. 이 초기 테스트는 두 실험에 대한 점들을 비교하는데 만약 두 실험이 동등하면 두 실험의 점들은 유사하게 나타나고 점들의 모든 쌍들 사이의 거리는 두 실험에 대해 같게 나온다. *Deseq2* 는 d_2를 d_1으로 변환하는 행과 열의 순서의 변경을 찾는다. 변환행렬 R을 찾는 단계는 다음과 같다. 우선, 가능한 행의 변환 $P_{n\times n}$행렬이 구성되는데, 여기서 $T_{d_1}T_{d_1}'$의 i번째 행이 $T_{d_2}T_{d_2}'$의 j번째 행의 요소와 같은 개수를 가지게 되면 $[P]_{ij} = 1$의 값을 가지며, 그 밖에는 $[P]_{ij} = 0$의 값을 갖는다. 그 다음 $[P]_{ij} = 1$을 만족하는 가장 작은 m이 선택되고 알고리즘은 $[R]_{1m} = 1$을 만든다. 이것은 T_{d_2}의 행이 T_{d_1}의 첫 번째 행으로 가능한 첫 번째 순서변경에 해당한다. $i \neq m$이면 $[R]_{1i}$는 0의 값을 갖고, $i \neq 1$인 경우에 $[R]_{im}$은 0의 값을 가지며, $j = 1, 2, \cdots, n$일 때 $[R]_{jl}$는 $[T_{d_1}T_{d_1}']_{1j} \neq [T_{d_2}T_{d_2}']_{ml}$인 모든 l에 대해 0의 값을 갖는다. 이 동일한 과정들은 R의 다음 행들에 실행되어 만일 어느 단계에서 R의 행이 0인 요소(원소)들로 나타나면 그 과정은 이전 행으로 되돌아가 P에 해당하는 행에서 그 다음으로 가능한 0 아닌 항목을 사용하고 R의 나머지는 단순화된다. 마지막 행에 이르러 R은 행 변환행렬이 된다. 만일 $H_{d_1}^q = R(H_{d_2}^{c_q})R'$를 만족하는 C가 존재하지 않는다면 다른 행 변환행렬은 그 행 안에서 다음으로 가능한 0이 아닌 항목을 갖는 P의 한 행을 위로 이동시킴으로 구성된다. 이 절차는 모든 행렬 R의 집합 $\mathcal{R}$을 찾을 수 있으며, 변환행렬의 완전한 집합을 반드시 검토하지 않고도 $T_{d_1} = RT_{d_2}CL$을 만족하는 행렬 C와 L에 해당하는 행렬을 찾을 수 있다. 만일 R이 만들어지지 않는다면 d_1과 d_2는 step 0에서 동등성을 만족하지 않으며, R이 만들어진다면 *Deseq2* 는 $H_{d_1}^q = R(H_{d_2}^{c_q})R'$ 를 만족하는 $R \in \mathcal{R}$이 적어도 하나 이상 존재하는지 조사한다.

알고리즘은 요인의 수준이 2 이상인 질적 요인을 사용하기 위해 거리행렬을 계산하는데 $H_{d_1}^q = R(H_{d_2}^{c_q})R'$를 대신해 $H_{d_1}^{1,2,\cdots,q} = R(H_{d_2}^{c_1,c_2,\cdots,c_q})R'$을 사용함으로써 적합시킬 수 있다.

양적 요인을 갖는 두 실험이 동등성을 만족하는지에 대한 필요충분조건은 질적 요인에서의 경우와 비슷하다. 그러나 동등성을 가진 실험을 그 실험의 처리행렬의 첫 번째 행이 모두 1이 되게 임의로 구성할 수 없다. 가장 간단한 해결책은 처리행렬 T_d에 모든 원소가 1인 특별한 행을 처리행렬의 첫 번째 행에 추가시키도록 수정하는 것

이다.

3.3. 예제

양적 요인을 가지는 두 비대칭형 실험 d_1과 d_2의 동등성 만족여부는 3.2.절에서 소개한 알고리즘을 그대로 적용 할 수 있다. 예로서 실험 d_1과 d_2의 처리행렬은 다음과 같다고 하자.

$$T_{D_1} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 1 & 1 \\ 0 & 1 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 1 \\ 2 & 0 & 0 & 0 & 1 \\ 2 & 0 & 0 & 1 & 0 \\ 2 & 1 & 1 & 0 & 1 \\ 2 & 1 & 1 & 1 & 0 \end{bmatrix} \qquad T_{D_2} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 1 & 1 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 0 \\ 2 & 0 & 0 & 0 & 0 \\ 2 & 0 & 1 & 1 & 0 \\ 2 & 1 & 0 & 1 & 1 \\ 2 & 1 & 1 & 0 & 1 \end{bmatrix}$$

<표 1> 양적 요인의 비대칭형 실험에 대한 알고리즘 결과

```
PASSES STAGE    0
PASSES STAGE    1
PASSES STAGE    4
PASSES STAGE    2
PASSES STAGE    3
** DESEQ1 PASSED !!!!!!!!!!

PASSES STAGE  5
EQUAL
COLUMN PERMUTATION  1  4  3  2  5
ROW PERMUTATION     1 10 13 12 11  7  9  6  8  2  4  3  5
[MESSAGE !!!] DESEQ2 PASSED >>> EQUIVALENT !!!!
```

요인의 수가 5이고 요인 수준을 살펴보면 한 요인의 요인수준은 3이며 나머지 네 요인의 요인수준은 2이다. 요인이 양적 요인이므로 실험횟수가 12인 두 실험 d_1과 d_2의 처리행렬의 첫 번째 행에 구성원소가 모두 1인 행을 추가시켜 처리행렬을 구성하였으며, 이 두 실험이 동등성을 만족하는지 알고리즘을 통해 알아본 결과는 <표 1>과 같다.

위 결과를 통해 두 실험은 동등하며 한 실험이 다른 실험으로 변환하는 행에 대한 순서변환은 1, 4, 3, 2, 5 이며 열에 대한 순서변환은 1, 10, 13, 12, 11, ⋯ 임을 알 수 있다.

참 고 문 헌

[1] Chen, J. and Lin, D.K.J. (1991). On the identity relationship of 2^{k-p} designs. *Journal of Statistical Planning and Inference*, Vol. 28, 95-98.

[2] Cheng, S.W. and Ye, K.Q. (2004). Geometric isomorphism and minimum aberration for factorial designs with quantitative factors. *The Annals of Statistics*, Vol. 32, 2168-2185.

[3] Cheng, S.W. and Wu, C.F.J. (2001). Factor screening and response surface exploration (with discussion). *Statistica Sinica*, Vol. 11, 553-604.

[4] Clark, J.B. and Dean A.M. (2001). Equivalence of fractional factorial designs. *Statistica Sinica,* Vol. 11, 537-547.

[5] Draper, N.R. and Mitchell, T.J. (1967). The construction of saturated 2_R^{k-p} designs. *Annals of Mathematical Statistics,* Vol. 38, 1110-1128.

[6] Draper, N.R. and Mitchell, T.J. (1968). Construction of the set of 256-run designs of resolution≥ 5 and the set of even 512-run designs of resolution≥ 6 with special reference to the unique saturated designs. *Annals of Mathematical Statistics,* Vol. 39, 246-255.

[7] Draper, N.R. and Mitchell, T.J. (1970). Construction of the set of 512-run designs of resolution≥ 5 and the set of even 1024-run designs of resolution≥ 6. *Annals of Mathematical Statistics,* Vol. 41, 876-887.

[8] Xu, H. (2005). A catalogue of three-level regular fractional factorial designs. *Metrika*, Vol. 62, 259-281.

[Received September 2006, Accepted November 2006]