Journal of the Korean Data & Information Science Society
2015, **26**(6), 1295–1303
http://dx.doi.org/10.7465/jkdi.2015.26.6.1295
한국데이터정보과학회지

심플렉스 중심배열법의 일부실시에 관한 연구[†]

김형순[1] · 박동권[2]

접수 2015년 8월 4일, 수정 2015년 11월 9일, 게재확정 2015년 11월 18일

요 약

혼합물 실험으로부터 실험자는 주효과와 저차의 상호작용 효과의 추정을 원한다. 이를 위해 심플렉스 중심배열법과 같은 적절한 실험을 통해 추정할 수 있다. 그러나, 요인의 수가 늘어나면 부득이 일부실시를 행하게 된다. 이 경우 각 성분의 혼합비율의 합이 일정하다는 제약 조건은 교락으로 인해 추정가능한 상호작용의 선택을 어렵게 한다. 이러한 문제를 해결하기 위해 본 논문에서는 Scheffé의 정준 모형 대신에 대수기하학을 기초로 한 동차다항식 (homogeneous polynomial)으로 구성된 모형을 도입하여 문제를 풀려고 한다. 이를 활용하여 심플렉스 중심배열법의 일부실시법에 대해 추정가능한 모형을 제시한다. 연산은 CoCoA 대수연산 소프트웨어를 이용하였다.

주요용어: 그뢰브너 기저, 동차다항식, 심플렉스 중심배열법, 아이디얼.

1. 서 론

혼합물 실험은 반응변수가 구성 성분의 비율에는 영향을 받지만 혼합물의 양에는 영향을 받지 않는 특수한 실험이다. q개의 성분의 혼합물에 있어서 x_i를 $i-$번째 성분의 혼합비율이라고 한다면, 혼합비율들은 다음과 같은 선형 제약을 받는다.

$$x_1 + x_2 + \cdots + x_q = 1. \tag{1.1}$$

이러한 혼합물계획의 모형과 분석에 관해서는 많은 연구가 이루어졌는데 대표적인 저서로는 Cornell (2002) 등이 있다.

대표적으로 많이 사용되는 혼합물 모형은 Scheffé (1958, 1963)에 의해 제안된 정준다항식 (canonical polynomial) 모형 (S-모형 이라 하자)이라 할 수 있다. 예로서 2차 S-모형은 다음과 같다.

$$E(Y) = \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_q x_q + \cdots + \beta_{12} x_1 x_2 + \cdots + \beta_{q-1,q} x_{q-1} x_q. \tag{1.2}$$

위의 S-모형은 절편을 없애고 대신에 모든 x_i를 모형에 포함시켜 표현한 것이다. 다른 모형으로는 Kronecker 곱하기를 사용하여 Draper와 Pukelsheim (1998)이 제안한 K-모형이 있다. 예로서 2차 K-모형은 다음과 같다.

$$E(Y) = \alpha_{11} x_1^2 + \alpha_{22} x_2^2 + \cdots + \alpha_{qq} x_q^2 + \cdots + \alpha_{12} x_1 x_2 + \cdots + \alpha_{q-1,q} x_{q-1} x_q. \tag{1.3}$$

식 (1.3)을 보면 모든 항의 차수가 2차로 동차다항식 (homogenuous polynomial)임을 볼 수 있다.

[†] 본 연구는 제1저자와 교신저자 각각의 2011학년도 연세대학교 학술연구비의 지원에 의하여 이루어진 것임.
[1] (220-710) 강원도 원주시 연세대길 1, 연세대학교 수학과, 교수.
[2] 교신저자: (220-710) 강원도 원주시 연세대길 1, 연세대학교 정보통계학과, 교수.
E-mail: statpdk@yonsei.ac.kr

위의 두 모형을 비교해보면 외견적으로 보면 S-모형이 주효과를 직접 포함하고 있어 더 적절해 보이나 사실상 (1.1)의 전제 조건에서는 서로 변환이 가능하여 (Cornell, 2002; Draper와 Pukelsheim, 1998) 동일하다고 할 수 있다. 반면에 추정에 있어 K-모형이 S-모형의 경우 보다 계획 행렬 X의 열사이에 존재하는 선형 종속의 정도 (ill-conditioning 혹은 collinearity)를 줄일 수 있는 모형이라고 알려져 있다 (Prescott 등, 2002).

실제로 많은 경우 실험의 수를 줄이기 위해 완전계획으로부터 일부 (fraction)를 택하여 실험점의 개수 (n)와 모형에 포함된 항들의 수가 일치하는 saturated design이 사용된다. 이때 어떻게 n개의 항을 선택할 수 있는가의 문제는 항들간의 교락구조 (generalized confounding relations)에 따라 결정될 수 있다.

이러한 문제의 해결을 위해 본 논문에서는 실험계획에서의 공간 내의 점들 즉, 기하학적 대상인 다양체 (variety)에 대한 문제를 다항식을 매개로 하여 아이디얼 (ideal) 즉, 대수적 문제로 전환하여 해결하는 계산가환대수 (computational commutative algebra; CCA) 접근을 하게 된다.

지금까지 다루어진 CCA 접근은 Design-아이디얼 (아이디얼, 그뢰브너 기저 등 기본 개념의 이해를 위해서는 Cox 등, 1992; Pistone과 Wynn, 1996; Kim과 Park, 2009 참조) 접근 방법으로 모형안에 추정가능항에 절편이 포함되게 되어 실제로는 모형에 $(k-1)$개만의 성분이 포함된 불완전 (slack) 모형이 된다. 이를 해결하기 위해 Maruri-Aguilar 등 (2007)에 의해 제안된 방법은 보통 다루어지는 실험점에서 0이 되는 다항식들의 집합인 아이디얼 (design ideal)이 아니라, 모든 동차다항식들의 집합인 콘아이디얼 (cone ideal)상에서 혼합물계획을 다루게 된다 (2절 참조). 이렇게 함으로서 혼합물계획의 일부실시법에 선택된 항들이 추정가능해지도록 선택할 수 있다.

2절에서는 CCA 방법을 약술하고, 3절에서는 이를 중심합성계획에 적용한다. 연산은 CoCoA 대수연산 소프트웨어를 이용하였다. CoCoA는 가환대수의 연산을 수행하기 위한 프로그램으로, 이태리 제노바 대학의 전산 대수 연구팀에서 개발되었다. http://cocoa.dima.unige.it에서 이에 대한 자세한 설명과 함께 무료로 프로그램을 다운 받을 수 있다.

2. 사영공간과 콘-아이디얼

이번 절에서는 먼저 사영공간 (projective space)과 콘-아이디얼 (cone ideal)에 대해 소개하고, 기존에 정립되어 있는 아핀 (affine) 공간에서의 design-아이디얼과의 관계를 조명한다.

혼합물 실험은 선형제약식 $\sum_{i=1}^{q} x_i = 1$에 의해 q-차원의 심플렉스인 집합 $\{(x_1, x_2, \cdots, x_q) \in R^q \mid \sum_{i=1}^{q} x_i = 1, 0 \le x_i \le 1\}$의 점들로 구성된다. 이 조건에 의해 다음과 같이 사영공간에서의 CCA 이론을 적용할 수 있다.

아핀공간 R^q의 원점 $\mathbf{0} = (0, 0, \cdots, 0)$을 제외한 두 점 $(x_1, x_2, \cdots, x_q)$와 $(x_1', x_2', \cdots, x_q')$사이에 다음과 같이 정의된 관계

$$(x_1, x_2, \cdots, x_q) \sim (x_1', x_2', \cdots, x_q')$$
$$\Leftrightarrow 0\text{이 아닌 실수 } \lambda\text{가 존재하여 } (x_1, x_2, \cdots, x_q) = \lambda(x_1', x_2', \cdots, x_q')$$

는 동치관계이며 기하학적으로는 주어진 두 점이 원점을 통과하는 하나의 직선상에 위치함을 의미한다. 이 동치관계에 의한 동치류들의 집합 $(R^q - \{\mathbf{0}\})/\sim$ 를 사영공간 $P^{q-1}(R)$이라 한다.

아핀공간 R^q의 원점 $\mathbf{0}$이 아닌 한 점 $(x_1, x_2, \cdots, x_q)$는 사영공간 $P^{q-1}(R)$의 한 원소 P를 결정하는데 이때 $(x_1, x_2, \cdots, x_q)$를 P의 동차좌표라 한다. 임의의 0이 아닌 실수 λ에 대해 $(\lambda x_1, \lambda x_2, \cdots, \lambda x_q)$도 마찬가지로 P의 동차좌표가 된다. 기하학적으로 사영공간 $P^{q-1}(R)$의 한 점 P는 아핀공간 R^q에서 원

점을 지나는 한 직선이라 할 수 있고 따라서 다음의 1-1 대응관계를 가짐을 알 수 있다.

$$P^{q-1}(R) \cong \{R^q\text{에서 원점을 지나는 직선들}\}$$

사영공간에서의 다양체와 동차 아이디얼에 관한 이론은 이미 아핀공간에서의 다양체와 design-아이디얼에 관한 이론과 유사하게 정립되어 있다 (Cox 등, 1992). 다항식의 각 항의 차수가 같은 경우 동차다항식이라 하는데 임의의 동차다항식 f에 대해 R^q의 한 점 d에서 $f(d) = 0$이면 $f(\lambda d) = \lambda^{\deg f} f(d) = 0$이 되어 f의 근은 사영공간 $P^{q-1}(R)$에서 잘 정의됨을 알 수 있다. 실계수를 갖는 변수 $x_1, \cdots, x_q$에 대한 모든 다항식들의 집합을 $R[x_1, \cdots, x_q]$로 나타내면, 이 집합은 다항식간의 덧셈과 곱셈에 의한 가환환 (commutative ring)이 된다. $R[x_1, \cdots, x_q]$에 속하는 임의의 다항식 $f_1, f_2, \cdots, f_n$에 의해 생성되는 아이디얼을 기호 $< f_1, f_2, \cdots, f_n >$로 쓰는데 이를 집합으로 표현하면 다음과 같다.

$$< f_1, f_2, \cdots, f_n > \quad = \quad \left\{ \sum_{i=1}^{n} h_i f_i | h_1, \cdots, h_n \in R[x_1, \cdots, x_q] \right\}$$

[정의 2.1] 임의의 동차다항식들의 집합 $\{f_1, f_2, \cdots, f_n\} \subseteq R[x_1, \cdots, x_q]$에 대하여

$$V(f_1, f_2, \cdots, f_n) = \left\{ P \in P^{q-1}(R) \mid 1 \le i \le n\text{인 임의의 } i\text{에 대해 } f_i(P) = 0 \right\}$$

를 $f_1, f_2, \cdots, f_n$에 의해 정의된 사영다항체 (projective variety)라 한다.

[정의 2.2] $R[x_1, \cdots, x_q]$에 속한 아이디얼 I에 대하여 I에 속한 각 다항식의 모든 동차성분들이 다시 I에 속하면 I를 동차아이디얼 (homogeneous ideal)이라 한다.

[보조정리 2.1] (Cox 등, 1992)
$R[x_1, \cdots, x_q]$에 속한 아이디얼 I에 대하여 다음은 동치명제이다.

(1) I는 동차아이디얼이다.
(2) 동차다항식 $f_1, f_2, \cdots, f_n$이 존재하여 $I \;=\; < f_1, f_2, \cdots, f_n >$
(3) 동차다항식으로 구성된 I의 축약G-기저 (reduced Gröbner basis)가 존재한다.

아핀공간 R^q의 심플렉스에 속하는 혼합물실험 $D = \{d_1, d_2, \cdots, d_n\}$의 각 점은 원점과 d_i를 지나는 직선과 1-1 대응이 되고 따라서 디자인 D는 사영공간 $P^{q-1}(R)$의 다양체인 아핀콘 $C_D = \{\lambda d | d \in D, \lambda \in R\}$와 대응하게 된다.

[정의 2.3] 아핀콘 C_D에 대하여 다음과 같이 정의된 다항식 아이디얼

$$Ideal(C_D) = \{f \in R[x_1, \cdots, x_q] \mid \forall(a_1, \cdots, a_q) \in C_D, f(a_1, \cdots, a_q) = 0\}$$

를 C_D에 의해 정의된 콘-아이디얼이라 한다.

[보조정리 2.2] 콘-아이디얼 $Ideal(C_D)$는 동차아이디얼이다.

[증명] $Ideal(C_D)$에 속한 임의의 다항식 f에 대해 차수 i인 성분을 f_i라 표현하면 $f = \sum_i f_i$이고 임의의 점 $(a_1, \cdots, a_q) \in C_D$와 임의의 실수 λ에 대해

$$0 = f(\lambda a_1, \lambda a_2, \cdots, \lambda a_q) = \sum_i \lambda^i f_i(a_1, \cdots, a_q)$$

이다. 위 항등식으로부터 각 동차성분 f_i는 $f_i(a_1, \cdots, a_q) = 0$를 만족한다. 따라서 각 f_i는 $Ideal(C_D)$에 속하고, 아이디얼 $Ideal(C_D)$는 동차아이디얼이다. □

지금까지 다루어진 CCA 접근은 Design-아이디얼 접근 방법으로 모형안에 추정가능항에 절편이 포함되게 되어 실제로는 모형에 $(k-1)$개만의 성분이 포함된 불완전 (slack) 모형이 된다. 이를 해결하기 위해 보통 다루어지는 실험점에서 0이 되는 다항식들의 집합인 아이디얼 (design-ideal)이 아니라, 모든 동차다항식들의 집합인 콘-아이디얼 (cone ideal)상에서 혼합물계획을 다루게 된다. 다음 정리는 두 아이디얼간의 관계를 잘 보여주는 중요한 정리가 되며 3절에서 심플렉스 중심배열의 일부실시실험에서 추정가능한 모형을 선택하는 기초 이론을 제공한다.

[**정리 2.3**] 아핀공간 R^q의 심플렉스에 속하는 혼합물실험 $D=\{d_1,d_2,\cdots,d_n\}$에 대해 다음이 성립한다.

$$Ideal(D) \;=\; <Ideal(C_D), \sum_{i=1}^{q} x_i - 1>$$

[**증명**] R^q의 심플렉스에 속하는 한 점 $d=(d_1,\cdots,d_q)$을 고려해 보자. 이 경우 원점과 점 d를 지나는 직선의 식에서

$$Ideal(C_{\{d\}}) \;=\; <d_2x_1-d_1x_2, d_3x_1-d_1x_3,\cdots,d_qx_{q-1}-d_{q-1}x_q>$$

이고 각 생성원 다항식 $d_ix_j-d_jx_i$는 점 d에서 값이 0이므로 $Ideal(\{d\})$에 속한다. 또 d는 심플렉스상의 점이므로 $\sum_{i=1}^{q} d_i - 1 = 0$이다. 따라서 다음의 포함관계가 성립한다.

$$<Ideal(C_{\{d\}}), \sum_{i=1}^{q} x_i - 1> \;\subseteq Ideal(\{d\}).$$

반대로 $Ideal(\{d\}) \;=\; <x_1-d_1,\cdots,x_q-d_q>$이고 각 생성원 다항식 x_j-d_j는

$$\begin{aligned} x_j-d_j &= (d_1+\cdots+d_q)x_j - d_j + \sum_{i=1}^{q}(d_j-d_j)x_i \\ &= \sum_{i=1}^{q}(d_ix_j-d_jx_i) + d_j\left(\sum_{i+1}^{q} x_i - 1\right) \end{aligned}$$

이므로 $Ideal(\{d\}) \subseteq \;<Ideal(C_{\{d\}}),\; \sum_{i=1}^{q} x_i - 1>$이 된다. 정리하면 R^q의 심플렉스에 속하는 한 점 $d=(d_1,\cdots,d_q)$에 대하여 다음 식이 성립한다.

$$Ideal(\{d\}) \;=\; <Ideal(C_{\{d\}}),\; \sum_{i=1}^{q} x_i - 1>$$

이를 $D=\{d_1,d_2,\cdots,d_n\}$로 확대하면 쉽게

$$Ideal(D) = \cap_{d\in D} I(\{d\}) = \cap_{d\in D} <Ideal(C_{\{d\}}),\; \sum_{i=1}^{q} x_i - 1> \;=\; <Ideal(C_D),\; \sum_{i=1}^{q} x_i - 1>$$

이 성립하므로 증명된다. □

3. 심플렉스 중심배열법의 일부실시

이번 절에서는 심플렉스 중심배열법의 일부실시의 경우 추정가능 상호작용을 선택할 수 있는 절차를 동차다항식을 이용하여 전개한다. 심플렉스 중심배열법은 심플렉스 격자배열법과는 달리 모든 실험에 참여한 성분의 비율이 같은 실험이다. 이 경우 k개의 성분을 갖는 심플렉스 중심배열법에서 얻어지는 (2^k-1)개의 실험점 D_k는 k개의 1차 성분, $\binom{k}{2}$개의 이차 성분, $\binom{k}{3}$개의 3차 성분 등으로 구성된다. 각

$i-$차 성분의 실험점을 $D^i(i=1,\ldots,k,\ D_k=\sum_{i=1}^{k} D^i)$라 하면, $2^k-1=\sum_{i=1}^{k}\binom{k}{i}$ 성립된다. 심플렉스 중심배열법 D_k는 2^k-완전요인실험에서 원점을 제외한 실험과 일대일 대응관계가 있음을 볼 수 있다.

앞에서 언급한 것처럼 D_k에서 다항식 $(\sum_{i=1}^{k} x_i - 1)$은 모든 실험점 D_k에서 0이 되어 $Ideal(D)$에 속하게 되어 추정이 제한받게 된다. 이를 해결하기 위해 좀 더 일반적인 접근이 필요하여 사영공간과 콘-아이디얼을 이용하게 되고 2절에서 보인 [정리 2.3]은 그 이론의 기초가 된다.

Maruri-Aguilar 등 (2007)은 동차다항식을 이용해서 혼합물실험의 추정가능항을 선택할 수 있는 다음 절차를 제공하였다. 이렇게 먼저 서론에서 언급한 K-모형의 기저 (추정가능항)를 구한 후, 필요시 S-모형의 기저로 변환되어 사용될 수 있다. 즉, 다음의 [단계 4]에서 구해진 SM은 K-모형의 기저 (추정가능항)이 되고, [단계 5]에서 변환된 δ는 S-모형의 기저가 된다.

[단계 1] : 콘-아이디얼의 G-기저를 구한다.
[단계 2] : 구해진 G-기저의 leading term (LT)을 구한다.
[단계 3] : 충분히 큰 s에 대해서 동차다항식을 나열한다.
[단계 4] : G-기저의 LT에 의해 나누어지지 않는 standard monomial (SM)을 구한다.
[단계 5] : (필요에 따라) 콘-아이디얼이 아닌 design-아이디얼의 기저 δ로 변 환한다.

본 논문에서는 심플렉스 중심배열법의 일부실시 중 다음의 두 가지 유형에 관해 고려하기로 한다. 물론 다음의 진행 절차는 다른 형태의 일부실시에도 적용할 수 있다.

$$[I-\text{유형}]\ F_m^I=\sum_{i=1}^{m} D^i, 1\le m\le k$$

$$[II-\text{유형}]\ F_m^{II}=D^1+D^m$$

F_m^I은 m-성분까지 실험하는 일부실시법이고, F_m^{II}은 모든 축점 (corner point; D^1)과 m-성분만 실험하는 요인실험이 된다. 당연히 $F_k^I=D_k$ 완전실험이고, $F_2^I=F_2^{II}$가 성립한다. 현실적으로 가능하면 m이 적을수록 효율적인 실험이 된다고 볼 수 있다. McConkey 등 (2000)은 $m=3$일 때 II-유형에서 D^3-실험을 더 줄여서 실험하는 일부실시를 제안하였다. 이러한 접근은 k가 클 때 실험을 줄이기 위한 좋은 제안이 된다.

위의 절차를 심플렉스 중심배열법의 일부실시에 적용시키는 과정에서 다음의 정리를 이용할 수 있다.

[정리 3.1] (Maruri-Aguilar 등, 2007)

① F_m^I의 콘-아이디얼은 다음과 같다. 단, $m>1$.

$$Ideal(C_{F_m^I}) \;=\; <x_1^2x_2-x_1x_2^2,\cdots,x_{k-1}^2x_k-x_{k-1}x_k^2,x_1x_2\cdots x_{m+1},\cdots,x_{k-m}\cdots x_{k-1}x_k>$$

② 단항식 순서를 $x_i>x_j(i>j)$로 하면 (즉, grlex 순서), LT $x_i^2x_j$와 $x_{i_1}\cdots x_{i_{m+1}}$에 의해 나누어지지 않는 차수 s를 갖는 $\sum_{j=1}^{m}\binom{k}{j}$개의 동차다항식이 있다.

위의 내용을 k=4 심플렉스 중심배열법의 일부실시에 적용시켜 보자.

[예 3.1] (k=4 심플렉스 중심배열법의 I-유형 일부실시)

$$\begin{aligned} D_4=&\{D^1:(1,0,0,0),(0,1,0,0),(0,0,1,0),(0,0,0,1)\}+\\ &\{D^2:(1/2,1/2,0,0),(1/2,0,1/2,0),(1/2,0,0,1/2),(0,1/2,1/2,0),(0,1/2,0,1/2),(0,0,1/2,1/2)\}+\\ &\{D^3:(1/3,1/3,1/3,0),(1/3,1/3,0,1/3),(1/3,0,1/3,1/3),(0,1/3,1/3,1/3)\}+\\ &\{D^4:(1/4,1/4,1/4,1/4)\}\end{aligned}$$

$F_2^I = D^1 + D^2, F_3^I = D^1 + D^2 + D^3, F_4^I = D^1 + D^2 + D^3 + D^4$로 일부실시 실험이 구성되는데 F_4^I는 완전 심플렉스중심배열로 콘-아이디얼은 아래와 같다.

$$Ideal(C_{F_4^I}) \;=\; < x_1^2x_2 - x_1x_2^2, x_1^2x_3 - x_1x_3^2, x_1^2x_4 - x_1x_4^2, x_2^2x_3 - x_2x_3^2, x_2^2x_4 - x_2x_4^2, x_3^2x_4 - x_3x_4^2 >$$

이제 일부실시 $F_2^I = D^1 + D^2$ (10개의 실험점)을 고려해 보자. [정리 3.1] 혹은 CoCoA를 사용하면 아래의 결과를 얻을 수 있다. 여기서 구해지는 K-모형의 standard monomial을 SM_s라 하자.

$$Ideal(C_{F_2^I}) \;=\; < Ideal(C_{F_4^1}), x_1x_2x_3, x_1x_2x_4, x_1x_3x_4, x_2x_3x_4 >$$

$$\mathrm{LT}(Ideal(C_{F_2^I})\text{의 G} - \text{기저}) = \{x_1^2x_2, x_1^2x_3, x_1^2x_4, x_2^2x_3, x_2^2x_4, x_3^2x_4, x_1x_2x_3, x_1x_2x_4, x_1x_3x_4, x_2x_3x_4\}$$

$$SM_2 = \{x_1^2, x_1x_2, x_1x_3, x_1x_4, x_2^2, x_2x_3, x_2x_4, x_3^2, x_3x_4, x_4^2\}$$

위에서 구해진 $Ideal(C_{F_3^I})$의 G-기저 SM_2는 [Step 5]에 따라 Design-아이디얼의 추정가능항 δ로 변환되어 사용될 수 있다. 이 과정은 software CoCoA의 IdealOfProjective 명령어를 사용하여 구할 수 있다 (부록 참조). Design-아이디얼의 G-기저는 다음과 같이 변환된다.

$$\delta = \{x_1, x_1x_2, x_1x_3, x_1x_4, x_2, x_2x_3, x_2x_4, x_3, x_3x_4, x_4\}$$

δ를 보면 모든 주효과와 2차-상호작용효과가 포함되어있음을 볼 수 있다.

일부실시 $F_3^I = D^1 + D^2 + D^3$ (14개의 실험점)의 경우도 동일한 과정을 거쳐 다음의 결과를 얻을 수 있다.

$$Ideal(C_{F_3^I}) \;=\; < Ideal(C_{F_4^1}), x_1x_2x_3x_4 >$$

$$\mathrm{LT}(Ideal(C_{F_3^I})\text{의 G} - \text{기저}) = \{x_1^2x_2, x_1^2x_3, x_1^2x_4, x_2^2x_3, x_2^2x_4, x_3^2x_4, x_1x_2x_3x_4\}$$

$$SM_3 = \{x_1^3, x_2^3, x_3^3, x_4^3, x_1x_2^2, x_1x_3^2, x_1x_4^2, x_2x_3^2, x_2x_4^2, x_3x_4^2, x_1x_2x_3, x_1x_2x_4, x_1x_3x_4, x_2x_3x_4\}$$

$$\delta = \{x_1, x_1x_2, x_1x_3, x_1x_4, x_2, x_2x_3, x_2x_4, x_3, x_3x_4, x_4, x_1x_2x_3, x_1x_2x_4, x_1x_3x_4, x_2x_3x_4\}$$

위의 사실을 일반화하여 다음 정리와 같이 모든 m에 대해 성립함을 증명할 수 있다.

[정리 3.2] k개의 성분을 갖는 심플렉스 중심배열법의 일부실시 F_m^I는 S-모형에서 k개의 주효과와 $\binom{k}{2}$개의 모든 2차-상호작용, $\cdots$, $\binom{k}{m}$개의 모든 m차-상호작용이 모두 추정가능하다.

[증명]: $m = 2$인 경우를 먼저 고려해 보자. $k = 2$이면 완전실험이 되므로 $k > 2$를 가정하면

$$Ideal(C_{F_2^I}) \;=\; < x_1^2x_2 - x_1x_2^2, x_1^2x_3 - x_1x_3^2, \cdots, x_{k-1}^2x_k - x_{k-1}x_k^2, x_1x_2x_3, \cdots, x_{k-2}x_{k-1}x_k >$$

$$\mathrm{LT}(Ideal(C_{F_2^I})\text{의 G} - \text{기저}) = \{x_1^2x_2, x_1^2x_3, \cdots, x_{k-1}^2x_k, x_1x_2x_3, \cdots, x_{k-2}x_{k-1}x_k\}$$

$$SM_3 = \{x_1^3, x_2^3, \cdots, x_k^3, x_1x_2^2, x_1x_3^2, \cdots, x_{k-1}x_k^2\}$$

임을 쉽게 알 수 있다. 이것으로부터 추정가능항들의 모임인 δ는 k개의 주효과와 $\binom{k}{2}$개의 모든 2차-상호작용인 $x_1, \cdots, x_k$와 $x_1x_2, x_1x_3, \cdots, x_{k-1}x_k$를 모두 포함함을 알 수 있다. 그런데 F_2^I의 원소의 개수가 $(k + \binom{k}{2})$개 이므로 아이디얼의 차원에 관한 정리 (Maruri-Aguilar 등, 2007)에 의해 추정가능항의 개수는 이 숫자를 넘을 수 없다. 따라서 추정가능항은

$$\delta = \{x_1, \cdots, x_k, x_1x_2, x_1x_3, \cdots, x_{k-1}x_k\}$$

이다. $m > 2$인 경우도 비슷하게 증명이 가능하다. □

[예 3.2] (k=4 심플렉스 중심배열법의 II-유형 일부실시)

일부실시 II-유형 $F_3^{II} = D^1 + D^3$ (8 실험점)을 고려했을 때, 다음의 결과를 구할 수 있다 (부록 참조).

$$\begin{aligned} Ideal(C_{F_3^{II}}) \;=\; &< x_1x_3 - x_2x_3 - x_1x_4 + x_2x_4, x_1x_2 - x_2x_3 - x_1x_4 + x_3x_4, x_3^2x_4 - x_3x_4^2, \\ & x_2x_3x_4 + 1/2x_1x_4^2 - 1/2x_2x_4^2 - 1/2x_3x_4^2, x_2^2x_4 - x_2x_4^2, x_1^2x_4 - x_1x_4^2, x_2^2x_3 - x_2x_3^2 > \end{aligned}$$

$$\text{LT}(Ideal(C_{F_3^{II}})\text{의 G} - \text{기저}) = \left\{x_1x_3, x_1x_2, x_3^2x_4, x_2x_3x_4, x_2^2x_4, x_1^2x_4, x_2^2x_3\right\}$$

$$SM_2 = \left\{x_1^2, x_1x_4, x_2^2, x_2x_3, x_2x_4, x_3^2, x_3x_4, x_4^2\right\}$$

$$\delta = \{x_1, x_1x_4, x_2, x_2x_3, x_2x_4, x_3, x_3x_4, x_4\}$$

[예 3.2]를 정리하면 심플렉스 중심배열 중 일부실시를 실험점

$$\{(1,0,0,0), (0,1,0,0), (0,0,1,0), (0,0,0,1), (1/3,1/3,1/3,0), (1/3,1/3,0,1/3), (1/3,0,1/3,1/3), (0,1/3,1/3,1/3)\}$$

에서 실험하면 $SM_2 = \left\{x_1^2, x_1x_4, x_2^2, x_2x_3, x_2x_4, x_3^2, x_3x_4, x_4^2\right\}$은 saturated design의 K-모형의 추정가능항이 되고, $\delta = \{x_1, x_1x_4, x_2, x_2x_3, x_2x_4, x_3, x_3x_4, x_4\}$는 S-모형의 추정가능항이 된다.

4. 결론

본 연구에서는 심플렉스 중심배열법의 일부실시의 경우 두 가지 유형에 관해 고려하였다. I-유형은 m-성분까지 실험하는 일부실시법이고, II-유형은 모든 축점(corner point; D^1)과 m-성분만 실험하는 요인실험이 된다. I-유형의 경우 S-모형의 모든 m차까지 상호작용이 추정가능하다. II-유형의 경우는 동차다항식을 이용하여 전개하여 구할 수 있다.

부록

[예 3.2]: F_3^{II}의 콘-아이디얼과 기저 변환을 위한 CoCoA 프로그램

```
"For the G-basis for Ideal(( C_{F_3^{II}}))"

Use V ::=Q[x[1..4]];
Points := [[1,0,0,0], [0,1,0,0], [0,0,1,0], [0,0,0,1],[1/3,1/3,1/3,0],[1/3,1/3,0,1/3], [1/3,0,1/3,1/3],
          [0,1/3,1/3,1/3]];
F_3 ^ II:= IdealOfProjectivePoints(Points); F_3 ^ II;

"For Leading terms of G-basis"

LT(F_3 ^ II);

"For change from SM_3 to δ"
NF(x[1]*(x[1]+x[2]+x[3]+x[4]), F3minus);
NF(x[2]*(x[1]+x[2]+x[3]+x[4]), F3minus);
NF(x[3]*(x[1]+x[2]+x[3]+x[4]), F3minus);
NF(x[4]*(x[1]+x[2]+x[3]+x[4]), F3minus);
NF(x[1]*(x[1]+x[2]+x[3]+x[4]) ^ 2, F3minus);
```

References

Cornell, J. A. (2002). *Experiments with mixtures*, Third Ed., Wiley, New York.

Cox, D., Little, J. and O'Shea, D. (1992). *Ideal, varieties, and algorithms*, Spring-Verlag, New York.

Draper, N. R. and Pukelsheim, F. (1998). Mixture models based on homogenuous polynomials. *Journal of Statistical Planning and Inference*, **71**, 303-311.

Kim, H. and Park, D. K. (2009). Gröbner basis vs Indicator function. *Journal of the Korean Data & Information Science Society*, **20**, 467-479.

Maruri-Aguilar, H., Notari, R. and Riccomagno, E. (2007). On the description and identifiability analysis of experiments with mixtures. *Statistica Sinica*, **17**, 1417-1440.

McConkey, B., Mezey, P., Dixon, D. and Greenberg, B. (2000). Fractional simplex designs for interaction screening in complex mixtures. *Biometrics*, **56**, 824-832.

Pistone, G. and Wynn, H. P. (1996). Generalized confounding with Gröbner bases. *Biometrika*, **83**, 653-666.

Prescott, P., Dean, A. M., Draper, N. R. and Lewis, S. M. (2002). Mixture experiments: Ill-conditioning and quadratic model specification. *Technometrics*, **44**, 260-268.

Scheffe, H. (1958). Experiments with mixtures. *Journal of Royal Statistical Society B*, **20**, 344-360.

Scheffe, H. (1963). The simplex-centroid designs for experiments with mixtures. *Journal of Royal Statistical Society B*, **25**, 235-263.

Journal of the Korean Data & Information Science Society
2015, **26**(6), 1295–1303

http://dx.doi.org/10.7465/jkdi.2015.26.6.1295
한국데이터정보과학회지

The fraction of simplex-centroid mixture designs†

Hyoung Soon Kim[1] · Dong Kwon Park[2]

[1]Department of Mathematics, Yonsei University
[2]Department of Information and Statistics, Yonsei University

Received 4 August 2015, revised 9 November 2015, accepted 18 November 2015

Abstract

In a mixture experiment, one may be interested in estimating not only main effects but also some interactions. Main effects and interactions may be estimated through appropriate designs such as simplex-centroid designs. However, the estimability problems, implied by the sum to one functional relationship among the factors, have strong consequences on the confounding and identifiability of models for such designs. To handle these problems, we address homogeneous polynomial model based on the computational commutative algebra (CCA) instead of using Scheffés canonical model which is typically used. The problem posed here is to give how to choose estimable main effects and also some low-degree interactions. The theory is tested using a fraction of simplex-centroid designs aided by a modern computational algebra package CoCoA.

Keywords: Gröbner basis, homogeneous polynomial, ideal, simplex-centroid design.

† Kim and Park's work was supported respectively by the Yonsei University Research Grant of 2011.
[1] Professor, Department of Mathematics, Yonsei University, Wonju 220-710, Korea.
[2] Corresponding author: Professor, Department of Information and Statistics, Yonsei University, Wonju 220-710, Korea. E-mail: statpdk@yonsei.ac.kr