Journal of the Korean Data & Information Science Society
2009, **20**(6), 1015–1027

한국데이터정보과학회지

그뢰브너 기저와 지시함수와의 관계

김형순[1] · 박동권[2]

[1]연세대학교 수학과 · [2]연세대학교 정보통계학과

접수 2009년 8월 2일, 수정 2009년 10월 29일, 게재확정 2009년 11월 3일

요 약

대수기하학적 접근이란 실험계획에서의 공간 내의 점들 즉, 기하학적 대상인 다양체에 대한 문제를 다항식을 매개로 하여 아이디얼 즉, 대수적 문제로 전환하고자 한 것이라 할 수 있다. 지금까지의 연구는 완전요인실험으로부터 효율적인 부분요인실험을 선택하는 절차에 집중되어 왔다. 본 논문에서는 지금까지 연구 방법의 역의 과정을 추정해 보기로 한다. 한 부분요인실험이 선택되었을 때, 그 실험의 교락구조를 그뢰브너 기저를 구한 후 해석한다. 다음으로 그뢰브너 기저를 생성자로 활용하여 선택된 부분실험의 집합을 구별하기 위한 다항함수인 지시함수를 구하는 절차를 알아보기로 한다. 실제로 몇 가지 부분요인실험을 예로 택하여 그 과정을 수행하였다. 연산은 CoCoA 대수연산 소프트웨어를 이용하였다.

주요용어: 교락, 그뢰브너 기저, 다항식, 대수기하학, 지시함수.

1. 서 론

대수기하학적 접근이란 실험계획에서의 공간 내의 점들 즉, 기하학적 대상인 다양체 (variety)에 대한 문제를 다항식을 매개로 하여 아이디얼 (ideal) 즉, 대수적 문제로 전환하고자 한 것이라 할 수 있다 (Pistone과 Wynn, 1996). 먼저 간략하게 대수기하학을 실험계획의 문제에 적용하여 정리하여 보자. m개의 요인을 가진 실험계획에서 모든 처리조합의 집합은 m-차원 벡터들로 구성된 아핀공간이라 할 수 있다. 각각의 처리조합은 이 공간 내의 점이 된다. 따라서 실험계획은 이 점들의 집합이라 할 수 있다. 이를 바탕으로 m개의 요인을 갖고 n번 실험한 계획 D를 다음과 같이 표기하자.

$$D = \left\{ \boldsymbol{x}^{(1)}, \cdots, \boldsymbol{x}^{(n)};\quad \text{단} \quad \boldsymbol{x} = (x_1, \cdots, x_m) \right\}$$

여기서 $d_1(x), \cdots, d_n(x)$를 D의 점 만을 해로 갖는 한 연립다항식 (디자인 다항식)이라면 D를 다음과 같이 나타낼 수 있다.

$$D = \{x \mid d_1(x) = 0, \cdots, d_n(x) = 0\}$$

이와 같은 다항식의 해들의 집합을 다양체 (variety)라고 한다. 쉬운 예로서 한 실험점 $\boldsymbol{x} = (a_1, \ldots, a_m)$은 다항식 $\{x_1 - a_1, x_{2-}a_2, \cdots, x_m - a_m\}$의 다양체가 된다. 또한 D상에서 0을 만족하는 모든 다항식들의 집합, 다시 말해서 다양체를 해로 갖는 모든 다항식들의 집합 $\boldsymbol{I}(D)$를 다음과 같이 정의하고

$$\boldsymbol{I}(D) = \{f(x) | f(x) = 0, \forall x \in D\}$$

D의 아이디얼 (ideal)이라고 부른다. 일반적으로 아이디얼 $\boldsymbol{I}$는 다음의 성질을 만족한다.

[1] (220-710) 강원도 원주시 흥업면 매지리 234, 연세대학교 수학과, 교수.
[2] 교신저자: (220-710) 강원도 원주시 흥업면 매지리 234, 연세대학교 정보통계학과, 교수.
E-mail: statpdk@yonsei.ac.kr

1. $0 \in I$
2. $f, g \in \boldsymbol{I}$이면 $f+g \in \boldsymbol{I}$
3. $f \ \in \ \boldsymbol{I}$이고 $h \in \ k[x_1, \cdots, x_m]$이면, $\ hf \ \in \ \boldsymbol{I}$

여기서 $k[x_1, \cdots, x_m]$는 체 (field) k의 원소를 계수로 갖는 미지수 (indeterminate) $x_1, \cdots, x_m$의 다항식의 집합을 말한다. 집합 $k[x_1, \cdots, x_m]$은 더하기와 곱하기에 닫혀 있는 환 (ring)이며 동시에 벡터 공간이 된다.

위에서 볼 수 있듯이 아이디얼은 선형대수의 부분공간 (subspace)과 유사하다는 것을 알 수 있다. 아이디얼은 다항식 $f_1, \cdots, f_s$에 의해 생성되고 다음과 같이 표현될 수 있다

$$< f_1, \cdots, f_s >= \left\{ \sum_{i=1}^{s} h_i f_i \ : \ h_1, \cdots, h_s \ \in \ k[x_1, \cdots, x_m] \right\}$$

여기서 집합 $\{f_1, \cdots, f_s\}$를 위 아이디얼의 기저 (basis)라 한다.

예를 들어 2^2-요인실험의 경우 모든 처리조합과 이에 대응하는 다양체, 아이디얼을 표 1.1에서 정리하였다. 2^2-요인실험의 각 처리조합을 2-차원 아핀공간 내의 네 개의 점으로 생각하면, 이 네 개의 점들은 각각 표 1.1의 다양체 및 아이디얼로 표현될 수 있다. $\boldsymbol{V}$(I)는 아이디얼 I에 의해 정의된 다양체를 말한다. 여기서 2^2-완전요인실험은 네 점의 합집합으로 구성된다. 이에 대한 다양체는 각 다양체의 합집합으로, 또한 아이디얼은 각 아이디얼의 교집합으로 표현되어 표 1.1의 마지막 행과 같이 나타난다.

표 1.1 2^2-요인실험의 다양체와 아이디얼

처 리 조 합	다 양 체	아 이 디 얼
(1, 1)	$\boldsymbol{V}(x_1-1, x_2-1)$	$< x_1-1,\ x_2-1 >$
(1, −1)	$\boldsymbol{V}(x_1-1, x_2+1)$	$< x_1-1,\ x_2+1 >$
(−1, 1)	$\boldsymbol{V}(x_1+1, x_2-1)$	$< x_1+1,\ x_2-1 >$
(−1, −1)	$\boldsymbol{V}(x_1+1, x_2+1)$	$< x_1+1,\ x_2+1 >$
2^2 -완전 요인실험	$\boldsymbol{V}(x_1^2-1, x_2^2-1)$	$< x_1^2-1,\ x_2^2-1 >$

완전요인실험이란 모든 처리조합에서 실험이 행해진다. 그러나 현실에서는 시간이나 비용의 문제로 그 일부의 조합만을 선택하여 실험하는 부분요인실험 (Fractional Factorial Design; FFD)이 필요하다. 지금까지는 주로 어떻게 좋은 FFD를 선택하는가에 연구의 목적이 집중되어 왔다. FFD는 실험의 수가 줄어들기 때문에 필연적으로 요인 효과간에 교락 (aliased)의 문제가 발생하게 된다. 요인실험 중 정규계획 (regular design)에서는 정의관계식 (defining relation)에 따라 교호작용 효과간에 완전 교락되거나 혹은 전혀 교락되지 않는 성질이 있어 고차 교호작용효과를 무시하고 관심의 대상이 되는 주효과 및 저차 교호작용효과를 추정하였다. 대표적인 정규 FFD는 2^{m-p}, 3^{m-p} 등으로 정의관계식을 구성하는 생성자 (generator)를 이용하면 선명도 (resolution)와 최소급변 (minimum aberration) 등 기준을 통해 효율적인 계획을 선택할 수 있다 (Zhang과 Park, 2000).

이에 반해 직교배열 (orthogonal array), Plackett와 Burman (PB)계획 등 비정규계획 (non-regular design)은 요인효과들의 관계를 간단한 정의관계식으로 표현할 수 없고, 실제로는 완전 교락이 아닌 부분적으로 교락되어 있어 저차 교호작용효과를 추정하는 데 어려움이 있다 (Plackett와 Burman, 1946). 예로서 12-run PB 계획은 주효과와 2-요인 교호작용효과의 교락 계수 값이 1/3또는 $-1/3$인 부분 교락된 계획으로 복잡한 교락 구조로 인해 전통적으로 주효과만을 선별 (screening)하는 데에만 이용되어 왔다. 그러나 이와 같이 주효과만을 위해 이 계획을 이용하는 것은 실험횟수에 비해 추정할 수 있는 요인이 적어 효율성이 떨어지고, 또한 모든 열에 교호작용이 퍼져있기 때문에 주효과만을 추정하는 데는 적절하지 못하다. 따라서 교호작용 효과를 추정할 수 있다면 보다 효율적으로 이용될 수 있을 것이다.

이 문제를 푸는 시도는 위에서 제안한 대수기하학의 그뢰브너 기저 (Gröbner basis; 앞으로 G-기저라고 함) 이론을 적용함으로서 해결 할 수 있다. 한 실험계획이 가지고 있는 교락구조의 설명을 G-기저를 통해 파악하고 이를 통해 추정가능함수를 구할 수 있다 (Park과 Kim, 2003).

이와 같이 선택된 FFD는 정규 계획이든 비정규 계획이든 간에 유일한 다항식인 지시함수 (indicator function)로 나타낼 수 있다. 수준이 2인 경우는 Fontana 등 (2000)과 Ye (2003)가 독립적으로 지수함수를 구하는 과정을 보이고 증명하였다. 이를 확대한 일반적인 경우는 Pistone과 Rogantin (2008)이 제안한 복수수지수함수를 이용할 수 있다. 그러나, 처리 수준의 복소수 코딩은 실제 적용시 해석에 어려움이 발생할 수 있다.

앞에서 언급하였듯이 지금까지의 연구는 완전요인실험으로부터 효율적인 부분요인실험을 선택하는 절차에 집중되어 왔다. 본 논문에서는 지금까지 연구 방법의 역의 과정을 추정해 보기로 한다. 한 부분요인실험이 선택되었을 때, 그 실험의 교락구조를 G-기저를 구한 후 해석한다. 다음으로 G-기저를 생성자로 활용하여 선택된 부분실험의 집합을 구별하기 위한 다항함수인 지시함수를 구하는 절차를 알아보기로 한다. 이를 위하여 '2,3절'에서는 G-기저와 지시함수에 대하여 정의한다. '4절'에서는 지시함수와 G-기저의 관계에 대해 살펴보고 G-기저의 생성자가 주어졌을 때 지시함수를 구하는 절차를 소개한다. 실제로 몇 가지 부분요인실험을 예로 택하여 그 과정을 수행하는데 연산은 CoCoA 대수연산 소프트웨어를 이용하였다. CoCoA는 가환 대수 (commutative algebra)의 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. 그뢰브너 기저

임의의 다항식 f가 아이디얼 $< f_1, \cdots, f_s >$에 속하는지 결정하는 문제 등 여러 관련 문제들은 나누는 순서와 관련되어 있다. 주목할 점은 단항식에서의 나누기와는 다르게 모든 기저에 대한 일반적인 나누기 알고리즘은 존재하지 않는다. 따라서, 다항식을 나누는 순서에 따라 다른 결과를 나타낼 수 있기 때문에 어떤 특별한 성질을 지닌 기저를 필요로 한다.

우선 단항식 순서 (monomial order)에 대해 설명할 필요가 있다. 단항식 순서는 나누기 알고리즘에서 변수들 간의 관계 및 차수의 크기를 결정한다는 데에 의미가 있다. 또한 그뢰브너 기저를 결정하는 데에 있어서도 중요한 역할을 한다. 단항식 순서란 다항식환 (polynomial ring) $k[x_1, \cdots, x_m]$상에서 아래의 조건들을 만족하는 m-차원 양의 정수계 ($Z^m_{\geq 0}$)에서의 관계 (relation), 또는 단항식 집합 $\{x^a, a \in Z^m_{\geq 0}\}$에서의 관계로서 다음 조건을 만족하는 것을 말한다.

1. $>$는 $Z^m_{\geq 0}$에서 완전 순서 (total order)이다.
2. $a > b$이고 $c \in Z^m_{\geq 0}$이면, $a + c > b + c$를 만족한다.
3. 공집합을 제외한 $Z^m_{\geq 0}$의 모든 부분집합은 $>$하에서 최소 원소를 갖는다.

단항식 순서에는 일반적으로 사전식, 차수가 붙은 사전식, 차수가 붙은 역사전식 순서 세 가지가 가장 많이 사용된다. 차수가 $a = (a_1, \cdots, a_m)$, $b = (b_1, \cdots, b_m) \in Z^m_{\geq 0}$일 때,

1. $a - b \in Z^m$에서 0이 아닌 가장 좌측의 원소가 양의 값을 가질 때, 사전식 순서 (LEXicographic order: LEX)는 $a > b$가 된다.
2. $|a| = \sum_{i=1}^m a_i > |b| = \sum_{i=1}^m b_i$, 또는 $|a| = |b|$이고 사전식 순서로 $a > b$이면, 차수가 붙은 사전식 순서 (Graded Lex Order: GRELEX)는 $a > b$이다.

3. $|a| = \sum_{i=1}^{m} a_i > |b| = \sum_{i=1}^{m} b_i$이고, $a - b \in Z^m$에서 0이 아닌 가장 우측의 원소가 음의 값을 가지면, 차수가 붙은 역 사전식 순서 (Graded Reverse Lex order: GREVLEX)는 $a > b$이다.

예를 들어 총 차수가 2보다 작거나 같은 단항식에 대하여 변수들 (indeterminates)의 기본 단항식 순서를 $x > y > z$로 하면, LEX, GRELEX, GREVLEX 순서는 각각 다음과 같다.

$$x^2 > xy > xz > x > y^2 > yz > y > z^2 > z > 1$$
$$x^2 > xy > xz > y^2 > yz > z^2 > x > y > z > 1$$
$$x^2 > xy > y^2 > xz > yz > z^2 > x > y > z > 1$$

위에서 보는 바와 같이 단항식 순서를 정하는 것에 따라 단항식 사이의 관계 혹은 크기가 다르게 정의된다.

일반적으로 변수가 하나인 경우와는 달리 변수가 두 개 이상일 경우에는 일반화된 알고리즘이 필요하다. 이해를 돕기 위해 예를 들어 설명하겠다. 다항식 $f = x^2y + xy^2 + y^2$를 $f_1 = xy - 1$, $f_2 = y^2 - 1$로 나누면 다음과 같이 나타낼 수 있다.

$$x^2y + xy^2 + y^2 = (x + y)(xy - 1) + 1(y^2 - 1) + x + y + 1$$

위의 방법을 $k[x_1, \cdots, x_m]$에서의 나누기 알고리즘 (division algorithm)이라고 한다. 이 알고리즘을 이용하여 일반적인 다항식을 다음과 같이 나타낼 수 있다.

$$f = a_1 f_1 + \cdots + a_s f_s + r \tag{2.1}$$

여기서 나머지 (remainder) r이 0일 경우에는 다항식 f는 $f_1, \cdots, f_s$에 의해 생성되는 아이디얼의 한 다항식이 된다. 그러나 다항식을 나누는 순서에 따라 나머지는 달라질 수 있다. 예를 들어, 다항식 $f = xy^2 - x$를 다항식 $f_1 = xy + 1$, $f_2 = y^2 - 1$로 나눈다면 다음과 같이 된다.

$$xy^2 - x = y \cdot f_1 + 0 \cdot f_2 + (-x - y)$$

그러나 순서를 달리하면 $xy^2 - x = x \cdot f_2 + 0 \cdot f_1 + 0$이 되어 이 경우에서는 $f \in < f_1, f_2 >$임을 알 수 있다. 따라서 같은 기저에 대해서도 나머지가 유일하게 결정되지 않음을 알 수 있다.

이를 해결하기 위해 단항식 (monomial)만을 고려해 보자. 주어진 다항식 f가 단항식 아이디얼에 속하는가는 다항식 f를 구성하는 단항식들에 의해 결정된다. 다시 말해 I를 단항식 아이디얼이라 하고, 다항식 $f \in k[x_1, \cdots, x_m]$에 대해 $f \in I$이면, f의 모든 항들은 I에 속하게 되고, 결국 f는 I내에 있는 단항식들의 k-선형결합으로 표현된다. 따라서, 단항식 아이디얼은 단항식들에 의해 유일하게 결정될 수 있다. 또한 어떤 두 단항식 아이디얼이 같다면 그 두 아이디얼은 같은 단항식들을 포함하고, 반대로 두 아이디얼이 같은 단항식들을 포함하고 있다면 그 두 아이디얼은 같은 아이디얼이 된다. 단항식 아이디얼은 유한 기저를 가지기 때문에, 이에 속하는 모든 단항식들은 이 기저들로 유일하게 결정된다. 즉 임의의 다항식 f가 단항식 아이디얼 $< x^{a(1)}, \cdots, x^{a(s)} >$에 속한다면 f를 $x^{a(1)}, \cdots, x^{a(s)}$로 나눈 나머지가 0이 된다. 또한 그 역도 성립한다. 따라서 여러 다항식을 나누는 순서에 따라 나머지가 달라지는 문제는 단항식 아이디얼로 해결될 수 있다. 단항식 순서를 정하면, 각 다항식 $f \in k[x_1, \cdots, x_m]$는 하나의 최고차항 (leading term; LT)을 갖는다. 이 때 LT(I)를 아이디얼 $I \subset k[x_1, \cdots, x_m]$에 속한 다항식들의 최고차항들의 집합이라고 한다면, $<LT(I)>$는 $LT(I)$의 원소들에 의해 생성된 아이디얼로 단항식 아이디얼이 된다. 그렇지만, $I =< f_1, \cdots, f_t >$일 때 일반적으로 $< LT(f_1), \cdots, LT(f_t) >$와 $LT(I)$가 같지는 않다. 실제로 $<LT(f_1), \cdots, LT(f_t)> \subset LT(I)$의 관계가 있다. 그러나 $<LT(I)> = <LT(g_1), \cdots, LT(g_s)>$를 만족하는 $g_1, \cdots, g_s \in I$가 존재한다.

정의 2.1 그뢰브너 기저 (Gröbner basis)

정해진 단항식 순서(monomial order) 아래서 아이디얼 I의 유한 부분집합 G=$\{g_1, \cdots, g_s\}$가 다음 식을 만족하면, G를 I의 그뢰브너 기저 (G-기저)라고 한다.

$$< LT(g_1), \cdots, LT(g_s) >=< LT(I) > \tag{2.2}$$

예를 들어, $g_1 = x^3 - 2xy, g_2 = x^2y - 2y^2 + x$에 의해 생성된 아이디얼 I를 고려해보자. LEX 순서에서 $y(x^3 - 2xy) - x(x^2y - 2y^2 + x) = -x^2$이므로 $-x^2$는 이 아이디얼의 한 원소다. 하지만 $LT(g_1) = x^3$, $LT(g_2) = x^2y$의 k -선형결합으로는 x^2을 생성하지 못한다. 즉 $< g_1, g_2 >$는 G-기저가 아니다. LEX 순서에서 G-기저는 다음과 같다.

$$G = \{xy, x^2, 2y^2 - x\}$$

G-기저는 특별한 성질을 가지고 있다. 첫째로, $G = \{g_1, \cdots, g_s\}$가 아이디얼 $I \subset k[x_1, \cdots, x_m]$의 G-기저라 하고 다항식 $f \in k[x_1, \cdots, x_m]$라 하면, $f = \sum_{i=1}^{s} q_i g_i + r$을 만족하는 나머지 $r \in k[x_1, \cdots, x_m]$은 유일하게 존재한다. 이때 나머지 r을 아이디얼 I에 대한 f의 정규형 (normal form)이라 한다. 만일 나머지가 0이라면 이것은 $f \in I$이기 위한 필요충분조건이 된다. 그러나 G-기저도 구성하는 원소의 순서, 즉 g_i의 순서를 달리하면 나머지는 변하지 않을지라도 몫 q_i는 달라진다. 둘째로, 나머지 r의 어떤 항도 $LT(g_1)$, $\cdots$, $LT(g_s)$에 의해 나눠지지 않는다. 마지막으로 정해진 단항식 순서에 따라 최고차항이 달라지기 때문에, 순서를 달리 정의하면 G-기저 또한 달라질 수 있다.

G-기저의 정의에 따르면 I의 원소를 하나 추가하여도 I의 G-기저임은 변하지 않는다. 즉 아이디얼 I에 대한 G-기저는 무수히 많이 존재하게 된다. 그러나 다음성질을 만족하는 G-기저는 유일하게 존재하며 이를 축약 (reduced) G-기저라 부른다.

1. $\forall p \in G$, $LT(p)$의 계수는 1이다.
2. $\forall p \in G$, p의 모든 단항식은 $<LT(G\text{-}\{p\})>$에 속하지 않는다.

[예 1] 2^{6-2}-부분실험

표 2.1에서 선택된 2^4개의 부분실험은 먼저 4개의 요인 $X_1, \cdots, X_4$으로 구성된 완전실험을 구성한 후 요인 X_5와 X_6은 정의관계식 $X_5 = X_1 \times X_2$, $X_6 = X_3 \times X_4$에 의해 구해진다.
이 부분실험을 CoCoA를 이용하여 LEX 순서로 G-기저를 구한 결과 다음과 같았다 ([부록 1] 참조).

$$< X_6 - X_3X_4, X_5 - X_1X_2, X_4^2 - 1, X_3^2 - 1, X_2^2 - 1, X_1^2 - 1 >$$

위의 G-기저를 보면 이 디자인이 어떻게 구성되어 있는지 명확하게 알 수 있다. 결과를 보면 요인 X_4, X_5, X_6가 다음과 같이 교락되어 있는 계획임을 알 수 있어 정규계획의 경우, 정의관계식과 일치함을 알 수 있다.

$$X_5 = X_1X_2,\ X_6 = X_3X_4$$

3. 지시함수

완전요인실험 D로부터 부분실험 F (즉 $F \subset D$)가 선택되었을 때, 선택된 해의 집합을 구별하기 위한 다항함수 I_F가 존재한다. 이를 지시함수라고 부르고 다음과 같이 정의 된다.

$$I_F(\boldsymbol{x}) = \begin{cases} 1 & \text{if } \boldsymbol{x} \in F \\ 0 & \text{if } \boldsymbol{x} \in D - F \end{cases}$$

표 2.1 2^6-2-부분요인실험

run	factor					
	X1	X2	X3	X4	X5=X1X2	X6=X3X4
1	-1	-1	-1	-1	1	1
2	-1	-1	-1	1	1	-1
3	-1	-1	1	-1	1	-1
4	-1	-1	1	1	1	1
5	-1	1	-1	-1	-1	1
6	-1	1	-1	1	-1	-1
7	-1	1	1	-1	-1	-1
8	-1	1	1	1	-1	1
9	1	-1	-1	-1	-1	1
10	1	-1	-1	1	-1	-1
11	1	-1	1	-1	-1	-1
12	1	-1	1	1	-1	1
13	1	1	-1	-1	1	1
14	1	1	-1	1	1	-1
15	1	1	1	-1	1	-1
16	1	1	1	1	1	1

Fontana 등 (2000)과 Ye (2003)는 독립적으로 요인 수준이 $\{-1, 1\}$로 구성된 2^m 완전요인실험 D의 부분실험 F의 지시함수를 다음과 같이 구할 수 있음을 증명하였다.

정리 3.1 (Ye, 2003)

n 개의 실험점을 가진 D의 부분실험 F의 지시함수는 아래와 같이 구해진다.

$$I_F(x) = a_0 + \sum_{l=1}^{m} \sum_{1 \le i_1 < \cdots < i_l \le m} a_{i_1 \ldots i_l} x_{i_1} \cdots x_{i_l} \tag{3.1}$$

식 (3-1)에서 계수는 $a_0 = \frac{n}{2^m}, a_{i_1 \ldots i_l} = \frac{1}{2^m} \sum_{x \in A} x_{i_1} x_{i_2} \cdots x_{i_l}$ *이다.*

[예 1 (계속)**]**

[예 1]의 $X_5 = X_1X_2, X_6 = X_3X_4$로 생성된 2^{6-2}-부분실험 F를 결정하는 다항식을 다음과 같이 정의할 수 있다.

$$I_F(x) = 1/4(1 + x_1x_2x_5)(1 + x_3x_4x_6)$$

식을 살펴보면 x_i는 1이나 -1의 값을 가지므로 x_i^2 =1 (i =1, $\cdots$,6)이 된다. 즉, $x_1x_2x_5 = x_3x_4x_6$ =1이 되므로 F의 모든 실험점에서 I_F(x)는 1의 값을 갖게 된다. 반면에, F가 아닌 실험점에서 $x_1x_2x_5 = x_3x_4x_6$ =-1이 되므로 I_F(x)의 값은 0을 갖게 된다.

동일한 부분실험 F에 대해[정리 3-1]에 의해 구한 지시함수는 다음과 같이 표현되고 앞에서 구한 $I_F(x)$와 같음을 확인할 수 있다.

$$I_F(x) = \frac{1}{4} + \frac{1}{4}x_1x_2x_5 + \frac{1}{4}x_3x_4x_6 + \frac{1}{4}x_1x_2x_3x_4x_5x_6$$

수준이 2 보다 큰 경우 혹은 혼합수준 실험의 경우 요인 s-수준은 복소수 s-제곱근에 의해 코드화 할 수 있다. 그러면 앞에서 본 2 요인 수준에서 소개된 이론이 혼합수준의 실험계획으로 일반화 될 수 있다. Pistone과 Rogantin (2008)은 아래와 같이 유일한 다항식으로 표현 될 수 있음을 보였다.

수준이 s인 요인 A를 복소방정식 $\zeta^s = 1$의 해집합

$$\Omega_s = \{\omega_k = exp(\frac{2\pi k}{s} i) : k = 0, 1, \cdots, s-1\}$$

으로 택하면 함수 $k \mapsto \omega_k$는 덧셈군 $Z_s = \{0, 1, \cdots, s-1\}$에서 곱셈군 Ω_s으로의 군동형사상 (group isomorphism)이다. 이제 위의 선택을 각 A_i에 적용하면혼합형 완전요인실험 $D = A_1 \times \cdots \times A_m$은 m-차원 복소공간 C^m의 부분집합으로 연립방정식계

$$\zeta_j^{s_j} = 1, j = 1, \cdots, m$$

의 해집합이다.

함수 $X_i : D \ni (\zeta_1, \cdots, \zeta_m) \mapsto \zeta_i$를 요인이라 하고 집합 $L = Z_{s_1} \times \cdots \times Z_{s_m}$의 원소 $\alpha = (\alpha_1, \cdots, \alpha_m)$에 대해 $X^\alpha = X_1^{\alpha_1} \cdots X_m^{\alpha_m}$를 주요인 또는 교호작용이라 부른다. D에서 정의된 모든 복소함수들의 집합 $C(D)$상에 내적 (inner product)을 $< f, g > = 1/|D| \sum_{\zeta \in D} f(\zeta)\overline{g(\zeta)}$로 정의하면 $C(D)$는 Hilbert 공간이 된다. 여기서 $|A|$는 A의 원소의 개수를 나타낸다. 이 때 D위의 모든 교호작용들의 집합 $\{X^\alpha \colon \alpha \in L\}$는 이 Hilbert 공간의 정규직교기저가 된다.

그리고 D의 부분요인실험 F의 지시함수 I_F는 다음 식을 만족한다.

$$I_F(\zeta) = \sum_{\alpha \in L} b_\alpha X^\alpha(\zeta), \quad b_\alpha = \frac{1}{|D|} \sum_{\zeta \in F} \overline{X^\alpha(\zeta)}$$

4. 그뢰브너 기저와 지시함수와의 관계

G-기저와 지시함수는 부분요인실험의 아이디얼을 다루는 데에 사용하는 매우 유용한 계산도구이다. 이번 절에서는 G-기저를 이용하여 지시함수를 구하는 과정을 살펴본다.

정의 4.1 소거 아이디얼 (elimination ideal)
아이디얼 $I = < f_1, \cdots, f_s > \subset k[x_1, \cdots, x_m]$*에 대하여* l*번째 소거 아이디얼을 다음과 같이 정의한다.*

$$I_l = I \cap k[x_{l+1}, \cdots, x_m]$$

위와 같이 정의된 I_l 은 $k[x_{l+1}, \cdots, x_m]$의 한 아이디얼임을 쉽게 보일 수 있다. 미지수 x_i들의 순서가 달라지면 그에 따라 다른 소거 아이디얼이 나오게 된다. 아이디얼 I에서 미지수 $x_1, \cdots, x_l$을 소거한다는 것은 l번째 소거 아이디얼 I_l에 속하는 0이 아닌 다항식을 찾는 것을 의미한다. 다음에 소개된 소거정리는 소거 아이디얼과 G-기저에 대한 중요한 정리로서 Cox 등 (1992)의 대수기하 관련 책 등에 소개되어 있다.

정리 4.1 소거정리 (the elimination theorem)
아이디얼 $I \subset k[x_1, \cdots, x_m]$의 LEX 순서에 따른 G-기저를 G라 할 때 집합

$$G_l = G \cap k[x_{l+1}, \cdots, x_m]$$

은 l번째 소거 아이디얼 I_l의 G-기저이다.

예로서 $C[x,y,z]$의 아이디얼 $I=< x^2+y+z-1, x+y^2+z-1, x+y+z^2-1 >$에 대하여 LEX 순서에 따른 G-기저는

$$\{g_1 = x+y+z^2-1, g_2 = y^2-y-z^2+z, g_3 = 2yz^2+z^4-z^2, g_4 = z^6-4z^4+4z^3-z^2\}$$

로 구해지고 소거 아이디얼은 다음과 같다.

$$I_1 = I \cap C[y,z] =< g_2, g_3, g_4 >, I_2 = I \cap C[z] =< g_4 >$$

다음의 정리들은 본 논문의 핵심이 되는 정리들로 Pistone 등 (2006)에 의해 간단히 소개된 것으로 이를 수정 확장하고 정확히 서술하고 증명하였다. 이 정리들은G-기저의 생성자가 주어졌을 때 어떻게 지시함수가 구해지는 지를 보여준다.

정리 4.2 완전요인실험 D와 부분실험 F의 아이디얼을 각각 $I(D) =< d_1, \cdots, d_p >$, $I(F) =< d_1, \cdots, d_p, g_1, \cdots, g_q >$라 하고, G가 D-축약 다항식일때 방정식 G = 0이 F의 생성방정식, 즉

$$< d_1, \cdots, d_p, g_1, \cdots, g_q >=< d_1, \cdots, d_p, G >$$

이기 위한 필요충분조건은 다음 (1)과 (2)를 만족하는 것이다.

(1) $j = 1, \cdots, q$에 대해 $h_j \in k[x_1, \cdots, x_m]$가 존재하여 $G - \sum_{j=1}^{q} h_j g_j \in I(D)$.
(2) 모든 g_i에 대하여 $s_i \in k[x_1, \cdots, x_m]$가 존재하여 $g_i - s_i G \in I(D)$.
그리고, $G = 1 - f$에 대해 만일 위의 (1)과 다음 (3)이 만족되면 f는 부분요인 F의 지시함수이다.
(3) 모든 g_i에 대하여 $fg_i \in I(D)$.

증명: 방정식 G=0이 F의 생성방정식이라 가정하자. 그러면 $G \in I(F)$이므로 $t_i, h_i \in k[x_1, \cdots, x_m]$가 존재하여 $G = t_1 d_1 + \cdots t_p d_p + h_1 g_1 + \cdots + h_q g_q$이고 따라서 $G - \sum_{j=1}^{q} h_j g_j = \sum_{i=1}^{p} t_i d_i \in I(D)$이다. 모든 g_i에 대하여 $g_i \in< d_1, \cdots, d_p, G >$이므로 $t_{i_k}, s_i \in k[x_1, \cdots, x_m]$가 존재하여 $g_i = t_{i1} d_1 + \cdots + t_{ip} d_p + s_i G$이고 따라서 $g_i - s_i G = \sum_{k=1}^{p} t_{i_k} d_k \in I(D)$이다.
역으로 조건 (1)은 $< d_1, \cdots, d_p, G >\subset I(F)$을, 조건 (2)는 $I(F) \subset< d_1, \cdots, d_p, G >$을 보장하므로 결국 방정식 G=0 이 F의 생성방정식이 된다.
그리고 $G = 1 - f$에 대해 (3)이 만족되면 (2)가 만족되므로 조건 (1), (3)이 만족되면 $I(F) =< d_1, \cdots, d_p, 1 - f >$이다. 이제 임의의 $a \in F$에 대해 $1 - f(a) = 0$ 이므로 $f(a) = 1$이다. 또 임의의 $a \in D - F$에 대해 조건 (3)에 의해 $f(a)g_i(a) = 0$ $(i = 1, \cdots, q)$이고 적어도 하나의 $g_i(a)$는 0이 아니어야 하므로 $f(a) = 0$이다. 즉 f는 부분실험 F의 지시함수이다. □

예로서 실험수준이 {-1, 0, +1}인 3^2-완전요인 실험에서 디자인 다항식은 다음과 같다.

$$d_1(x) = x^3 - x, d_2(y) = y^3 - y$$

부분실험을 $g(x,y) = (x^2-1)y$로 생성한다고 할 때 f가 정리 4.2의 (1), (3)을 만족한다면

$$\left\{\begin{array}{c} x^3 - x = 0 \\ y^3 - y = 0 \\ 1 - f - h(x^2-1)y = 0 \\ f(x^2-1)y = 0 \end{array}\right\}$$

가 된다. 여기서 우리는 x와 y를 사용하여 h는 제거하고 f를 결정하고자 한다. 위의 세 번째 방정식의 양 변에 $(x^2-1)y$를 곱하면 $(x^2-1)y - h(x^2-1)^2y^2 = 0$이 되고, 다른 방정식을 사용하여 우리는 $f - x^2y^2 + y^2 - 1 = 0$을 구할 수 있다. 그러므로 지시함수 $f = x^2y^2 - y^2 + 1$이 된다. 위의 연립방정식계와 동치인 다음 방정식계

$$\left\{\begin{matrix} x^3 - x = 0 \\ y^3 - y = 0 \\ f - x^2y^2 + y^2 - 1 = 0 \\ hx^2y - hy + x^2y^2 - y^2 = 0 \end{matrix}\right\}$$

는 변수의 순서가 $x < y < f < h$로 주어진 단항식들의 LEX 순서에 따른 하삼각형 형태의 연립방정식계이고 가장 작은 최고차항 (Leading Term; LT)을 가진다. 즉, 위의 다항식들은 사전식 순서에 의한 G-기저이다. 위의 예를 통해 다음의 정리 4.3을 유추해 낼 수 있다.

정리 4.3 완전요인실험 D와 부분실험 F의 아이디얼을 각각 $I(D) =< d_1, \cdots, d_p >$, $I(F) =< d_1, \cdots, d_p, g >$이라 하자. 다항식환 $k[h, f, x_1, \cdots, x_m]$에서 아이디얼 $< d_1, \cdots, d_p, 1-f-hg, fg >$의 첫번째 소거아이디얼

$$< d_1, \cdots, d_p, 1 - f - hg, fg > \cap k[f, x_1, \cdots, x_m]$$

의 LEX 순서에 따른 축약 G-기저는 $f - \sum_{a\in L} b_a x^a$의 형태의 유일한 다항식을 포함한다. 이 때 부분실험 F의 지시함수 I_F는 $\sum_{a\in L} b_a x^a$가 된다.

증명: 먼저 아이디얼 $J = < d_1, \cdots, d_p, 1 - f - hg, fg >$의 다양체 $\mathrm{V}(J) \subseteq k^{m+2}$를 구하여 보자.

$$f(1 - f - hg) + h(fg) = f - f^2 \in J$$

이므로 $\mathrm{V}(J)$에 속한 점들의 f 성분은 0 또는 1이다. f성분이 0인 경우 D - F에 속한 x에 대해 $h = 1/\ g(\ x)$가 되고, f성분이 1인 경우 F에 속한 x에 대해 $g(x) = 0$이므로 h는 임의상수가 될 수 있다. 즉 J의 다양체는 다음과 같다.

$$V(J) = \{(1/g(x), 0, x) : x \in D - F\} \cup \{(h, 1, x) : h \in k, x \in F\}$$

이제 첫 번째 미지수 h를 소거한 소거아이디얼 $J_1 =< d_1, \cdots, d_p, 1-f-hg, fg > \cap k[f, x_1, \cdots, x_m]$의 다양체 $V(J_1)$은 집합 $\{(0, x) : x \in D - F\} \cup \{(1, x) : h \in k, x \in F\}$이 되고 LEX 순서에 따른 축약 G-기저는 D - F에서 0이고 F에서 1인 값을 갖는 미지수 $x_1, \cdots, x_m$들의 한 다항식 $\sum_{a\in L} b_a x^a$에 대해 $f - \sum_{a\in L} b_a x^a$형태의 유일한 다항식을 포함하게 된다. 따라서 다항식 $\sum_{a\in L} b_a x^a$는 F의 지시함수 I_F이다. □

정리 4.4 완전요인실험 D와 부분실험 F의 아이디얼을 각각 $I(D) =< d_1, \cdots, d_p >, I(F) =< d_1, \cdots, d_p, g_1, \cdots, g_q >$이라 하자. 각 다항식 $g_i(i = 1, \cdots, q)$에 의해 생성된 다양체와 아이디얼을 각각 $F_i = V(d_1, \cdots, d_p, g_i)$, $I(F_i) =< d_1, \cdots, d_p, g_i >$라 놓으면 $F = \cap_{i=1}^q F_i$이고 지시함수사이의 관계는 $I_F = NF_D(I_{F_1} \cdots I_{F_q})$이다.

증명: 다양체와 아이디얼의 관계에 의해

$$F = V(d_1, \cdots, d_p, g_1, \cdots, g_q) = \bigcap_{i=1}^{q} V(d_1, \cdots, d_p, g_i) = \bigcap_{i=1}^{q} F_i$$

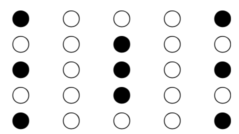
임은 명백하다. 이제 함수 $f = NF_D(I_{F_1} \cdots I_{F_q})$라 놓고 f가 부분실험 F의 지시함수임을 보이자. 정규형의 정의에 의해 다항식 $h \in I(D)$가 존재하여 $I_{F_1} \cdots I_{F_q} = h + f$가 성립한다. 임의의 $x \in F$에 대해 $I_{F_i}(x) = 1 (i = 1, \cdots, q)$이고 D의 점 x가 F에 속하지 않는다면 적어도 하나의 F_i에 속하지 않게 되고 $I_{F_i}(x) = 0$이 된다. 따라서

$$f(x) = I_{F_1}(x) \cdots I_{F_q}(x) - h(x) = \begin{cases} 1 \cdots 1 - 0 = 1 & \text{if } x \in F \\ 0 - 0 = 0 & \text{if } x \in D - F \end{cases}$$

□

[예 2] (CoCoA를 활용한 5^2-부분실험의 G-기저와 지시함수)
5^2완전요인실험에서 수준을 $\{-2, -1, 0, 1, 2\}$로 놓고 다음의 9점 부분실험을 선택하였다.

$$F = \{(-2, 2), (-2, 0), (-2, -2), (0, 1), (0, 0), (0, -1), (2, 2), (2, 0), (2, -2)\}$$



수준의 집합이 $\{-2, -1, 0, 1, 2\}$인 완전요인실험 D의 디자인 다항식은 아래와 같다.

$$d_1 = x(x^2 - 1)(x^2 - 4), \quad d_2 = y(y^2 - 1)(y^2 - 4)$$

먼저 아이디얼 $\boldsymbol{I}(F)$의 G-기저를 구하면

$$g_1 = x^3 - 4x, g_2 = x^2 y - \frac{4}{3}y^3 + \frac{4}{3}y, \quad g_3 = xy^3 - 4xy, g_4 = y^5 - 5y^3 + 4y$$

이다. 다항식 $g_4 = d_2$이므로 축약 생성 다항식계 (reduced generation set)는 g_1, g_2, g_3만을 사용하면 된다. 각각의 g_i $(1 \le i \le 3)$에 의해 생성된 부분요인을 F_i라 놓고 그 지시함수를 각각 구하면 [정리 4-4]에 의해 그것들의 곱의 정규형을 구함으로써 처음 주어진 부분요인실험 F의 지시함수를 구할 수 있다. 정리 4.3에 의해 LEX 순서에 의한 G-기저가 제공하는 지시함수를 각각 구하여 보면 다음과 같다. G-기저와 지시함수를 구하는 과정은 첨부된 CoCoA 프로그램 (참조 [부록 2])을 참고한다.

$$\begin{aligned} I_{F_1} &= -\frac{1}{12}x^4y^4 + \frac{1}{3}x^4y^2 + \frac{5}{12}x^2y^4 - \frac{5}{3}x^2y^2 + 1 \\ I_{F_2} &= -\frac{11}{144}x^4y^4 + \frac{47}{144}x^4y^2 + \frac{59}{144}x^2y^4 - \frac{239}{144}x^2y^2 - \frac{1}{12}y^4 + \frac{1}{12}y^2 + 1 \\ I_{F_3} &= \frac{1}{3}x^4 - \frac{4}{3}x^2 + 1 \end{aligned}$$

그리고 F의 지시함수 I_F는 다음과 같다.

$$\begin{aligned} I_F &= NF_D(I_{F_1} \cdots I_{F_3}) \\ &= \frac{1}{144}x^4y^4 - \frac{13}{144}x^4y^2 + \frac{1}{3}x^4 + \frac{11}{144}x^2y^4 + \frac{1}{144}x^2y^2 - \frac{4}{3}x^2 - \frac{1}{12}y^4 + \frac{1}{12}y^2 + 1 \end{aligned}$$

5. 결 론

본 논문에서는 임의의 부분요인실험이 선택되었을 때, 그 실험의 교락구조를 G-기저를 구하여 해석한 후 구해진 G-기저를 생성자로 활용하여 지시함수를 구하는 절차를 알아보았다. 실제로 몇개의 부분요인실험을 예로 택하여 CoCoA 대수연산 소프트웨어를 이용하여 수행하였다. 이러한 과정을 통해 임의로 선택된 실험의 구조를 명확하게 해석할 수 있었다. 본 논문에서 소개된 접근 방법은 혼합물계획 등 보다 복잡한 계획에서의 부분요인실험의 구조 해석을 용이하게 하여 효율적인 선택을 할 수 있도록 확대될 수 있으리라 기대된다.

부 록

[부록 1]

[예 1: 2^{6-2} 부분요인실험의 G-기저와 지시함수 생성을 위한 CoCoA 프로그램]

```
Use R ::= Q[h,f,x[1..6]], Lex; Set Indentation;
S := Intersection(
Ideal(x[1]-1, x[2]-1, x[3]-1, x[4]-1, x[5]-1, x[6]-1),
Ideal(x[1]-1, x[2]-1, x[3]-1, x[4]+1, x[5]-1, x[6]+1),
Ideal(x[1]-1, x[2]-1, x[3]+1, x[4]-1, x[5]-1, x[6]+1),
Ideal(x[1]-1, x[2]-1, x[3]+1, x[4]+1, x[5]-1, x[6]-1),
Ideal(x[1]-1, x[2]+1, x[3]-1, x[4]-1, x[5]+1, x[6]-1),
Ideal(x[1]-1, x[2]+1, x[3]-1, x[4]+1, x[5]+1, x[6]+1),
Ideal(x[1]-1, x[2]+1, x[3]+1, x[4]-1, x[5]+1, x[6]+1),
Ideal(x[1]-1, x[2]+1, x[3]+1, x[4]+1, x[5]+1, x[6]-1),
Ideal(x[1]+1, x[2]-1, x[3]-1, x[4]-1, x[5]+1, x[6]-1),
Ideal(x[1]+1, x[2]-1, x[3]-1, x[4]+1, x[5]+1, x[6]+1),
Ideal(x[1]+1, x[2]-1, x[3]+1, x[4]-1, x[5]+1, x[6]+1),
Ideal(x[1]+1, x[2]-1, x[3]+1, x[4]+1, x[5]+1, x[6]-1),
Ideal(x[1]+1, x[2]+1, x[3]-1, x[4]-1, x[5]-1, x[6]-1),
Ideal(x[1]+1, x[2]+1, x[3]-1, x[4]+1, x[5]-1, x[6]+1),
Ideal(x[1]+1, x[2]+1, x[3]+1, x[4]-1, x[5]-1, x[6]+1),
Ideal(x[1]+1, x[2]+1, x[3]+1, x[4]+1, x[5]-1, x[6]-1));
T:=ReducedGBasis(S); T;
D:= [x[1]^2-1, x[2]^2-1, x[3]^2-1, x[4]^2-1, x[5]^2-1, x[6]^2-1];
G:= [-x[3]+x[4]x[6], -x[1]+x[2]x[5]];
S:=[]; N:=[7,7];
For I:=1 To Len(G) Do
PrintLn "I=" , I;
L:=ConcatLists([D, [1-f-hG[I]], [fG[I]]]);
Id:=Saturation(Ideal(L), Ideal(f, x[1], x[2], x[3], x[4], x[5], x[6]));
GB:=ReducedGBasis(Id); GB;
S:=Concat([f-GB[N[I]]], S);
```

```
PrintLn " ";
EndFor; S;
SF:=NF(Product(S),Ideal(D)); SF;
```

[부록 2]
[예 2: 5^2 부분요인실험의 G-기저와 지시함수 생성을 위한 CoCoA 프로그램]

```
Use R ::= Q[h, f, x, y], Lex; Set Indentation;
S := Intersection(Ideal(x+2, y-2), Ideal(x+2, y), Ideal(x+2, y+2),
Ideal(x, y-1), Ideal(x, y), Ideal(x, y+1), Ideal(x-2, y-2),
Ideal(x-2, y), Ideal(x-2, y+2));
B:=ReducedGBasis(S); B;
D:=[x(x^2-1)(x^2-4), y(y^2-1)(y^2-4)];
G:=[x^3 - 4x, x^2y - 4/3y^3 + 4/3y, xy^3 - 4xy];
S:=[]; N:=[2, 4, 3];
For I:=1 To Len(G) Do
PrintLn "I=" , I;
L:=ConcatLists([D, [1-f-hG[I]], [fG[I]]]);
Id:=Saturation(Ideal(L), Ideal(f, x, y));
GB:=ReducedGBasis(Id); GB;
S:=Concat([f-GB[N[I]]], S);
PrintLn " ";
EndFor; S;
SF:=NF(Product(S), Ideal(D)); SF;
```

참고문헌

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

Fontana, R., Pistone, G. and Rogantin, M. P. (2000). Classification of two-level factorial fractions. *Journal of Statistical Planning and Inference*, **87**, 149-172.

Park, D. K. and Kim, H. (2003). A New approach for selecting fractional factorial designs. *Journal of the Korean Data & Information Science Society*, **14**, 707-714.

Pistone, G., Riccomagno, E. and Rogantin, M. P. (2006). *Algebraic statistics method for DOE*, Unpublished Manuscript.

Pistone, G. and Rogantin, M. P. (2008). Indicator function and complex coding for mixed fractional factorial designs. *Journal of Statistical Planning and Inference*, **138**, 107-121.

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

Plackett, R. L. and Burman, J. P. (1946). The design of optimum multifactorial experiments. *Biometrika*, **33**, 305-325.

Ye, K. Q. (2003). Indicator functions and its application in two-level factorial designs. *Annals of Statistics,* **31**, 984-994

Zhang, R. and Park, D. K (2000). Optimal blocking of two-level fractional factorial designs. *Journal of Statistical Planning and Inference,* **91**, 107-121.

Journal of the Korean Data & Information Science Society
2009, **20**(6), 1015–1027

한국데이터정보과학회지

Gröbner basis versus indicator function

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

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

Received 2 August 2009, revised 29 October 2009, accepted 3 November 2009

Abstract

Many problems of confounding and identifiability for polynomial models in an experimental design can be solved using methods of algebraic geometry. The theory of Gröbner basis is used to characterize the design. In addition, a fractional factorial design can be uniquely represented by a polynomial indicator function. Gröbner bases and indicator functions are powerful computational tools to deal with ideals of fractions based on each different theoretical aspects. The problem posed here is to give how to move from one representation to the other. For a given fractional factorial design, the indicator function can be computed from the generating equations in the Gröbner basis. The theory is tested using some fractional factorial designs aided by a modern computational algebra package CoCoA.

Keywords: Algebraic geometry, Aliased, Gröbner basis, indicator function, polynomial.

[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