The Korean Communications in Statistics
Vol. 13 No. 1, 2006, pp. 1-10

Strategical Issues in Multiple-Objective Optimal Experimental Design[1)]

Young-Il Kim[2)] and Myung-Wook Kahng[3)]

Abstract

Many of statistical experimental designs have multiple goals. It is often impractical to use the single-objective criterion for this purpose. It is necessary to modify the existing optimum experimental design criteria. There exist three criteria handling this problem in general: compound, constrained, maxi-min approach. This paper extends Kahng and Kim's idea to develop another approach to incorporate several experimental design criteria in accordance of their importance in practical way. Furthermore this paper investigate its relationship with the maxi-min approach. It shows logically that the often realized infeasibility can be still avoided with the rank of importance of the objectives intact.

Keywords : Φ-optimality; Compound design; Constrained design; maxi-min design.

1. 서론

Box와 Draper (1975)가 제안했던 실험계획 시 고려하여야 할 14가지의 기준은 주어진 모형과 가정에 의존한 단일목적(single objective)의 최적실험계획법(optimal experimental design)에 대한 비판을 이끌기에 충분하다고 본다. 통계적인 실험의 많은 문헌은 단일목적을 가진 실험계획을 염두에 두고 진행된 것이 사실이다. Fedorov (1972)나 Silvey (1980) 그리고 Pukelsheim (1993)의 책들은 이러한 관점에서 쓰여졌다고 본다. 여러 목적을 가지고 있는 최적실험계획법은 산발적으로 소개 되어오다가 Cook과 Wong (1994)의 논문 이래 많은 연구가 소개되고 있다. 그 중에서도 Atkinson과 Bogacka (1997), Zhu, Ahn과 Wong (1998), Huang과 Wong (1998a, b)의 논문을 들 수 있다. 다중목적을 가지고 있는 실험계획은 특별히 어려운 점 없이 단일목적을 가진 실험계획의 기존의 알고리즘을 이용하여 구성할 수 있으며 주어진 모형, 가정 등이 실험결과에 의해 잘못되었다고 판단하더라도 강건성(robustness)을 유지하는 특징을 가지고 있기 때문에 이에 대한 연구가 최근에 들어 많은 각광을 받고 있다.

---

1) This research is supported by the 2005 ChungAng University research fund.

2) Professor, Department of Information System, ChungAng University, KyunggiDo 456-756, Korea. Correspondence : yik01@cau.ac.kr

3) Professor, Department of Statistics, Sookmyung Women's University, Seoul 140-742, Korea.

단일목적을 가진 실험계획은 대부분 실험자의 욕구를 충족시키지 못하는 경우가 많다. 특히 최적실험계획법이론은 모형을 가정하기 때문에 모형에 대한 불확실성이 존재한다면 문제가 된다. 또한 모형에 대한 가정이 올바르다 하더라도 실험기준이 여럿이라면 다목적 실험을 연상하여 볼 수 있을 것이다. 최적실험계획에서 제일 각광받는 실험기준은 $D-$최적이나 이는 모수추정에 국한한 목적인 것이다. 만약 실험자가 다른 목적의 실험기준을 동시에 염두에 둔다면 부적합한 것이다. 모형에 부수한 가정도 역시 문제가 된다. 대부분의 경우는 등분산을 염두에 두고 실험을 실시하나 이는 올바르지 않다. 이런 경우에도 다중목적의 실험이 요구된다. 더욱이 이러한 현상들이 복합적으로 나타나는 문제인 경우에는 근본적으로 하나의 실험기준으로 다루는 것은 매우 어려울 것이다. 본 논문의 2절에서는 단일목적을 둔 실험기준에 대한 내용 및 표현방법을 알아보고 다중목적 실험의 필요성을 좀더 기술적으로 알아본다. 그리고 3절에서는 본 논문이 제시하는 다중목적을 염두에 둔 실험계획 하에서의 전략을 모형화 하고 예제들을 통해 이를 살펴본다. 마지막으로 4절에서는 향후 연구과제 및 결론을 제시한다.

2. 표현방법

최적실험문제는 대부분 회귀모형을 정한 다음 발생되는 기대 Fisher의 정보행렬(expected Fisher information matrix: 앞으로는 단순히 정보행렬)의 convex 함수로서 실험기준이 되는 목적함수를 만든 다음 이를 최적화하는 문제로 귀결된다. 목적함수 혹은 최적실험기준(optimal criterion)은 편의상 Φ로 표현한다. 본 논문에서는 최적실험에서 많이 쓰이는 연속실험(continuous design)만을 염두에 둔다. 연속실험은 실험영역(design space)에 부여된 확률 메저(probability measure)로 처리되며 이는 받힘점(support point)과 받힘점에 대한 질량분포(mass distribution)로 성격이 정해진다. 이러한 실험을 본 논문에서는 ξ라 한다. 이를 위해 다음과 같은 회귀모형을 설정한다.

$$y(x) = f^T(x)\beta + \epsilon, \qquad x \in X \tag{1}$$

여기서 $y(x)$는 x에서 관측 가능한 일변반응변수(univariate response variable)이며 $f(x)$는 사용자가 지정한(user-specified) 실험영역인 X에서의 알려진 회귀함수이다. X의 구간은 편의상 $X = [-1, 1]$로 정한다. 오차ϵ는 평균이 0이고 등분산인 정규분포를 따른다고 가정한다. 오차의 구조가 이분산이고 오차의 분산이 특정한 양의 함수(positive function)에 비례한다고 가정하더라도 위의 모형은 간단한 변환을 통해 최적이론을 적용할 수 있다. 이 모형에 대한 임의의 실험계획 ξ의 정보행렬은 다음과 같이 보일 수 있다.

$$M(\xi) = \int_X f(x) f^T(x)\, \xi(dx)$$

여기서 불필요한 상수는 무시한다. 그리고 본 논문에서는 비특이(non-singular)인 정보행렬에 관련된 실험계획만 고려한다.

흔히 많이 쓰이는 실험기준에는 $D-$,$E-$ 그리고 $A-$최적(optimality)이 있다. $D-$최적은 주어진 모형의 모수 β에 대한 추정에 관련된 기준으로 $\Phi(\xi) = -\ln|M(\xi)|$ 이

다. $E-$ 최적은 $\Phi(\xi)=\max_{c^Tc=1}c^TM(\xi)^{-1}c$ 이다. 다른 흥미로운 기준으로 추정반응표면(estimated response surface)의 분산의 최대값을 최소화하는 $G-$최적이 있다. 즉, $\bar{d}(\xi)=\max_{x\in X}d(x,\xi)=f^T(x)M(\xi)^{-1}f(x)$을 최소화하는 기준이다. 이 기준은 위에서 정의된 등분산 모형 하에서는 $D-$최적과 동격이다. 또 많이 쓰이는 다른 유형의 기준은 $\Phi(\xi)=tr\,LM(\xi)^{-1}$으로 정의되는 $L-$최적이다. 여기서 L은 비음의 정부호(non-negative definite)행렬이다. 만약 L이 단위행렬인 경우는 $A-$최적을 얻는다. 주어진 점이 z이고 $L=f(z)f(z)^T$이면 주어진 점 z에서의 추정반응(estimated response)의 분산을 최소화하는 기준이 된다. 이러한 모든 실험기준은 정보행렬 공간(space of information matrices) 에서 볼록한(convex) 함수이다. 이에 관련된 참고문헌으로서는 Pukelsheim (1993)이 있다.

위에서 언급한 기준들은 모두 모형 $f(x)$ 이 알려져 있다고 가정하는데 현실적으로는 그렇지 않다. 이를 위해 $f(x)$ 를 $f^T(x)=(f_1^T(x),f_2^T(x))$의 두 부분으로 분해하여 보자. 예를 들어 실험자가 차수 r 의 다항회귀모형에 대해 자신을 가지고 있다 하더라도 만약 참의 모형이 차수가 $m(>r)$ 인 다항회귀모형일 가능성에도 대비한다면 실험자는 r 차수의 회귀모형만을 염두에 둔 실험기준을 생각하지는 않을 것이다. 차수가 높은 모형에 대해서도 좋은 모수 추정값을 제공하는 실험을 하려고 할 것이다. 이 경우는 $f_1^T(x)=(1,x,\cdots,x^r)$, $f_2^T(x)=(x^{r+1},\cdots,x^m)$ 그리고 $\beta=(\beta_1^T,\beta_2^T)$이다. 다시 언급하면 실험은 β_1뿐 아니라 모형이 확장될 것에 대비해 β_2 에 대한 좋은 추정도 제공하는 측면을 가져야 한다. 이러한 접근방법은 다양한 문헌에서 확인할 수 있다. Atwood (1969), Läuter (1974), Atkinson (1972) 등을 들 수 있다. 분해된 모형 $Ey(x)=f_1^T(x)\beta_1+f_2^T(x)\beta_2$ 하에서의 정보행렬은 다음과 같이 쓰여 질 수 있다.

$$M_m(\xi)=\begin{pmatrix}M_{11}(\xi) & M_{12}(\xi)\\ M_{21}(\xi) & M_{22}(\xi)\end{pmatrix}$$

여기서 $M_{11}(\xi)$ 는 크기가 $r+1$ 그리고 $M_{22}(\xi)$ 는 크기가 $m-r$ 이다. 만약 r 차수의 다항회귀모형이 적합되는 경우는 최소제곱 추정량 β_1의 공분산은 $M_{11}^{-1}(\xi)$ 이다. 한편으로 차수 m 인 다항회귀모형이 적합된다면 β_2 의 최소제곱 추정량의 공분산 행렬은 $M_{22.1}^{-1}(\xi)$ 이다. 여기서 $M_{22.1}(\xi)=M_{22}(\xi)-M_{21}(\xi)M_{11}^{-1}(\xi)M_{12}(\xi)$ 이다. 이와 같은 상황에서는 $|M_{22.1}(\xi)|$과 $|M_{11}(\xi)|$의 균형을 요하는 실험기준이 요구될 것이다. 이러한 문제에 대한 논의는 Studden (1982)에 의해 제기되었다.

또한 모형에 대한 불확실성 뿐 아니라 실험기준도 문제가 된다. 같은 모형에 대해서도 여러 개의 실험기준을 시도하는 경우가 있다. $D-$최적이 많이 쓰이는 것이 사실이나 간혹 다른 최적과 결합하여 좀 더 나은 실험을 주곤 한다. Wong (1995)은 차수가 낮은 다항회귀모형에 대해 $D-$최적과 $A-$최적을 결합하여 모수추정을 시도하였다. 그리고 모형에 포함된 모수는 실험자의 관점에 따르면 다 같은 중요성을 가지는 것이 아니다. 이러한 상황에서 $D-$최적을 쓴다는 것은 별로 바람직하지 않을 것이다.

이와 같이 여러 목적을 가진 실험기준의 필요성은 실험자의 관점에서 보게 되면 자

연스럽게 대두가 된다. 다중목적 실험계획법은 이유가 어찌되었든 다중목적 실험기준에 대한 중요성에 대한 우선순위를 실험자가 부여할 수 있다면 다중목적 실험계획은 자연스럽게 제약조건식으로 표현이 가능하다. 중요한 실험기준에 대한 효율성을 보존하면서 부차적인 실험기준에 대한 효율성을 극대화 하는 실험을 생각할 수 있다. 이러한 개념을 바탕으로 태어난 것이 제약조건 실험계획법이다. 제약조건 실험계획은 Stigler (1971), Lee (1987, 1998), Cook과 Fedorov (1995), Dette와 Franke (2000) 등에 의해 연구가 진행되었다. 3절에서는 문헌에서 대두된 제약조건 실험과 기타 다중목적 실험과의 연계성을 언급한 다음 다중목적 실험계획이 가지고 있는 문제점을 지적하고 이를 해결하는 방법에 대해 언급하겠다.

3. 제약조건 실험계획법

임의의 실험계획 ξ이 $\Phi-$최적실험 ξ_{Φ}^{*}에 대해 가지는 $\Phi-$효율성(efficiency)은 $e(\xi,\xi_{\Phi}^{*})=\Phi(\xi_{\Phi}^{*})/\Phi(\xi)$로 정의된다. 만약 위에서 정의된 모형이 차수가 m인 다항회귀 모형인 경우는 ξ_{Φ}^{*}이 $D-$최적이라면 임의의 실험 ξ의 $D-$와 $G-$효율성은 각각 $\{|M_m(\xi)/|M_m(\xi^{*})|\}^{1/(m+1)}$과 $(m+1)/\overline{d_m}(\xi)$이다. 어떤 실험의 효율성이 0.5라면 최적실험과 같은 효율성을 가지기 위해서는 실험이 효율성의 역수인 2회의 반복실험이 되어야 한다는 의미이다. 문제의 단순성을 위하여 실험자가 Φ_1, Φ_2이라는 두개의 convex 목적함수를 가지고 있다고 가정한다. 그리고 Φ_1이 Φ_2에 비해 더 중요하다고 판단하다면 제약조건 실험계획법은 $\Phi_1(\xi)\le c$ 하에서 $\Phi_2(\xi)$를 최소화하는 것이다. 그러나 실험자가 지정하여야 하는 상수 c는 명기하기가 불편하므로 편의상 제약조건은 $e(\xi,\xi_{\Phi_1}^{*})\ge e$로 바꾸어진다. 여기서 $e(\xi,\xi_{\Phi_1}^{*})=\Phi_1(\xi_{\Phi_1}^{*})/\Phi_1(\xi)$이며 e는 0과 1사이의 값으로 최소한의 Φ_1-효율성이며 실험자가 지정하여야 하는 값이다. 이러한 제약조건 실험계획법은 Φ_1, Φ_2의 선형결합으로 이루어지는 목적함수, $\Phi(\xi|\lambda)=\lambda\Phi_1(\xi)+(1-\lambda)\Phi_2(\xi)$, $\lambda\in[0,1]$를 최소화 하는 Läuter (1976)의 복합최적실험(compound optimal design), ξ_λ 방법과 동격임이 Cook과 Wong (1994)에 의해 증명되었다. 대략적으로 말하면 적절한 제약조건이 형성이 되어 있는 실험계획법에서 복합최적실험계획법으로 목적이 구현된다는 것이다. Cook과 Wong (1994)은 효율성그림(efficiency plot)이라는 도구를 이용하여 제약조건 실험계획법의 구성을 시도하였다. 이는 x축은 λ로, y축은 ξ_λ의 Φ_1, Φ_2-효율성을 그리는 그림이다. 제약조건 실험계획법에서 주어진 상수 e를 y축에서 수평으로 움직여 $e(\xi_\lambda,\xi_{\Phi}^{*})$와 만나면 바로 밑으로 수직선을 그려 λ를 확인하고 해당되는 실험계획을 구하는 방법이다. 제약조건 실험계획법을 직접적으로 구하는 방법보다 이 방법을 취하는 이유는 복합최적실험 ξ_λ을 확인하는 방법은 이미 수치적으로 잘 알려져 있기 때문이다. 만약 두개의 실험기준만 존재 한다면 ξ_λ의 Φ_1, Φ_2-효율성은 어느 순간 교차한다고 증명되어 있다. 또한 교차되는 지점에서는 두 실험기준 하에서의 최소의 효율성을 최대화 하는 점임이 밝혀졌다. 따라서 이러한 방법은 김영일 (1993)이

제안한 maxi-min 접근방법에 대한 새로운 알고리즘을 제안한다고 볼 수 있다. 자세한 내용은 Imhof와 Wong (1999)을 참조 바란다.

실험기준이 3개 이상인 다중목적을 염두에 둔 경우를 고려하여 보자. 구체적으로 m 개의 실험기준을 $\Phi_1, \Phi_2, \cdots, \Phi_m$ 로 표현한다. 그리고 사전에 명시된 상수 $e_1, \cdots, e_{m-1}$ 에 대해 $\Phi_i(\xi)$ 효율성은 적어도 이보다 큰 조건하에서 $\Phi_0(\xi)$ 을 최소화 하는 실험기준이 순수한 의미에서 다중목적 실험계획이다.

$$\min \Phi_m(\xi) \quad \text{subject to} \quad e(\xi, \xi^*_{\Phi_i}) \geq e_i,\ i = 1, \cdots, m-1 \qquad (2)$$

그러나 실험기준이 3개 이상인 다중목적을 염두에 둔 경우에는 문제가 복잡해진다. 간혹 문헌에서 이러한 경우가 취급되어 왔으나 구체적인 해를 제시하는 경우는 매우 드물다. 그리고 다중목적 실험에서 m 개의 실험기준들이 달성하여야 하는 효율성을 실험자가 사전에 명시하는 것은 그렇게 쉬운 일은 아니다. 따라서 이론적인 해를 구하는 것은 거의 불가능하기 때문에 많은 경우 수치해석적인 알고리즘을 적용하여 구하여야 하는데 이 또한 대부분 소수의 제약조건하에서만 해를 찾을 수 있다. 왜냐하면 제약조건식이 2개 이상으로 늘어나는 경우는 모든 제약조건을 다 만족하는 실험방법이 존재하지 않는 경우가 있어 시작 타당점(starting feasible design)을 찾는 문제가 대두되기 때문이다. Pukelsheim (1993)에 의하면 이론적으로는 제약조건식이 2개 이상인 경우 제약최적실험에 대한 구체적인 해(explicit solution)가 일반적으로 존재하지 않는다고 알려졌다. 따라서 김영일과 강명욱 (2002)은 타당치 않은 해의 위험을 피해갈 수 있는 순차적인 알고리즘을 제안하였다. 이 방법은 항상 시작 타당점(starting feasible design)이 주어지므로 해를 구하는데 어려움이 없다. 또한 우선순위가 높은 실험기준이 주 제약조건으로 순차적으로 들어감으로써 실험자가 명시한 효율성을 충족시킬 수 있는 장점이 있다. 따라서 주어진 효율성을 달성시키지 못할 수 도 있다는 제약성을 염두에 두고 실험을 하는 이와 같은 방법은 일시에 모든 제약조건에 대한 효율성을 부여하는 것보다는 낫다고 판단된다.

그러나 이와 같은 방법도 몇 가지 문제점을 지니고 있다고 판단된다. 첫째 우선순위가 제일 높은 실험기준이 가져야 하는 효율성에 대한 하한값을 설정하는데 많은 어려움이 있을 수밖에 없다. 모형과 실험기준의 특징을 모르는 상태에서 이를 실험자에게 요구하는 것은 무리이다. 따라서 실험자가 e_1 를 명기하였다 하더라도 이러한 점을 반영하여 실험이 설정되어야 한다. 두 번째로는 우선순위가 높은 실험기준에 대해 설정한 효율성보다 우선순위가 낮은 실험기준의 효율성이 결과적으로 더 높게 나올 수 있다. 따라서 실험기준의 우선순위에 따라 실험에 따른 효율성의 순서가 정해져야 한다. 예제를 통해 이 문제를 짚어보기로 한다. 예제의 모든 계산은 염준근과 남기성 (2000)의 유전자 알고리즘을 이용하여 구하였다.

예제 1 실험자는 2차 회귀모형을 염두에 두고 있다. 그러나 참의 모형이 일차 모형일 경우에 대비해 $D-$효율성, $e(\xi, \xi_1^D)$ 는 적어도 e_1 만큼 보장받아야 한다고 생각한다. 이러한 조건하에서 2차 모형에 대한 $e(\xi, \xi_2^D)$ 을 최대화 하는 식 (2)의 실험계획을 생각하여 본다. e_1 의 값이 81.7%($=\sqrt{2/3}$)보다 낮게 설정되어 있으면 $e(\xi, \xi_2^D)$은

100%를 보장 받는다. 즉, 받힘점(supporting point), -1, 0, 1에 균일질량 1/3을 부여하는 실험계획이 바로 우리가 원하는 실험이 된다. 그러나 e_1 의 값이 $\sqrt{2/3}$ 보다 높게 설정되어 있다면 $e(\xi, \xi_2^D)$ 는 100% 보장 받지 못한다. 이와 같은 경우 목적함수가 가질 수 있는 최대값은 $[27(1-\xi(0))^2\xi(0)/4]^{1/3}$ 로 나타난다. 여기서 $\xi(0)=1-e_1^2$ 이며 받힘점 0에 대한 질량이 된다. 문제는 위에서 언급하였듯이 실험자가 e_1 값을 너무 낮게 설정하게 되는 경우에는 덜 중요하다고 판단한 실험기준의 효율성이 e_1 보다 높게 나온다는 점이다. 다시 말해 실험자가 실험의 특성을 무시하고 e_1 의 값을 명시하도록 하기에는 무리가 따른다. 따라서 e_1 값이 너무 낮게 설정되어 있다 하더라도 다음과 같은 추가적인 제약조건을 만들어 이를 보완할 필요가 있다.

$$\max e(\xi, \xi_2^D) \quad \text{subject to} \quad e(\xi, \xi_1^D) \geq e_1 \text{ and } e(\xi, \xi_1^D) \geq e(\xi, \xi_2^D) \tag{3}$$

따라서 이러한 e_1 값이 낮게 설정되어 있다하더라도 실험은 이를 무시하고 두 번째 조건을 만족하는 범위 내에서 새로운 실험을 구성할 것이다. 이 예제에서는 $\xi(0)$ 의 값은 $e(\xi, \xi_1^D) = e(\xi, \xi_2^D)$ 이 일치할 때까지 계속 감소한다. 이때 효율성은 0.915523 이다. 이러한 실험디자인은 두 효율성이 일치하는 실험하는 점에서 발생되므로 이는 김영일 (1993)이 언급한 maxi-min 방법과 일치한다. 만약 e_1 값이 두 효율성이 일치하는 값 0.915523보다 높게 설정되어 있다면 두 번째 조건식은 무시된다. 왜냐하면 이는 자연스럽게 만족이 되기 때문이다. 예를 들면 e_1 이 0.92 라면 목적함수의 최대값은 식 (3)에 의해 0.905626이 된다. 두 번째 조건식 즉 중요한 실험기준의 효율성이 덜 중요한 실험기준의 효율성보다 더 높게 나와야 한다는 자연스러운 조건을 만족하게 되는 것이다.

따라서 본 논문에서는 김영일과 강명욱 (2002)이 제안한 순차적인 실험을 수정 보완한다.

(1) 실험자가 m 개의 실험기준을 가지고 있다고 가정한다. 실험자는 실험기준에 대해 우선순위를 부여할 수 있다. 편의상 $m=1$ 에 해당하는 실험기준을 실험자가 제일 우선하여 만족시켜야 하는 주 제약조건이라 하자. 즉, 높은 m 값은 우선순위가 낮은 제약조건이다.

(2) 첫 번째 주 제약조건을 만족하는 범위 내에서 두 번째 실험기준의 효율성인 목적함수를 최대화하는 실험을 아래와 같은 조건 하에서 구한다. 구해진 목적함수의 최대값을 $\max e(\xi, \xi_{\Phi_2}^*) = e_2^*$ 로 놓는다. 여기서 e_1 은 실험자가 임의로 설정한다.

$$\max e(\xi, \xi_{\Phi_2}^*) \text{ subject to } e(\xi, \xi_{\Phi_1}^*) \geq e_1 \text{ and } e(\xi, \xi_{\Phi_1}^*) \geq e(\xi, \xi_{\Phi_2}^*)$$

(3) 세 번째 실험기준을 목적함수로 하는 3 개의 제약조건식을 설정한다.

$$\max e(\xi, \xi_{\Phi_3}^*) \text{ subject to } e(\xi, \xi_{\Phi_1}^*) \geq e_1,\ e(\xi, \xi_{\Phi_2}^*) \geq e_2 \text{ and}$$
$$e(\xi, \xi_{\Phi_1}^*) \geq e(\xi, \xi_{\Phi_2}^*) \geq e(\xi, \xi_{\Phi_3}^*)$$

여기서 e_2 는 e_2^* 와 $e_2^* > e_2$ 의 관계를 가지게 함으로서 타당한 해가 존재토록 한다. 만약 e_2 가 낮게 설정된다 하더라도 역시 위에서 설명한 이유로 3번 째 조건식에서 이

를 조정할 것이다.

(4) 이와 같은 방법으로 모든 제약조건식을 소화한다.

다음의 두 예제는 3개의 실험기준을 가지고 있는 경우이다.

예제 2 Huang (1996)은 2차 다항회귀모형에 대해 3가지 실험기준을 설정하였다. 하나는 $D-$최적이며 두 번째는 실험영역 X에 대해 균일분포를 부여한 $I-$최적 (integrated optimal), 그리고 마지막인 실험기준은 $z=2$인 점에서의 외삽 (extrapolation)이다. 이 세 가지 최적에 대한 Φ_i는 다음과 같다.

$$\Phi_1(\xi)=-\ln\left[\frac{|M(\xi)|}{|M(\xi_1^*|}\right],\ \Phi_2(\xi)=\frac{tr\,M^{-1}(\xi)\,W}{tr\,M^{-1}(\xi_2^*)\,W},\ \ \Phi_3(\xi)=\frac{v(z,\xi)}{v(z,\xi_3^*)}$$

여기서 $v(x,\xi)=f^T(x)\,M(\xi)^{-1}f(x)$ 이며 $W=\int_X f(x)f^T(x)\,U(dx)$ 이다. 그리고 각 Φ_i에 대해 실험 ξ의 효율성은 각각 $e(\xi,\xi_{\Phi_i}^*)$로 표기한다. 실험자는 e_1에 대한 값으로 0.95를 설정하였다. 첫 두개의 최적기준으로 나온 실험은 $\xi(-1)=0.281$, $\xi(-0.042)=0.431, \xi(1)=0.288$ 이다. $D-$최적에 대한 효율성은 97.89%를 달성하였으며 $I-$최적에서도 같은 97.89%를 달성하였다. 다음 단계로서 e_1, e_2를 각각 95%, 90%로 조정한 다음 세 번째 실험기준을 고려한 경우, Φ_1,Φ_2,Φ_3-효율은 모두 96.13%로 같게 나타났다. 즉, maxi-min 방법에 의해 효율성이 구해 졌다고 본다. 이때 실험은 $\xi(-1)=0.220$, $\xi(-0.043)=0.435$, $\xi(1)=0.345$ 이다. 이는 e_1, e_2을 각각 95%, 90%로 설정한 다음 Φ_3-최적을 추구한 식 (2)에서의 실험보다 낫다. 물론 직접적인 해를 구하는 과정이 3차원의 효율성그림을 이용하는 번거로움도 있지만 Huang (1996)이 구한 실험, $\xi(-1)=0.199, \xi(-0.033)=0.411$, $\xi(1)=0.390$은 세 실험기준에 대한 효율성이 각각 95.1%, 92.4%, 98.3%로 나타나 제일 중요하지 않은 실험기준에 대한 효율성이 필요 없이 높게 나타났기 때문이다. 이와 같이 본 논문에서 제시한 순차적인 방법을 이용하면 비타당성(infeasibility)이 존재할지 모르는 위험성을 회피하면서 여러 실험기준의 중요성을 반영할 수 있다. 또한 제안된 알고리즘은 실험자가 실험의 특징을 잘못 파악하여 낮게 e_i를 설정하였다 하더라도 조건식에 의거 이러한 점을 실험에 반영한다.

예제 3 여기에서도 실험자는 이차회귀모형을 가정한다. 그러나 분산에 대한 가정은 등분산을 확신하지 못한다. 다만 가장 작은 분산의 값과 가장 큰 값의 분산의 값의 비는 고정된 값 $\gamma(\geq 1)$이며 실험영역에서는 분산은 증가한다는 사실이다. 이는 $var(\epsilon)=[(\gamma-1)x+(\gamma+1)]/2$로 표현할 수 있다. $\gamma=1$은 등분산을 의미한다. 실험자는 γ가 1, 3 혹은 5의 값을 가질 수 있다고 판단한다. 물론 $\gamma=1$이면 $D-$최적인 $\xi(\pm 1)=\xi(0)=1/3$이 $G-$최적이 되나 $\gamma>1$인 경우는 그렇지 않다. 이럴 경우 $G-$최적은 $\xi(+1)/\xi(-1)=\gamma$, $\xi(0)=1/3$, $\sum\xi(x_i)=1$ 이다. 여기서 x_i는 받힘점, -1, 0,

1을 의미한다. Φ_i, $i=1,2,3$ 을 각각 $\gamma=1,3,5$ 에 대해 $G-$최적이라 하자. 실험자는 실험기준의 중요성을 역순으로 명시하였다. 즉, 실험자는 $\gamma=5$ 인 경우에 대비해 효율성의 하한값, e_1 을 95%로 정하였다면 본 논문에서 제안한 방법에 의하면 두 번째 실험 기준의 효율성, 즉 $\gamma=3$ 에 대한 $G-$효율성은 최대 83.33%로 나온다. 따라서 e_2 는 이보다 작은 80%로 설정하면 제일 중요하지 않은 실험기준 $\gamma=1$ 에 대한 $G-$효율성은 65.97%로 나온다. 실험은 $\xi(-1)=0.139$, $\xi(1)=0.527$, $\xi(0)=0.333$ 으로 계산된다. 또한 $\xi(+1)/\xi(-1)$ = 3.8로 확인된다. 그러나 $\gamma=3$ 에 대한 $G-$효율성은 여전히 83.33%이 된다. 이는 첫 번째 실험기준에 대한 효율성을 95% 유지하려면 두 번째 실험기준이 83.33% 이하로 효율성이 설정되더라도 무시된다. 따라서 세 번째 실험기준이 추가된다 하더라도 더 이상 실험계획은 바뀌지 않는다. 이와 같은 현상은 γ 값이 커지면 커질수록 받힘점 +1에 대한 질량이 올라가는 성질에 기인한다. 즉, 이 예제의 경우는 $\gamma=5$ 에 대비한 95%의 $G-$효율성을 가지는 실험이 바로 다중제약 실험이 되는 것이다.

4. 결론

본 논문은 다중제약조건 실험문제에서 흔히 나타날 수 있는 비타당성의 문제를 순차적으로 해결하는데 있어 실험기준의 우선순위를 실험의 효율성으로서 표현되도록 추가적인 조건식을 제안한데 그 의의가 있다. 이는 Wong (1995, 1999)이 제시한 순차적인 방법을 개량한 것으로서 효율성 그림 도구를 이용하지 않아도 된다. 더욱이 우선순위를 조건식으로 부여함으로서 여러 실험기준의 중요성을 반영하였다. 비록 본 연구는 특정 모형에 대한 해석적인 해를 구한 것은 아니나 다중목적의 실험에 대한 필요성을 실험기준의 순위에 따라 적절히 반영하는 방법을 제시하였다고 본다. 이러한 방법은 선형모형뿐 아니라 비선형의 모형에도 적용시켜볼 수 있을 것이다.

참고문헌

[1] 김영일 (1993). D-와 이분산 G-최적을 중심으로 한 오차-로버스트적 실험계획법. 응용통계연구, 제 6권 2호, 303-309.

[2] 김영일, 강명욱 (2002). Multiple Constrained Optimal Experimental Design. 한국통계학회논문집, 제 9권 3호, 619-627.

[3] 염준근, 남기성 (2000). A Study on D-Optimal Design Using the Genetic Algorithm. 한국통계학회논문집, 제 7권 1호, 357-370.

[4] Atkinson, A.C. (1972). Planning experiments to detect inadequate regression models. *Biometrika*, Vol. 59, 275-293.

[5] Atkinson, A.C. and Bogacka, B. (1997). Compound $D-$and D_{s-}-optimum designs for determining the order of a chemical reaction. *Technometrics*, Vol. 39. 347-356.

[6] Atwood, C.L. (1969). Optimal and Efficient Designs of Experiments. *The Annals of Mathematical Statistics*, Vol. 40, 1570-1602.

[7] Box, G.E.P. and Draper, N.R. (1959). A basis for the selection of a response surface design. *Journal of the American Statistical Association*, Vol. 54, 622-653.

[8] Box, G.E.P. and Draper, N.R. (1975). Robust Design. *Biometrika*, Vol. 62, 347-352.

[9] Cook, R.D. and Wong, W.K. (1994). On the equivalence between constrained and compound optimal designs. *Journal of the American Statistical Association*, Vol. 89, 687-692.

[10] Cook, R.D. and Fedorov, V.V. (1995). Constrained optimization of experimental design with discussion. *Statistics*, Vol. 26, 129-178.

[11] Dette, H, and Franke, T. (2000). Constrained D_1- and $D-$optimal designs for polynomial regression. *Ruhr-Universitat Bochum technical paper.*

[12] Fedorov, V.V. (1972). *Theory of Optimal Experiemnts.* Academic Press, New York.

[13] Huang, Y.C. (1996). *Multiple-objective optimal designs.* Doctor of Public Health Dissertation, Department of Biostatistics, School of Public Health, UCLA.

[14] Huang, Y.C, and Wong, W.K. (1998a). Sequential construction of multiple -objective designs. *Biometrics*, Vol. 54, 188-197.

[15] Huang, Y.C. and Wong, W.K. (1998b). Multiple-objective designs. *Journal of Biopharmaceutical Statistis*, Vol. 8, 635-643.

[16] Imhof, L. and Wong, W.K. (1999). A graphical method for finding maximin designs. *Biometrics*, Vol. 54, 188-197.

[17] Läuter, E. (1974). Experimental planning in a class of models. *Mathematishe Operationsforshung und Statistik*, Vol. 5, 673-708.

[18] Läuter, E. (1976). Optimal multipurpose designs for regression models. *Mathmatische Operationsforsh und Statistics*, Vol. 7, 51-68.

[19] Lee, C.M.S. (1987). Constrained optimal designs for regression models. *Communications in Statistics, Part A-theory and Methods*, Vol. 16, 765-783.

[20] Lee, C.M.S. (1998). Constrained optimal designs. *Journal of Statistical Planning and Inference*, Vol. 18, 377-389.

[21] Pukelsheim, F. (1993). *Optimal Design of Experiments.* Wiley, New York.

[22] Silvey, S.D. (1980). *Optimal Design.* Chapman Hall, New York.

[23] Stigler, S.M. (1971). Optimal experimental design for polynomial regression. *Journal of the American Statistical Association*, Vol. 66, 311-318.

[24] Studden, W.J. (1982). Some robust-type $D-$optimal designs in polynomial

regression. *Journal of the American Statistical Association*, Vol. 66, 311-318.

[25] Wong, W.K. (1995). A graphical approach for constructing constrained $D-$and $L-$optimal designs using efficiency plots. *Journal of Statistical Simulation and Computations*, Vol. 53, 143-152.

[26] Wong, W.K. (1999). Recent advances in multiple-objective design strategies. *Statistica Neerlandica*, Vol. 53, 257-276.

[27] Zhu, W., Ahn, H. and Wong, W.K. (1998). Multiple-objective optimal designs for the logit model. *Communications in Statistics-Theory and Methods*, Vol. 27, 1581-1592.

[Received July 2005, Accepted November 2005]