

Some Examples of Constrained Optimal Experimental Design for Nonlinear Models

Youngil Kim^{a,1} · Dae-Heung Jang^b · Seongbaek Yi^b

^aSchool of Business and Economics, Chung-Ang University

^bDepartment of Statistics, Pukyong National University

(Received October 11, 2014; Revised November 28, 2014; Accepted November 28, 2014)

Abstract

Despite the fact that the optimal design for nonlinear model depends on the unknown quantity of parameter to estimate basically, its popularity is growing in bio and engineering statistics area since all those models in the area are virtually nonlinear. In this paper we have dealt with the case when the researcher has multiple objectives in experimentation, decision among the competing models, protection against the departure from the assumed model, and the conflicting interests among design criteria. To tackle these issues we attempted several new approaches which are taking advantage of the easiness of constrained optimal design. Several nonlinear models were tested.

Keywords: Nonlinear model, constrained optimal design, multi-objective design strategy.

1. 서론

비선형회귀모형에서 반응변수 $n \times 1$ 벡터 $y(x)$ 는 식 (1.1)과 같이 x 의 비선형 함수로 연결되어 있다.

$$y = \eta(x, \theta) + \epsilon. \quad (1.1)$$

여기서 $\theta^T = (\theta_0, \theta_1, \dots, \theta_p)$ 는 추정하여야 할 모수벡터이다. 실험설계 ξ 는 실험영역 Ω 내의 유한 k 개의 점 $x_i \in \Omega, i = 1, 2, \dots, k$ 에 대한 확률질량함수 $\xi(x_i), i = 1, 2, \dots, k$ 로 기술되는 데 전체 실험에 쓰이는 관측값의 개수는 n 이고 각 받힘점(support point) x_i 에서 반복되는 관측값의 개수가 $n_i, i = 1, 2, \dots, k$ 라 한다면 개개의 $\xi(x_i), i = 1, 2, \dots, k$ 는 n_i/n 의 형태를 가지며 $\sum_i^k \xi(x_i) = 1.0$ 의 제약조건을 갖는다. $n \times \xi(x_i)$ 가 정수라는 제약조건이 주어지면 정확실험설계(exact design)라 하고 제약조건이 없는 경우를 근사실험설계(approximate design)라 한다. 본 연구에서는 근사실험계획만 고려한다. 식 (1.1)에서 오차항 ϵ 의 구조로서 기댓값 0 그리고, 일반성의 손실 없이 $\sigma^2 = 1$ 인 비상관의 정규분포를 가정한다면 실험 ξ 에 대한 Fisher의 정보행렬(information matrix)은 식 (1.2)와 같이 주어진다.

$$M(\xi, \theta) = \int_{\Omega} \frac{\partial \eta(x, \theta)}{\partial \theta} \frac{\partial \eta(x, \theta)}{\partial \theta^T} d\xi(\mathbf{x}). \quad (1.2)$$

¹Corresponding author: Professor, School of Business and Economics, Chung-Ang University, 84 Heukseok-ro, Dongjak-gu, Seoul 156-756, Korea. E-mail: yik01@cau.ac.kr

여기서 $\frac{\partial \eta(x, \theta)}{\partial \theta}$ 는 미분벡터(derivative vector)라 하며 차원은 $p \times 1$ 이다. 이러한 정보행렬은 받힘점의 개수가 유한이므로 식 (1.3)과 같이 표기가 가능하다.

$$M(\xi, \theta) = V^T W V. \quad (1.3)$$

여기서 $V = \frac{\partial \eta}{\partial \theta}$ 는 i -번째 행이 받힘점 x_i 에서의 반응함수의 경사도(gradient) $\frac{\partial \eta(x_i, \theta)}{\partial \theta}$ 인 $k \times p$ 자코비안(Jacobian)행렬이다. 또한 W 는 대각원소가 x_1, x_2, \dots, x_k 에 부여된 질량함수 $\xi(x_i)$ 인 $k \times k$ 대각행렬이다. 또한 받힘점 x 에서의 예측분산 함수는 식 (1.4)와 같이 표기된다.

$$d(x, \xi, \theta) = \frac{\partial \eta(x, \theta)}{\partial \theta^T} M^{-1}(\xi, \theta) \frac{\partial \eta(x, \theta)}{\partial \theta}. \quad (1.4)$$

여기서 $\frac{\partial \eta(x, \theta)}{\partial \theta}$ 의 차원은 $p \times 1$ 이므로 식 (1.4)는 특정한 x 에서 하나의 값을 제공한다. 모수 θ 에 해당하는 최대우도추정량(MLE)의 공분산의 일차(first order)근사가 $M^{-1}(\xi, \theta)$ 이므로 식 (1.4)의 분산 역시 반응변수의 예측분산에 대한 일차근사이다.

기존 비선형모형에서 많이 선호하는 실험기준으로서 식 (1.2)의 정보행렬의 행렬식(determinant)을 최대화하는 실험기준 국지(local) $D(\theta)$ -최적과 식 (1.4)의 최대값을 최소화하는 국지 $G(\theta)$ -최적 실험기준이 있다. 그러나 이 두 실험기준은 동격이며 최적여부는 식 (1.4)의 분산의 최대값이 p 와 일치하는지의 여부로 판단되기 때문에 오차항의 이분산구조와 같이 특별한 경우가 아니면 두 최적실험기준을 동시에 고려할 필요는 없다. 비선형모형인 경우 최적실험은 미지의 모수에 의해 결정되므로 본 연구에서는 편의상 국지란 용어는 생략한다.

자코비안 행렬 V 를 $V = [V_1 | V_2]$ 로 분할표기하면 V_1 은 $k \times p_1$ 인 $V = \frac{\partial \eta}{\partial \theta_1}$ 이 되고 V_2 는 $k \times p_2$ 인 $V = \frac{\partial \eta}{\partial \theta_2}$ 가 된다. 여기서 θ_1^T 는, $\theta = [\theta_1 | \theta_2]$ 로 표기할 때의 $p_1 \times 1$ 모수벡터이며 θ_2^T 는 $p_2 \times 1$ 모수벡터이다. 또한 정보행렬 $M(\xi, \theta)$ 는 식 (1.5)와 같은 표기가 가능하다.

$$M(\xi, \theta) = \begin{bmatrix} M_{11} & M_{12} \\ M_{21} & M_{22} \end{bmatrix} = \begin{bmatrix} V_1^T W V_1 & V_1^T W V_2 \\ V_2^T W V_1 & V_2^T W V_2 \end{bmatrix}. \quad (1.5)$$

θ_2 에 해당되는 정보행렬은 $M_{22} - M_{21} M_{11}^+ M_{21}^T$ 이다. 여기서 M_{11}^+ 는 M_{11} 의 일반화 역행렬(generalized inverse)이다. 기존문헌에서는 $M_{22} - M_{21} M_{11}^+ M_{21}^T$ 의 행렬식을 최대화하는 실험기준을 $D_s(\theta_2)$ -최적이라 한다. 참고로 $M_{22} - M_{21} M_{11}^+ M_{21}^T$ 의 행렬식은 $\frac{|M(\xi, \theta)|}{|M_{11}(\xi, \theta)|}$ 과 일치한다. 이러한 $D_s(\theta_2)$ -최적은 아래 식 (1.6)에서 정의된 $d_s(x, \xi, \theta)$ 의 최대값인 p_2 와의 일치여부로 확인이 가능하다.

$$d_s(x, \xi, \theta) = \frac{\partial \eta(x, \theta)}{\partial \theta^T} M^{-1}(\xi, \theta) \frac{\partial \eta(x, \theta)}{\partial \theta} - \frac{\partial \eta(x, \theta)}{\partial \theta_1^T} M_{11}^{-1}(\xi, \theta) \frac{\partial \eta(x, \theta)}{\partial \theta_1}. \quad (1.6)$$

이와 같은 동격성에 관한 내용은 Kiefer (1961)를 참조하기 바란다. 그리고 일반적인 선형 및 비선형모형에 관한 Silvey (1980)이나 Atkinson 과 Donev (1992), Berger와 Wong (2009), Fedorov와 Leonov (2013) 등의 책을 참조하기 바란다.

기본적으로 $D(\theta)$ -최적은 모형의 모수에 대한 일반적인 분산인 행렬식에 관한 기준이며 $D_s(\theta_2)$ -최적은 부분 모수에 대한 추정에 관한 기준이 되는 것이다. 따라서 이 두 최적기준의 선택은 실험자가 실험의 특성에 따라 결정을 하여야 한다.

이러한 실험기준을 이용하여 최적 실험을 구하는 문제는 주어진 모형이 참이라고 가정하여야 한다. 그러나 이는 현실적으로 가정하기 힘든 상황이며 또한 실험 기준 역시 실험자의 선호도에 따라 복수의 실험기준을 설정하여야 하는 문제가 발생된다. 이를 해결하기 위하여 기존문헌에서는 기본적인 몇 가지

방법이 존재하는데 이에 대한 설명 전에 임의의 실험 ξ 가 최적실험에 대하여 가지는 효율을 먼저 정의할 필요가 있다.

$D(\theta)$ -효율: $D(\theta)$ -최적에 대한 실험 ξ 의 $D(\theta)$ -효율 $\phi_\xi(D(\theta))$ 은 다음과 같이 정의된다.

$$\phi_\xi(D(\theta)) = \left\{ \frac{\det(M(\xi, \theta))}{\det(M(\xi_D^*, \theta))} \right\}^{1/p}.$$

여기서 ξ_D^* 는 $\eta(x, \theta)$ 에 대한 $D(\theta)$ -최적실험이며 \det 는 행렬식을 의미한다.

$G(\theta)$ -효율: $G(\theta)$ -최적에 대한 실험 ξ 의 $G(\theta)$ -효율 $\phi_\xi(G(\theta))$ 은 다음과 같이 정의된다.

$$\phi_\xi(G(\theta)) = \frac{p}{\max_{x \in \Omega} d(x, \xi, \theta)}.$$

여기서 $G(\theta)$ -최적의 경우 분자의 값은 동격성에 의하여 p 가 된다.

$D_s(\theta_2)$ -효율: $D_s(\theta_2)$ -최적에 대한 실험 ξ 의 $D_s(\theta_2)$ -효율 $\phi_\xi(D_s(\theta_2))$ 은 다음과 같이 정의된다.

$$\phi_\xi(D_s(\theta_2)) = \left\{ \frac{\det(M(\xi, \theta))/\det(M_{11}(\xi, \theta))}{\det(M(\xi_{D_s}^*, \theta))/\det(M_{11}(\xi_{D_s}^*, \theta))} \right\}^{1/p_2}.$$

여기서 $\xi_{D_s}^*$ 는 $\eta(x, \theta)$ 에 대한 $D_s(\theta_2)$ -최적실험이며 \det 는 행렬식을 의미한다. 그리고 편의상 $D_s(\theta_2)$ -최적은 θ_2 에 대해서만 정의한다.

$G(\theta_2)$ -효율: $G(\theta_2)$ -최적에 대한 실험 ξ 의 $G(\theta_2)$ -효율 $\phi_\xi(G(\theta_2))$ 은 다음과 같이 정의된다.

$$\phi_\xi(G(\theta_2)) = \frac{p_2}{\max_{x \in \Omega} d(x, \xi, \theta)}.$$

여기서 $G(\theta_2)$ -최적의 경우 분자의 값은 동격성에 의하여 p_2 가 된다.

위와 같은 효율은 어느 실험기준 ψ 이든지 정의가 될 수 있다. 간단히 표현하면 실험 ξ 의 ψ -효율이란 실험기준 ψ 에 대하여 실험 ξ 가 ψ -최적실험 ξ_ψ^* 에 대하여 가지는 효율 $\phi_\xi(\xi_\psi^*)$ 이라 정의가 된다.

모형이 불확실하거나 실험기준이 복수인 경우에는 이를 해결하기 위하여 다음과 같은 기본적인 방법이 기준문헌에 존재 한다.

방법 1 복합실험기준: Läuter (1974)가 제안한 실험 기준으로서 k 개의 실험기준 $\psi_i, i = 1, 2, \dots, k$ 가 존재하는 경우 다음과 같은 선형 결합을 최대화하는 실험기준 ξ 이다.

$$\max_{\xi} \alpha_1 \phi_\xi(\xi_{\psi_1}^*) + \alpha_2 \phi_\xi(\xi_{\psi_2}^*) + \dots + \alpha_k \phi_\xi(\xi_{\psi_k}^*), 0 \leq \alpha_i \leq 1, \sum_{i=1}^k \alpha_i = 1.0.$$

방법 2 제약조건실험기준: Stigler 등 (1971)이 제안한 실험기준으로서 k 개의 실험기준 $\psi_i, i = 1, 2, \dots, k$ 가 존재하는 경우 이 중 하나를 제외한 나머지를 제약조건식으로 두고 다른 하나를 목적함수에 위치하여 목적함수를 최대화하는 실험기준 ξ 이다.

$$\max_{\xi} \phi_\xi(\xi_{\psi_j}^*) \text{ subject to } \phi_\xi(\xi_{\psi_i}^*) \geq c_i, i = 1, 2, \dots, i \neq j, \dots, k.$$

그러나 $k > 2$ 인 경우에 제약조건식을 모두 만족하지 못하는 비타당(infeasible)해의 가능성 때문에 이러한 방법은 특수한 경우를 제외하고는 $k = 2$ 에 국한 하는 경우가 대부분이다.

방법 3 최소-최대(maxi-min) 실험기준: k 개의 실험기준 $\psi_i, i = 1, 2, \dots, k$ 가 존재하는 경우 실험 ξ 가 k 개의 실험 기준에 대하여 가지는 효율을 $\phi_\xi(\xi_{\psi_i}^*), i = 1, 2, \dots, k$ 라 한다면 최소-최대 실험은 다음과 같은 실험 ξ 를 의미한다.

$$\max_{\xi} \min_i \phi_\xi(\xi_{\psi_i}^*), i = 1, 2, \dots, k.$$

이 실험은 매우 보수적인 방법으로 Imhof와 Wong (2000)에 의하면 $k = 2$ 인 경우 $\phi_\xi(\xi_{\psi_1}^*) = \phi_\xi(\xi_{\psi_2}^*)$ 가 되어 버린다.

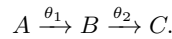
기존문헌에서의 비선형문제는 복합기준으로 방법 1을 다루고 있다. 그러나 방법 1은 각 실험기준에 대해서 가중치 α_i 를 설정하는 주관적인 문제가 대두가 되는바 본 연구에서는 방법 2와 3에 국한하여 내용을 전개할 것이다. 왜냐하면 실험기준에 같은 가중치 $\alpha_i = 1/k$ 가 부여되더라도 실제 각 실험기준에 대한 효율 $\phi_\xi(\xi_{\psi_i}^*)$ 효율, $i = 1, 2, \dots, k$ 은 동일한 값을 가진다고 보장을 하지 못하기 때문이다. 오히려 ψ -효율의 틀에서 판단하는 방법 2와 방법 3이 실험자에게는 더 직관적이라 판단된다. 본 연구에서는 주로 방법 2인 제약조건실험 위주로 예제를 제시할 것이다.

제 2절에서는 비선형을 염두에 둔 실험을 실시할 경우 나타나는 불확실성을 시나리오 별로 설명하고 앞에서 언급한 방법들이 비선형모형에 적용되는 예제를 언급하여 본다. 일부 예제에서는 Kim과 Lim (2007)의 제안에 따라 방법 2의 제약조건의 실험과 방법 3의 최대최소실험의 혼합형의 실험기준을 비선형 모형에 시험적으로 적용하여 볼 것이다. 그리고 제 3절에서는 비선형모형에 대한 향후 과제를 제안하여 본다.

2. 시나리오별 예제

2.1. 제약조건실험기준

많은 비선형 모형은 화학 동역학(chemical kinetics)연구에서 나오는데 다음과 같은 연속적인 일차 반응(first order reaction)모형을 생각하여 본다.



시간 x 에서 성분(component) B 의 농도(concentration)는 Hill과 Hunter (1974)에 의하면 다음 식 (2.1)과 같은 모형으로 설정된다.

$$\eta(x, \theta) = [B] = \frac{\theta_1}{\theta_1 - \theta_2} (\exp(-\theta_2 x) - \exp(-\theta_1 x)), x \geq 0. \quad (2.1)$$

이 모형은 Hill과 Hunter (1974)이후로 O'Brien (2005) 등에 의하여 다루어 졌다. 식 (2.1) 모형의 (θ_1, θ_2) 의 초기값으로 $(\theta_1^0, \theta_2^0) = (0.2, 0.7)$ 으로 한다면 $D(\theta)$ -최적의 실험은 $\xi_D : \xi(1.5987) = \xi(5.9055) = 0.5$ 로 나온다. 그러나 $D_s(\theta_2)$ -최적은 $x = 6.7619$ 에 모든 질량을 부여하는 실험이 나오므로 이런 실험으로 나온 결과는 주어진 모형을 적합시키는데 바람직하지 않다. 따라서 주어진 모형의 모수에 대한 추정과 더불어 다목적의 실험을 실시할 필요가 있다. 따라서 다음과 같은 제약 조건 실험을 제안한다.

$$\max_{\xi} \frac{|M(\xi, \theta)|}{|M_{11}(\xi, \theta)|} \text{ subject to } \left\{ \frac{|M(\xi, \theta)|}{|M(\xi_D^*, \theta)|} \right\}^{1/p} \geq c.$$

여기서 c 는 0과 1사이의 값으로 실험자가 명시하여야 하는 상수이다.

이러한 실험은 모형의 전체 모수 추정에 대한 정보량을 보전해 주는 조건으로 θ_2 에 대한 추정을 하는 그런 실험이 된다. $c = 0.9$ 로 할 경우 $D_s(\theta_2)$ -최적과 달리 2개의 받힘점을 가진다.

$$\xi_{c=0.9} : \xi(1.5088) = 0.2873, \xi(6.1764) = 0.7127.$$

그리고 $D(\theta)$ -최적과 $D_s(\theta_2)$ -최적에 대한 효율을 구하여 보면 다음과 같다.

$$\phi_{\xi_c}(D(\theta)) = 0.9, \phi_{\xi_c}(D_s(\theta_2)) = 0.7300.$$

제약조건의 조건을 충족하면서 목적함수를 최대화하는 전형적인 두 개의 목적함수의 균형을 맞추는 실험이 된다. 향후 모든 예제에서의 제약조건 실험은 ξ_c 이라 하였다.

2.2. 세 가지 이상의 목적을 가지고 있는 제약조건 실험

성분(component) A 가 모두 다 성분 B 로 변환(conversion) 되지 못하는 경우 화학공학에서는 식 (2.1)의 모형보다 식 (2.2)의 모형이 선호된다.

$$\eta(x, \theta) = \theta_3(\exp(-\theta_2 x) - \exp(-\theta_1 x)), x \geq 0. \tag{2.2}$$

위의 모형은 Fresen (1984)에 의해 자료가 분석이 되었는데 총 18개의 자료 값에서 다음과 최소제곱추정값을 구하였다.

$$\theta_1^0 = 0.05884, \theta_2^0 = 4.298, \theta_3^0 = 21.80.$$

이를 모수의 초기값으로 하여 $D(\theta)$ -최적을 구하여 보면 다음과 같은 $D(\theta)$ -최적실험이 나온다. 이 모형 역시 다른 많은 비선형 모형처럼 받힘점은 3개로 등질량을 가진다.

$$\xi_D(0.2288) = \xi_D(1.3886) = \xi_D(18.417) = 1/3.$$

또한 Atkinson과 Donev (1992)에 의하면 실험자는 모형의 모수 추정뿐 아니라 몇 가지 다른 관심 사항을 가진다고 한다. 첫째로 함수 밑의 영역(area under the curve; AUC), 두 번째로 최대 농축까지 걸리는 시간(time to maximum concentration; TMC, x_{max}) 그리고 마지막으로 최대농축량(maximum concentration; MC)이다. 각각의 관심사항은 $\theta_1, \theta_2, \theta_3$ 의 함수로 적어보면 다음과 같다.

$$AUC : \theta_3 \left(\frac{1}{\theta_1} - \frac{1}{\theta_2} \right) = g_1(\theta), x_{max} : \frac{\log \theta_2 - \log \theta_1}{\theta_2 - \theta_1} = g_2(\theta), MC : \eta(x_{max}, \theta) = g_3(\theta)$$

선형모형이며 $g(\theta)$ 가 θ 의 선형의 형태인 경우에는 $g(\theta)$ 의 분산은 적절한 상수 c 에 대하여 $var(g(\theta)) = var(c^T \beta) = c^T M(\theta)^{-1} c$ 로 나타나는데 비선형 모형이며 $g(\theta)$ 가 θ 의 비선형의 형태인 경우에는 테일러 전개에 의하여 c 에 들어가는 각 원소 c_i 는 $c_i = \frac{\partial g(\theta)}{\partial \theta_i}$ 로 표기된다. 즉, $var(g(\theta)) \approx c^T M(\theta)^{-1} c$ 으로 표기된다. 자세한 내용은 Silvey (1980)나 Atkinson과 Donev (1992)의 책을 참조하기 바란다.

여기서는 $D(\theta)$ -최적 뿐 아니라 3가지 $g_i(\theta), i = 1, 2, 3$ 의 분산 값을 최소화 하는 부수적인 목적함수가 주어져 있다. 기존문헌에서는 이런 기준을 c -최적이란 한다. 참고로 이 3가지 목적에 대한 개별적인

c -최적은 다음과 같다. $g_1(\theta)$ 와 $g_2(\theta)$ 인 경우는 두 받힘점에서 그리고 $g_3(\theta)$ 인 경우는 한 받힘점에서 만 질량을 가진다.

$$\begin{aligned} \text{AUC} : \xi(0.2327) &= 0.0135, \xi(17.633) = 0.9865, \\ \text{xmax} : \xi(0.1793) &= 0.6062, \xi(3.5671) = 0.3938, \\ \text{MC} : \xi(1.0122) &= 1.0. \end{aligned}$$

주어진 모형의 모수추정에 필요한 $D(\theta)$ -최적이 각 $g_i(\theta)$ 에 대하여 가지는 효율은 각각 $\phi_{\xi_D}(\xi_{g_1}^*) = 0.3431$, $\phi_{\xi_D}(\xi_{g_2}^*) = 0.6602$, $\phi_{\xi_D}(\xi_{g_3}^*) = 0.3910$ 로 나타나는데 효율의 양극화는 세 가지 $g_i(\theta)$, $i = 1, 2, 3$ 의 c -최적 실험기준간의 서로 상충된 결과로 해석된다. 여기서 $\xi_{g_i}^*$, $i = 1, 2, 3$ 은 각 $g_i(\theta)$, $i = 1, 2, 3$ 에 대한 c -최적이다. 따라서 Atkinson과 Donev (1992)는 식 (2.3)과 같은 복합적인 실험기준을 제안하였다.

$$\Psi\{M(\xi, \theta)\} = \sum_{j=1}^3 \log \left\{ c_j^T(\theta) M^{-1}(\xi, \theta) c_j(\theta) \right\}. \quad (2.3)$$

식 (2.3)를 최소화하는 실험을 AD -최적이라 하자. 다음이 AD -실험기준에 의한 결과이다.

$$\xi_{AD} : \xi(0.2704) = 0.2991, \xi(1.3784) = 0.2029, \xi(15.753) = 0.4908.$$

그러나 불행히도 이러한 실험이 각 c -최적에 대해 가지는 실험은 효율은 그리 높지 않다. 역시 서로의 c -최적의 실험기준이 가지는 상충되는 목적 때문으로 파악된다.

$$\phi_{\xi_{AD}}(\xi_{g_1}^*) = 0.4953, \phi_{\xi_{AD}}(\xi_{g_2}^*) = 0.7810, \phi_{\xi_{AD}}(\xi_{g_3}^*) = 0.6837.$$

그러나 이러한 실험의 D -최적에 대하여 가지는 효율은 0.9167로 상대적으로 높다. 각 c -최적의 실험은 비교적 용이하게 D -효율을 충족하다는 의미이다. 따라서 본 연구에서는 다음과 같은 제약실험을 제안한다.

$$\max_{\xi} \left\{ \phi_{\xi_c}(\xi_{g_1}^*), \phi_{\xi_c}(\xi_{g_2}^*), \phi_{\xi_c}(\xi_{g_3}^*) \right\} \text{ subject to } \left\{ \frac{|M(\xi, \theta)|}{|M(\xi_D^*, \theta)|} \right\}^{1/p} \geq c.$$

여기서 c 는 0과 1사이의 값으로 실험자가 명시하여야 하는 상수이다. 여기서는 임의적으로 c 의 값을 0.9로 설정하여 그 차이점을 본다. 이러한 실험은 3개의 받힘점을 가진다.

$c = 0.9$,

$$\begin{aligned} \xi_c : \xi(0.2435) &= 0.2201, \xi(1.4204) = 0.2201, \xi(18.3138) = 0.5598, \\ \phi_{\xi_c}(\xi_{g_1}^*) &= 0.5760, \phi_{\xi_c}(\xi_{g_2}^*) = 0.6750, \phi_{\xi_c}(\xi_{g_3}^*) = 0.6105, \phi_{\xi_c}(\xi_D^*) = 0.9. \end{aligned}$$

그리고 $D(\theta)$ -최적과 $g(\theta)$ 에 대한 c -최적에 대한 효율을 구하여 보았다. AD -최적과 엇비슷한 결과가 나왔으나 c 의 값을 통하여 $D(\theta)$ -최적과 세 가지 c -최적간의 균형을 실험자가 확인하여 볼 수 있을 것이다. 이러한 실험은 Kim과 Lim (2007)의 방법을 취한 것인데 다목적의 실험인 경우 매우 효과적이다.

2.3. 외삽에 관한 제약조건 실험: Michaelis-Menten 모형

식 (2.4)의 모형은 생명과학분야에서 효소역학(enzyme-kinetics)을 연구하는데 있어 인기 있는 모형이다. Raaijmakers (1997)에 의하여 이 모형을 이용한 통계적인 분석이 많이 이루어지고 있음을 알 수 있다.

$$\eta(x, \theta) = \frac{\theta_1 x}{\theta_2 + x}. \quad (2.4)$$

여기서 θ_1 은 반응률(reaction rate)이 도달할 수 있는 최대값이고 θ_2 는 Michaelis-Menten 상수이다. θ_2 에 따라 이 함수의 성장률이 정해진다. 여기서는 $\theta_1 = \theta_2 = 1$ 로 가정한다.

다항회귀모형인 경우 외삽문제는 오래 전부터 문헌에서 다루어진 바 있지만 비선형 모형인 경우는 외삽의 문제는 잘 소개가 되어 있지 않다. Berger 등 (2009)의 책 210 페이지에서도 예제가 나와 있듯이 외삽문제는 비선형의 문제라 하더라도 실험자가 관심을 가질 수 있다. 식 (2.4)에서 실험영역을 $[0, 10]$ 으로 하는 경우 외삽점 $x_0 = 12$ 에서의 분산의 값을 최소화 하는 식 (2.5)와 같은 외삽-최적실험 ξ_{ext} 은

$$\min_{\xi} d(x_0, \xi, \theta) = \frac{\partial \eta(x_0, \theta)}{\partial \theta^T} M^{-1}(\xi, \theta) \frac{\partial \eta(x_0, \theta)}{\partial \theta}. \quad (2.5)$$

$\xi_{ext} : \xi(0.6040) = 0.0582, \xi(10) = 0.9418$ 과 같이 나타난다. 실험영역의 우측의 극단점에 많은 질량이 배치됨을 알 수 있다. 실험자는 $x_0 = 12$ 이외에 다른 외삽점에도 관심을 가지는 경우가 있다. $x_0 = 20$ 인 경우 $\xi_{ext} : \xi(0.6040) = 0.1536, \xi(10) = 0.8464$ 로 나타난다.

외삽뿐 아니라 모형에 관한 모수 추정을 요구하는 실험인 경우에는 제약조건 실험도 타당할 것이다. 왜냐하면 외삽인 경우 모형에 대한 가정이 변하지 않아야 한다는 가정이 있는데 이는 매우 불안정한 가정이다. 그리고 외삽-최적은 D -최적에 대하여 가지는 효율은 매우 작다. 참고로 D -최적은 $\xi(0.8333) = \xi(10) = 0.5$ 이며 $x_0 = 10$ 인 경우 외삽최적은 $\phi_{\xi_{ext}}(\xi_D^*) = 0.4357$ 의 낮은 효율을 제공한다. 이런 경우 Kim과 Jang (2014)의 방법을 적용하여 모형추정과 외삽-최적에 관한 균형을 맞추는 것이 바람직하다.

실험영역을 벗어나는 2개의 외삽점을 고려하여 보자. $x_{01} = 12, x_{02} = 20$ 인 경우 제약조건 실험을 다음과 같이 설정하여 보자.

$$\max_{\xi} \min \left\{ \phi_{\xi}(\xi_{ext}^{x_{01}}), \phi_{\xi}(\xi_{ext}^{x_{02}}) \right\} \text{ subject to } \phi_{\xi}(\xi_D^*) \geq c.$$

여기서 $\phi_{\xi}(\xi_{ext}^{x_0})$ 는 x_0 에서의 외삽-최적이다. 이러한 실험은 전체 모형에 대한 $D(\theta)$ -효율을 c 이상 유지하면서 다른 두 외삽점에 대한 최소효율을 끌어 올리는 실험기준이 된다. 그리고 받힘점의 위치는 0.6040과 0.8333사이에 재정립되었다. 외삽점이 $x_0 = 20$ 인 경우 받힘점 10에 부여된 질량은 줄어드는 것을 생각하면 당연한 결과이다.

$$\xi_{c=0.9} : \xi(0.8000) = 0.2825, \xi(10) = 0.7175.$$

그리고 $D(\theta)$ -최적과 각 외삽점에서의 외삽최적에 대한 효율을 다음과 같이 구하여 보았다.

$$\phi_{\xi_c}(\xi_D^*) = 0.9, \phi_{\xi_c}(\xi_{ext}^{x_{01}}) = 0.7957, \phi_{\xi_c}(\xi_{ext}^{x_{02}}) = 0.9102$$

2.4. 모형 차별에 관한 제약조건 실험: Michaelis-Menten 모형의 변형

다음과 같은 2개의 선형 모형이 주어져 있다고 보자.

$$\begin{aligned}\eta(x, \theta) &= \theta_0 + \theta_1 x + \theta_2 x^2, \\ \eta(x, \theta) &= \theta_0 + \theta_1 e^x + \theta_2 e^{-x}.\end{aligned}$$

그러나 실험자는 어느 모형이 더 참의 모형인지 확신을 하지 못하는 경우다. 실험영역이 $[-1, 1]$ 라면 각 개별 모형에 대한 D -최적은 똑같이 $\xi(\pm 1) = \xi(0) = 1/3$ 이다. 이러한 D -최적은 결과적으로 위에 언급한 두 개의 모형 중 어느 하나가 참의 모형으로 밝혀진다 하더라도 효율 면에서 보면 전혀 문제가 없어 보인다. 그러나 아직 어느 모형이 참인지를 모르는 상태에서는 좋은 선택은 아니다. 따라서 모형차별차원에서의 실험은 다르게 고려하여야 한다. 따라서 기존문헌에서는 식 (2.6)과 같이

$$\eta(x, \theta) = \theta_0 + \theta_1 x + \theta_2 x^2 + \theta_3 e^x + \theta_4 e^{-x}. \quad (2.6)$$

두 모형을 결합하여 $D_s(\theta_1, \theta_2)$ 와 $D_s(\theta_3, \theta_4)$ 의 선형결합 형태의 $-$ 최적을 고려하곤 한다. Atkinson과 Donev (1992)는 다음과 같은 실험기준을 제시하였다.

$$\log(|M|/|M_{11}|)/s_1 + \log(|M|/|M_{22}|)/s_2$$

여기서 M 은 식 (2.6)의 전체 모형이고 M_{11} 은 θ_1 과 θ_2 에 대한 M 의 부분행렬이고 M_{22} 는 θ_3 과 θ_4 에 대한 M 의 부분행렬이다. 그리고 s_1 과 s_2 는 해당하는 부분모수의 개수가 되므로 각각 2이다. 이러한 기준은 선형결합형태로 본 부분모수의 행렬식을 최대화하는 실험기준이다. 편의상 이런 복합적인 실험기준을 AD 기준이라 하자. 이는 1 절에서 언급한 방법 1의 변형이다. 그러면 다음과 같은 균형적(symmetrical)인 AD -최적 실험이 얻어진다.

$$\xi_{AD} : \xi(\pm 1) = 0.140, \xi(\pm 0.649) = 0.259, \xi(0) = 0.202.$$

Lopez-Fidalgo 등 (2008)은 식 (2.7)과 같은 비선형 모형을 제안하였다. 2.3절에서 언급한 Michaelis-Menten 모형의 변형적인 형태이다.

$$\eta(x, \theta) = \frac{\theta_1 x}{\theta_2 + x} + Fx. \quad (2.7)$$

$F = 0$ 이면 위의 모형은 Michaelis-Menten 모형이 된다. 실질적인 문제는 실험자가 참의 모형으로서 이탈되는 F 의 크기에 대한 확신이 없는 경우이다. Michaelis-Menten 모형이 올바른지 아니면 위 식 (2.7)의 모형이 올바른지에 대한 실험을 실시하려고 한다. 즉, 모수 θ_1 과 θ_2 에 대한 추정과 F 에 대한 추정간의 균형을 이루는 문제가 생긴다.

이 문제 역시 Atkinson과 Donev (1992)의 방법을 적용하여 보았는데 Atkinson과 Donev (1992)의 방법은 두 실험기준에 가중치를 균일하게 두는 관계로 각 $D_s(\theta)$ 와 $D_s(F)$ 에 대한 효율은 다음과 같이 유사하게 나옴을 미리 짐작하여 볼 수 있다.

$$\begin{aligned}\xi(10) &= 0.2307, \xi(3.4480) = 0.4324, \xi(0.4802) = 0.3369, \\ \phi_{\xi_{AD}}(D_s(\theta)) &= 0.9138, \phi_{\xi_{AD}}(D_s(F)) = 0.9292.\end{aligned}$$

또한 제약 조건실험을 다음과 같이 적용하여 보자.

$$\max_{\xi} \frac{|M(\xi, \theta)|}{|M_{22}(\xi, \theta)|} \text{ subject to } \left\{ \frac{|M(\xi, \theta)|}{|M_{11}(\xi, \theta)|} \right\}^{1/s_1} \geq c.$$

여기서 c 는 0과 1사이의 상수로 실험자가 명시하여야 한다. 현재 예에서는 c 의 값을 편의상 0.8와 0.9로 한다.

$c = 0.8$,

$$\xi_c : \xi(10) = 0.2946, \xi(3.7334) = 0.4463, \xi(0.4216) = 0.2590, \\ \phi_{\xi_c}(D_s(\theta)) = 0.8, \phi_{\xi_c}(D_s(F)) = 0.9932.$$

$c = 0.9$,

$$\xi_c : \xi(10) = 0.2400, \xi(3.461) = 0.4345, \xi(0.4722) = 0.3254, \\ \phi_{\xi_c}(D_s(\theta)) = 0.9, \phi_{\xi_c}(D_s(F)) = 0.9421.$$

결과를 비교하면 c 에 따라 AD 기준과 차이가 뚜렷하다. c 값을 낮출수록 $\phi_{\xi_c}(D_s(F))$ 값은 매우 높아진다. 그러나 두 실험기준의 균형을 이루는 것은 별 차이가 없어 보인다. 다만 Atkinson과 Donev (1992)의 방법과 비교하면 상수를 유기적으로 설정하는 것이 실험자의 입장에서는 더 유연성을 가져다 줄 것이다. 실험자는 항상 주요하다고 생각되는 최적기준에 그 효율성을 먼저 설정을 하고 싶어 하기 때문이다.

3. 토의

기본적으로 비선형 모형이라도 최적실험이론은 선형이론과 동일하게 적용이 된다. 다만 모수의 초기값에 의존하는 특징이 있기 때문에 선형모형과 달리 다루어야 한다. 그러나 선형모형과 마찬가지로 비선형 모형 역시 모형이 가지고 있는 모수추정에 국한하여 실험이 실시되지 않는다. 본 연구에서는 비선형모형에 대한 실험을 실시하는 경우 발생하는 부수적인 이슈를 유형별로 알아보았는데 본 연구에서는 기존문헌에서 잘 다루지 않았던 제약조건실험을 비선형 모형에 적용시켜보았다. 경우에 따라서 기본적인 제약조건의 변형을 통하여 실험자가 가지고 있는 복수의 실험기준 등을 수용하는 하이브리드형의 실험기준을 마련하였다. 이슈별로 제약조건의 가능성을 예제를 통하여 그 가능성을 종합적으로 열어 둔 것에 대하여 본 연구의 의의가 있다고 본다.

References

- Atkinson, A. C. and Donev, A. N. (1992). *Optimal Experimental Designs*, Oxford Press, Oxford.
- Berger, M. P. F. and Wong, W. K. (2009). *An Introduction to Optimal Designs for Social and Biomedical Research*, Wiley, New York.
- Dette, H., Kiss, C., Bevanda, M. and Bretz, F. (2010). Optimal designs for the emax, log-linear and exponential models, *Biometrika*, **97**, 513-518.
- Fedorov, V. V. and Leonov, A. N. (2013). *Optimal Design for Nonlinear Response Models*, Chapman & Hall/CRC, reprint in kindle edition.

- Fresen, J. (1984). Aspects of bioavailability studies, M.Sci. Thesis, *Department of mathematical Statistics, University of Capetown*.
- Hill, W. J. and Hunter, W. G. (1974). Design of experiments for subsets of parameters, *Technometrics*, **16**, 425-434.
- Imhof, L. and Wong, W. K. (2000). A graphical method for finding maximin designs, *Biometrics*, **56**, 113-117.
- Kiefer, J. (1961). Optimum designs in regression problem II, *Annals of mathematical Statistics*, **32**, 298-325.
- Kim, Y. I. and Lim, Y. B. (2007). Hybrid approach when multiple objectives exist, *The Korean Communications in Statistics*, **14**, 531-540.
- Kim, Y. I. and Jang, D. H. (2014). Some criteria for optimal experimental design at multiple extrapolation points, *The Korean Journal of Applied Statistics*, **27**, 693-703.
- Kitos, C. P. (2001). Design aspects for the Michaelis-Menten model, *Biometrical Letters*, **38**, 53-66.
- Läuter, E. (1974). Experimental planning in a class of models, *Mathematische Operationsforschung und Statistik*, **5**, 673-708.
- Lopez-Fidalgo, J., Thommasi, C., and Trandafir, P. C. (2008). Optimal designs for discriminating between some extensions of the Michaelis-Menten model, *Journal of Statistical Planning and Inference*, **138**, 3797-3804.
- O'Brien, T. (2005). Designing for parameter subsets in gaussian nonlinear regression models, *Journal of Data Science*, **3**, 179-197.
- Raaijmakers, J. G. W. (1997). Statistical analysis of the Michaelis-Menten equation, *Biometrics*, **43**, 780-793.
- Silvey, S. D. (1980). *Optimum Design : An Introduction to the Theory for Parameter Estimation*, Chapman and Hall, London.
- Stigler, S. M. (1971). Optimal experimental design for polynomial regression, *Journal of the American Statistical Association*, **66**, 311-318.

비선형모형에 적용한 제약조건 최적실험의 예제들

김영일^{a,1}, 장대흥^b, 이성백^b

^a중앙대학교 경영학부, ^b부경대학교 통계학과

(2014년 10월 11일 접수, 2014년 11월 28일 수정, 2014년 11월 28일 채택)

요약

비선형모형에 대한 최적실험은 주어지는 모수의 초기값에 의존하는 특징이 있음에도 불구하고 비선형모형에 대한 최적실험은 비선형모형이 주류인 생물이나 화학공학 통계분야에서는 끊임없이 연구되어 왔다. 본 연구에서는 실험자가 다수의 실험목적 을 가지고 있는 환경에서 상충되는 목적간의 균형을 맞추는 관점에서 비선형 모형에 대한 최적실험 기준을 제약 실험으로 살펴보았다. 문헌에서 존재하는 기존의 방법 뿐 아니라 새로운 융합기준도 적용하였는데 시나리오 별로 그 적용 형태를 몇 가지 비선형 모형을 통하여 알아보았다.

주요어: 비선형모형, 제약 최적실험, 다목적 실험전략.

¹교신저자: (156-756) 서울특별시 동작구 흑석동 84, 중앙대학교 경영경제대학 경영학부, 교수.
E-mail: yik01@cau.ac.kr