Journal of Korean
Data & Information Science Society
2005, Vol. 16, No. 2, pp. 255 ~261

Loss Function Approach to Multiresponse Robust Design[1)]

Duk-Joon Chang[2)] · Yong-Man Kwon[3)]

Abstract

Many designed experiments require the simultaneous optimization of multiple responses. In this paper, we propose how to simultaneously optimize multiple responses for robust design when data are collected from a combined array. The proposed method is based on the quadratic loss function. An example is illustrated to show the proposed method.

Keywords: Combined array, Multiple responses, Quadratic loss function, Robust design, Simultaneous optimization

1. 서론

많은 실험계획에서는 반응변수가 한 개인 경우 설계인자들의 최적조건을 찾고 있으나 실제 실무에서는 동시에 고려하여야 할 반응변수가 여러 개인 경우 즉, 다중 반응(multiple responses)인 경우가 대부분이다. 다중 반응에서 설계인자의 최적조건을 찾는 방안은 Derringer와 Suich(1980), Khuri와 Conlon(1981) 그리고 Vining(1998)등에 의하여 이루어졌다. Derringer와 Suich(1980)는 기대함수를 이용한 다중 반응 동시 최적화 방안을 제시하였다. 이 방법은 사용하기 편리한 점이 있으나 반응변수간의 분산-공분산을 고려하지 않았다. 그리고 Khuri와 Conlon(1981)은 거리함수(distance function)을 이용한 다중 반응 동시 최적화 방안을 제시하였다. 이 방법은 반응변수간의 분산-공분산을 고려하였으나 단순히 목표치와의 거리만을 고려할 뿐 실제 현장에서 필요한 경제적인 측면을 전혀 고려하지 않았다. 또한 Vining(1998)은 손실함수(loss function)를 이용한 다중 반응 동시 최적화 방안을 제시하였다. 그러나 앞서 제시된 모든 방안들은 잡음인자의 영향에 의한 품질변동을 고려하지 않은 동시 최적화

1) This study was supported by Chosun University Research Funds, 2004.

1) First Author : Professor, Department of Statistics, Changwon National University, Changwon, 641-773, Korea.

3) Corresponding Author : Associate Professor, Department of Computer and Statistics, Chosun University, Gwangju, 501-759, Korea.
E-mail : ymkwon@chosun.ac.kr

방안이다.

이전의 실험계획법에서는 단지 품질특성치의 평균을 개선하는데 초점을 두고 최적 조건을 찾는 경향이 있었으나 다구찌 품질공학(Taguchi [1987])에서는 품질특성의 평균뿐만 아니라 변동(분산 혹은 표준편차)을 가능한 줄이는 것을 목적으로 한다는 점에서 차이가 있다. 다구찌 파라미터 설계에서 직교배열표를 이용한 교차배열(product array)은 제어인자와 잡음인자의 모든 교호작용을 고려함으로써 실험수가 지나치게 많을 뿐 아니라 축차실험을 고려하지 않는 등 많은 단점을 가지고 있다. 따라서 실험수를 줄일 수 있을 뿐 아니라 기존의 잘 정립된 실험계획법 이론을 이용한 대체방안이 연구되고 있다. 그 중에서 통합배열(combined array)접근법이 Welch, Yu, Kang, 와 Sacks(1990)에 의해 처음으로 제안되었다. 통합배열이란 잡음인자를 제어인자와 같이 하나의 실험배열에 배치하는 실험방법을 말한다.

Vining(1998)은 제어인자만으로 이루어진 실험계획에서 손실함수를 이용한 다중 반응 동시 최적화 방안을 제안하였으나 본 논문에서는 손실함수를 이용하되 통합배열에 의한 실험배치에서 로버스트 설계를 위한 다중 반응 동시 최적화 방안을 제안하고자 한다.

2. 다중 반응 로버스트 설계에서 손실함수를 이용한 동시 최적화

반응변수 y는 제어인자들($\boldsymbol{x}$)과 잡음인자들($\boldsymbol{z}$)에 의해 값이 정해진다고 가정하자. 제어인자들과 잡음인자들을 각 각 $\boldsymbol{x}=(x_1, x_2, \cdots, x_l)'$와 $\boldsymbol{z}=(z_1, z_2, \cdots, z_m)'$로 나타내기로 하자. 실험수는 N이고 반응변수들의 개수를 r 이라고 하면 i 번째 반응변수의 이차회귀모형은 다음과 같다.

$$y_i(\boldsymbol{x}, \boldsymbol{z}) = \beta_{i0} + \boldsymbol{x}'\ \boldsymbol{\beta}_i + \boldsymbol{x}'B_i\boldsymbol{x} + \boldsymbol{z}'R_i\boldsymbol{z} + \boldsymbol{z}'\ \boldsymbol{\gamma}_i + \boldsymbol{z}'D_i\boldsymbol{x} + \varepsilon_i,\ \ i=1,2,\cdots,r \tag{2.1}$$

여기서 $\boldsymbol{\beta}_i$는 $l\times 1$, $\boldsymbol{\gamma}_i$는 $m\times 1$, $B_i' = B_i$는 $l\times l$, $R_i' = R_i$는 $m\times m$, D_i는 $m\times l$인 모집단의 회귀계수들의 벡터 혹은 행렬이고 ε_i는 i번째 반응변수에서의 실험오차이다. 식 (2.1)을 행렬로 나타내면 다음과 같다.

$$\boldsymbol{y}_i = X\ \boldsymbol{\theta}_i + \ \boldsymbol{\varepsilon}_i, \qquad i=1,2,\cdots,r \tag{2.2}$$

여기서 $\boldsymbol{y}_i$는 i 번째 반응변수에서의 측정치의 $N\times 1$ 벡터이고, X는 $N\times p$ 행렬인 계획행렬(design matrix)이고 계수는 p, i번째 반응변수에서의 $\boldsymbol{\theta}_i$ 는 $p\times 1$인 모집단의 회귀계수들의 벡터이고 $\boldsymbol{\varepsilon}_i$는 실험오차의 벡터이다. 식 (2.2)와 관련된 가정으로

$$E(\ \boldsymbol{\varepsilon}_i) = 0,\ Var(\ \boldsymbol{\varepsilon}_i) = \sigma_{ii}I_N,\ Cov(\ \boldsymbol{\varepsilon}_i,\ \boldsymbol{\varepsilon}_j) = \sigma_{ij}I_N \qquad i,j=1,2,\cdots,r,\quad i\neq j$$

이 성립된다고 하자. Σ는 $r\times r$행렬이고 (i,j)번째 원소는 σ_{ij}이다. Σ 의 불편추정량은 다음과 같다.

$$\widehat{\Sigma} = Y'\,[I_N - X(X'X)^{-1}X']\,Y\,/\,(N-p) \tag{2.3}$$

여기서 $Y=(\,y_1,\ y_2,\cdots,\ y_r)$이고 I_N은 모든 원소가 1인 $N\times N$ 행렬이다. $\widehat{\theta}_i$의 분산-공분산행렬은 다음과 같다.

$$Var(\,\widehat{\theta}\,) = (X'X)^{-1}\Sigma.$$

i번째 품질특성의 추정식은 다음과 같다.

$$\widehat{y}_i(x,z) = g'(x,z)\,\widehat{\theta}_i, \qquad i=1,2,\cdots,r \tag{2.4}$$

여기서 $g'(x,z)$는 (x,z)점에서 계산되어진 X행렬의 행과 같은 형태의 벡터이다. 식 (2.4)로부터 다음이 성립된다.

$$\Sigma_{\widehat{y}}(x,z) = Var[\,\widehat{y}\,(x,z)] = g'(x,z)(X'X)^{-1}g(x,z)\Sigma, \tag{2.5}$$

여기서 $\widehat{y}(x,z)=(\,\widehat{y}_1(x,z),\ \widehat{y}_2(x,z),\cdots,\ \widehat{y}_r(x,z))'$은 점 (x,z)에서 추정된 품질특성의 벡터이다. 식(2.5)의 불편추정량은 다음과 같다.

$$\widehat{\Sigma}_{\widehat{y}}(x,z) = \widehat{Var}[\,\widehat{y}(x,z)] = g'(x,z)(X'X)^{-1}g(x,z)\,\widehat{\Sigma}\,. \tag{2.6}$$

식 (2.4)는 다시 다음과 같은 적합된 이차회귀모형으로 나타낼 수 있다.

$$\widehat{y}_i(x,z) = b_{i0} + x'\,b_i + x'\,\widehat{B}_i x + z'\,\widehat{R}_i z + z'\,r_i + z'\,\widehat{D}_i x,\ \ i=1,2,\cdots,r.$$

잡음인자들 z는 실험할 때는 제어할 수 있으나 실제 현장에서는 제어할 수 없는 확률변수들이다. z에 관한 정보가 없는 경우 흥미영역 R_z에서 균일분포(uniform distribution)를 가정하자. 잡음인자 z의 범위에서의 어떤 x에서 평균을 $\widehat{m}_i(x)$로 나타내기로 하고 이를 " i번째 추정된 평균모형식"이라 하면 다음과 같다.

$$\widehat{m}_i(x) = E[\,\widehat{y}(x,z)] = \int_{R_z}\widehat{y}_i(x,z)p(z)dz, \qquad i=1,2,\cdots,r$$

여기서 $p(z)$는 z의 확률밀도함수이며 z는 $R_z(-1 \le z \le 1)$에서 균일분포를 한다. Box와 Jones(1992)는 평균모형식은 다음과 같다고 하였다.

$$\widehat{m}_i(x) = b_{i0} + x'\ b_i + x'\ \widehat{B}_i x + (tr\ \widehat{R}_i)/3, \quad i = 1, 2, \cdots, r \tag{2.7}$$

여기서 $tr(\widehat{R}_i)$는 행렬 $\widehat{R}_i$의 대각선 원소들의 합이다.

반응변수가 한 개인 경우 $y(x, z)$의 추정식은 $\hat{y}(x, z)$이고 목표치를 τ라 할 때 손실함수는 다음과 같다.

$$L = c[\hat{y}(x, z) - \tau]^2,$$

여기서 c는 비용계수(cost coefficient)이다. 다중 반응변수인 경우 즉, 반응변수가 여러 개인 경우 $\mathbf{y}(x, z)$의 추정식은 $\widehat{\mathbf{y}}(x, z)$이고 목표치를 $\boldsymbol{\tau} = (\tau_1, \tau_2, \cdots, \tau_r)'$라 할 때 손실함수는 다음과 같다.

$$L = [\widehat{\mathbf{y}}(x, z) - \boldsymbol{\tau}]' C [\widehat{\mathbf{y}}(x, z) - \boldsymbol{\tau}], \tag{2.8}$$

여기서 C는 비용행렬(cost matrix)인데 양정치(positive definite)행렬이다. c_{ii}는 각 각의 추정된 반응변수들이 대응되는 목표치와 편차 그리고 $c_{ij}(i \ne j)$는 추정된 반응식 $\hat{y}_i$와 $\hat{y}_j$ 대응되는 목표치와 편차에 따라 발생하는 경제적인 손실비용을 나타낸다. 식 (2.8)의 기대손실함수(expected loss function)를 구하면 다음과 같다(Graybill(1976)을 보시오).

$$E[L] = [E[\widehat{\mathbf{y}}(x, z)] - \boldsymbol{\tau}]' C [E[\widehat{\mathbf{y}}(x, z)] - \boldsymbol{\tau}] + tr[C\boldsymbol{\Sigma}_{\widehat{\mathbf{y}}}(x, z)] \tag{2.9}$$

만약 식 (2.7)으로부터 $\widehat{\mathbf{m}}(x) = (\widehat{m}_1(x), \widehat{m}_2(x), \cdots, \widehat{m}_r(x))'$이라하면 식 (2.9)는 다음과 같이 나타낼 수가 있다.

$$E[L] = [\widehat{\mathbf{m}}(x) - \boldsymbol{\tau}]' C [[\widehat{\mathbf{m}}(x) - \boldsymbol{\tau}] + E[tr[C\boldsymbol{\Sigma}_{\widehat{\mathbf{y}}}(x, z)]] \tag{2.10}$$

식 (2.10)을 최소로 하는 제어인자의 최적조건을 구하기 위하여 $Var[\widehat{\mathbf{y}}(x, z)]$의 불편추정량 식 (2.6)을 이용하여 추정하면 다음과 같다.

$$\widehat{E}[L] = [\widehat{\mathbf{m}}(x) - \boldsymbol{\tau}]' C [[\widehat{\mathbf{m}}(x) - \boldsymbol{\tau}] + E[tr[C\ \widehat{\boldsymbol{\Sigma}}_{\widehat{\mathbf{y}}}(x, z)]] \tag{2.11}$$

이제 우리는 추정된 기대손실함수 식 (2.11)을 이용하여 통합배열에서 반응변수가 여러 개인 경우 로버스트 설계를 하기 위한 동시최적화 방안으로 기대손실함수를 최소로 하는 제어인자의 최적조건 즉, 다중 반응 동시최적화 방안을 다음과 같이 제안한다.

$$SL = \underset{\boldsymbol{x} \in R_x}{Min} \ \widehat{E}[L] \tag{2.12}$$

여기서 R_x는 제어인자 $\boldsymbol{x}$의 흥미영역이다. 위에서 제안한 다중 반응 동시최적화 방안 SL은 윈도우용 MATLAB 6.0의 도구 박스에 있는 “const” 함수를 사용하여 최적조건을 구할 수 있다.

3. 예제

통합배열에서 제어인자 x_1, x_2 그리고 잡음인자 z를 <표 1>과 같이 실험배치 한다. <표 1>은 중심합성계획에 의한 실험배치이며 y_1은 망목특성 그리고 y_2는 망대특성인 가상적인 자료이다.

<표 1> 통합배열에 의한 실험자료

실험수	x_1	x_2	z	y_1	y_2
1	-1	-1	-1	80.6	81.4
2	-1	-1	1	74.9	95.9
3	-1	1	-1	83.1	105.0
4	-1	1	1	71.2	103.0
5	1	-1	-1	66.8	74.0
6	1	-1	1	74.2	76.8
7	1	1	-1	38.1	81.2
8	1	1	1	36.8	76.9
9	-1.41	0	0	80.9	100.0
10	-1.41	0	0	42.4	50.5
11	0	-1.41	0	73.4	71.2
12	0	-1.41	0	45.0	101.0
13	0	0	0	77.4	102.0
14	0	0	0	74.6	104.0

<표 1>로부터 제어인자와 잡음인자간의 교호작용을 고려한 추정된 이차회귀모형은 각 각 다음과 같다.

$$\hat{y}_1(x, z) = 76.00 - 12.37x_1 - 8.96x_2 - 7.22x_1^2 - 8.45x_2^2 \\ - 8.11x_1x_2 + 5.38z^2 - 1.44z + 2.96x_1z - 1.86x_2z. \tag{3.1}$$

$$\hat{y}_2(x, z) = 103.00 - 12.21x_1 + 6.68x_2 - 13.96x_1^2 - 8.50x_2^2 \\ - 2.93x_1x_2 + 6.23z^2 - 1.38z + 1.75x_1z - 2.95x_2z. \tag{3.2}$$

식 (3.1)과 식 (3.2)로부터 식 (2.7)을 이용하면, 다음과 같은 평균모형식을 얻는다.

$$\widehat{m_1}(x) = 77.79 - 12.37x_1 - 8.96x_2 - 7.22x_1^2 - 8.45x_2^2 - 8.11x_1x_2.$$
$$\widehat{m_2}(x) = 105.08 - 12.21x_1 + 6.68x_2 - 13.96x_1^2 - 8.50x_2^2 - 2.93x_1x_2$$

홍미영역 R_x는 $-1 \le x_1, x_2 \le 1$이며, y_1은 망목특성이므로 $\widehat{m_1}(x)$의 목표치 τ_1을 특정값인 75.00으로 두고 y_2는 망대특성이므로 R_x 상에서 $\widehat{m_2}(x)$의 최대값 109.65를 목표치 τ_2로 둔다. 식 (2.3)에 따라 추정된 분산-공분산행렬은 다음과 같다.

$$\hat{\Sigma} = \begin{bmatrix} 5.4367 & 2.3361 \\ 2.3361 & 68.2638 \end{bmatrix}.$$

비용행렬은 두 반응변수의 손실비용을 고려하여

$$C = \begin{bmatrix} 49 & 21 \\ 21 & 9 \end{bmatrix}$$

와 같이 결정하였다. 비용계수를 정할 때 개별적인 손실만을 고려한다면, 소비자 허용한계점 $\tau_i \pm \Delta_i$에서 손실이 A_i이라면

$$c_{ii} = A_i / \Delta_i^2$$

이다. 비용행렬 C는 반응변수가 목표값을 벗어났을 때 발생하는 손실비용뿐만 아니라 분석자가 사전 정보를 이용하여 반응변수간의 손실비용의 크기 차이 그리고 소비자, 생산자, 기술적인 문제 등을 고려하여 결정함으로써 각 각 반응변수의 가중치 역할을 한다. 분석자는 충분한 사전 정보를 이용하여 비용행렬의 변화에 따른 제어인자의 최적조건의 변화를 살펴볼 필요가 있다.

MATLAB 6.0의 도구 박스에 있는 “const” 함수를 사용하여 최적조건을 구할 때 국소(local) 최소값 일 수도 있으므로 초기치를 달리하여 비교해 보아야 한다. 동시최적화 방안인 식(2.12)를 적용하면 제어인자의 최적조건은 다음과 같다.

$$x_1 = 0.03, \qquad x_2 = 0.29.$$

제어인자의 최적조건에서 $\widehat{m}_1(x)$, $\widehat{m}_2(x)$는 각 각 74.92와 106.68이고 기대손실비용은 100.44 이다. 소비자 손실은 89.68이고 반응함수 추정에 따른 오차는 소비자 손실에 12%에 해당하는 10.76이다.

4. 결론

본 논문에서는 고려하여야할 반응변수가 여러 개인 경우에 비용함수를 이용한 다중반응 동시최적화 방안을 제안하였다. 실험배치는 로버스트 설계에서 통합배열을 이용함으로써 다구찌가 제안한 교차배열보다 실험수를 크게 줄일 수 있으며 기존의 잘 정립된 실험계획법 이론을 활용할 수 있어 다구찌 파라미터 설계에서 단점으로 지적된 제반 문제점을 해결할 수 있는 대체방안이 될 수 있을 것이다. 이 논문에서 제안한 동시최적화방안은 반응변수간의 분산-공분산을 고려하였으며 비용행렬을 이용하여 경제적인 손실비용 등을 고려하였다.

참고문헌

1. Box, G. E. P. and Jones, S. P. (1992). Designing Products that Are Robust to the Environment, *Total Quality Management*, Vol. 3, 265-282.
2. Derringer, G. and Suich, R. (1980). Simultaneous Optimization of Several Response Variables, *Journal of Quality Technology,* Vol. 12, pp. 214-219.
3. Graybill, F. A. (1976). *Theory and Application of the Linear Model*, Duxbury Press, Boston, MA.
4. Khuri, A. I. and Conlon, M. (1981). Simultaneous Optimization of Multiple Responses Represented by Polynomial Functions, *Technometrics*, Vol. 23, 363-375.
5. Taguchi, G. (1987). *System of Experimental Design*: Engineering Methods to Optimize Quality and Minimize Cost, White Plains, NY: UNIPUB / Kraus International.
6. Vining, G. G. (1998). A Compromise Approach to Multiresponse Optimization, *Journal of Quality Technology*, Vol. 30, 309-313.
7. Welch, W. J., Yu, T. K., Kang, S. M. and Sacks, J. (1990). Computer Experiments for Quality Control by Parameter Design, *Journal of Quality Technology*, Vol. 22, 15-22.

[2005년 3월 접수, 2005년 4월 채택]