

로버스트 파라미터 설계에서 인자분석을 이용한 동시 최적화 방안에 관한 연구

권용만¹ · 홍연웅²

요 약

본 논문에서는 로버스트 파라미터(robust parameter) 설계에서 다특성(multiple quality characteristics)인 경우 제어인자의 동시 최적화 조건을 찾는 방안으로 인자분석(factor analysis)에 의한 최적화 방안을 제시한다. 또한 하나의 사례를 들어 제안한 방안과 기존의 방안을 비교 연구하였다.

주요용어 : 로버스트 파라미터 설계, 다특성, 동시 최적화, 인자분석.

1. 서론

다구찌의 로버스트 파라미터 설계에서 직교배열표를 이용한 교차배열(product array)은 내측배열(inner array)에 있는 제어인자(혹은 설계인자)와 외측배열(outer array)에 있는 잡음인자(혹은 비제어인자)의 모든 교호작용을 고려한 실험배치를 하여 신호 대 잡음비(signal-to-noise ratio ; SN비)를 이용한 자료분석을 하였다. 파라미터 설계에서는 품질특성치(혹은 반응변수)를 분석하는데 있어서 평균과 변동을 하나로 묶은 성능측도(performance measure)인 SN비를 사용한다. SN비는 모든 특성치에서 크면 클수록 좋게 만들어진 성능측도이다. 교차배열에서 잡음인자는 품질특성치의 품질변동을 유발시키는 역할을 함으로써 변동에 둔감하면서 동시에 품질특성치의 평균을 목표치에 접근하는 제어인자의 최적조건을 찾을 수 있는 로버스트 설계를 가능하게 한다. 전통적인 실험계획법에서는 주로 품질특성치의 평균을 개선하는데 초점을 두고 최적조건을 찾는 경향이 있었으나 다구찌 품질공학(Taguchi 1986, 1987)에서는 품질특성의 평균뿐만 아니라 변동을 가능한 줄이는 것을 목적으로 한다는 점에서 차이가 있다.

실험계획에서 주로 품질특성이 한 개인 경우 제어인자들의 최적조건을 찾고 있으나 실제 실무에서는 어떤 경우의 실험계획에 관한 문제이든 동시에 고려하여야 할 품질특성치가 여러 개인 경우 즉, 다특성인 경우가 대부분이다. 다특성에서 제어인자의 최적조건을 찾는 방안은 Derringer와 Suich(1980) 그리고 Khuri와 Conlon(1981)등에 의하여 연구되었다. 그러나 앞서 연구된 방안들은 품질변동을 고려하지 않은 동시 최적화 방안이다.

¹501-759 광주광역시 동구 서석동 375, 조선대학교 전산통계학과 부교수. ymkwon@chosun.ac.kr

²750-711 경북 영주시 풍기읍 교촌동 1, 동양대학교 인터넷산업공학부 교수

인자분석(factor analysis)은 상호연관된 여러 품질특성(다특성 혹은 다변량 확률변수)들 간의 내부적 상호의존관계를 그 저변에 내재하는 가설적 개념인 소수 몇 개의 공통인자(common factor, 혹은 줄여서 인자(factor))를 통해 재현, 해석하고자 하는데 그 목적이 있다. 따라서 이는 변수중심의 통계적 기법으로서 인자를 통해 차원축약을 기하고, 이에 기초하여 전체 정보를 요약 내지 단순화시키는 과정을 거친다(김기영과 전명식 2000).

본 논문에서는 인자분석(factor analysis)을 이용한 성능측도를 이용하여 로버스트 설계에서 다특성인 경우 동시 최적화하는 두 가지 방안을 제시하고 기존의 다특성 실험의 분석결과(박석현 1993, p. 255-272)와 비교하고자 한다.

2. 로버스트 파라미터 설계에서 동시 최적화 방안

2.1 인자분석을 이용한 성능측도

어떤 시스템이나 제품에서 고려하여야 할 p 개 종류의 품질특성이 있다면 그들에 대응하여 품질 특성별로 p 개 종류의 SN비 SN_1, SN_2, \dots, SN_p 가 계산된다. 여기서 주어진 실험배치에서 실험 번호의 개수를 n 이라 하자. 총 계산된 SN비는 $p \times n$ 개가 존재할 것이다. p 개 종류의 품질특성별로 n 개만큼 존재하는 이들의 평균과 표준편차를 고려하여 표준화한 p 개의 SN비를 $SN_1^*, SN_2^*, \dots, SN_p^*$ 하자. 예를 들어 p 번째 품질특성의 SN비인 SN_p 에서 n 개의 SN비를 $SN_{p1}, SN_{p2}, \dots, SN_{pn}$ 이라고, 이들 n 개의 평균과 표준편차를 각각 $\overline{SN_p}$ 과 $s.d(SN_p)$ 라 할 때 표준화된 SN_p^* 는

$$SN_p^* = (SN_p - \overline{SN_p}) / s.d(SN_p)$$

이 된다.

인자분석을 하는데 있어서 인자(factor)를 추출하는 방법으로 주성분인자법(principal component factoring), 주축인자법(principal axis factoring), 최대우도법(maximum likelihood method) 등 많은 방법이 있으나 이 논문에서는 인자모형을 적합하는 방법으로 주성분분석으로부터 유도되는 주성분인자법을 사용하기로 한다. 따라서 우리는 먼저 계산된 p 개의 SN비를 사용하여 주성분분석을 다음과 같은 순서로 한다.

- (1) SN_1, SN_2, \dots, SN_p 을 표준화하여 $SN_1^*, SN_2^*, \dots, SN_p^*$ 라 하자.
- (2) $SN_1^*, SN_2^*, \dots, SN_p^*$ 의 상관행렬(correlation matrix) R 을 구한다.
- (3) R 의 고유값(eigen values)과 고유벡터(eigen vectors)를 구한다.
- (4) 생성된 p 개의 인자(혹은, 여기서는 주성분)중에 전체 고유값의 총합에서 차지하는 비율 큰

인자 첫 k 개를 택한다.

(5) 인자부하행렬 즉, 성분행렬 L_1 을 구한다.

한편, 주성분인자법을 사용하여 인자분석에서 보유되는 인자의 갯수를 결정하기 위한 몇 가지 기준을 보면 대체로 다음과 같다(허명희와 양경숙, 2001).

(a) 고유값이 1이상인 인자만 고려한다.(Kaiser의 규칙 혹은 0.7이상(Jolliffe, 1986))

(b) 고유값의 누적점유율 즉, 인자의 공헌도에 일정한 기준을 적용한다.(예를 들어 누적점유율 $\geq 80\%$)

(c) 산비탈 그림(scree plot)에서 비탈이 급격하게 꺾이는 점부터 주성분을 무시한다.

앞서 기준에 의하여 인자의 갯수가 정해진 상태하에서 인자가 가지는 개념의 해석을 수월하게 하기 위하여 단순한 구조로 변경시키기 위한 방법으로 인자의 회전을 고려하여 보자. 인자회전방법으로 직교회전방법과 사교회전(oblique rotation)등 많은 방법이 있으나 이 논문에서는 직교회전방법에서 대표적인 배리맥스 회전(varimax rotation, Kaiser 1958)방법을 사용하기로 하자. 배리맥스 방법은 인자부하행렬(loading matrix of factor)의 열내 제곱 부하값들의 최대화 한다. 그럼으로써 인자부하행렬의 열 요소가 0 또는 ± 1 에 가까운 값으로 단순하게 한다. 배리맥스 방법을 사용하여 인자회전이후의 인자부하행렬 즉, 성분행렬을 L_2 라 하자.

인자모형은 표준화된 변수 $SN_1^*, SN_2^*, \dots, SN_p^*$ 를 인자 F_1, F_2, \dots, F_k 로 표현하였다. 이에 대한 역수로 F_1, F_2, \dots, F_k 를 $SN_1^*, SN_2^*, \dots, SN_p^*$ 로 표현할 수 있는데 이를 인자점수(factor score)라고 한다. 회귀방법에 의한 인자점수의 변환식은 다음과 같다(Johnson과 Wichem, 1992, 9.5 절).

$$F = L_2' R^{-1} S \quad (2.1)$$

여기서, $F = (F_1, F_2, \dots, F_k)$, $S = (SN_1^*, SN_2^*, \dots, SN_p^*)$, L_2 는 배리맥스 방법을 사용하여 인자회전이후의 인자부하행렬이고 R 은 $SN_1^*, SN_2^*, \dots, SN_p^*$ 의 상관행렬이다.

인자회전에 의한 첫 k 개의 회전인자 모형식을 F_1, F_2, \dots, F_k 라 할 때, 첫 k 개의 인자에 의한 성능측도를 다음과 같이 제안하기로 한다.

$$TF = \sum_{i=1}^k \rho_i \sigma_i F_i \quad (2.2)$$

여기서, $\rho_i = 1$ 인 경우는 인자 F_i 의 계수가 양의 경향, $\rho_i = -1$ 인 경우는 음의 경향, $\rho_i = 0$ 인 경우는 모호한 경향이다. σ_i 는 F_i 계수의 제곱합이 선택된 첫 k 개 인자 계수들의 제곱합에서 차지하는 비율이다. σ_i 는 첫 k 개 인자에서 회전인자 F_i 의 공헌도이다. $\rho_i = 1$ 인 경우

는 p 개 변수에 대한 계수에서 큰 양의 계수(positive coefficient) 몇 개가 큰 음의 계수(negative coefficient)의 절대값 보다 월등히 큰 경우이며, 반대인 경우는 $\rho_i = -1$, 그리고 분명하지 않는 경우는 $\rho_i = 0$ 로 한다. 이렇게 하는 이유는 SN비는 크면 클수록 더 좋은 성능측도이기 때문에 p 개 품질특성치의 SN비를 가지고 인자분석을 이용하고 있는 성능측도 TF 도 또한 클수록 더 좋은 성능측도이기 때문이다.

2.2 기존의 동시 최적화 방안

로버스트 파라미터 설계에서 다특성인 경우 동시최적화하는 기존의 방안은 다음과 같은 순서로 한다.

- (1) 모든 품질특성의 SN을 구한다.
- (2) 품질특성별로 수준별 SN의 합계와 분산분석표를 작성한다.
- (3) 모든 품질특성에 대하여 유의한 설계인자를 수준별로 정리한 인자효과에 대한 종합결과표를 작성한다.

3. 사례연구

로버스트 파라미터 설계에서 다특성 동시 최적화 방안으로 인자분석에 의한 방안과 기존의 방안을 통한 제어인자의 최적수준을 정하는 방안등 두 가지 방안을 1959년도에 일본 국유철도의 송입공장에서 있었던 사례(Taguchi (1987, 189-191))를 들어 비교 연구하고자 한다. 사례로 든 실험에서는 두 개의 철판을 전기용접할 때 작업성, 외관, X선 검사, 용접부위의 기계적 인장강도 그리고 신도등 5개의 품질특성치에 대한 설계인자의 최적조건을 찾으려고 한다. 내측배열에 $L_{16}(2^{15})$ 직교배열표를 사용하여 모두 2수준의 9개 제어인자를 <표 3.1>과 같이 배치하였으며 본 연구에서는 주효과만 고려하기로 한다. <표 3.2>는 5개의 품질특성치의 실험데이터와 SN이다. 작업성, 외관, X선 검사는 계수분류치이며 각 각의 SN비(박성현 1993, p.258, p.261, p.263)를 S_1 , S_2 , S_3 라 하자. 인장강도와 신도는 잡음인자에 대한 반복 실험을 실시하지 않아서 SN비의 계산을 할 수 없으나 두 특성 모두 망대특성이므로 편의상 원래의 실험데이터를 SN비로 취급하여 S_4 와 S_5 라 하고 분석한다.

3.1 제안한 최적화 방안

본 논문에서는 인자분석을하는데 있어서 인자(factor)를 추출하는 방법으로 주성분분석으로부터 유도되는 주성분인자법을 사용하기로 한다. 따라서 <표 3.2>에서 구하여진 5개 품질특성의 SN비(

$SN_1, SN_2, SN_3, SN_4, SN_5$)을 사용하여 주성분분석을 다음과 같은 순서로 하여 보자.

- (1) $SN_1, SN_2, SN_3, SN_4, SN_5$ 을 표준화하여 $SN_1^*, SN_2^*, SN_3^*, SN_4^*, SN_5^*$ 라 하자.
- (2) $SN_1^*, SN_2^*, SN_3^*, SN_4^*, SN_5^*$ 의 상관행렬(correlation matrix) R 을 구한다.
- (3) R 의 고유값(eigen values)과 고유벡터(eigen vectors)를 구한다
- (4) 인자부하행렬 즉, 성분행렬(component matrix)를 구한다.

위의 절차에 의하여 <표 3.3>과 <표 3.4>와 같은 결과를 얻는다. <표 3.3>은 5개 항목간의 상관 계수를 통해서 구한 각 축의 고유값과 누적점유율이다. 인자의 차원을 결정하는 몇 가지 기준에 의하여 고유값이 1이상이면서 누적점유율이 80%이상인 되는 첫 3개의 주성분을 택하기로 한다. 주 성분과 인자와의 관계로부터 인자부하행렬 즉, 성분행렬 L_1 을 구하면 <표 3.4>와 같다. 인자가 가지는 개념의 해석을 수월하게 하기 위하여 단순한 구조로 변경시키기 위한 방법으로 인자의 회전을 고려하여 보자. 인자회전방법으로 배리맥스 회전방법을 사용하여 성분행렬 L_2 를 구하면 <표 3.5>와 같다. 식(2.1)로부터 인자회전이후의 인자점수행렬을 구하면 <표 3.6>과 같다. 인자점수행렬 <표 3.6>로부터 첫 3개의 회전인자 점수의 모형식을 구하면 다음과 같은 선형결합식이 된다.

$$\begin{aligned} F_1 &= -0.565 SN_1^* - 0.052 SN_2^* + 0.601 SN_3^* + 0.086 SN_4^* - 0.003 SN_5^* \\ F_2 &= 0.185 SN_1^* + 0.661 SN_2^* + 0.160 SN_3^* + 0.436 SN_4^* + 0.151 SN_5^* \\ F_3 &= -0.091 SN_1^* + 0.274 SN_2^* - 0.105 SN_3^* - 0.352 SN_4^* + 0.785 SN_5^* \end{aligned} \quad (3.1)$$

앞서 제안한 성능측도 식(2.2)로부터, TF 는 다음과 같이 된다.

$$TF = \sum_{i=1}^3 \rho_i \sigma_i F_i$$

먼저 식(3.1)의 인자점수의 계수로부터 F_1 은 계수의 경향이 모호함으로 $\rho_1 = 0$ 로 두고 F_2 와 F_3 는 양의 경향이므로 $\rho_2 = \rho_3 = 1$ 로 하자. <표 3.5>으로부터 첫 3개 인자에서 회전인자 F_i 의 공헌도 $\sigma_1, \sigma_2, \sigma_3$ 를 구하기 위하여 먼저 회전된 성분행렬의 계수의 제곱을 각각 v_1, v_2, v_3 라 하면

$$v_1 = (-0.082)^2 + (-0.053)^2 + (0.864)^2 + (0.089)^2 + (0.078)^2 = 1.4555$$

이 되고 비슷하게, $v_2 = 1.4384, v_3 = 1.2309$ 가 된다. 또한 각 각의 공헌도는

$$\sigma_1 = v_1/(v_1 + v_2 + v_3) = 0.3529$$

이 되고, $\sigma_2 = v_2/(v_1 + v_2 + v_3) = 0.3487, \sigma_3 = v_3/(v_1 + v_2 + v_3) = 0.2984$ 이 된다. 따라서 식

(2.2)로부터 성능측도 TF 는 다음과 같다.

$$TF = 0.3487 F_2 + 0.2984 F_3 \tag{3.2}$$

식(3.1)과 식(3.2)로부터, $L_{16}(2^{15})$ 직교배열표에서 실험번호(i , 1번-16번)에 따른 TF 값을 구하면 <표 3.7>와 같다. 마지막으로 <표 3.1>과 <표 3.7>로부터 각 인자수준별 TF 값의 합을 구하면 <표 3.8>가 되고, 분산분석을 하면 유의확률 10%에서도 통계적으로 유의한 인자가 존재하지 않는다. 따라서 성능측도 TF 의 인자수준별 합계 <표 3.8>로부터, 최종적으로 설계인자의 다특성 동시 최적수준은 $A_1 B_0 C_0 D_0 F_0 G_0 H_0 I_1 J_1$ 이 된다.

<표 3.1> 제어인자 배치

열	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
인자	A	H		I			B	D	F	G	J				C

<표 3.2> 실험데이터와 SN비

실험 번호	실험데이터						SN비				
	작업성	외 관		X 선	인장강도	신도	S_1	S_2	S_3	S_4	S_5
	용보곤 이통관	앞 면 상중하	뒷 면 상중하	상중하	kg/cm ²	%					
1	010	100	100	310	43.7	33.6	-6.02	0.00	-2.43	43.7	33.6
2	010	010	001	310	40.2	40.2	-6.02	-8.13	-2.43	40.2	40.2
3	010	010	010	400	42.4	30.5	-6.02	-6.02	0.00	42.4	30.5
4	100	100	010	220	44.7	23.7	0.00	-3.98	-3.98	44.7	23.7
5	100	010	001	220	42.4	34.7	0.00	-8.13	-3.98	42.4	34.7
6	001	001	001	400	45.9	21.8	-9.54	-9.54	0.00	45.9	21.8
7	100	100	010	220	42.2	24.8	0.00	-3.98	-3.98	42.2	24.8
8	001	010	001	310	40.6	29.8	-9.54	-8.13	-2.43	40.6	29.8
9	010	010	100	211	42.4	33.7	-6.02	-3.98	-5.74	42.4	33.7
10	010	010	010	310	45.5	25.5	-6.02	-6.02	-2.43	45.5	25.5
11	010	100	010	310	43.6	36.9	-6.02	-3.98	-2.43	43.6	36.9
12	010	010	100	400	40.6	29.0	-6.02	-3.98	0.00	40.6	29.0
13	100	100	001	031	44.0	30.3	0.00	-6.99	-7.20	44.0	30.3
14	001	001	001	301	40.2	39.0	-9.54	-9.54	-4.77	40.2	39.0
15	010	100	010	130	42.5	27.9	-6.02	-3.98	-5.12	42.5	27.9
16	010	100	100	400	46.5	40.8	-6.02	0.00	0.00	46.5	40.8

<표 3.3> 설명된 총분산

성분	초기 고유값		
	전체	분산(%)	누적(%)
1	1.582	31.635	31.635
2	1.447	28.939	60.574
3	1.097	21.930	82.504
4	0.489	9.775	92.279
5	0.386	7.721	100.000

<표 3.4> 성분행렬

	성분		
	1	2	3
SN_1^*	0.678	-0.578	0.148
SN_2^*	0.573	0.378	0.616
SN_3^*	-0.141	0.880	-0.157
SN_4^*	0.738	0.440	-0.120
SN_5^*	-0.478	0.053	0.810

<표 3.5> 회전된 성분행렬

	성분		
	1	2	3
SN_1^*	-0.832	0.288	-0.202
SN_2^*	-0.053	0.898	0.204
SN_3^*	0.864	0.247	-0.102
SN_4^*	0.089	0.696	-0.510
SN_5^*	0.078	0.064	0.937

<표 3.6> 회전된 인자점수 계수행렬

	성분		
	1	2	3
SN_1^*	-0.565	0.185	-0.091
SN_2^*	-0.052	0.661	0.274
SN_3^*	0.601	0.160	-0.105
SN_4^*	0.086	0.436	-0.352
SN_5^*	0.003	0.151	0.785

<표 3.7> 성능측도 TF값

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
TF	0.69	0.07	-0.10	-0.14	-0.10	-0.85	-0.14	-0.46	0.21	-0.30	0.43	0.00	-0.18	-0.20	-0.06	1.13

<표 3.8> 인자수준별 TF값 합계

인자 수준	PC값 합계	인자 수준	PC값 합계	인자 수준	PC값 합계	인자 수준	PC값 합계	인자 수준	PC값 합계
A_0	-1.03	B_0	0.20	C_0	0.65	D_0	0.75	F_0	0.99
A_1	1.03	B_1	-0.20	C_1	-0.65	D_1	-0.75	F_1	-0.99
G_0	0.84	H_0	0.87	I_0	-0.66	J_0	-1.26	0	1.35
G_1	-0.84	H_1	-0.87	I_1	0.66	J_1	1.26	1	-1.35

3.2 기존의 최적화 방안

사례로 든 실험에 대하여 박성현(1993, p. 255-272)은 앞서 2.2절에서 언급한 기존의 방안을 이용하여 다특성인 경우 제어인자의 동시 최적화 수준을 구하였다. 모든 품질특성치에 대하여 유의한 제어인자를 수준별로 정리한 인자효과에 대한 종합결과표(박성현(1993, p. 270-271)에서 최적수준을 $A_1 B_1 C_0 D_0 F_0 G_1 H_0 I_1 J_1$ 라 하였다.

3.3 비교

앞서 3.1절과 3.2절로부터 우리가 제안한 성능측도 TF를 이용한 다특성 동시최적화 방안과 기존의 방안을 비교할 수 있을 것이다. 기존의 방안에서 동시최적화 조건은 $A_1 B_1 C_0 D_0 F_0 G_1 H_0 I_1 J_1$, 제안한 성능측도 TF를 이용한 동시최적화 조건은 $A_1 B_0 C_0 D_0 F_0 G_0 H_0 I_1 J_1$ 이다. 기존의 방안과 제안한 성능측도 TF에서 최적조건의 차이는 B인자와 G인자에 있음을 알 수 있다. 그런데 일치하지 않는 B인자와 G인자는 기존의 방안에서 B인자는 인장강도, 신도등 2가지 품질특성에서 J인자는 작업성, X 선, 인장강도, 신도등 4가지 품질특성에서 동시에 유의하면서 품질특성별로 최적수준이 서로 달라 최적수준을 택하는데는 실험자의 주관적인 선택에 따르기 때문이다. 따라서 두 가지 방안에서 제어인자의 최적조건은 거의 비슷한 결과를 얻을 수 있었다.

4. 결론

로버스트 파라미터 설계에서 다특성인 경우 SN을 사용하여 인자분석을 이용한 성능측도 TF를 제안하였다. 성능측도 TF를 실제 사례에 적용한 결과 기존의 방법과 비슷한 최적조건을 찾을 수 있었다. 우리가 새로이 제안한 최적화 방안은 기존 방안 보다 훨씬 간편하게 최적수준을 구할 수 있으며 인자분석의 특징인 차원의 축소를 통한 공통인자의 의미도 파악할 수가 있어 좋은 방안이 될 수 있을 것이다.

참고문헌

- [1] 김기영, 전명식 (2000). *다변량 통계자료분석*, 자유아카데미.
- [2] 박성현 (1993). *품질공학*, 민영사.
- [3] 허명희, 양경숙 (2001). *SPSS 다변량자료분석*, SPSS아카데미.
- [4] Derringer, G. and Suich, R.(1980). "Simultaneous Optimization of Several Response Variables," *Journal of Quality Technology* 12, 214-219.
- [5] Johnson, R. A. and Wichern D. W. (1992). *Applied Multivariate Statistical Analysis*, Prentice Hall.
- [6] Jolliffe, I. T. (1986). *Principal Component Analysis*, Springer-Verlag.
- [7] Kaiser, H. F. (1958). "The Varimax Criterion for Analytic Rotation in Factor Analysis," *Psychometrika* 23, 187-200.
- [8] Khuri, A. I. and Conlon, M.(1981). "Simultaneous Optimization of Multiple Responses Represented by Polynomial Functions," *Technometrics* 23, 363-375.
- [9] Taguchi, G. (1986). *Introduction to Quality Engineering: Designing Quality Into Products and Processes*, White Plains, NY : UNIPUB / Kraus International.
- [10] Taguchi, G. (1987). *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost*, White Plains, NY: UNIPUB / Kraus International.