

Robust Design for Multiple Quality Characteristics using Principal Component Analysis

Yong-Man Kwon¹⁾ · Yeon-Woong Hong²⁾

Abstract

Robust design is to identify appropriate settings of control factors that make the system's performance robust to changes in the noise factors that represent the source of variation. In this paper we propose how to simultaneously optimize multiple quality characteristics using the principal component analysis of multivariate statistical analysis. An example is illustrated to compare it with already proposed method.

Keywords : Robust design, Principal component analysis, Optimize multiple quality characteristics

1. 서론

종래의 실험계획법에서는 주로 품질특성치의 평균을 개선하는데 초점을 두고 최적 조건을 찾는 경향이 있었으나 다구찌 품질공학(Taguchi (1986, 1987))에서는 품질특성의 평균뿐만 아니라 변동을 가능한 줄이는 것을 목적으로 한다는 점에서 차이가 있다. 다구찌 파라미터 설계(로버스트 설계의 일종)에서 직교배열표를 이용한 교차배열(product array)은 내측배열(inner array)에 있는 설계인자(혹은 제어인자)와 외측배열(outer array)에 있는 잡음인자(noise factor)의 모든 교호작용을 고려한 실험배치를 하여 신호 대 잡음비(signal-to-noise ratio ; SN)를 이용한 자료분석을 하였다. 파라미터 설계에서는 품질특성치(혹은 반응변수)를 분석하는데 있어서 평균과 변동을 하나로 묶은 수행측도(performance measure)인 SN을 사용한다. SN은 모든 특성치에서 크면 클수록 좋게 만들어진 수행측도이다. 교차배열에서 잡음인자는 품질특성치의 품

-
- 1) Associate Professor, Dept. of Computer Science and Statistics, Chosun University, Gwangju, 501-759, Korea.
E-mail : ymkwon@chosun.ac.kr
 - 2) Associate Professor, Division of Internet Industrial Engineering, Dongyang University Yongju, Kyungpook, 750-012, Korea
E-mail : ywhong@dyu.ac.kr

질변동을 유발시키는 역할을 함으로써 변동에 둔감하면서 동시에 품질특성치의 평균을 목표치에 접근하는 제어인자의 최적조건을 찾을 수 있는 로버스트 설계를 가능하게 한다.

대부분의 실험계획에서는 반응변수가 한 개인 경우 설계인자들의 최적조건을 찾고 있으나 실제 실무에서는 어떤 경우의 실험계획에 관한 문제이든 동시에 고려하여야 할 품질특성치가 여러 개인 경우 즉, 다특성(multiple quality characteristics)인 경우가 대부분이다. 다특성에서 설계인자의 최적조건을 찾는 방안은 Derringer와 Suich(1980) 그리고 Khuri와 Conlon(1981)등에 의하여 연구되었다. 그러나 앞서 연구된 방안들은 품질변동을 고려하지 않은 동시 최적화 방안이다.

본 논문의 목적은 주성분분석(principal component analysis)을 이용한 수행측도를 이용하여 로버스트 설계에서 다특성인 경우 동시최적화하는 방법을 제시하고 기존의 다특성 실험의 분석결과(박성현(1993), 255-272)와 비교하고자 한다. 2절에서는 다특성 동시 최적화 방안을 제안하고 3절에서는 본 논문에서 제안한 최적화 방안을 예제를 들어 비교하였고 4절에서는 본 논문의 이점을 제시하였다.

2. 로버스트 설계를 위한 다특성 동시 최적화

2.1 주성분분석을 이용한 수행측도

다변량 자료분석기법인 주성분분석은 여러개의 특성치들을 변환시켜 주성분(principal component ; PC)이라고 부르는 서로 독립적인 새로운 인공변수들을 유도한다. 이때 각 주성분이 보유하는 변이의 크기를 기준으로 중요도의 순서를 생각할 수 있는데, 그들중 첫 k 개($k < p$)의 주성분에 의해 원래 자료에 내재하는 전체변이중 되도록 많은 부분을 차지하도록 변환시킴으로서 정보의 손실을 최소화하는 차원의 축소를 할 수 있다.

로버스트 설계에서 어떤 시스템이나 제품에서 고려하여야할 p 개의 품질특성치가 있다면 그들에 대응하여 p 개의 SN 즉, S_1, S_2, \dots, S_p 이 생성된다. 우리는 생성된 p 개의 SN을 사용하여 주성분분석을 다음과 같은 순서로 한다.

- (1) S_1, S_2, \dots, S_p 을 표준화하여 $S_1^*, S_2^*, \dots, S_p^*$ 라 하자.
- (2) $S_1^*, S_2^*, \dots, S_p^*$ 의 상관행렬(correlation matrix) R 을 구한다.
- (3) R 의 고유값(eigenvalues)과 고유벡터(eigenvectors)를 구한다.
- (4) 생성된 p 개의 주성분중에 전체 고유값의 총합에서 차지하는 비율 큰 주성분 첫 k 개를 택한다.

한편, 주성분공간의 차원을 결정하는 몇 가지 기준을 보면 대체로 다음과 같다.

첫째, 고유값이 0.7이상인 주성분만 고려한다(Jolliffe, 1986)).

둘째, 고유값의 누적점유율에 일정한 기준을 적용한다.(예를 들어 누적점유율 $\geq 80\%$)

먼저 i 번째 주성분 즉, 제 i 주성분을 PC_i 라고 할 때, PC_1 이 전체 분산에 대해 차지하는 비율 즉, PC_1 의 고유값이 전체 고유값의 총합에서 차지하는 비율이 아주 큰

경우는 PC_1 을 수행측도로 선택하여 설계인자의 최적조건을 찾는다. 그러나 PC_1 이 전체분산에 대해 차지하는 비율이 아주 크지 않아서 첫 k 개의 주성분이 선택되는 경우는 다음과 같이 수행측도를 제안하기로 한다.

$$PC = \sum_{i=1}^k \rho_i \sigma_i PC_i \quad (2.1)$$

여기서, PC_i 가 $\rho_i = 1$ 인 경우는 주성분의 계수가 양의 경향, $\rho_i = -1$ 인 경우는 음의 경향, $\rho_i = 0$ 인 경우는 모호한 경향이다. σ_i 는 PC_i 의 고유값이 전체 고유값의 총합에서 차지하는 비율이다. $\rho_i = 1$ 인 경우는 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 p 개 품질특성치의 SN를 가지고 주성분분석을 하고 있으며 또한 SN은 크면 클수록 더 좋은 수행측도이기 때문이다. 그런데 ρ_i 를 정할 때 반드시 3가지 경우 중에 1가지를 택하여야 하는 것은 아니다. 때로는 명백한 양의 경향이 아니고 대체로 양의 경향이라면 예를 들어 $\rho_i = 0.7$ 를 택할 수도 있는 것이다.

2.2 기존의 방안

종래에 로버스트 설계에서 다특성인 경우 동시최적화 방안은 다음과 같다.

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

마지막으로, 인자효과에 대한 종합결과표를 보고 최적수준을 판단한다.

3. 예제

예제로 든 실험은 1959년도에 일본 국유철도의 송입공장에서 있었던 사례이다 (Taguchi (1987, 189-191)). 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 설계인자를 <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 SN(박성현(1993), 258, 261, 263)을 S_1, S_2, S_3 라 하자. 인장강도와 신도는 잡음인자에 대한 반복 실험을 실시하지 않아서 SN의 계산을 할 수 없으나 두 특성 모두 망대특성이므로 편의상 원래의 실험데이터를 SN으로 취급하여 S_4 와 S_5 라 하고 분석한다.

<표 3.1> 설계인자 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

로버스트 설계를 위한 다특성 동시 최적화 방안에서 주성분분석에 의한 방안과 기존의 방안 즉, 5개 품질특성치 각 각의 SN을 사용하여 분산분석을 통한 설계인자의 최적수준을 정하는 방안등 2가지 방안을 예를 들어 비교 연구하고자 한다.

<표 3.2> 실험데이터와 SN

실험번호	실험데이터						SN				
	작업성	외 관		X 선	인장강도	신도	S ₁	S ₂	S ₃	S ₄	S ₅
	용보콘 이통란	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

3.1 수행측도 PC에 의한 최적화 방안

<표 3.2>에서 구하여진 5개 품질특성의 SN(S₁, S₂, S₃, S₄, S₅)을 사용하여 주성분분석을 다음과 같은 순서로 하여 보자.

- (1) S₁, S₂, S₃, S₄, S₅을 표준화하여 S₁^{*}, S₂^{*}, S₃^{*}, S₄^{*}, S₅^{*}라 하자.
- (2) S₁^{*}, S₂^{*}, S₃^{*}, S₄^{*}, S₅^{*}의 상관행렬(correlation matrix) R을 구한다.
- (3) R의 고유벡터(eigenvectors)와 고유값(eigenvalues)를 구한다. 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 상관계수를 통해서 구한 각 축의 고유값과 누적점유율이다. 주성분공간의 차원을 결정하는 몇 가지 기준에 의하여 고유값이 0.7이상이면서 누적점유율이 80%이상인 되는 첫 3개의 주성분을 택하기로 한다. <표 3.4>는 제3주성분까지 고려된 주성분계수이다. 5개 품질특성에 대한 주성분을 3개까지 고려하면 다음과 같은 선형결합식이 된다.

$$\begin{aligned} \text{제1주성분} : PC_1 &= 0.678S_1^* + 0.573S_2^* - 0.141S_3^* + 0.738S_4^* - 0.478S_5^* \\ \text{제2주성분} : PC_2 &= -0.578S_1^* + 0.378S_2^* + 0.880S_3^* + 0.440S_4^* + 0.053S_5^* \end{aligned} \quad (3.1)$$

$$\text{제3주성분} : PC_3 = 0.148S_1^* + 0.616S_2^* - 0.157S_3^* - 0.120S_4^* + 0.810S_5^*$$

앞서 제안한 수행측도 식(2.1)로부터, PC 는 다음과 같이 된다.

$$PC = \sum_{i=1}^3 \rho_i \sigma_i PC_i$$

먼저 식(3.1)의 주성분의 계수로부터 $\rho_1 = \rho_2 = \rho_3 = 1$, <표 3.3>으로부터 $\delta_1 = 0.3164$, $\delta_2 = 0.2894$, $\delta_3 = 0.2193$ 이 된다. 따라서 식(2.1)로부터 수행측도 PC 는 다음과 같다.

$$PC = 0.3164 PC_1 + 0.2894 PC_2 + 0.2193 PC_3 \quad (3.2)$$

식(3.1)과 식(3.2)로부터, $L_{16}(2^{15})$ 직교배열표에서 실험번호(i , 1번-16번)에 따른 PC 값을 구하면 <표 3.5>와 같다. 마지막으로 <표 3.1>과 <표 3.5>로부터 각 인자수 준별 PC 값의 합을 구하면 <표 3.6>이 되고, 분산분석표를 작성하여 보면 <표 3.7>이 된다. <표 3.7> 분산분석표는 유의확률이 10%에 가까운 인자 B, C, I만을 고려하고 나머지 A, D, F, G, H, J 인자는 오차항에 풀링한 후 분산분석표를 작성하였다.

<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
S_1^*	0.678	-0.578	0.148
S_2^*	0.573	0.378	0.616
S_3^*	-0.141	0.880	-0.157
S_4^*	0.738	0.440	-0.120
S_5^*	-0.478	0.053	0.810

<표 3.5> 수행측도 PC 값

i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16
PC	0.94	-0.77	0.02	0.48	-0.43	-0.05	0.07	-0.86	-0.12	0.31	0.37	0.01	-0.27	-1.25	-0.09	1.65

<표 3.6> 인자수준별 PC값 합계

인자 수준	PC값 합 계	인자 수준	PC값 합 계	인자 수준	PC값 합 계	인자 수준	PC값 합 계	인자 수준	PC값 합 계
A_0	-0.60	B_0	-1.77	C_0	3.49	D_0	0.50	F_0	1.31
A_1	0.60	B_1	1.77	C_1	-3.49	D_1	-0.50	F_1	-1.31
G_0	0.70	H_0	1.24	I_0	-1.64	J_0	0.01	0	3.24
G_1	-0.70	H_1	-1.24	I_1	1.64	J_1	-0.01	1	-3.24

<표 3.7> 분산분석표

요인	제공합	자유도	평균제공	F	유의확률
A	0.091	1	-	-	-
B	0.787	1	0.787	3.444	0.088
C	3.048	1	3.048	13.341	0.003
D	0.061	1	-	-	-
F	0.432	1	-	-	-
G	0.121	1	-	-	-
H	0.382	1	-	-	-
I	0.673	1	0.673	2.948	0.112
J	0.000	1	-	-	-
오차 (오차)	1.654 2.741	6 12	- 0.228	-	-
합계	7.249	15			

<표 3.7>로부터, 분산분석 결과 유의수준 10%에서 B, C는 유의하고 I는 대체로 유의한 인자라 할 수 있다. SN을 사용하여 주성분분석을 이용한 수행측도 PC의 수준별 평균합 <표 3.6>로부터, 설계인자의 최적수준은 $A_1 B_1 C_0 D_0 F_0 G_0 H_0 I_1 J_0$ 가 된다.

3.2 기존의 최적화 방안

예를 든 실험에 대하여 박성현(1993, 255-272)은 기존의 방안을 이용하여 다특성인 경우 설계인자의 최적수준을 구하였다. 모든 품질특성치에 대하여 유의한 설계인자를 수준별로 정리한 인자효과에 대한 종합결과표(박성현(1993, 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 결과 비교

3.1절과 3.2절로부터 기존의 방안과 우리가 제안한 수행측도 PC를 이용한 다특성 동시최적화 방안과 비교할 수 있을 것이다.

기존의 방안에서 동시최적화 조건은 $A_1 B_1 C_0 D_0 F_0 G_1 H_0 I_1 J_1$ 이고 제안한 수행측도 PC 를 이용한 동시최적화 조건은 $A_1 B_1 C_0 D_0 F_0 G_0 H_0 I_1 J_0$ 이다. 두 가지 방안에서 최적조건의 차이는 G 인자와 J 인자에 있음을 알 수 있다. 그런데 일치하지 않는 G 인자와 J 인자는 기존의 방안에서 G 인자는 작업성, 외관, 인장강도, 신도등 4가지 품질특성에서 J 인자는 작업성, 인장강도, 신도등 3가지 품질특성에서 동시에 유의하면서 품질특성별로 최적수준이 서로 달라 최적수준을 택하는데는 실험자의 주관적인 선택에 따르기 때문이다. 또한 수행측도 PC 를 이용한 분산분석에서는 G 인자와 J 인자 경우는 수준별 PC 의 합계에서 거의 차이가 없어 유의하지 않는 것을 <표 3.6>과 <표 3.7>에서 알 수 있다. 따라서 두 가지 방안에서 설계인자의 최적조건은 거의 비슷한 결과를 얻을 수 있었다.

4. 결론

로버스트 설계에서 고려하여야할 품질특성이 여러 개인 경우 즉 다특성인 경우 SN을 사용하여 주성분분석을 이용한 수행측도 PC 를 제안하였다. 수행측도 PC 를 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 주성분분석의 특징인 차원의 축소를 통한 주성분의 의미도 파악할 수가 있어 좋은 방안이 될 수 있을 것이다.

참고문헌

1. 박성현 (1993). 품질공학, 민영사.
2. Derringer, G. and Suich, R.(1980). "Simultaneous Optimization of Several Response Variables," *Journal of Quality Technology*, Vol. 12, pp. 214-219.
3. Jolliffe, I. T. (1986). *Principal Component Analysis*, Springer-Verlag.
4. Khuri, A. I. and Conlon, M.(1981). "Simultaneous Optimization of Multiple Responses Represented by Polynomial Functions," *Technometrics*. Vol. 23, pp. 363-375.
5. Taguchi, G. (1986). *Introduction to Quality Engineering: Designing Quality Into Products and Processes*, White Plains, NY : UNIPUB / Kraus International.
6. Taguchi, G. (1987). *System of Experimental Design: Engineering Methods to Optimize Quality and Minimize Cost*, White Plains, NY: UNIPUB / Kraus International.

[2003년 6월 접수, 2003년 7월 채택]