대한산업공학회/한국경영과학회 2002 춘계공동학술대회
한국과학기술원(KAIST) 2002년 5월 3일-4일

새로운 손실함수 적용을 통한 다중 반응표면분석

(A New Loss Function Approach To Multiresponse Optimization)

고영현, 나석희, 김광재, 전치혁

포항공과대학교 산업공학과
(Department of Industrial Engineering, POSTECH)

Abstract

It is often required to choose the optimum operating conditions for several responses simultaneously. In solving this multiresponse problem, the correlation of several responses, quality of prediction and the robustness of each response variable are must be considered. This paper proposes a new loss function approach that allows to consider these three important aspects. A numerical example illustrates the proposed methodology.

1. 연구배경

반응표면분석방법(response surface methodology: RSM)은 통상 어떤 하나의 반응변수 (response variable)를 최적화하기 위한 독립변수(independent variable) 혹은 작업환경변수(operating condition)를 구하는 것이다. 하지만 대부분의 실제 문제에서는 동시에 여러 개의 반응변수를 최적화 해야 하는 경우가 많다. 예를 들어, 고무제품의 인장강도 (tensile strength)와 탄성 (elasticity)을 동시에 최적화하는 문제의 경우 인장강도의 증가는 탄성의 감소를, 탄성의 증가는 인장강도의 감소를 가져오므로, 반응표면 분석방법을 통한 개별적인 최적화는 문제가 있음을 알 수 있다. 이 같은 문제를 해결하고 동시에 여러 개의 반응변수를 최적화하기 위하여 다중반응표면분석(multiple response surface methodology: MRS) 방법이 연구되었다. 다중반응표면분석은 실험에서 얻어진 여러 개의 예측식을 통하여 최적치를 찾아내는 방법이므로, 다중반응표면분석의 목적함수를 정의하는데 있어 각 반응 변수의 상관관계는 물론 예측식의 예측품질 그리고 로버스트성을 모두 고려할 수 있어야 한다.

다중 반응표면분석 방법은 만족도함수 (desirability function)을 기반으로 한 Derringer and Suich (1980)의 연구를 시작으로 활발히 연구되었다 [1]. 이 방법은 각각 개별 반응변수에 대한 0~1 사이의 만족도 값을 구하고 이를 통해 여러 반응변수에 대한 전체적인 만족도 값을 개별 만족도의 기하평균 (geometric mean)을 통해 얻고 이를 최적화 하는 방법으로 전문가의 의견을 쉽게 반영할 수 있다는 장점이 있다 [2,3]. 그러나, 만족도 함수를 이용한 방법은 각 반응 변수를 사이의 상관관계(correlation)를 고려하지 못하며 예측품질 (quality of prediction)을 고려하지 못한다는 단점을 가지고 있다.

다중 반응표면분석 문제의 또 다른 접근 방법으로 Pignatiello(1993)에 의해 제안된 손실함수(loss function) 방법이 있다 [5]. 본 연구에서는 기존 손실함수 방법의 단점을 분석하고, 개선을 통하여 상관관계, 예측품질, 로버스트성을 모두 고려할 수 있는 새로운 방법을 제안하고자 한다. 2절에서는 손실함수를 이용한 다중 반응표면분석방법에 대한 기존 연구 방법 및 문제점에 대해 살펴 보고 3절에서 이를 해결하기 위한 새로운 방법을 제안한다. 4절에서는 수치예제를 통해 본 연구의 결과의 유효성을 살펴보고, 5절에서 결론을 맺는다.

2. 기존 연구 및 한계점

본 연구는 $y(x)=x'\beta+\varepsilon$, ε 이 평균이 0, 분산이 $\Sigma_y(x)$ 인 오차항을 가지는 모델에 대하여 이루어졌으며, $\hat{y}(x)=x'\hat{\beta}$ 로 추정이 가능하다. 본 연구에서 이용되는 기호를 대략적으로 소개하면, 반응변수가 k 개 있는 경우 $y(x)$, $\hat{y}(x)$ 는 $k\times 1$ 벡터이고, θ는 각 반응변수 y 의 목표치를 나타내는 $k\times 1$ 벡터, $\Sigma_y(x)$ 는 x 에서의 y 의 $k\times k$ 공분산 행렬을, C 는 각 반응변수 및 교호작용에 대한 $k\times k$ 의 비용행렬 (cost matrix)을 의미한다.

단일 반응 변수에 대하여 손실함수는 일반적으로 식(1)과 같이 정의된다.

$$L(x)=c(y(x)-\theta)^2 \tag{1}$$

c 는 비용상수이며, 목표치 θ로부터 벗어난 정도에 따라 손실이 커지는 성질을 가지고 있다. 궁극적으로는 식(2)와 같이 정의되는 손실함수의 평균을 최소화하는 것이 목표이다.

$$\begin{aligned}\min E[L(x)] &= E[c(y(x)-\theta)^2] \\ &= c\sigma_y^2(x)+c(E[y(x)]-\theta)^2\end{aligned} \tag{2}$$

식(2)의 우측 항에서 알 수 있듯이 손실함수의 평균이 특정 x 에서의 분산, 목표치 θ 로부터 벗어난 정도(bias) 이렇게 2개의 항목으로 나뉘어짐을 알 수 있다. 따라서 손실함수를 최소화하는 것은 각 반응변수를 최적화하는 동시에 각 반응 변수의 분산이 작은 안정적인 최적 조건을 찾아내는 작업이라 할 수 있다.

Pignatiello(1993)은 이러한 손실함수의 개념을 다중반응표면분석에 적용한 방법으로써 식 (3)과 같은 손실함수를 정의하고, 식(4)의 손실함수의 평균을 최소화한다.

$$L(x,y(x))=(y(x)-\theta)'C(y(x)-\theta) \tag{3}$$

$$\begin{aligned}\min E[L(x)] = (\hat{y}(x)-\theta)'C(\hat{y}(x)-\theta) \\ +trace[C\Sigma(x)]\end{aligned} \tag{4}$$

일반적으로 C 는 어떤 반응변수가 더 중요하거나 손실 비용이 더 크다는 추가적인 정보가 없다면 Σ^{-1} 를 사용한다. C 의 구성 또한 중요한 문제이지만 본 연구의 범위를 넘어가므로 Vining(1998)을 참조하기 바란다[6]. 위에서 언급되었듯이 이 접근 방법은 반응 변수의 상관관계 뿐 아니라 로버스트성을 고려하였다는 장점을 가지고 있으나, 예측품질, 즉 모델의 분산에 대한 반영이 부족하다.

Vining(1998)은 모델의 예측품질을 고려한 보다 일반적인 형태의 손실함수 방법을 제시하였다. Vining은 식(5)처럼 여러 반응변수의 예측치에 대한 손실을 평가하는 손실함수를 정의하고,

$$L[\hat{y}(x),\theta]=(\hat{y}(x)-\theta)'C(\hat{y}(x)-\theta) \tag{5}$$

다음과 같이 손실 함수의 평균을 최소화하는 방법을 제안했다.

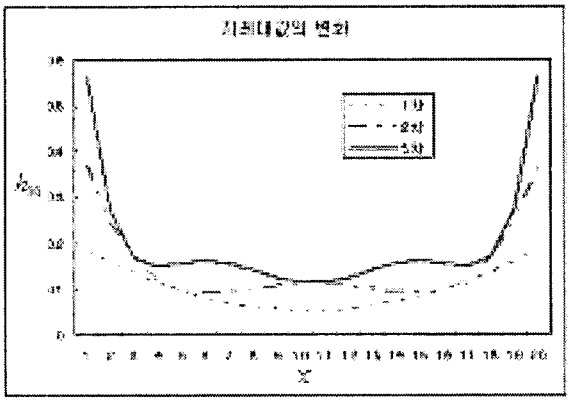
$$\begin{aligned}\min E[L(x)] \\ = (E[\hat{y}(x)]-\theta)'C(E[\hat{y}(x)]-\theta) \\ +trace[C\Sigma_{\hat{y}}(x)]\end{aligned} \tag{6}$$

평균 손실을 최소화하는데 위 식의 우변 첫번째 항은 목표값에 대해 벗어난 정도에 대한 손실을 의미하고, 두 번째 항은 예측품질에 대한 벌점을 의미한다. Pignatiello의 방법과의 차이는 $\Sigma_y(x)$ 과 $\Sigma_{\hat{y}}(x)$ 의 차이로 좁혀질 수 있고, 이는 $\Sigma_y(x)$ 가 각 반응 변수의 분산을 뜻하므로 로버스트성을 반영하는데 반해, $\Sigma_{\hat{y}}(x)$ 는 예측치의 분산을 뜻하므로 모델의 분산, 즉 예측 품질을 반영한다고 생각할 수 있다. 따라서, 이 방법은 반응 변수의 상관관계 뿐 아니라 예측품질을 고려하였다는 장점을 가지고 있다. 예를 들어, 각 반응변수를 예측하기 위해 OLS(ordinary least square)를 사용하였다면 어떤 x_0 에서의 예측치의 분산은

$$\Sigma_{\hat{y}}(x_0)=x_0'(X'X)^{-1}x_0\Sigma=h_{00}\Sigma \tag{7}$$

와 같이 나타날 수 있고, 지레대값(leverage value)라 불리는 h_{00} 의 값은 x 의 경계선

부분에서 급속히 커진다는 사실이 알려져 있다. [그림 1]은 x 의 변화에 따른 h_{00} 를 그려본 것으로 경계 부근에서 h_{00} 와 식 (7)에 의해 모델의 분산이 커지고 중심부분에서는 작아지는 경향이 있음을 알 수 있고 모델의 차수가 커질수록 경계선 부분의 분산이 급격히 커지고 있음을 알 수 있다.



[그림 1] 지레대값의 변화

이 경우는 OLS를 적용한 경우지만 대부분의 예측 모델은 경계 부분에서의 분산이 급격히 커지는 경향이 있고 Vining의 방법은 그러한 문제를 해결 하려는 노력으로 볼 수 있다. 반면에, 이 방법은 모델의 분산을 중요시 여겨 Pinatiello가 중요시 여긴 로버스트성을 완벽히 고려하지 못한다는 단점이 있다.

3. 새로운 다중반응표면분석 방법

3.1. 새로운 손실함수를 이용한 다중 반응표면분석

본 연구에서 새로운 손실함수를 정의하였다. 이 손실함수의 평균을 목적식으로 사용하면 상관관계 뿐만 아니라 예측품질 그리고 로버스트성을 모두 고려할 수 있다. 본 연구에서 제안하는 손실함수는 식(8)과 같다.

$$L[\hat{y}_{new}(x),\theta=(\hat{y}_{new}(x)-\theta)'C(\hat{y}_{new}(x)-\theta) \quad (8)$$

여기서 $\hat{y}_{new}(x)$ 는 새롭게 주어진 데이터에 대한 특정 x 에서의 예측값을 의미한다. 이 $\hat{y}_{new}(x)$ 의 분산은 예측 모델에 의한 분산과 자체 내에 존재하는 분산으로 나뉘어진다([그림 2]).

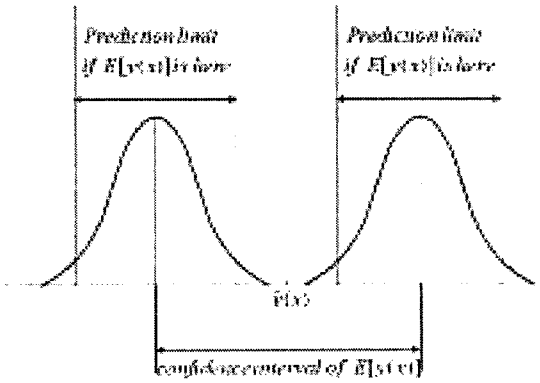
[그림 2]에서 두 정규분포의 중심간의 거리는 $\hat{y}(x)$ 의 분산, 즉 예측 모델의 신뢰구간을 의미하고, 종내에서의 신뢰구간은 $y(x)$ 자체내에 존재하는 분산에 의한 것임을 알 수 있다. 즉, 예측품질과 관련된 $\hat{y}(x)$ 의 분산과 로버스트성과 관련된 $y(x)$ 의 분산이 모두 포함되었음을 알 수 있다. 단변량 관점에서의 $\hat{y}_{new}(x)$ 의 분산은 식(9)와 같이 계산될 수 있고,

$$\sigma^2(\hat{y}_{new}(x))=\sigma^2(y(x))+\sigma^2(\hat{y}(x)) \quad (9)$$

이의 다변량으로의 확대를 생각하면 새롭게 제안된 손실함수를 최소화 하는 것은 다음의 식(10)을 최소화하는 문제로 변환이 가능하다.

$$\min E(L)=[\hat{y}(x)-\theta]'C[\hat{y}(x)-\theta] + trace[C\Sigma_{\hat{y}}(x)]+trace[C\Sigma_{y}(x)] \quad (10)$$

식 (10)의 의미를 자세히 살펴보면 첫째 항은 목표치와의 차이를 의미하고, 둘째 항은 Vining의 목적식 식(6)에 나타난 예측품질을 의미하는 예측값의 분산, 셋째 항은 Pignatiello의 목적식 식(5)에 나타난 로버스트성을 의미하는 반응변수의 분산으로 구성이 됨을 알 수 있다. 따라서 식 (10)의 최소화는 반응 변수의 상관관계 뿐만 아니라 모델의 예측 품질 그리고 실험의 분산을 각각 균형있게 최적화하기 위한 방향으로 진행되며 그에 따라 보다 안정적인 최적 조건을 구할 수 있게 된다.



[그림 2] $\hat{y}_{new}(x)$ 분산의 분해

3.2 반응변수별 예측품질의 차이를 반영하는

비용 행렬

앞에서 언급된 예측품질은 각 반응변수를 구성하는 모델 내의 특정 x에서의 예측값의 분산으로 설명이 되었다. 하지만 각 반응변수를 나타내는 모델간의 예측성능의 차이는 반영되지 못한 것을 할 수 있다. 예를 들어, y_1에 대한 모델의 결정계수 (r_1^2)가 0.99, y_2에 대한 모델의 결정계수 (r_2^2)가 0.7이라면 보다 정확한 모델인 y_1에 대한 모델에 더 비중을 주어야 할 것이다. 따라서 본 연구에서는 비용행렬의 조정을 통해 각 모델간의 예측 성능을 반영하는 방법을 제안한다.

기존의 비용행렬을 C, 그 행렬의 i,j 요소를 c_{ij} 라 하면, 각 변수모델간 예측성능을 반영한 새로운 C'는 아래와 같이 정의될 수 있다.

$$C' = \begin{bmatrix} c_{11}r_{11} & c_{12}r_{12} & c_{1p}r_{1p} \\ c_{21}r_{21} & c_{21}r_{22} & c_{2p}r_{2p} \\ \\ c_{p1}r_{p1} & c_{p2}r_{p2} & c_{pp}r_{pp} \end{bmatrix} \quad (11)$$

where $r_{ij} = r_i \cdot r_j$, r_i: 반응변수 i의 결정계수

새로운 비용행렬 C'는 각 변수가 포함된 모델의 예측성능에 대하여 그 모델의 성능에 따라 비중을 다르게 하여 변수모델간의 예측품질을 가능하게 해준다.

4. 수치예제

Vining(1998) 논문에서 사용된 Polymer Experiment 데이터를 이용하여 실제 분석을 실시하였다. 이 데이터는 Central Composite Design(CCD)를 사용하고 각 반응변수의 모델은 각각 2차의 다항회귀모델(full second order polynomial model)이고 OLS(ordinary least square) 방법을 사용하여 추정하였다. 이 데이터는 반응변수 y_1 (conversion)은 망대특성을 가지며, 반응 변수 y_2(thermal activity)는 망목특성을 가지고 목표치는 57.5이다. 즉, y_1은 크게, y_2는 57.5에 가깝게 만들면서 각 반응변수의 분산이 작고 모델이 안정적인 실험조건 x_1, x_2, x_3의 값을 얻어내는데 목적이 있으며, 목표를 맞추기 위해 $\theta = (100, 57.5)'$를 사용하였다. 또한 이 실험은 반복이 없어 각 반응변수의 분산을 추정하기 힘들고 따라서 본 연구에서는 로버스트성의 반영여부를 확인하기 위하여 y_1, y_2 각각에 대해 정규난수(Normal random variate)를 발생시켜 5개의 반복 자료를 만들었다.

$$y_{1i} = y_1 + N(0,2)$$
$$y_{2i} = y_2 + N(0,1), \quad i = 1,2,...,5$$

각 실험 조건 X와 반복 실험을 통한 Y의 평균 및 분산, 공분산은 부록 [표 A1]와 같다. 또한, 비용행렬은 일반적으로 많이 사용하는 $C = \Sigma^{-1}$를 사용하였고, 이렇게 만들어진 데이터를 통하여 각 방법론에 필요한 목적식을 구성하는 $trace[C\Sigma_y(x)]$, $trace[C\Sigma_{\hat{y}}(x)]$ 와 $(\hat{y}(x)-\theta)'C(\hat{y}(x)-\theta)$ 를 구해보았다(부록 [표 A2]). 여기서 $trace[C\Sigma_{\hat{y}}(x)]$은 Vining의 목적식에 포함된 예측 품질을 나타내는 부분으로 [그림 1]에서 나타나는 데로 실험의 중심으로 갈수록 작아지는 경향이 있다는 것을 볼 수 있고, 이 세가지의 상대적 크기를 통해 목적식의 값, 로버스트성, 예측 품질의 상대적 중요성을 짐작할 수 있다. 또한 최종적으로 각 방법에 대한 최적 반응변수와 그의 환경변수를 구하였다([표 1]). T는 상관관계나 로버스트성, 예측 품질을 전혀 고려하지 않고 θ에 가장 가까운 값을 보이는 최적 반응값을 나타낸다.

[표 1] 각 방법별 최적 반응변수 및 환경변수

	$\hat{y}_1(x)$	$\hat{y}_2(x)$	x_1	x_2	x_3
T	95.491	56.715	-0.727	1.682	-0.838
P	95.651	56.489	-0.784	1.682	-0.877
V	95.016	58.305	-0.362	1.682	-0.571
(1)	95.582	56.591	-0.761	1.682	-0.858
(2)	101.3	53.192	-1.682	1.682	-1.55

T: 최적반응값; P:Pignatiello; V:Vining
(1): 제안된 방법; (2): 비용행렬을 고려

반응변수의 목표값을 최적화 하는 방법(T)의 결과와 Pignatiello(P)의 결과가 차이를 보이는 것은 각 실험 조건에 의한 Y의 분산을 줄이려는 즉 Robustness를 확보하려는 성질에 의한 것임을 예상할 수 있다. Vining(V)의 최적 실험 조건이 Pignatiello(P) 의 최적 실험 조건 보다 원점으로 이동한 것은 예측 품질에 대한 신뢰도를 높이려는 특성에 의한 것임을 확인할 수 있다.

Vining의 방법을 적용한 경우에서는 [그림 1]에서 보듯이 실험의 중심쪽으로 최적 실험 건을 이동시켜주는 효과를 준다는 것을 확인 할 수 있다. 이는 Vining이 제안한 가장 큰 문제점 중의 하나로 판단된다. 그에 따라 $trace[C \cdot h_{00}\Sigma_y(x)]$ 의 값이 실험의 중심에 에서 가장 작아지는 경향 때문에 예측 품질의 향상을 위해 로버스트성을 희생할 수 있다는 것을 의미한다.

본 연구에서 제안한 방법(1) 을 적용한 결과는 Vining의 예측 품질과 Pignatiello의 로버스트성을 동시에 고려한 방법으로서 위 결과는 최적 실험 조건에서의 반응 변수의 분산을 감소시키려는 Pignatiello의 최적 실험 조건과 예측 품질에 대한 비용(penalty)이 고려된 Vining의 최적 실험 조건을 타협한 최적 실험 조건을 제시해 준다([표 1]의(1)).

또한 y_1에 대한 모델의 결정계수 (r_1^2)가 0.99, y_2에 대한 모델의 결정계수 (r_2^2)가 0.7로 가정한 경우에 대해서 앞의 예제를 사용하여 제안된 방법 (2) 의 결과를 구하였다. y_1 의 예측식의 예측 품질이 더 좋으므로 y_1 의 결과를 더 좋은 방향, 즉 크게 만들고 y_2 를 희생시키는 결과가 얻어지는 것을 확인할 수 있다.

5. 토의 및 결론

본 연구에서는 새로운 손실함수를 정의하고 그 손실함수의 최소화를 통하여 상관관계, 예측품질 그리고 로버스트성까지 모두 반영하는 최적값을 구해낸다는 것을 보였다. 또한 비용함수의 조정을 통해 모델간의 예측 품질 또한 반영할 수 있는 방법을 제안하였다. 그리고 마지막으로 수치예제를 통해서 제안된 방법과 Pignatiello와 Vining의 방법을 비교하고 그 유효성을 살펴보았다 ([표 2]).

본 연구에서 제안된 방법은 많은 부분을 반영하는 대신 반응변수의 분산, 즉 로버스트성을 고려하기 위해, Pignatiello의 경우와 마찬가지로 반복실험이 필요하다. 본 연구에서 제안한 방법은 상관관계, 예측품질 그리고 로버스트성에 대하여 각각의 중요도가 정해져 있으므로 이들에 대한 중요도를 특정 문제에 적합하게 조정 해주는 방법에 대한 추후 연구가 필요할 것으로 판단된다. 또한 본 연구의 성능을 보다 잘 설명하기 위한 더 많은 실험이 이루어져야 할 것으로 판단된다.

[표 2] 손실함수를 이용한 연구의 비교

	Pignatiello (1993)	Vining (1998)	제안 (2002)
상관관계	O	O	O
로버스트성	O	X	O
모델내의 예측품질	X	O	O
모델간의 예측품질	X	X	O

Reference

[1] Derringer, G., Suich, R., "Simultaneous optimization of several responses variables", *Journal of Quality Technology*, 12, pp.214-219, 1980

[2] Harrington, E., "The desirability function", *Industrial Quality Control*, 21, pp.494-498, 1965

[3] Kim, Kwang-Jae, Lin, K.J. Dennis, "Simultaneous optimization of mechanical properties of steel by maximizing exponential desirability function", *Applied Statistics*, 49(3), pp.311-325, 2000.

[4] Pignatiello, Jr., J.J., "Strategies for Robust Multiresponse Quality Engineering", *IIE Transactions* ,25, pp.5-15, 1993.

[5] Vining. G. Geoffrey, Myers, H. Raymond, "Combining Taguchi and response surface philosophies: A dual response approach", *Journal of Quality Technology*, 22(1), pp.38-45, 1990.

[6] Vining, G. Geoffrey, "A compromise approach to multiresponse optimization", *Journal of Quality Technology*, 30(4), pp.309-313, 1998.

부록

[표 A1] 예제 실험 데이터의 표본 평균, 표본 분산, 표본 공분산

	x_1	x_2	x_3	$\hat{E}[y_1(x)]$	$\hat{\text{var}}(y_1(x))$	$E[y_2(x)]$	$\hat{\text{var}}(y_2(x))$	$\text{cov}(y_1, y_2)$
1	-1	-1	-1	74.02	4.81	53.54	0.423	-0.948
2	1	-1	-1	52.55	4.009	63.36	2.169	0.9417
3	-1	1	-1	88.95	3.839	53.96	0.696	0.7418
4	1	1	-1	71.15	1.274	62.53	1.859	0.8105
5	-1	-1	1	71.55	3.229	57.26	1.18	-1.45
6	1	-1	1	89.49	0.344	67.38	1.05	-0.304
7	-1	1	1	66.68	0.669	60.38	0.618	0.128
8	1	1	1	95.82	4.616	67.85	0.278	0.0154
9	-1.682	0	0	75.55	1.916	58.58	1.153	-0.266
10	1.682	0	0	79.73	2.85	65.51	0.511	0.2674
11	0	-1.682	0	84.89	0.798	59.77	1.167	0.2499
12	0	1.682	0	95.67	2.067	61	0.272	0.2273
13	0	0	-1.682	54.25	2.013	57.12	0.949	-0.361
14	0	0	1.682	80.13	1.421	63.35	0.733	0.6101
15	0	0	0	80.75	4.386	59.39	0.451	0.8591
16	0	0	0	75.11	5.371	61.15	1.652	0.325
17	0	0	0	74.93	4.063	59.33	1.703	-2.017
18	0	0	0	82.11	4.548	60.97	0.867	1.3325
19	0	0	0	79.85	4.116	61.16	2.148	-0.922
20	0	0	0	91.83	0.414	59.28	0.465	0.3044

[표 A2] 목적 함수식의 분석

	x_1	x_2	x_3	$trace[C\Sigma_y(x)]$	$trace[C\cdot\Sigma_{\hat{y}}(x)]$	$(\hat{y}(x)-\theta)'C(\hat{y}(x)-\theta)$
1	-1	-1	-1	0.0749	0.0312	4.885843
2	1	-1	-1	0.1719	0.0412	21.73411
3	-1	1	-1	0.0689	0.0237	1.345294
4	1	1	-1	0.1316	0.0250	9.385165
5	-1	-1	1	0.1223	0.0345	5.795268
6	1	-1	1	0.0799	0.0151	8.806394
7	-1	1	1	0.0470	0.0099	9.662536
8	1	1	1	0.0528	0.0251	8.162316
9	-1.68	0	0	0.0981	0.0214	4.697762
10	1.682	0	0	0.0538	0.0168	9.25895
11	0	-1.68	0	0.0854	0.0150	2.389361
12	0	1.682	0	0.0316	0.0110	1.16381
13	0	0	-1.68	0.085	0.0200	14.98711
14	0	0	1.682	0.0554	0.0119	6.531198
15	0	0	0	0.0542	0.0057	3.328145
16	0	0	0	0.1520	0.0108	6.408453
17	0	0	0	0.1714	0.0115	5.283996
18	0	0	0	0.0797	0.0067	3.837645
19	0	0	0	0.1913	0.0119	4.685192
20	0	0	0	0.0325	0.0015	0.86424