다중 반응 변수 문제 해결을 위한 손실 함수 방법에서 비용 행렬의 보정에 관한 연구

The Study of Adjusting the Cost Matrix in Loss Function Approach for Multiresponse Optimization

이대원, 김소희, 김광재, 이재욱

경상북도 포항시 남구 효자동 산31번지 포항공과대학교 산업공학과

e-mail: {woosuhan,eunmorai,kjk,jaewookl}@postech.ac.kr

Abstract

For solving multiresponse problems, a variety of loss function approaches have been proposed assuming that a cost matrix is known and fixed. However a cost matrix is also an important factor in loss function approaches, because the optimal solution is very sensitive to the cost matrix. In this paper, we propose a novel method for adjusting the cost matrix by considering the predictive ability of the estimated response models. Simulation results for the generated data set show that the proposed method is competitive with previously reported methods.

1. 연구배경

제품이나 공정의 설계에서 부딪치게 되는 어려운 문제 중 하나가 동시에 달성할 수 없는 목표값을 가지는 여러 개의 반응 변수를 동시에 고려해서 전체 품질이 최적이 되도록 하는 입력 변수를 찾는 것이다. 이러한 다중 반응 변수 문제(Multiresponse System)를 해결하기 위해 쓰이는 대표적인 방법이 손실 함수(Loss Function) 접근법이다. 손실 함수 접근법은 목표치와 여러 개의 반응 변수 사이의 오차를 하나의 손실 함수로 나타내고 이러한 손실 함수의 기대값을 최소화하는 입력 변수를 찾는 방법이다. 손실함수의 기대값은 반응 변수가 얼마나 목표값에 가까운가를 나타내는 편차(Bias)와 반응 변수 자체가 지닌 고유한 분산(Variance)의 두 항으로 구성되어 있다. Pignatiello(1993)가 제안한 방법은 아래 식 (1)과 같이 실제 반응 변수(y)를 이용하여 그 기대손실을 최소화하는 것이다. 이 방법은 통계적으로는 가장 엄밀한 방법이나 일반적으로 반응 변수에 대한 확률분포를 모르기 때문에 $E[y(x)]$와 $Cov[y(x)]$를 구하기가 어렵다는 단점이 있다. Vining(1998)은 이러한 선행 연구의 단점을 보완하기위해 기대손실을 나타내는 식(1)에서 $y(x)$ 대신에 $\hat{y}(x)$을 추정하여 이용하였다. 본 논문에서는 Vining(1998)이 제안한 방법을 기본적으로 이용하였다.

$$
\begin{aligned}
E(L(y(x),\theta)) &= Bias + Variance \\
&= (E[y(x)]-\theta)^T C (E[y(x)]-\theta) \\
&\quad + C \cdot trace[Cov(y(x))]
\end{aligned}
\qquad (1)
$$

기대손실을 구하고자 할 때, 또 다른 어려운 문제 중에 하나가 비용 행렬(C)을 결정하는 것이다. 비용 행렬의 의미는 반응 변수가 목표값에 비해 Δ만큼 오차가 발생했을 때 발생하는 비용(/단위)으로, 반응 변수들 간의 상대적 중요도를 나타내는 척도로 해석할 수 있다. 이러한 비용 행렬은 일반적으로 현장 실무자의 경험에 의해 결정된다. 하지만 비용 행렬의 각 항 간의 상대적인 크기가 다중 반응 표면법의 최적해에 민감하게 영향을 미치므로 비용 행렬의 값을 좀 더 객관적인 값으로 보정해 줄 필요가 있다. 비용 행렬에 영향을 줄 수 있는 요소는 반응 변수간의 상대적 중요도(손실비용), 추정 모델의 정확성, 반응 변수 간의 상관 관계 등으로 생각할 수 있다. 위 항목 중 반응 변수 간의 상대적 중요도는 의사결정자의 주관적인 판단에 의해 결정할 수 밖에 없다. 따라서 본 논문에서는 좀 더 엄밀하게 기대손실을 추정하고 이를 통해 좀 더 좋은 최적해를 구하기 위해, 주관적 비용 행렬이 주어졌을 때 추정 모델의 정확성을 바탕으로 비용 행렬을 보정하는 방법을 제안한다.

2. 기존 연구에서의 비용 행렬의 이용 사례 및 문제점

2.1 기존 연구에서 비용 행렬의 이용 사례

Pignatiello(1993)는 손실에 대한 실제 공정 데이터를 이용하여 비용 행렬을 결정하는 방법을 제안하였다. 공정 데이터는 각각의 반응 변수가 목표 값에서 얼마만큼 벗어났는지에 대한 오차(Δ)와 이로 인해 발생한 실제 공정의 손실 비용(L)으로 이루어져 있다. 이들 데이터를 이용하여 미지수가 비용 행렬의 성분인 연립 방정식을 해를 세우고 이를 풀어 비용행렬을 결정할 수 있다. 하지만, Pignatiello(1993) 방법은 비용 행렬을 구하기 위해 $(\Delta_{1,i}, \Delta_{2,i}, L_i)$인 공정 데이터가 최소한 (반응 변수 개수)·(반응 변수 개수+1)/2 개가 필요하지만, 이러한 공정 데이터를 구하기가 어렵다는 단점이 있다.

Khuri & Conlon(1981)는 손실함수 접근법의 특별한 경우인 Generalized Distance를 이용하여 손실을 구하였으며 비용 행렬은 아래와 같이 통계적 방법으로 결정하였다.

$$C = \hat{\Sigma}^{-1} / z^T(x)(X^TX)^{-1}z(x) \tag{2}$$

Ames (1997), Vining(1998), Ko(2002)의 연구에서는 모두 비용 행렬은 주어진 값으로 가정하거나 반응 변수의 추정 모델의 분산-공분산 행렬의 역행렬인 $\hat{\Sigma}_y^{-1}$를 이용하고 있다.

2.2 기존의 비용 행렬의 이용 사례의 문제점

기존의 손실함수 접근법의 연구에 있어서, 비용 행렬은 주관적으로 주어진 값으로 가정하거나, 또는 데이터 자체만을 고려한 통계적 수치인 $\hat{\Sigma}^{-1}$를 비용 행렬로 이용하였다. 하지만 주관적인 수치만 반영하거나 또는 데이터만으로 의해 결정된 수치만을 반영하는 것은 표 1. 과 같은 오류를 범할 가능성이 있다.

표 1. 비용 행렬의 결정에서 주관적 또는 객관적 수치 하나만을 반영했을 때의 문제점

주관적 수치만 반영	통계적 수치만 반영
• 변수간의 중요도의 정확한 상대적 비율을 반영하지 못한다. • 각 반응 변수 모델의 추정 설명력이 아주 낮은 반응 변수를 중요한 반응 변수로 강조할 수 있다. • 작업자가 알지 못하는 반응 변수 간의 상관관계를 고려 못한다.	• 공정 전문가의 경험적 지식을 반영하지 못한다. • 중요한 변수보다는 추정 설명력이 높은 변수가 강조된다.

3. 새로운 비용 행렬 추정 방안

3.1 제안한 비용 행렬 추정 방안의 배경

본 논문에서 제안하는 비용 행렬 추정의 기본 아이디어는, 주관적으로 결정된 비용 행렬 수치를 반응 변수 모델의 추정 설명력을 반영하는 R^2 수치로 보정하여 주관적, 객관적 수치를 동시에 반영한 비용 행렬을 구하는 것이다.

MRS는 여러 개의 상충되는 반응 변수들의 상대적 중요도에 따라 반응 변수의 추정값이 목표값에 가깝도록 하는 입력변수 x^*를 찾아 낸다. 하지만, 반응 변수 추정 모델의 설명력이 낮은 경우, MRS를 통해 구한 최적 입력값(x^*)을 실제 공정에 적용하였을 때, 예측한 차이(추정 반응값−목표값) 보다 더 큰 오류가 발생할 수 있다(그림 1의 Performance Assessment 참조). 이는 최적 입력값을 찾기 위해 MRS에 이용되는 것은 반응 변수 추정 모델이지만, 구한 최적 입력값을 실제 공정에 적용하였을 때의 반응값은 실제의 반응 함수에 의해 결정되기 때문이다.

따라서 MRS procedure 단계에서 이러한 오차를 줄이기 위해, '*추정 설명력이 높은 반응 변수에 상대적 더 큰 중요도가 할당*'되도록 한다면 실공정에서 좀 더 좋은 성능을 낼 수 있는 최적 입력값을 구할 수 있을 것이다.

3.2 새로운 비용 행렬 추정 방법론

비용 행렬 추정 방안의 제안에 앞서, 제안된 방안에서 n개의 다중 반응 변수 문제에 이용되는 R_{mat}행렬과 Array multiply operator를 아래 식(3), (4)과 같이 정의한다.

$$R_{mat} = \begin{bmatrix} (1+R_1)^2 & \cdots & (1+R_1)(1+R_n) \\ \vdots & \ddots & \vdots \\ (1+R_n)(1+R_1) & \cdots & (1+R_n)^2 \end{bmatrix} \tag{3}$$

R_{mat}행렬에서 R_i의 의미는 i번째 반응 모델의 추정 설명력을 나타내는 R-square의 제곱근을 의미한다. 따라서, R_{mat}행렬의 모든 성분 1~2사이의 값을 가지며, 해당항의 반응 변수들의 추정 설명력이 높을수록 그 값은 2에 가까워지고 낮을수록 1에 가까워진다. 모든 성분에 1을 더한 이유는 각 반응 변수들 간의 상대적 차이를 좀 더 증가 시키기 위해서이다.

Array multiply operator인 $[\ ].\times[\ ]$는 각 행렬의 각 성분끼리의 곱을 의미한다. 이는 아래

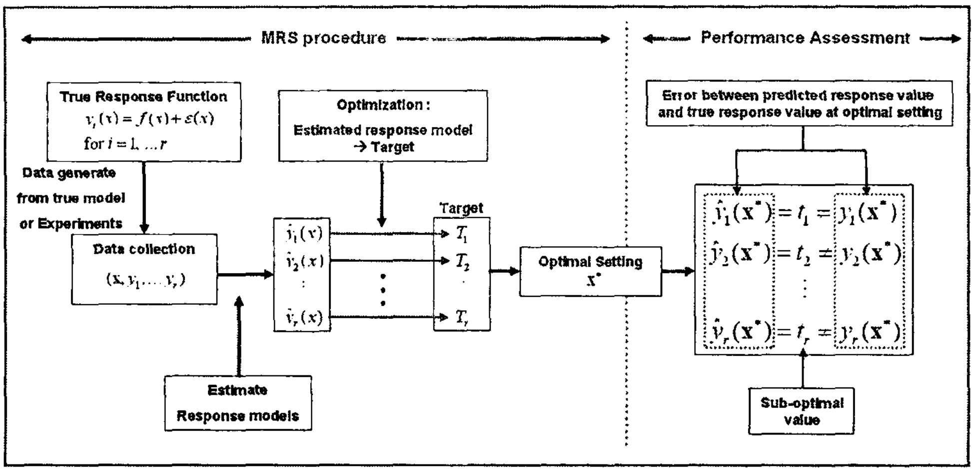

그림 1. MRS의 최적해 결정 과정과 예측성능과 실제 성능과의 비교

식(4)과 같이 정의한다.

$$
\begin{aligned}
&A.\times B \\
&= \begin{bmatrix} a_{11} & \cdots & a_{1j} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ a_{i1} & \cdots & a_{ij} & \cdots & a_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nj} & \cdots & a_{nn} \end{bmatrix} .\times \begin{bmatrix} b_{11} & \cdots & b_{1j} & \cdots & b_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ b_{i1} & \cdots & b_{ij} & \cdots & b_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ b_{n1} & \cdots & b_{nj} & \cdots & b_{nn} \end{bmatrix} \qquad (4) \\
&= \begin{bmatrix} a_{11}b_{11} & \cdots & a_{1j}a_{1j} & \cdots & a_{1n}b_{1n} \\ \vdots & \ddots & \vdots & & \vdots \\ a_{i1}b_{i1} & \cdots & a_{ij}b_{ij} & \cdots & a_{in}b_{in} \\ \vdots & & \vdots & \ddots & \vdots \\ a_{n1}b_{n1} & \cdots & a_{nj}b_{nj} & \cdots & a_{nn}b_{nn} \end{bmatrix}
\end{aligned}
$$

위의 정의를 이용하여, 본 논문에서 제안하는 '*반응 변수 모델의 추정 설명력을 고려한 비용 행렬 추정 방안*'은 아래 식(5)과 같다.

$$[C'] = [C].\times[R_{mat}] \qquad (5)$$

여기서, C는 공정작업자에 의해 주관적으로 결정된 비용 행렬이고 C'는 반응 변수 모델의 추정 설명력을 고려하여 보정된 비용 행렬이다.

제안한 반응 변수 모델의 추정 설명력을 고려하여 보정한 비용 행렬은 다음과 같은 성질을 가진다.

- 보정된 비용 행렬의 각 성분은 각각의 반응 변수가 가진 고유의 중요도를 반응 변수 추정 설명력에 따라 일부 재할당한 의미이다.
- 제안한 방법론에 따라, 추정 모델의 설명력이 높은 반응 변수에 상대적 더 큰 중요도가 할당된 효과가 나타난다.
- 만일 추정 반응 모델이 실제 반응 함수와 일치한다면, 주관적으로 결정한 비용 행렬이 최적이다.

제안된 비용 행렬 추정 방법론은 R^2행렬에 의해, 기존의 비용 행렬의 각 성분이 [0~1] 사이의 서로 다른 비율로 감가상각 된다. 제안한 방법으로 보정된 비용 행렬의 의미는 추정 정확도가 낮은 반응 변수에 할당된 비용을 추정 정확도가 높은 반응 변수에 재할당 함으로써 ***"MRS procedure 효과를 높이는 것"***이다. 즉, 같은 비용으로 가장 정확하게 제어할 수 있는 반응 변수에 더 많은 비용을 할당한다고 볼 수 있다.

4. 수치 예제 실험 결과

수치 예제를 통해 제안한 방법론의 성능을 평가하여 보았다. 실험은 '데이터 생성', 'MRS 최적화', '최적해의 평가'의 3단계로 구성하였다. 데이터 생성은 3개의 입력 변수와 2개의 반응 변수로 이루어진 아래 식(6)과 같은 반응 변수 함수를 이용하여 375개의 sample을 생성하였다.

$$
\begin{aligned}
y_1(x) &= \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} 0.5 \\ 1.5 \\ 2.5 \end{bmatrix}\right)^T \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} 3 \\ 1.5 \\ 1 \end{bmatrix}\right) + \epsilon_1(x) \\
y_2(x) &= \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} 1.5 \\ 2 \\ 0.5 \end{bmatrix}\right)^T \left(\begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} - \begin{bmatrix} 1.5 \\ 0.5 \\ 2.5 \end{bmatrix}\right) + \epsilon_2(x)
\end{aligned} \qquad (6)
$$

여기서 $\epsilon_1(x) \sim N(0, 0.1)$, $\epsilon_2(x) \sim N(0, 3)$인 정규 분포에서 추출된 noise이다. 실험의 2단계에서는 MRS를 최적화하였다. 생성한 데이터를 통해 구한 반응 변수 추정 모델은 아래와 같다.

$$
\begin{aligned}
&\hat{y}_1(x) = -6.24 + 2.50x_1 - 1.49x_3 + 1.00x_1^2 + \\
&1.00x_2^2 + 1.00x_3^2 \\
&(R^2 = 99.9\%) \\
&\hat{y}_2(x) = 4.62 - 3.01x_1 - 2.46x_2 - 2.96x_3 + 0.93x_1^2 \\
&+ 1.02x_2^2 + 1.02x_3^3 + 0.10x_1x_2 + 0.06x_1x_3 - 0.03x_2x_3 \\
&(R^2 = 86.9\%)
\end{aligned} \qquad (7)
$$

최적화 단계에서 이용된 주관적으로 주어진 비용 행렬 값과 목표값은 아래와 같이 정의하였으며, 실험은 목표값을 T_1, T_2, T_3의 3개의 값으로 임의로 정하여 3번 수행하였다.

$$C = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix},\ T_1 = \begin{bmatrix} 10 \\ 20 \end{bmatrix},\ T_2 = \begin{bmatrix} 30 \\ 20 \end{bmatrix},\ T_3 = \begin{bmatrix} 15 \\ 10 \end{bmatrix}$$

실험 결과는 아래와 같다.

표 2. $T_1 = [10\ \ 20]^T$일 때의 결과

비용 행렬	x^*			$\hat{y}(x^*)$		$\bar{y}(x^*)$	
	x_1	x_2	x_3	$\hat{y}_1(x^*)$	$\hat{y}_2(x^*)$	$\bar{y}_1(x^*)$	$\bar{y}_2(x^*)$
C	-2.00	3.27	1.19	14.53	15.47	13.77	15.41
C'	1.05	5.03	0.35	13.01	16.40	12.80	19.26

$T_2 = [30\ \ 20]^T$일 때의 결과

비용 행렬	x^*			$\hat{y}(x^*)$		$\bar{y}(x^*)$	
	x_1	x_2	x_3	$\hat{y}_1(x^*)$	$\hat{y}_2(x^*)$	$\bar{y}_1(x^*)$	$\bar{y}_2(x^*)$
C	3.20	3.69	-2.18	33.61	16.39	33.57	14.94
C'	2.27	3.49	-2.76	30.00	20.00	28.92	22.64

$T_3 = [15\ \ 10]^T$일 때의 결과

비용 행렬	x^*			$\hat{y}(x^*)$		$\bar{y}(x^*)$	
	x_1	x_2	x_3	$\hat{y}_1(x^*)$	$\hat{y}_2(x^*)$	$\bar{y}_1(x^*)$	$\bar{y}_2(x^*)$
C	2.10	0.09	-1.92	13.47	11.53	13.57	15.83
C'	0.73	1.10	-2.12	14.07	11.11	14.40	10.41

여기서 $\hat{y}(x^*)$는 구한 최적 입력 값인 x^*에 대한 반응 변수 추정 모델의 값이며 $\bar{y}(x^*)$는 실제 반응 변수의 기댓값이다. 실험 결과, 반응 변수의 추정 값인 $\hat{y}(x^*)$는 비용 행렬을 보정하기 전과 보정한 후 모두 목표값에 거의 유사하게 나왔다. 하지만 x^*를 실제 공정에 적용한 값인 실제 반응 변수의 기대값 $\bar{y}(x^*)$는 비용 행렬을 보정한 값이 보정하지 않은 값보다 상대적으로 목표값에 더 가깝게 결과가 나왔다. 이는 반응 변수 추정 모델의 R^2값이 $\hat{y}_1$보다 $\hat{y}_2$가 작기 때문에 $\hat{y}_2$에 할당될 비용이 상대적으로 추정 설명력이 높은 $\hat{y}$에 재할당됨으로써 보정하기 전보다 좀 더 효율적인 비용할당이 되었기 때문이다. 이는 "설명력이 떨어지는 y_2의 중요도를 낮추고, 설명력이 큰 y_1의 상대적 중요도를 높였다는 의미"를 지닌다.

5. 토의 및 결론

본 논문에서는 주관적으로 결정된 비용 행렬을 반응 변수 모델의 추정 설명력을 반영하여 보정하는 방법론을 제안하였다. 제안한 비용 행렬 보정의 목적은 추정 설명력이 낮은 반응 변수에 할당된 비용을 설명력이 높은 반응 변수에 재할당함으로써 MRS optimization의 효율을 높이는 것이다. 즉, 비용 행렬 보정을 통해 같은 비용으로 가장 정확하게 제어할 수 있는 반응 변수에 더 많은 비용(상대적 중요도)을 할당한다고 볼 수 있다. 추정 설명력이 좋은 반응 변수($R^2 = 99.9\%$)와 좋지 않은 반응 변수($R^2 = 89.9\%$)로 이루어진 데이터를 이용하여 제안한 비용 행렬 보정 방법론을 적용한 결과, 제안한 보정된 비용 행렬을 이용한 경우, MRS Optimization의 효율이 높아졌음을 알 수 있었다.

본 논문에서는 제안한 방법론은 임의로 정의한 함수에서 생성된 데이터만을 이용하여 실험하였기때문에, 실제 공정 데이터에 대한 추가 실험으로 제안한 방법론의 성능을 일반화하여 검증할 필요가 있다. 아울러, 주관적 비용 행렬의 체계적 결정 단계와 추정 설명력을 반영할 새로운 방안은 중요한 추후 연구 과제라 할 수 있다.

Reference

[1] Pignatiello, "Strategies for Robust Multi-response Quality Engineering", IIE Transactions, Vol.25, No.3, 1993.

[2] Kim & Lin, "Simultaneous optimization of mechanical properties of steel by maximizing exponential desirability functions", 2000

[3] Vining and Bohn, "Response Surface for the mean and Variance Using a Nonparametric Approach", Joanal of Quality Technology, Vol.30, No.3, 1998.

[4] Khuri and Colon, "Simultaneous optimization of multiple responses represented by polynomial regression functions", Technometrics, Vol. 23.

[5] Ko, Kim and Jun., "A New Loss Function -Based Method for Multiresponse Optimization", Journal of Quality Technology, 2004 (To appear).