⊠연구논문

Simultaneous Optimization of Multiple Responses Using Weighted Desirability Function *

Sung Hyun Park · Jun Oh Park
Dept. of Computer Science and Statistics, Seoul National University

Abstract

The object of multiresponse optimization is to determine conditions on the independent variables that lead to optimal or nearly optimal values of the response variables. Derringer and Suich (1980) extended Harrington's (1965) procedure by introducing more general transformations of the response into desirability functions. The core of the desirability approach condenses a multivariate optimization into a univariate one. But because of the subjective nature of this approach, inexperience on the part of the user in assessing a product's desirability value may lead to inaccurate results. To compensate for this defect, a weighted desirability function is introduced which takes into consideration the variances of the responses.

1. Introduction

In many experimental situations, it is quite common that several responses, rather than a single response, are measured from each setting of a group of input variables. The analysis of data from a multiresponse experiment requires careful consideration of the multivariate nature of the data. In other words, the response variables should not be investigated individually and independently of one another. Interrelationship that may exist among the responses can render univariate investigation meaningless. For example, if we desire to optimize several response functions simultaneously, it would be futile to obtain separate individual

* This study was partially supported by Korean Ministry of Education through Research Fund, BSRI-96-1415.

optima. So in a multivariate situation, the optimization problem is more complex than in the single response case. The main difficulty stems from the fact that two or more response variables are under investigation simultaneously, and the meaning of optimum becomes unclear since there is no unique way to order multivariate values of a multiresponse function. Furthermore, the optimal condition for one response may be far from optimal or even physically impractical conditions for the remaining responses.

The object of the multiresponse optimization is to determine conditions on the input variables that lead to optimal or nearly optimal values of the response variables. In an effort to find optimal conditions on several responses, the desirability function approach will be introduced in the next section. The other methods are briefly outlined as follows.

(1) Graphical superimposition method

In case there are only two or three input variables, this method is not only easy to understand and use, but also simple and straightforward. First of all, for each response, a response contour is obtained by fitting, in general, the second-order response model. Then by superimposition of response contours, we arrive at optimal conditions. Even though this procedure is practically very useful, it is difficult to apply when the number of input variables exceeds three, and to identify one set of conditions or one point in the experimental region as being optimal.

(2) Primary and secondary function

Myers and Carter (1973) introduced an algorithm for determining conditions on the input variables that maximize or minimize a primary response function subject to having an equality constant on a secondary response function. In other words, the secondary response function imposes certain constraints on the optimization of the primary response function. Biles (1975) extended this approach to include several secondary response functions within specified ranges. Biles's procedure employs a modification of the method of steepest ascent described by Box and Wilson (1951). In many cases, it is necessary to optimize the responses simultaneously rather than to optimize one response with the other constraints. Therefore, this method may be adapted to the restricted cases. Recently Vining and Myers (1990) proposed a dual response approach which combines Taguchi and response surface methods. Lin and Tu (1995) developed Vining and Myers's approach using the MSE(Mean Squared Error) criterion.

(3) Distance function approach

Khuri and Conlon (1981) presented several distance functions that measure the overall closeness of the response functions to achieving their respective optimal values at the same set of operating conditions. Multiresponse optimization is thus reduced to minimizing an appropriate distance function with respect to the input variables. This approach permits the user to account for the variances and covariances of the estimated responses and for the random error variation associated with the estimated ideal optimum.

(4) P_M and P_V measures

Park, Kwon, and Kim (1995) studied simultaneous optimization of multiple responses for robust design. They suggested P_M and P_V measures. The $P_M$measure can be used without a prior knowledge about the estimated mean responses. The P_V measure is reasonable to minimize the variances when we have a prior knowledge about the mean responses. P_V is simple and easy to compute. P_V also allows the user to make a decision on the range of the estimated mean responses. However, because of inappropriate decision on the range of mean responses, it may lead to inaccurate results.

2. Desirability function approach

2.1 Desirability function

Suppose each of the k response variables is related to the p independent variables by

$$y_i = f_i(x_1, x_2, \cdots, x_p) + \varepsilon_i, \quad i = 1, 2, \cdots, k$$

where f_i denotes the functional relationship between y_i and $x_1, x_2, \cdots, x_p$. If we make the usual assumption that $E(\varepsilon_i) = 0$ and $Var(\varepsilon_i) = \sigma_i^2$ for each i, then $E(y_i) = \eta_i = f_i(x_1, x_2, \cdots, x_p)$, $i = 1, 2, \cdots, k$, where η_i is represented by second order models within a certain region of interest in general.

The desirability function involves transformation of each estimated response variable $\hat{y}_i(=\hat{\eta}_i)$ to a desirability value d_i, where $0 \le d_i \le 1$. The value of d_i increases as the desirability of the corresponding response increases. The individual desirabilities are then combined using the geometric mean G,

$$G = (d_1 \times d_2 \times \cdots \times d_k)^{\frac{1}{k}}.$$

When k is large, the variation of d_i has much influence on G. So Park (1981) suggested harmonic mean H,

$$H = \frac{k}{\frac{1}{d_1} + \frac{1}{d_2} + \cdots + \frac{1}{d_k}} = \frac{k \prod_{j=1}^{k} d_j}{\sum_{i=1}^{k} \prod_{j \ne i}^{k} d_j}.$$

This single value of G or H gives the overall assessment of the desirability of the combined response levels. If any $d_i = 0$ (that is, if one of the response variables is unacceptable), then G or H is 0 (that is, the overall product is unacceptable).

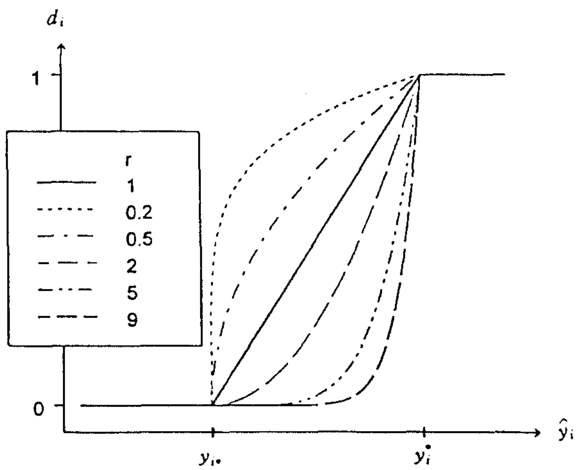
2.2 One-sided transformation (maximization of $\hat{y}_i$)

Let d_i be the ith individual desirability function, which is usually defined by

$$d_i = \begin{cases} 0 & \hat{y}_i \le y_{i*} \\ \left[\dfrac{\hat{y}_i - y_{i*}}{y_i^* - y_{i*}} \right]^r & y_{i*} \le \hat{y}_i \le y_i^* \\ 1 & \hat{y}_i \ge y_i^* \end{cases} \qquad (1)$$

where y_{i*} is the minimum acceptable value of $\hat{y}_i$ and y_i^* is the satisfactory value of $\hat{y}_i$ for $i = 1, 2, \cdots, k$ and r is an arbitrary positive constant. A large value of r would be specified if it were desirable for the value of $\hat{y}_i$ to increase rapidly above y_{i*}. On the other hand, a small value of r would be specified if having values of $\hat{y}_i$ considerably above y_{i*} were not of critical importance. For this reason the desirability function approach permits the user to make subjective judgements on the importance of each response. This is attractive to an experienced user. However, because of the subjective nature of the desirability approach, inexperience on the part of the user in assessing a product's desirability value may lead to inaccurate results. That is, the choice of r value contains user's subjective judgements and an inappropriate r value may result in the improper optimum condition.

Minimization of $\widehat{y}_i$ is equivalent to maximization of $-\widehat{y}_i$. Therefore, minimization of $\widehat{y}_i$ is not elaborated here.



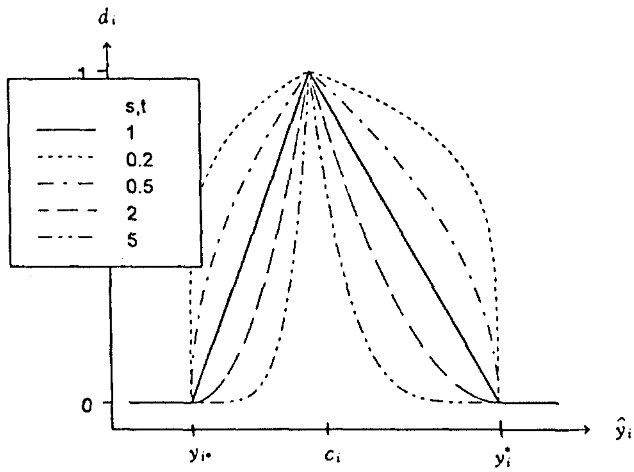
< Figure 1 > Transformation (1) for Various Values of r

2.3 Two-sided transformation

When the response variable y_i has both a minimum acceptable value and a maximum acceptable value, the individual desirability function is defined by

$$d_i = \begin{cases} \left[\dfrac{\widehat{y}_i - y_{i*}}{c_i - y_{i*}} \right]^s & y_{i*} \le \widehat{y}_i \le c_i \\ \left[\dfrac{\widehat{y}_i - y_i^*}{c_i - y_i^*} \right]^t & c_i \le \widehat{y}_i \le y_i^* \\ 0 & \widehat{y}_i \le y_{i*} \text{ or } \widehat{y}_i \ge y_i^* \end{cases} \tag{2}$$

where c_i is the target value for the ith response, and s and t are arbitrary positive constants. In this situation, y_{i*} is the minimum acceptable value of $\widehat{y_i}$ and y_i^* is the maximum acceptable value. The values of s and t in the two-sided transformation play the same role as that of r does in the one-sided transformation.



< Figure 2 > Transformation (2) for Various Values of s and t

Since $\widehat{y_i}$ is a continuous function of the $x_i, i=1,2,\cdots,p$, both G and H are continuous functions of d_i's, respectively. Therefore, it follows that both G and H are continuous of the $x_i, i=1,2,\cdots,p$. As a result, existing univariate search techniques can be used to maximize G or H over the independent variables domain.

3. Weighted desirability function approach

The desirability functions (1) and (2) assume that, when y_{i*} and y_i^* are determined by the user, each y_i has the same degree of importance whether y_i's have different degrees of importance or different variances. Such assumption is not practical in real situations. Therefore, we propose here the weighted desirability function

$$\mathrm{WG}=\left(d_1^{\ w_1}\times d_2^{\ w_2}\times\cdots\times d_k^{\ w_k}\right)^{\frac{1}{k}},\quad \sum_{i=1}^{k} w_i=k,\quad w_i > 0,\quad i=1,2,\cdots,k$$

where w_i is the weight for y_i, and the average of w_i's is 1. If y_i's are all equally important or they have equal variances, we may use $w_i=1$ for each y_i.

If we have no prior information for the importance of each y_i, a good choice of w_i is to make w_i proportional to the coefficient of variation of y_i,

$$\mathrm{CV}_i=\frac{\sqrt{MSE_i}}{\bar{y}_i}$$

where $\bar{y}_i$ is the sample mean of the ith response, and $\sqrt{MSE_i}$ is the estimate of σ_i obtained from fitting a regression model. We may say that WG redistributes each weight in G according to each CV_i. Note that WG decreases as its weight w_i increases since d_i varies in [0, 1]. If any $d_i=0$, WG is 0 like G. That is, if one of the response is unacceptable, then the overall product is unacceptable.

Each weighted desirability function is in one-sided transformation

$$d_i^{w_i}=\begin{cases} 0 & \hat{y}_i\le y_{i*} \\ \left[\dfrac{\hat{y}_i-y_{i*}}{y_i^*-y_{i*}}\right]^{rw_i} & y_{i*}\le\hat{y}_i\le y_i^* \\ 1 & \hat{y}_i\ge y_i^* \end{cases} \tag{3}$$

and in two-sided transformation

$$d_i^{w_i}=\begin{cases} \left[\dfrac{\hat{y}_i-y_{i*}}{c_i-y_{i*}}\right]^{sw_i} & y_{i*}\le\hat{y}_i\le c_i \\ \left[\dfrac{\hat{y}_i-y_i^*}{c_i-y_i^*}\right]^{tw_i} & c_i\le\hat{y}_i\le y_i^* \\ 0 & \hat{y}_i\le y_{i*}\ \text{or}\ \hat{y}_i\ge y_i^*. \end{cases} \tag{4}$$

The constants r, s and t play the same role as those of G.

The performance of this weighted desirability function is illustrated in the following example.

4. Example

In the development of a tire tread compound, the optimal combination of three ingredient(independent) variables - hydrated silica level x_1, silane coupling agent level x_2, and sulfur level x_3 - was sought. The properties to be optimized and constraint levels were as follows.

PICO abrasion Index, y_1	$120 < y_1$	$y_{1*} = 120$		
200% Modulus, y_2	$1000 < y_2$	$y_{2*} = 1000$		
Elongation at Break, y_3	$400 < y_3 < 600$	$y_{3*} = 400$	$y_3^* = 600$
Hardness, y_4	$60 < y_4 < 75$	$y_{4*} = 60$	$y_4^* = 75$

For y_1 and y_2, the one-sided transformation given by (3) was used and for y_3 and y_4, the two-sided transformation given by (4) was used.

In the example given in Derringer and Suich(1980), they employed the rotatable central composite design with six center points in three variables. <Table 1> shows the data which were then fitted to the second degree polynomial models,

$$\widehat{y_i} = b^i_0 + \sum_{s=1}^{3} b^i_s x_s + \sum_{s=1}^{3}\sum_{t=s}^{3} b^i_{st} x_s x_t \quad i = 1, 2, 3, 4.$$

The resultant fitted equations are

$$\widehat{y_1} = 139.12 + 16.49x_1 + 17.88x_2 + 10.91x_3 - 4.01x_1^2 - 3.45x_2^2 - 1.57x_3^2$$
$$+ 5.13x_1x_2 + 7.13x_1x_3 + 7.88x_2x_3$$

$$\widehat{y_2} = 1261.11 + 268.15x_1 + 246.50x_2 + 139.48x_3 - 83.55x_1^2 - 124.79x_2^2 + 199.17x_3^2$$
$$+ 69.38x_1x_2 + 94.13x_1x_3 + 104.38x_2x_3$$

$$\widehat{y_3} = 400.38 - 99.67x_1 - 31.40x_2 - 73.92x_3 + 7.93x_1^2 + 17.31x_2^2 + 0.43x_3^2 + 8.75x_1x_2 + 6.25x_1x_3 + 1.25x_2x_3$$

$$\widehat{y_4} = 68.91 - 1.41x_1 + 4.32x_2 + 1.63x_3 + 1.56x_1^2 + 0.06x_2^2 - 0.32x_3^2 - 1.63x_1x_2 + 0.13x_1x_3 - 0.25x_2x_3.$$

< Table 1> Experimental Design

x_1	x_2	x_3	y_1	y_2	y_3	y_4
-1	-1	1	102	900	470	67.5
1	-1	-1	120	860	410	65
-1	1	-1	117	800	570	77.5
1	1	1	198	2294	240	74.5
-1	-1	-1	103	490	640	62.5
1	-1	1	132	1289	270	67
-1	1	1	132	1270	410	78
1	1	-1	139	1090	380	70
-1.633	0	0	102	770	590	76
1.633	0	0	154	1690	260	70
0	-1.633	0	96	700	520	63
0	1.633	0	163	1540	380	75
0	0	-1.633	116	2184	520	65
0	0	1.633	153	1784	290	71
0	0	0	133	1300	380	70
0	0	0	133	1300	380	68.5
0	0	0	140	1145	430	68
0	0	0	142	1090	430	68
0	0	0	145	1260	390	69
0	0	0	142	1344	390	70

The $\overline{y}_i, \sqrt{MSE_i}$, $CV_i \times 100$, and w_i for each y_i are given in <Table 2>.

< Table 2 > Each weight proportional to CV

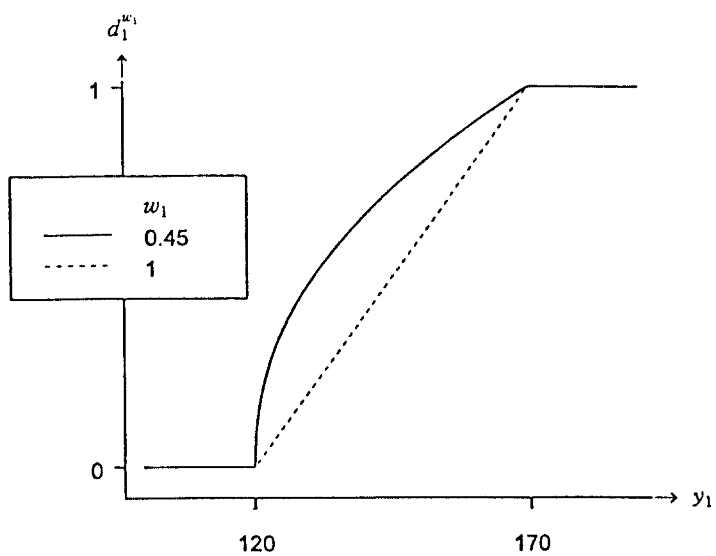
i	$\overline{y}_i$	$\sqrt{MSE_i}$	$CV_i \times 100$	w_i
1	133.1	5.61	4.22	0.45
2	1255.0	328.69	26.19	2.82
3	417.5	20.55	4.92	0.53
4	69.8	1.27	1.82	0.20

We use the grid-search method to find an optimum formulation. <Table 3> shows the difference in the optimum condition and the estimated value of each response between G and WG.

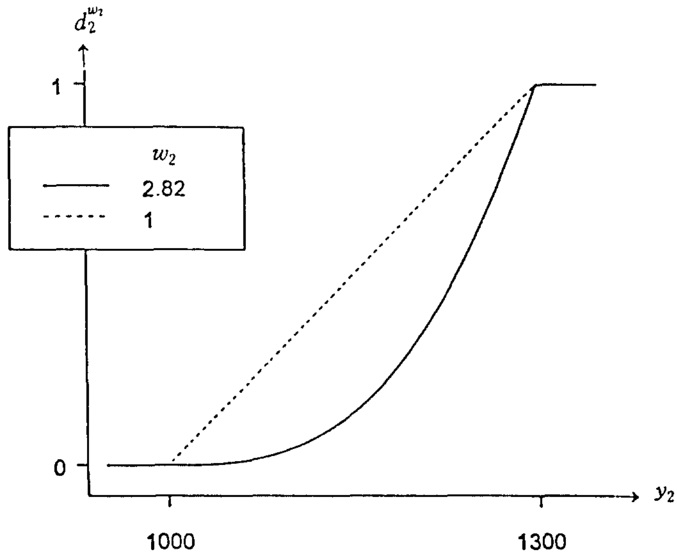
< Table 3 > Different optimal conditions under G and WG

	G	WG
	x_1=-0.050 x_2=0.145 x_3=-0.868	x_1=-0.158 x_2=0.437 x_3=-0.879
$120 < y_1$	129.5	130.38
$1000 < y_2$	1300.0	1300.02
$400 < y_3 < 600$	465.7	471.00
$60 < y_4 < 75$	68.0	69.62

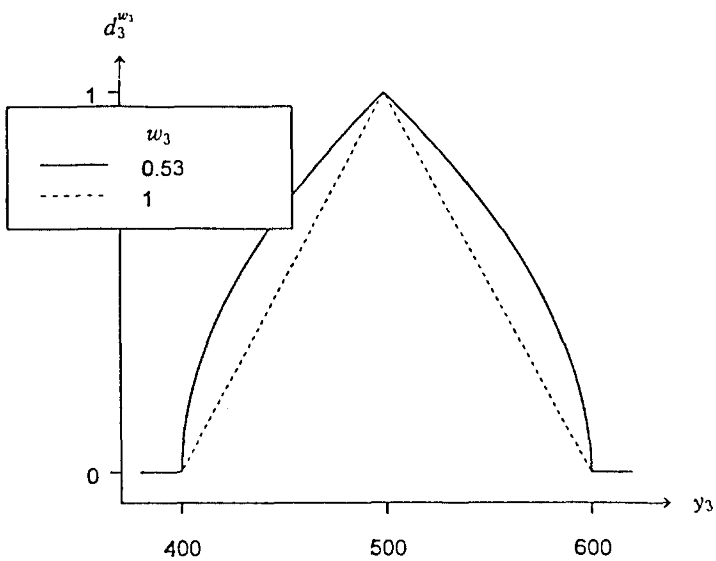
Figures 3-6 show the performance of the $d_i^{w_i}$, i=1, 2, 3, 4. We assume that r=1 and s=t=1 in (3) and (4).



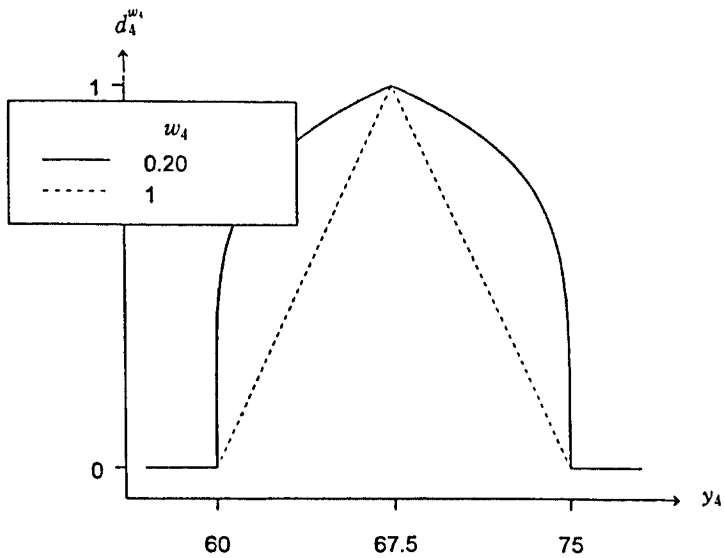
< Figure 3 > The performance of d_1



< Figure 4 > The performance of d_2



< Figure 5 > The performance of d_3



< Figure 6 > The performance of d_4

5. Conclusion

Since the desirability function condenses a multivariate optimization problem into a univariate one and WG is continuous of the x_i, $i=1,2,\cdots,p$ as G is, we have only to consider the univariate techniques to find the maximum of weighted desirability function. In this paper, the grid-search method was used to find the maximum value.

We can see the different optimum conditions between WG and G. The estimate value of the response of which CV_i is large is moved to the desirable point (maximum value in one-sided transformation or target value in two-sided transformation) in WG. To the contrary, the estimate value of the response whose CV_i is small is moved to the opposite direction. Therefore, if the response whose CV_i is relatively large is more important than others, it is useful to employ WG instead of using G.

References

[1] Biles, W.E.(1975), "A response surface method for experimental optimization of multiresponse process," *Industrial and Engineering Chemistry Process Design and Development*, Vol. 14, pp. 152-158.

[2] Box, G.E.P. and Wilson, K.B.(1951) "On the experimental attainment of optimum conditions," *Journal of Royal Statistics Society*. Vol. 13, pp. 1-45.

[3] Derringer, G. and Suich, R.(1980), "Simultaneous optimization of several response variables," *Journal of Quality Technology*, Vol. 12, pp. 214-219.

[4] Harrington, E.C. Jr.(1965), "The Desirability Function," *Industrial Quality Control*, Vol. 21, No. 10, pp. 494-498.

[5] Khuri, A.I. and Conlon, M.(1981), "Simultaneous optimization of multiple response represented by polynomial regression function," *Technometrics*, Vol. 23, pp. 363-375.

[6] Khuri, A.I. and Cornell, J.A.(1987), *Response Surfaces*, Marcel Dekker, New York.

[7] Lin, D.J. and Tu, W.(1995), "Dual Response Surface Optimization," *Journal of Quality Technology*, Vol. 27, pp. 34-39.

[8] Myers, R.H. and Carter, W.H., Jr.(1973), "Response surface techniques for dual response systems," *Technometrics*, Vol. 15, pp. 301-317.

[9] Myers, R.H., Carter, W.H., Jr. and Khuri, A.I.(1989), "Response Surface Methodology: 1966-1988," *Technometrics*, Vol. 31, pp. 137-157.

[10] Park, S.H.(1981), "Simultaneous Optimization Techniques for Multi-purpose Response Functions," *Journal of Military Operations Research Society of Korea*, Vol. 7, pp. 118-138.

[11] Park, S.H., Kwon, Y.M. and Kim, J.J.(1995), "Simultaneous optimization of multiple responses for robust design," *Total Quality Management*, pp. 381-390, Chapman & Hall, London.

[12] Park, S.H.(1991), *Regression Analysis* , Minyoungsa, Seoul.

[13] Vining, G.G. and Myers, R.H.(1990), "Combining Taguchi and Response Surface Philosophies: A Dual Response Approach," *Journal of Quality Technology*, Vol. 22, pp. 38-45.