▨응용논문

편최소제곱 반응표면함수를 이용한 공정 최적화에 관한 연구*

박성현
서울대학교 통계학과
최엄문
서울대학교 통계연구소
박창순
중앙대학교 응용통계학과

A Study on Process Optimization Using Partial Least Squares Response Surface Function

Sung-Hyun Park
Dept. of Statistics, Seoul National University
Um-Moon Choi
Statistical Research Institute, Seoul National University
Chang-Soon Park
Dept. of Aplied Statistics, Chungang University

Abstract

Response surface analysis has been a popular tool conducted by engineers in many processes. In this paper, response surface function, named partial least squares response surface function is proposed. Partial least squares response surface function is a function of partial least squares components and the response surface modeling is used in either a first-order or a second-order model. Also, this approach will have the engineers be able to do the response surface modeling and the process optimization even when the number of experimental runs is less

* This study was partially supported by Korean Ministry of Education through Research Fund, 1998-015-D00046.

than the number of model parameters. This idea is applied to the nondesign data and an application of partial least squares response surface function to the process optimization is considered.

1. Introduction

When there are exact dependencies among the columns of a matrix, that is, when one or more columns can be exactly expressed as linear combinations of other columns, it is said that there exists collinearity. In statistics the term multicollinearity is used in situations where the variables are mutually collinear because of high correlations among the variables. If there is multicollinearity among the variables in regression analysis, the least squares estimates can not be obtained or, even if they are obtained, they are unstable.

Some well-devised techniques can be used to reduce the effects of multicollinearity in regression analysis. For example, there are such techniques as principal component regression(Massy, 1965) and ridge regression(Hoerl and Kennard, 1970). A method, called *partial least squares* (PLS) regression, has been developed in the field of chemometrics to reduce the effect of multicollinearity of explanatory variables to the results of regression analysis. It has worked well in many chemical problems and has become one of the most popular regression methods in chemometrics. In recent simulation studies for PLS regression its good performance has been known to statisticians (see, e.g., Frank and Friedman, 1993). As the PLS regression algorithm originally proposed by Wold(1975) was difficult to be understood compared with other regression methods such as ordinary least squares, principal component regression, and ridge regression, several algorithms have been developed to look at PLS regression in a new and easy light (see, e.g., Martens, 1985; Helland, 1988).

This paper discusses PLS response surface function which is written as

$$\eta = g(T_1, \cdots, T_q), \quad q \leq p,$$

where T_i's are PLS components and p is the number of independent variables. If we suppose the first-order model, the function g is the first-order function. If we suppose the second-order model, the function g is the second-order function. The case of single response variable is also considered. This idea is applied to the

cases of the actual process, nondesign data and $n \leq p$ in the first-order model, $n < (p+1)(p+2)/2$ in the second-order model, and the general case, nondesign data and $n > p$ in the first-order model, $n \geq (p+1)(p+2)/2$ in the second-order model. As for PLS response surface function, the response surface modeling and the process optimization are considered.

2. Partial Least Squares Response Surface Function

In this section we consider PLS response surface function which is written as

$$\eta = g(T_1, \cdots, T_q), \quad q \leq p, \tag{2.1}$$

where T_i's are PLS components and either a first-order or a second-order model is used. At first, the proposed algorithm using algorithm of Wold for composing PLS components is as follows.

1.1 Coding formula:

$$x_i = \frac{2X_i - (X_{iL} + X_{iH})}{X_{iH} - X_{iL}}, \quad i = 1, \cdots, p,$$

where X_{iL} and X_{iH} are the low and high values of X_i, respectively, X_i is the $i-$ th process variable.

1.2 Initialize(Centering):

$$\mathbf{X}_0 \leftarrow (\mathbf{X} - \mathbf{1}\ \bar{\mathbf{x}}'), \qquad \mathbf{y}_0 \leftarrow (\mathbf{y} - \bar{y}\ \mathbf{1}),$$

where n and p indicate the number of observations and process variables, respectively, $\mathbf{X}$ is an $n \times p$ matrix of process variables, $\mathbf{y}$ is an $n-$dimensional vector of response variable, $\mathbf{1}$ is a unit vector with 1 in its all elements, $\bar{\mathbf{x}}$ is the mean vector of $\mathbf{X}$, and $\bar{y}$ is the mean of $\mathbf{y}$.

1.3 For $i=1,2,\cdots,q$, compute

$$\mathbf{w}_i = \mathbf{X}'_{i-1}\ \mathbf{y}_{i-1}$$

$$\mathbf{t}_i = \mathbf{X}_{i-1}\ \mathbf{w}_i,$$

where T_i's are the PLS components and T_i is called the ith component variable.

$$\mathbf{p}_i = \frac{\mathbf{X}'_{i-1}\ \mathbf{t}_i}{\mathbf{t}'_i\ \mathbf{t}_i} \quad \left(= \frac{\mathbf{X}'\ \mathbf{t}_i}{\mathbf{t}'_i\ \mathbf{t}_i}\right)$$

$$q_i = \frac{\mathbf{y}'_{i-1}\ \mathbf{t}_i}{\mathbf{t}'_i\ \mathbf{t}_i} \quad \left(= \frac{\mathbf{y}'\ \mathbf{t}_i}{\mathbf{t}'_i\ \mathbf{t}_i}\right)$$

$$\mathbf{X}_i = \mathbf{X}_{i-1} - \mathbf{t}_i\ \mathbf{p}'_i$$

$$\mathbf{y}_i = \mathbf{y}_{i-1} - \mathbf{t}_i\ q_i\ .$$

Suppose that the engineer is concerned with a process involving a response Y that depends on the PLS components $T_1,\cdots,T_q$. The PLS components, T_i, represent the linear combination of x_j,

$$T_i = c_{i0} + \sum_{j=1}^{p} c_{ij}\, x_j, \quad i=1,\cdots,q, j=1,\cdots,p,$$

where c_{i0} and c_{ij} are constants. The relationship is

$$Y = g(T_1,\cdots,T_q) + \varepsilon, \tag{2.2}$$

where the form of the true response function g is unknown and perhaps very complicated, and ε is the term that represents other sources of variability not accounted for in g.

The case is considered where there is one response variable, Y, and p process variables, $X_1,\cdots,X_p$. The PLS form of a first-order model in p input variables $X_1,\cdots,X_p$ is

$$Y=\delta_0+\sum_{i=1}^{q}\delta_i T_i+\varepsilon, \tag{2.3}$$

where $q\leq p$, Y is an observable response variable, $\delta_0, \delta_1, \cdots, \delta_q$ are unknown parameters, and ε is a random error term. If ε has a zero mean, then the nonrandom portion of the model in (2.3) represents the true mean response, η, that is,

$$\eta=\delta_0+\sum_{i=1}^{q}\delta_i T_i \tag{2.4}$$

and ε in (2.3) is regarded as the experimental error.

As an introduction to the construction of the first-order model, let us write the first-order model, over n observations, in matrix form as

$$\mathbf{y}= \mathbf{T}_q \ \boldsymbol{\delta}_q+ \boldsymbol{\varepsilon}, \tag{2.5}$$

where $\mathbf{y}$ is a vector of n observations, $\boldsymbol{\delta}_q=(\delta_0, \delta_1, \cdots, \delta_q)'$ is a $(q+1)\times 1$ vector of unknown parameters, $\boldsymbol{\varepsilon}=(\varepsilon_1, \cdots, \varepsilon_n)'$ is an $n\times 1$ vector of errors, and

$$\mathbf{T}_q=\begin{bmatrix} 1 & T_{11} & T_{12} & \cdots & T_{1q} \\ 1 & T_{21} & T_{22} & \cdots & T_{2q} \\ \vdots & \vdots & \vdots & \cdots & \vdots \\ 1 & T_{n1} & T_{n2} & \cdots & T_{nq} \end{bmatrix},$$

is an $n\times(q+1)$ matrix of settings of the PLS components. More specifically, the $\mathbf{T}_q$ matrix is of the form $\mathbf{T}_q=[\ \mathbf{1}:\ \mathbf{C}_T]$ where $\mathbf{1}$ is an $n\times 1$ column vector of ones and $\mathbf{C}_T$ is an $n\times q$ matrix. The matrix $\mathbf{C}_T$ will be referred to as the *component matrix*. We assume that the random errors are independently distributed as normal variables with zero mean and common variance σ^2.

Then the least squares estimator of $\boldsymbol{\delta}_q$ is

$$\mathbf{b}_q=(\ \mathbf{T}_q'\ \mathbf{T}_q)^{-1}\ \mathbf{T}_q'\ \mathbf{y}. \tag{2.6}$$

Hence, the fitted response surface model is

$$\hat{Y}= b_0+ b_1\ T_1+\cdots+ b_q\ T_q. \tag{2.7}$$

In the full model,

$$\mathbf{y}=\ \mathbf{T}\ \boldsymbol{\delta}+\ \boldsymbol{\varepsilon},$$

where $\mathbf{T}$ matrix is partitioned by $\mathbf{T}=[\ \mathbf{T}_q,\ \mathbf{T}_r]$, hence, model is rewritten by

$$\mathbf{y}=\ \mathbf{T}_q\ \boldsymbol{\delta}_q+\ \mathbf{T}_r\ \boldsymbol{\delta}_r+\ \boldsymbol{\varepsilon},$$

where $\mathbf{T}_q$ includes intercept and q components, and $\mathbf{T}_r$ includes $p-q$ components. The MSE of $\boldsymbol{\delta}_q$ is

$$\mathrm{MSE}(\ \mathbf{b}_q)=(\ \mathbf{T}_q'\ \mathbf{T}_q)^{-1}\ \sigma^2+\ \mathbf{A}\ \ \boldsymbol{\delta}_r\ \ \boldsymbol{\delta}_r'\ \mathbf{A}',$$

because

$$\begin{aligned}
E(\ \mathbf{b}_q) &= E[(\ \mathbf{T}_q'\ \mathbf{T}_q)^{-1}\ \mathbf{T}_q'\ \mathbf{y}] \\
&= (\ \mathbf{T}_q'\ \mathbf{T}_q)^{-1}\ \mathbf{T}_q'\ \mathbf{T}\boldsymbol{\delta} \\
&= (\ \mathbf{T}_q'\ \mathbf{T}_q)^{-1}\ \mathbf{T}_q'[\ \mathbf{T}_q\ \boldsymbol{\delta}_q+\ \mathbf{T}_r\ \boldsymbol{\delta}_r] \\
&= \boldsymbol{\delta}_q+(\ \mathbf{T}_q'\ \mathbf{T}_q)^{-1}\ \mathbf{T}_q'\ \mathbf{T}_r\ \boldsymbol{\delta}_r \\
&= \boldsymbol{\delta}_q+\ \mathbf{A}\ \boldsymbol{\delta}_r,
\end{aligned}$$

$$\begin{aligned}
\mathrm{Var}(\ \mathbf{b}_q) &= \mathrm{Var}[(\ \mathbf{T}_q'\ \mathbf{T}_q)^{-1}\ \mathbf{T}_q'\ \mathbf{y}] \\
&= (\ \mathbf{T}_q'\ \mathbf{T}_q)^{-1}\ \mathbf{T}_q'\ \mathbf{I}\sigma^2(\ \mathbf{T}_q'\ \mathbf{T}_q)^{-1}\ \mathbf{T}_q \\
&= (\ \mathbf{T}_q'\ \mathbf{T}_q)^{-1}\sigma^2,
\end{aligned}$$

where $\mathbf{A}=(\mathbf{T}_q'\mathbf{T}_q)^{-1}\mathbf{T}_q'\mathbf{T}_r$,

$$\mathbf{T}_q=\begin{bmatrix}1 & T_{11} & T_{12} & \cdots & T_{1q}\\ 1 & T_{21} & T_{22} & \cdots & T_{2q}\\ \vdots & \vdots & \vdots & \cdots & \vdots\\ 1 & T_{n1} & T_{n2} & \cdots & T_{nq}\end{bmatrix},$$

$$\mathbf{T}_r=\begin{bmatrix}T_{1(q+1)} & T_{1(q+2)} & \cdots & T_{1p}\\ T_{2(q+1)} & T_{2(q+2)} & \cdots & T_{2p}\\ \vdots & \vdots & \cdots & \vdots\\ T_{n(q+1)} & T_{n(q+2)} & \cdots & T_{np}\end{bmatrix}.$$

The least squares estimator, $\mathbf{b}$, is rewritten by

$$\mathbf{b}=\begin{bmatrix}\mathbf{b}_q\\ \mathbf{b}_r\end{bmatrix}=(\mathbf{T}'\mathbf{T})^{-1}\mathbf{T}'\mathbf{y},$$

because

$$\begin{aligned}E(\hat{Y}_q) &= \mathbf{t}'_q\boldsymbol{\delta}_q+\mathbf{t}'_q\mathbf{A}\boldsymbol{\delta}_r,\\ \mathrm{Var}(\hat{Y}_q) &= \mathbf{t}'_q(\mathbf{T}_q'\mathbf{T}_q)^{-1}\mathbf{t}_q\sigma^2,\\ \mathrm{MSE}(\hat{Y}_q) &= \mathbf{t}'_q(\mathbf{T}_q'\mathbf{T}_q)^{-1}\mathbf{t}_q\sigma^2+(\mathbf{t}'_q\mathbf{A}\boldsymbol{\delta}_r-\mathbf{t}'_r\boldsymbol{\delta}_r)^2.\end{aligned}$$

We can consider several criteria for selection of the number of components, q. We propose the residual mean square,

$$\mathrm{MSE}_q=\frac{\mathrm{SSE}_q}{n-q-1},$$

the coefficient of determination.

$$R^2_q = 1 - \frac{SSE_q}{SST},$$

the adjusted coefficient of determination,

$$R^2_{aq} = 1 - (1 - R^2_q)\frac{n-1}{n-q-1},$$

and the total squared error,

$$C_q = \frac{SSE_q}{\hat{\sigma}^2} + 2(q+1) - n,$$

where $\hat{\sigma}^2$ is MSE in the full model.

We can consider a sequential F-test to select the number of components. When we select a pertinent response surface model, we add in the model important component variable one by one. Hence, the test of the null hypothesis, $H_0: \delta_{q+1} = 0$, is performed by calculating the value of the test statistic

$$F = \frac{SSR(b_{q+1}/b_0, b_1, \cdots, b_q)}{MSE},$$

where

$SSR(b_{q+1}/b_0, b_1, \cdots, b_q) = SSR(b_0, b_1, \cdots, b_{q+1}) - SSR(b_0, b_1, \cdots, b_q)$ and MSE is calculated in the model, $Y = \delta_0 + \delta_1 T_1 + \cdots + \delta_{q+1} T_{q+1} + \varepsilon$. Assuming normality of the errors, if the null hypothesis is true, the F-statistic follows F distribution with 1 and $n-q-1$ degrees of freedom. If the value of F-statistic exceeds $F_{\alpha;1,n-q-1}$, then the null hypothesis is rejected at the α level of significance.

In the absence of sufficient knowledge concerning the shape of the true response surface, generally the experimenter's first attempt at approximating the shape is by fitting a first-order model to the response values. When, however, the first-order model suffers from lack of fit arising from the existence of surface curvature, the first-order model is upgraded by adding higher order terms to it. The next higher order model is the second-order model

$$Y=\delta_0+\sum_{i=1}^{q}\delta_i T_i+\sum_{i=1}^{q}\delta_{ii} T_i^2+\sum_{i=1}^{q-1}\sum_{j=2}^{q}\delta_{ij} T_i T_j+\varepsilon, \tag{2.8}$$

where $T_1,\cdots,T_q$ are the PLS components which influence the response Y; δ_0, $\delta_i(i=1,\cdots,q)$, $\delta_{ij}(i=1,\cdots,q;\ j=1,\cdots,q)$ are unknown parameters, and ε is the random error.

Then the least squares estimator of $\boldsymbol{\delta}_q$ is

$$\mathbf{b}_q=(\mathbf{T}_q'\mathbf{T}_q)^{-1}\mathbf{T}_q'\mathbf{y}, \tag{2.9}$$

where

$$\mathbf{T}_q=\begin{bmatrix} 1 & T_{11} & T_{12} & \cdots & T_{1q} & T_{11}^2 & \cdots & T_{1q}^2 & T_{11}T_{12} & \cdots & T_{1(q-1)}T_{1q} \\ 1 & T_{21} & T_{22} & \cdots & T_{2q} & T_{21}^2 & \cdots & T_{2q}^2 & T_{21}T_{22} & \cdots & T_{2(q-1)}T_{2q} \\ \vdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots & \vdots & \cdots & \vdots \\ 1 & T_{n1} & T_{n2} & \cdots & T_{nq} & T_{n1}^2 & \cdots & T_{nq}^2 & T_{n1}T_{n2} & \cdots & T_{n(q-1)}T_{nq} \end{bmatrix}.$$

Hence, the fitted response surface model is

$$\begin{aligned}\hat{Y} = {} & b_0+b_1T_1+\cdots+b_qT_q+b_{11}T_1^2+\cdots+b_{qq}T_q^2 \\ & +b_{12}T_1T_2+\cdots+b_{(q-1)q}T_{(q-1)}T_q.\end{aligned} \tag{2.10}$$

3. Application of Partial Least Squares Response Surface Function to Process Optimization

In this section we consider procedures that can be used to find the settings of the input variables which produce the most desirable response values. These response values may be the maximum yield or the highest level of quality coming off the production line. Similarly, we may seek the variable settings that minimize the cost of making the product. In any case, the set of values of the input variables which result in the most desirable response values is called the set of optimum conditions. An application of PLS response surface function to process

optimization is considered.

The first step in the process of seeking optimum conditions is to identify the input variables that have the greatest influence on the response. Generally, the fewer the number of variables that have an effect on the response, the easier it is to identify them. Once the important variables are discovered, the next step is to postulate a model which expresses the response of interest as a function of the variables. If nothing, or even if very little, is known of the relationship between the response variable and the important input variables, then the simplest form of model equation is postulated. The first-order model provides the basis for performing an initial set of experiments, which, upon completion, may suggest the fitting of a different model form along with performing further experimentation. If at any time in the process of model developing it is discovered that further experimentation appears uneconomical, the procedure is terminated. The sequence of fitting and testing the model forms and the eventual selection of a model are the prelude to the determination of the optimum operating conditions for a process.

In the case of single response variable, for searching optimum conditions dual quasi-Newton optimization is used. Let us consider the fitting of a second-order model in q component variables of the form

$$Y = \delta_0 + \sum_{i=1}^{q} \delta_i T_i + \sum_{i=1}^{q} \delta_{ii} T_i^2 + \sum_{i=1}^{q-1} \sum_{j=2}^{q} \delta_{ij} T_i T_j + \varepsilon. \tag{3.1}$$

For our purposes, let us assume that observed response values are collected at the points of a second-order design and the fitted second-order polynomial is

$$\hat{Y} = b_0 + \sum_{i=1}^{q} b_i T_i + \sum_{i=1}^{q} b_{ii} T_i^2 + \sum_{i=1}^{q-1} \sum_{j=2}^{q} b_{ij} T_i T_j. \tag{3.2}$$

After the fitted model in (3.2) is checked for adequacy of fit in the region defined by the coordinates of the design and is found to be adequate, the model is then used to discover the optimum condition inside the experimental region.

The dual quasi-Newton optimization technique works well from medium to moderately large optimization problems where the objective function and the gradient are much faster to compute than the Hessian. The dual quasi-Newton optimization technique does not need to compute second-order derivatives, but in general it requires more iterations than the techniques which compute second-order derivatives.

4. Applied Example

This example is the data of the process in a S oil refinery which is given in Park (1990). The response variable is the color tone of the outlet of the reactor in the sweetening process and the process variables of effecting the color tone are six process variables. The six process variables are feed mecaptan (X_1), feed temperature (X_3), feed ending point (X_4) and oil content (X_6) in the feed condition and air injection content (X_2), activity injection content (X_5) in the reactor handing condition. The data set is a good example of what often happens in practice. The data is presented in <Table 1>.

< Table 1 > Raw Data of Color Tone

X_1	X_2	X_3	X_4	X_5	X_6	Y	X_1	X_2	X_3	X_4	X_5	X_6	Y
111.6	130	46	247	62	100.0	21	118.6	139	52	265	52	13.8	17
91.5	135	50	251	70	94.2	21	117.7	139	50	270	40	13.8	17
90.8	135	50	245	70	94.1	21	117.5	150	50	273	40	13.8	21
80.8	175	43	247	80	16.3	13	114.3	173	40	259	40	1.9	18
82.8	145	46	248	80	22.0	20	133.2	115	58	270	60	9.6	19
78.6	145	46	255	80	22.0	20	129.5	107	53	274	40	9.6	23
93.2	120	43	254	80	22.0	21	128.4	128	59	262	40	9.6	21
93.7	120	38	253	80	16.4	23	121.4	128	57	268	40	11.0	21
95.9	120	38	253	80	10.9	21	150.0	110	51	269	40	2.7	23
97.6	115	42	256	70	10.8	27	150.8	110	50	276	40	3.0	23
100.0	115	42	254	70	10.8	26	150.0	110	49	275	40	3.0	24
100.0	115	50	260	50	10.0	26	106.9	110	49	271	40	3.0	26
100.0	115	51	264	40	0.9	26	124.3	120	53	271	40	3.0	20
100.0	119	47	268	40	0.9	27	139.2	105	52	274	40	2.2	24
98.7	119	46	259	40	1.0	28	139.7	119	53	270	40	1.2	22
100.0	119	44	268	40	1.4	27	140.4	119	52	266	40	1.2	22
94.0	135	45	261	50	2.8	23	140.4	119	48	276	40	1.2	23
138.7	160	53	268	70	96.9	17	138.9	119	50	270	40	1.3	26
157.0	143	56	264	36	95.1	20	139.5	119	51	271	40	1.8	23
160.0	150	55	265	36	94.4	18	105.9	133	52	260	40	4.9	20
160.0	137	47	261	65	94.4	21	110.0	133	52	269	40	3.9	23
119.0	128	59	274	40	86.2	21	101.6	133	53	273	40	3.9	22
116.0	139	53	273	52	13.8	19							

At first, the correlation matrix and multicollinearity measure are presented <Tables 2>.

< Table 2 > Correlation Matrix and Multicollinearity Measure

	X_1	X_2	X_3	X_4	X_5	X_6	Y	VIF	Condition Number
X_1	1.00	-0.17	0.52	0.63	-0.55	0.15	-0.07	2.3563	
X_2	-0.17	1.00	-0.07	-0.34	0.25	0.39	-0.74	1.2794	
X_3	0.52	-0.07	1.00	0.55	-0.56	0.18	-0.18	2.0827	
X_4	0.63	-0.34	0.55	1.00	-0.76	-0.33	0.17	3.8467	4.1652
X_5	-0.55	0.25	-0.56	-0.76	1.00	0.28	-0.26	2.7667	
X_6	0.15	0.39	0.18	-0.33	0.28	1.00	-0.35	2.1549	
Y	-0.07	-0.74	-0.18	0.17	-0.26	-0.35	1.00		

The components, $T_1, \cdots, T_6$, are obtained as

$$T_1 = -46.0 - 5.8314 x_1 - 48.647 x_2 - 12.751 x_3 + 13.548 x_4 - 26.101 x_5 - 34.564 x_6,$$

$$T_2 = -3.35 - 16.443 x_1 - 33.860 x_2 - 20.542 x_3 - 13.800 x_4 + 4.2203 x_5 + 4.5939 x_6,$$

$$T_3 = 1.27 - 0.0548 x_1 - 26.074 x_2 - 5.5019 x_3 - 3.0171 x_4 - 12.124 x_5 + 23.024 x_6,$$

$$T_4 = 0.11 - 4.0011 x_1 + 1.4505 x_2 - 3.1761 x_3 + 0.0958 x_4 - 4.9825 x_5 + 2.1048 x_6,$$

$$T_5 = 0.18 - 0.0786 x_1 + 0.1298 x_2 - 1.1366 x_3 + 1.3942 x_4 + 0.3081 x_5 + 0.4672 x_6,$$

$$T_6 = -0.03 - 0.4225 x_1 - 0.0754 x_2 + 0.2322 x_3 + 0.3917 x_4 + 0.1447 x_5 + 0.1249 x_6.$$

And for each model, model selection criterions are displayed in <Table 3>.

< Table 3 > Model Selection Criterions for Color Tone Data

	R^2_q	R^2_{aq}	MSE_q	C_q	Sequential $F-$test
Coded Variable	0.9432	0.8530	1.4784		
$q=1$	0.4342	0.4073	5.9601	51.98	
$q=2$	0.6744	0.6327	3.6937	24.81	$F=9.59>F_{0.05;3,39}$
$q=3$	0.8089	0.7598	2.4152	4.64	$F=6.17>F_{0.05;4,35}$
$q=4$	0.8218	0.7387	2.6275	3.37	$F=0.43<F_{0.05;5,30}$
$q=5$	0.8713	0.7641	2.3721	5.03	
$q=6$	0.9432	0.8530	1.4784	7.00	

The number of components, $q=3$, is selected by the sequential $F-$test. By fitting a second-order model with the first three components,

$$\begin{aligned}\hat{Y}(T_1,\cdots,T_3) = {} & 21.912085+0.050037\,T_1+0.085157\,T_2+0.006440\,T_3 \\ & +0.000160\,T_1^2-0.000766\,T_2^2+0.000264\,T_3^2 \\ & +0.000841\,T_1\,T_2-0.001113\,T_1\,T_3-0.000172\,T_2\,T_3.\end{aligned}$$

The engineers wish to maximize the response variable. <Table 4> shows optimum condition, predicted value under the optimum condition, $\widehat{\mathrm{Var}}(\hat{Y})$ and $\widehat{\mathrm{Bias}}^2(\hat{Y})$.

< Table 4 > Optimum Condition for Color Tone Data

	Optimum Condition	Predicted Value	$\widehat{\mathrm{Var}}(\hat{Y})$	$\widehat{\mathrm{Bias}}^2(\hat{Y})$
Coded Variable	$x_1 = 0.019, x_2=-0.874$ $x_3 = -0.051, x_4=-0.093$ $x_5 = 0.028, x_6= 0.470$	36.0236	16.9027	·
$q=1$	$x_1 = -0.085, x_2=-0.715$ $x_3 = -0.187, x_4= 0.199$ $x_5 = -0.383, x_6=-0.508$	22.5239	0.0481	446.2972
$q=2$	$x_1 = -0.224, x_2=-0.833$ $x_3 = -0.335, x_4=-0.000$ $x_5 = -0.222, x_6=-0.304$	24.7671	0.1787	371.0129
$q=3$	$x_1 = -0.214, x_2=-0.805$ $x_3 = -0.319, x_4= 0.031$ $x_5 = -0.232, x_6=-0.385$	24.9931	1.5063	5.7829
$q=4$	$x_1 = -0.323, x_2=-0.769$ $x_3 = -0.410, x_4=-0.030$ $x_5 = -0.285, x_6=-0.228$	25.9058	1.9522	47.5665
$q=5$	$x_1 = 0.012, x_2=-0.759$ $x_3 = 0.232, x_4=-0.598$ $x_5 = -0.012, x_6=-0.102$	27.9518	7.4621	0.0421
$q=6$	$x_1 = 0.019, x_2=-0.874$ $x_3 = -0.051, x_4=-0.093$ $x_5 = 0.028, x_6= 0.470$	36.0236	16.9027	·

Because we select $q=3$, we now decide optimum condition,

$x_1=-0.214,\ x_2=-0.805,\ x_3=-0.319,\ x_4=0.031,\ x_5=-0.232,\ x_6=-0.385.$

Hence, the optimum conditions of six process variables are

$X_1=110.5,\ X_2=111.8,\ X_3=45.1,\ X_4=260.9,\ X_5=52.8,\ X_6=31.3.$

References

[1] Frank, I. E. and Friedman, J. H.(1993), "A Statistical View of Some Chemometrics Regression Tools," *Technometrics*, Vol. 35, pp. 109-135.

[2] Helland, I. S.(1988), "On the Structure of Partial Least Squares Regression," *Communications in Statistics-Simulation and Computation*, Vol. 17, pp. 581-607.

[3] Hoerl, A. E. and Kennard, R. W.(1970), "Ridge Regression: Biased Estimation for Nonorthogonal Problems," *Technometrics*, Vol. 12, pp. 55-67.

[4] Martens, H.(1985), Multivariate Calibration, Dr. Techn. thesis. Technical University of Norway, Trondheim.

[5] Massy, W. F.(1965), "Principal Components Regression in Exploratory Statistical Research," *Journal of the American Statistical Association*, Vol. 60, pp. 234-246.

[6] Park, S. H.(1990), *Applied Experimental Design*, Youngji Moonhwasa, Seoul.

[7] Wold, H.(1975), *Soft Modeling by Latent Variables; The Non-linear Iterative Partial Least Squares Approach, In Perspectives in Probability and Statistics*, Papers in Honour of M. S. Bartlett, Edited by J. Gani, Academic Press. Inc.. London.