Selection of active effects based on a ridge regression in supersaturated response surface designs

Shun Matsuura*1, Shogo Tajima2, Hideo Suzuki1
1. Keio University : 3-14-1 Hiyoshi, Kohoku-ku, Yokohama-shi, Kanagawa-ken, 223-8522, Japan
2. Yokohama Plant, NISSAN MOTOR CO.,LTD. : 2 Takara-cho, Kanagawa-ku, Yokohama-shi, Kanagawa-ken, 220-8623, Japan
*contact author’s e-mail address: matsuura@ae.keio.ac.jp

Abstract:
Response surface method has been a commonly used statistical method for quality improvement. As the importance of short-term and low-cost product development becomes higher nowadays, conventional response surface designs may need too large number of experimental runs to be accepted. Recently, supersaturated response surface designs were constructed by using several columns of Hadamard matrices (two-level orthogonal arrays), which considerably reduce the number of experimental runs, and a stepwise regression based on $F$-tests was applied to their data analysis. In this paper, a criterion is introduced for constructing supersaturated response surface designs, and a ridge regression is used for the data analysis of supersaturated response surface designs. Several comparisons through numerical simulations are also presented.

Keywords
Response surface method, Ridge regression, Supersaturated design, Two-level orthogonal array, Variable selection

1. Introduction
In manufacturing industries, response surface method (Myers et al., 2009) has been a commonly used statistical method for obtaining a response surface model such as a second-order model that describes relationships between the response of interest (the quality characteristic) and the design factors. In the application of response surface method, experiments are first conducted according to a response surface design (an experimental design for a response surface model); then, a response surface model is estimated by analyzing data obtained from the experiments.

Myers et al. (2004) presented a review of recent developments in response surface method. Central composite designs and Box and Behnken designs have been commonly used as response surface designs for second-order response surface models (see, for example, Morris, 2000). Response surface designs having some optimal and/or desirable properties have been proposed by many papers (e.g., Khuri, 1992; Draper and John, 1998; Parker et al., 2007; Anderson-Cook et al., 2009; Macharia and Goos, 2010). Kolaiti and Koukouvinos (2007) and Androulakis et al. (2013) investigated the properties of response surface designs constructed from three-level orthogonal arrays when observations possibly have some correlations. Letsinger et al. (1996), Vining and Kowalski (2008), and Ikeda et al. (2014) discussed methods for estimating response surface models based on split-plot experiments.

Recently, the period of new product development becomes shorter, as multi-functionalized products increase and their variety becomes diversified. The importance of short-term and low-cost product development becomes higher. In addition, in some cases, it involves high cost and takes a long time to conduct only one experiment. Hence, product development with as small number of experiments as possible is needed. The conventional response surface designs such as central composite designs and Box and Behnken designs may need too large number of experimental runs to be accepted.

Matsuura et al. (2011) constructed supersaturated designs for response surface models (supersaturated response surface designs). An experimental design is called a supersaturated design if the number of experimental runs (i.e., the number of rows) is smaller than the number of parameters to be estimated (i.e., the number of columns). Supersaturated designs are effectively used for identifying ‘active factors’ (i.e., factors that actually affect the response) among many factors, especially when the ‘effect sparsity’ assumption holds (i.e., when only a few factors are active). The construction and analysis of supersaturated designs have been studied in
many papers. Allen and Bernshteyn (2003) constructed supersaturated designs maximizing the probability of identifying active factors. Yamada et al. (1999) and Georgiou et al. (2003) discussed the construction of three-level supersaturated designs. Yamada and Lin (2002), Koukouvinos et al. (2007), Chen and Liu (2008), and Chatterjee et al. (2011) studied construction methods for mixed-level supersaturated designs. See also Niki et al. (2011) and Gupta et al. (2012). Yamada (2004) used a stepwise regression for the data analysis of supersaturated designs and examined type I and type II error rates. Recently, Li et al. (2010) proposed variable selection methods based on cluster analysis, Edwards and Mee (2011) constructed a randomization test method, and Koukouvinos et al. (2011) proposed using entropy measures for variable selection. Other methods for the data analysis of supersaturated designs have been discussed by, for example, Abraham et al. (1999), Beattie et al. (2002), Li and Lin (2003), Cossari (2008), and Marley and Woods (2010). See Georgiou (2014), in which a review of the construction and analysis of supersaturated designs was provided.

In Matsuura et al. (2011), a supersaturated response surface design (SRSD) was constructed by combining three experimental designs: a set of several columns of a \((8m + 4) \times (8m + 4)\) Hadamard matrix (two-level orthogonal array) with a natural number \(m\), a set of with \(2p\) axial points, and a set of \(n_0\) center points (which is a modification of central composite design); then, a response surface model was estimated using a stepwise regression based on \(F\)-tests. The SRSDs were applied to a robust parameter design problem, and it was shown that the SRSDs may be effective for finding the effects (i.e., the main effects, the two-factor-interaction effects, and the quadratic effects) of active factors under the ‘effect sparsity’ assumption; more specifically, the SRSDs are effective when the number of active factors (i.e., the number of the effects of active factors) is smaller than approximately one-third of the number of experimental runs. However, no criterion has been proposed to determine which columns should be selected for constructing a SRSD. In addition, only a stepwise variable selection based on \(F\)-tests was used for the data analysis of SRSDs and has not been compared to other variable selection methods.

In this paper, a criterion is introduced to determine which columns should be selected for constructing a SRSD, where the criterion can be seen as a modification of a criterion that Wu (1993) gave for supersaturated designs. For the data analysis of SRSDs, we use a variable selection by \(t\)-tests based on a ridge regression. We compare the stepwise regression and the ridge regression by numerical simulations, and also discuss how the numbers of center points in SRSDs affect their precisions.

In the next section, we review response surface method and SRSDs. In Section 3, a criterion is given for determining which columns should be selected for constructing a SRSD. Section 4 presents a variable selection method by \(t\)-tests based on a ridge regression for SRSDs. Section 5 gives several numerical results to compare the precisions of the stepwise regression and the ridge regression. Concluding remarks are given in Section 6.

2. Review of response surface method and supersaturated response surface designs

Let \(p\) denote the number of design factors. Consider the following second-order response surface model:

\[
y = b_0 + \sum_{i=1}^{p} b_i x_i + \sum_{i<j}^{p} b_{ij} x_i x_j + \sum_{i=1}^{p} b_{ii} x_i^2 + \epsilon. \tag{2.1}
\]

Here, we let \(y\) denote the response of interest (the quality characteristic), \(x_1, x_2, \ldots, x_p\) denote the levels of the \(p\) design factors, \(b_0\) denote the intercept, \(b_1, b_2, \ldots, b_p\) denote the main effects of the \(p\) design factors, \(b_{1,2}, b_{1,3}, \ldots, b_{p-1,p}\) denote the two-factor-interaction effects, \(b_{1,3}, b_{2,2}, \ldots, b_{p,p}\) denote the quadratic effects of the \(p\) design factors, and also let \(\epsilon\) denote an experimental error distributed as a normal distribution \(N(0, \sigma^2)\).

We further let \(x = (x_1, x_2, \ldots, x_p)'\), \(f(x) = (1, x_1, x_2, \ldots, x_p, x_1 x_2, x_1 x_3, \ldots, x_{p-1} x_p, x_1^2, x_2^2, \ldots, x_p^2)'\), and \(\beta = (b_0, b_1, b_2, \ldots, b_p, b_{1,2}, b_{1,3}, \ldots, b_{p-1,p}, b_{1,1}, b_{2,2}, \ldots, b_{p,p})'\)

\((f(x)\) and \(\beta\) are \((p+1)(p+2)/2 \times 1\) vectors since \(1 + p + \frac{p(p-1)}{2} + p = \frac{(p+1)(p+2)}{2}\). Then, we can rewrite (2.1) as

\[
y = f(x)' \beta + \epsilon.
\]

Suppose that the values of the parameter vector \(\beta\) and the error variance \(\sigma^2\) are unknown and experiments are conducted according to a response surface design for estimating \(\beta\). Let \(x_i = (x_{i1}, x_{i2}, \ldots, x_{ip})'\), \(i = 1, 2, \ldots, n\) denote the settings of the levels of the \(p\) design factors in the response surface design, where \(n\) denotes the number of experimental runs. We let \(X = (f(x_1), f(x_2), \ldots, f(x_n))'\). Then, \(X\) is the designed matrix \((n \times \frac{(p+1)(p+2)}{2})\) matrix for the model (2.1). The data vector \(y = (y_1, y_2, \ldots, y_n)'\) obtained from the experiments can be expressed as

\[
y = X \beta + \epsilon,
\]

where \(\epsilon\) denotes the vector of experimental errors distributed as \(N_n(0, \sigma^2 I_n)\) (\(0_n\) denotes the zero vector of length \(n\) and \(I_n\) denotes the \(n \times n\) identity matrix).
If \( XX' \) is full rank, then \( \beta \) can be estimated by the ordinary least-squares estimator
\[
\hat{\beta}_{OLS} = (XX')^{-1}XX'y.
\] (2.2)
However, if \( XX' \) is not full rank, for example, if \( X \) is a supersaturated design (i.e., \( n < \frac{(p+1)(p+2)}{2} \) holds), then the ordinary least-squares estimator is not available.

Matsuura et al. (2011) constructed supersaturated response surface designs (SRSDs) based on Hadamard matrices (two-level orthogonal arrays). Note that the designs were intended for the application to a robust parameter design problem and the quadratic effects of noise factors were not taken into account. Hence, we modify their designs to enable the estimation of all quadratic effects. Here, a SRSD based on a two-level orthogonal array is constructed by combining the following three experimental designs:

(i) Select \( p \) columns from the columns of a two-level orthogonal array with \( 8m + 4 \) experimental runs, where \( m \) is a natural number such that the number of columns of the two-level orthogonal array is not smaller than \( p \). Then, \( p \) design factors are assigned to the selected \( p \) columns.
(ii) Construct a set of axial points: a design with \( 2p \) experimental runs in which the level of each design factor is sequentially set to be \( \pm 1 \) and the levels of all other design factors are set to be \( 0 \).
(iii) Construct a set of center points: a design with \( n_0 \) experimental runs in which the levels of all \( p \) design factors are set to be \( 0 \).

Here, if \( 8m + 4 + 2p + n_0 < \frac{(p+1)(p+2)}{2} \), the designed matrix \( X \) for the model (2.1) is a supersaturated design. An example of the settings of the levels of \( p = 6 \) design factors for a SRSD is shown in Table 1, where we use the first six columns (from the 1st column to the 6th column) of a Plackett-Burman design (Plackett and Burman, 1946) with 12 experimental runs (P.B.12) and set \( n_0 = 2 \). Plackett-Burman designs with 12 experimental runs (P.B.12) and with 20 experimental runs (P.B.20) are given, for example, in Wu and Hamada (2009, p.397,398). For reference, Table 2 shows P.B.12 and P.B.20. The designed matrix for the model (2.1) obtained from the settings of Table 1 is shown in Table 3. The designed matrix of Table 3 is a \( 26 \times 28 \) matrix and hence is a supersaturated design (and is not full rank).

Table 1: An example of the settings of the levels of \( p = 6 \) design factors for a SRSD

|   |   |   |   |   |   |
|---|---|---|---|---|---|
| 1 | 1 | -1 | 1 | 1 | 1 |
| -1| 1 | 1 | -1 | 1 | 1 |
| 1 | -1| 1 | 1 | -1 | 1 |
| -1| 1 | -1| 1 | 1 | -1 |
| -1| 1 | -1| 1 | -1| 1  |
| 1 | -1| 1 | -1| 1 | -1 |
| 1 | 1 | 1 | 1 | -1| -1 |
| -1| 1 | 1 | 1 | -1| 1  |
| 1 | -1| 1 | 1 | 1 | -1 |
| -1| 1 | -1| 1 | -1| 1  |
| 1 | 0 | 0 | 0 | 0 | 0  |
| -1| 0 | 0 | 0 | 0 | 0  |
| 0 | 1 | 0 | 0 | 0 | 0  |
| 0 | -1| 0 | 0 | 0 | 0  |
| 0 | 0 | -1| 0 | 0 | 0  |
| 0 | 0 | 0 | 1 | 0 | 0  |
| 0 | 0 | -1| 1 | 0 | 0  |
| 0 | 0 | 0 | 0 | 1 | 0  |
| 0 | 0 | 0 | 0 | -1| 0  |
| 0 | 0 | 0 | 0 | 0 | -1 |
| 0 | 0 | 0 | 0 | 0 | 0  |
Table 2: Plackett-Burman designs with 12 and 20 experimental designs (P.B.12 and P.B.20) (given in Wu and Hamada (2009, p.397,398))

| P.B.12 | P.B.20 |
|--------|--------|
| + + + + + + - - - - | + + + + + + - - - - + + + + |
| - + + + + + + - - + + | - + + + + + + - - + + + + - + |
| + - + + - + + + - - + | + - + + - + + + - + + - - - + | + - - - + + + - - - + + + |
| - - + + + + + + + - - + | - - - - + + + - + + + + - + |
| + - + - + + + + + + + | + + - + - + + + + + + + + + + |
| + + - + + + + + + + + + | + + - + + + + + + + + + + + + + |

Table 3: An example of SRSD (26 x 28 matrix) for p = 6 design factors based on P.B.12

| b_0  | b_1  | b_2  | b_3  | b_4  | b_5  | b_6  |
|------|------|------|------|------|------|------|
| 1    | -1   | 1    | -1   | 1    | 1    | 1    |
| 1    | -1   | 1    | -1   | -1   | 1    | 1    |
| 1    | -1   | 1    | -1   | -1   | -1   | 1    |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |
| 1    | -1   | 1    | -1   | -1   | -1   | -1   |

[DOI : 10.17929/jqc.2.1] Copyright © 2016 Journal of the Japanese Society for Quality Control. All rights reserved.
3. Criterion for constructing supersaturated response surface designs

We let $X \ (n \times \{(p + 1)(p + 2)/2\}$ matrix) be a SRSD. Then, if $k$ design factors among the $p$ design factors are active but we do not know which $k$ design factors are active, then the following geometric mean of the maximum values of D-optimality may be an expected performance of the SRSD $X$:

$$
\left\{ \prod_{i=1}^{\binom{p}{k}} \det (X'_i X_i) \right\}^{1/\binom{p}{k}}. \quad (3.1)
$$

Here, $X_1, ..., X_{\binom{p}{k}}$ are the response surface designs of $n \times \{(k + 1)(k + 2)/2\}$ matrix, where each $X_i (i = 1, 2, ..., \binom{p}{k})$ is a response surface design that is obtained by selecting $(k + 1)(k + 2)/2$ columns from $X$ corresponding to the intercept, the main effects, the two-order interactions effects, and the quadratic effects of each combination of selection of $k$ design factors from the $p$ design factors.

This criterion (3.1) can be seen as a modification of Wu (1993). Wu (1993) gave a criterion for supersaturated designs that is an arithmetic (not geometric) mean of the values of D-optimality. However, this criterion may give a high evaluation to a supersaturated design including a combination of $k$ design factors such that $\det (X'_i X_i) = 0$ in which the effects of $k$ design factors are not estimable. When we do not know which $k$ design factors are active, it is desirable that the effects of $k$ design factors are always estimable whichever $k$ design factors are active. The criterion (3.1) gives the worst evaluation (zero value) to a supersaturated design if the design includes a combination of $k$ design factors such that $\det (X'_i X_i) = 0$.

For computation, we note that choosing a SRSD having a maximum value of (3.1) is equivalent to choosing a SRSD having a maximum value of

$$
1 \sum_{i=1}^{\binom{p}{k}} \log (\det (X'_i X_i)).
$$

Using the criterion (3.1), we can compare different SRSDs. For example, suppose that $p = 8$ and we construct SRSDs of $36 \times 45$ matrix by selecting $p = 8$ columns from the 19 columns of P.B.20. Numerical calculations reveal that, in all $\binom{19}{8} = 75582$ combinations of selection of $p = 8$ columns from the 19 columns of P.B.20, when $k = 3$, 2052 combinations attain a maximum value of (3.1); when $k = 4$, 855 combinations attain a maximum value of (3.1); when $k = 5$, 171 combinations attain a maximum value of (3.1). Numerical calculations also reveal that some combinations give the maximum values of (3.1) for $k = 3, 4, 5$ simultaneously, one of which is $(8,10,12,13,16,17,18,19)$. Hence, if we expect that about 3 ~ 5 design factors of the $p = 8$ design factors are active, then the SRSD based on $(8,10,12,13,16,17,18,19)$ (or other combinations giving the maximum values of (3.1) for $k = 3, 4, 5$ simultaneously) is recommended.

We note that Evangelaras et al. (2003, 2007) discussed non-isomorphic of two-level orthogonal arrays. Their results may be applied in some cases of our setting. Using isomorphism will clarify the properties of SRSDs in more detail, which is an important future issue.

4. Variable selection by ridge regression

For the data analysis of SRSDs, Matsuura et al. (2011) used a stepwise regression that has been commonly used for multivariate regression analysis: starting with no selected variable and alternately conducting forward selection and backward elimination based on $F$-tests with a significance level $\alpha$. In this paper, the following variable selection method by $t$-tests is used for the ridge regression is used for the data analysis of SRSDs.

Let $X$ be a SRSD of $n \times \{(p + 1)(p + 2)/2\}$ matrix. In this section, we let $q = \frac{(p + 1)(p + 2)}{2}$ for simplicity. Let $\lambda$ be a ridge parameter. Then, the ridge estimator of $\beta$ is given by

$$
\hat{\beta}_\lambda = (X'X + \lambda I_p)^{-1}X'y.
$$

Although Hoerl et al. (1975) recommended that the value of $\lambda$ is set to be

$$
\lambda = \frac{q\delta^2}{\hat{\beta}_{OLS} \hat{\beta}_{OLS}}.
$$

$\hat{\beta}_{OLS}$ (2.2) is not available since $X$ is a SRSD. Hence, we use a modification of the method by Hoerl and Kennard (1976), as the following:

(i) Set an initial value of ridge parameter $\lambda_0 > 0$, and also set a calculation error bound $\eta > 0$.
(ii) Compute the ridge estimator
(iii) Compute the ridge parameter

$$\hat{\beta}_d(\lambda) = (X'X + \lambda I_n)^{-1}X'y.$$  

(iv) If $|\lambda_1 - \lambda_0| < \eta$ holds, then finish the iteration. Otherwise, set $\lambda_0 = \lambda_1$ and go back to step (ii).

Note that Hoerl and Kennard (1976) used $\lambda_0 = 0$. However, $\hat{\beta}_d(0)$ is equivalent to $\hat{\beta}_{OLS}$ and thus, is not available. Hence, we use $\lambda_0 > 0$.

For conducting the above algorithm, we need an estimator $\hat{\sigma}^2$ of the error variance $\sigma^2$. If $X'X$ is full rank, then $\hat{\sigma}^2 = \frac{1}{n-q}(y - X\hat{\beta}_{OLS})'(y - X\hat{\beta}_{OLS})$ is a commonly used unbiased estimator. However, in this case, $X$ is a SRSD and $\hat{\beta}_{OLS}$ is not available. Hence, we obtain an unbiased estimator $\hat{\sigma}^2$ of $\sigma^2$ by using only data from the $n_0$ center-point experiments, as follows:

$$\hat{\sigma}^2 = \frac{\sum_{i=n-n_0+1}(y_i - \bar{y}_c)^2}{n_0 - 1}$$ with $\bar{y}_c = \frac{\sum_{i=n-n_0+1}y_i}{n_0}$

For variable selection based on the above ridge estimator, the approximate $t$-test in Halawa and Bassioni (2000) can be used, as the following. We let $\beta_i$ be the $i$th element of $\beta$ and $\hat{\beta}_{d,i}(\lambda)$ be the $i$th element of $\hat{\beta}_d(\lambda)$. Then,

$$t_i = \frac{\hat{\beta}_{d,i}(\lambda)}{\sqrt{\text{Var}[\hat{\beta}_{d,i}(\lambda)]}}$$

is approximately distributed as a $t$-distribution with $n_0 - 1$ degrees of freedom under the null hypothesis $H_0: \beta_i = 0$, where we let $\text{Var}[\hat{\beta}_{d,i}(\lambda)]$ be the $i$th diagonal element of the following approximate covariance matrix of $\hat{\beta}_d(\lambda)$:

$$\text{Var}[\hat{\beta}_d(\lambda)] = \hat{\sigma}^2(X'X + \lambda I_n)^{-1}X'X(X'X + \lambda I_n)^{-1}.$$  

From this, the null hypothesis $H_0: \beta_i = 0$ is rejected at the significance level of $\alpha$ if $|t_i| > t(n_0 - 1, \alpha)$ holds, where $t(n_0 - 1, \alpha)$ is the upper 100$\alpha$/2 percentile of the $t$-distribution with $n_0 - 1$ degrees of freedom.

Note that when $n_0$ is small, the standard error of $\hat{\sigma}^2$ may be large. Hence, in the next section, we discuss the impact of the number of $n_0$ on the precision of the variable selection by $t$-tests.

5. Numerical comparison of the stepwise regression and the ridge regression

Finally, we present several results of numerical simulations to compare the stepwise regression and the ridge regression.

Let $p = 8$ ($\frac{(p+1)(p+2)}{2} = 45$). We construct SRSDs ($36 + n_0 \times 45$ matrix) by selecting 8 columns from the 19 columns of P.B.20. As noted in Section 3, the combination of $k = 3,4,5,6,7$ values of (3.1) for $k = 3,4,5$ simultaneously. Hence, for numerical simulations, we use $(8,10,12,13,16,17,18,19)$ for constructing SRSDs.

Here, we set the number of center points to be $n_0 = 3,4,5,6,7$, the significance level to be $\alpha = 0.05$, and the error variance to be $\sigma^2 = 1$. For simplicity, we set the intercept $\beta_0$ to be 100, all effects of active factors to be a constant $\beta = 2$ or $\beta = 3$, and the other effects to be $\beta = 0$.

We consider the following cases.

Case 1-1: Suppose that three of the $p = 8$ design factors are active and the values of the active effects are $\beta = 2$. The number of active effects (except the intercept $\beta_0$) is 9. Let

$$\beta = (100,2,2,2,0,0,0,0,0,2,2,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
Case 2-2: Suppose that four of the \( p = 8 \) design factors are active and the values of the active effects are \( \beta = 3 \). The number of active effects (except the intercept \( \beta_0 \)) is 14. Let

\[
\beta = (100,3,3,3,0,0,0,0,3,3,0,0,0,0,3,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0)'.
\]

We used the statistical software R for numerical simulations, and we obtained each power value (ratio that an active effect is selected by variable selection) for the main effects, the two-factor interaction effects, and the quadratic effects, respectively, and each type I error rate (ratio that a non-active effect is selected by variable selection) for the main effects, the two-factor interaction effects, and the quadratic effects, respectively, by 10000 numerical simulations for each setting. The obtained results are summarized in Tables 4 and 5. In Tables 4 and 5, ‘Power’ in the ‘Main effects’, ‘Two-factor-interaction effects’, and ‘Quadratic effects’ columns means the average of the power values of active main effects, active two-factor interaction effects, and active quadratic effects, respectively. Similarly, ‘Type I error’ in the ‘Main effects’, ‘Two-factor-interaction effects’, and ‘Quadratic effects’ columns means the average of the type I error rates of non-active main effects, non-active two-factor interaction effects, and non-active quadratic effects, respectively. The units of the values in Tables 4 and 5 are %.

Table 4 indicates that, although the stepwise regression shows high power values in the cases 1-1 and 1-2 (the number of active factors is 3), it shows low power values, especially for the main effects and the two-factor-interaction effects, in the cases 2-1 and 2-2 (the number of active factors is 4). The power values of the quadratic effects are larger than those of the main and two-factor-interaction effects in the cases 2-1 and 2-2. The difference may arise from the fact that the columns to which the main effects and the two-factor-interaction effects are assigned in the SRSD are confounding (not orthogonal) and hence the multicollinearity inflates as the number of selected effects increases in the variable selection, while the degrees of confounding between the quadratic effects and the other effects are small due to the existence of axial points in the SRSD. Table 4 also shows that the type I error rates are far from 5% (the significance level) in the cases 2-1 and 2-2. In particular, the type I error rates of the main and two-factor-interaction effects are too large.

Next, Table 5 indicates that the precision of the ridge regression largely increases as the number of center points \( n_0 \) increases, and it shows high power values in the cases 1-2 and 2-2 (the values of the active effects are \( \beta = 3 \)) when \( n_0 \) is more than about 4 ~ 6. It is also shown that the type I error rates in the ridge regression are closer to 5% than the stepwise regression for the cases 2-1 and 2-2 (the number of active factors is 4). The results indicate that the precision of the ridge regression seems to be more robust to the number of active effects compared to the stepwise regression. This may be because the stepwise regression employs the ordinary least-squares estimator using the columns selected by the variable selection and the ordinary least-squares estimator is largely affected by the degrees of confounding as the selected columns increase, while the ridge regression does not select the columns for the estimation of the values of effects and selects active effects after the estimation. In contrast, the number of center points \( n_0 \) largely affects the performance of the ridge regression unlike the stepwise regression. This may be due to the fact that the ridge regression in this paper always uses only data from the \( n_0 \) center-point experiments for estimating the error variance \( \sigma^2 \) (hence the degrees of freedom are \( n_0 - 1 \), while the degrees of freedom for the error variance estimation are \( n - (\text{number of selected effects}) \) in the stepwise regression.

From these results, it seems that the stepwise regression is better than the ridge regression when the number of active factors is 3 and the ridge regression is better than the stepwise regression for large \( n_0 \) when the number of active factors is 4. Hence, it may be expected that, when the number of active factors is large, the ridge regression will be performs well compared to the stepwise regression.

Finally, as mentioned in Section 1, various methods for the data analysis of supersaturated designs have been proposed. In addition, Lasso (Tibshirani, 1996) has become to be considered a promising method. Applying other methods to the data analysis of SRSDs and comparing them with the stepwise regression and the ridge regression will be an important future issue.
Table 4: Results of numerical simulations for the stepwise regression

| Case | $n_0$ | Stepwise regression |     | Main effects | Two-factor-interaction effects |     | Quadratic effects |     |
|------|-------|---------------------|-----|--------------|-------------------------------|-----|-------------------|-----|
|      |       |                     |     | Power        | Type I error                 |     | Power             | Type I error |
| Case 1-1 |       |                     | 3   | 95.780       | 7.778                        |     | 91.667           | 6.561         |
|        |       |                     | 4   | 95.893       | 7.548                        |     | 91.427           | 6.484         |
|        |       |                     | 5   | 95.913       | 7.482                        |     | 91.713           | 6.512         |
|        |       |                     | 6   | 95.990       | 7.600                        |     | 91.770           | 6.377         |
|        |       |                     | 7   | 96.353       | 7.490                        |     | 92.113           | 6.393         |
| Case 1-2 |       |                     | 3   | 98.043       | 7.048                        |     | 95.807           | 5.875         |
|        |       |                     | 4   | 98.097       | 6.694                        |     | 96.043           | 5.782         |
|        |       |                     | 5   | 98.303       | 6.794                        |     | 96.430           | 5.577         |
|        |       |                     | 6   | 98.257       | 6.508                        |     | 96.313           | 5.566         |
|        |       |                     | 7   | 98.363       | 6.402                        |     | 96.360           | 5.605         |
| Case 2-1 |       |                     | 3   | 40.435       | 15.615                       |     | 38.405           | 15.467         |
|        |       |                     | 4   | 42.530       | 16.045                       |     | 39.092           | 15.898         |
|        |       |                     | 5   | 43.908       | 16.528                       |     | 39.642           | 16.389         |
|        |       |                     | 6   | 45.125       | 16.678                       |     | 40.292           | 16.582         |
|        |       |                     | 7   | 46.735       | 16.900                       |     | 40.637           | 16.933         |
| Case 2-2 |       |                     | 3   | 50.153       | 14.800                       |     | 42.283           | 17.540         |
|        |       |                     | 4   | 52.435       | 14.805                       |     | 43.353           | 18.097         |
|        |       |                     | 5   | 53.423       | 15.165                       |     | 43.928           | 18.415         |
|        |       |                     | 6   | 54.210       | 15.473                       |     | 44.432           | 18.773         |
|        |       |                     | 7   | 55.245       | 15.660                       |     | 44.822           | 18.941         |

*Note. The values in the columns of ‘Power’ in ‘Main effects’, ‘Two-factor-interaction effects’, and ‘Quadratic effects’ denote the averages of the power values of active main effects, active two-factor interaction effects, and active quadratic effects, respectively. The values in the columns of ‘Type I error’ in ‘Main effects’, ‘Two-factor-interaction effects’, and ‘Quadratic effects’ denote the averages of the type I error rates of non-active main effects, non-active two-factor interaction effects, and non-active quadratic effects, respectively. The units of the values are %.
Table 5: Results of numerical simulations for the ridge regression

| Case 1-1        | $n_0$ | Power | Type I error | Power | Type I error | Power | Type I error |
|-----------------|-------|-------|--------------|-------|--------------|-------|--------------|
|                  | 3     | 35.760| 5.120        | 33.913| 7.134        | 38.903| 5.060        |
|                  | 4     | 48.927| 4.862        | 45.823| 7.460        | 53.430| 4.894        |
|                  | 5     | 57.497| 5.192        | 53.840| 8.272        | 63.010| 4.926        |
|                  | 6     | 62.740| 5.090        | 58.750| 8.722        | 67.670| 5.202        |
|                  | 7     | 66.667| 4.912        | 61.663| 8.780        | 71.840| 5.022        |
| Case 1-2        | 3     | 61.200| 5.056        | 57.453| 9.370        | 66.027| 4.984        |
|                  | 4     | 79.867| 4.888        | 74.377| 10.730       | 84.220| 4.814        |
|                  | 5     | 88.523| 5.230        | 82.190| 12.296       | 91.630| 4.962        |
|                  | 6     | 92.157| 4.926        | 86.460| 12.856       | 94.997| 5.060        |
|                  | 7     | 94.087| 4.704        | 88.343| 13.152       | 96.353| 4.960        |
| Case 2-1        | 3     | 36.205| 5.035        | 39.327| 7.858        | 39.680| 5.158        |
|                  | 4     | 48.693| 4.715        | 52.210| 8.520        | 53.570| 4.748        |
|                  | 5     | 57.583| 4.923        | 60.478| 9.406        | 62.628| 4.938        |
|                  | 6     | 62.218| 4.745        | 64.323| 9.650        | 67.893| 4.790        |
|                  | 7     | 66.200| 4.975        | 67.598| 10.172       | 71.783| 5.085        |
| Case 2-2        | 3     | 60.423| 4.873        | 62.482| 10.716       | 64.873| 5.110        |
|                  | 4     | 79.570| 5.043        | 78.817| 13.259       | 83.955| 4.923        |
|                  | 5     | 88.153| 5.093        | 85.650| 14.823       | 91.645| 5.108        |
|                  | 6     | 92.113| 5.068        | 88.977| 15.683       | 94.953| 5.055        |
|                  | 7     | 94.348| 4.848        | 90.817| 16.504       | 96.150| 5.095        |

*Note. See the note in Table 4.

6. Conclusion

This paper introduced a criterion for the selection of columns of a Hadamard matrix (two-level orthogonal array) for constructing a supersaturated response surface design (SRSD), and proposed the use of a variable selection based on a ridge regression for the data analysis of SRSDs. The ridge regression was compared to the stepwise regression through numerical simulations for a SRSD. Numerical results indicated that the ridge regression may show high power values for active effects compared with the stepwise regression when the number of active effects is large and the number of center-point experiments is about more than $4 \sim 6$.

As noted in Sections 3 and 5, investigating the properties of SRSDs using isomorphism and comparing the ridge regression with other methods including Lasso will be important future issues.

References

Abraham, B., Chipman, H., and Vijayan K. (1999), “Some risks in the construction and analysis of supersaturated designs”, *Technometrics*, Vol. 41, No. 2, pp. 135-141.

Allen, T. T. and Bernshteyn, M., (2003), “Supersaturated designs that maximize the probability of identifying active factors”, *Technometrics*, Vol. 45, No. 1, pp. 90-97.

Anderson-Cook, C. M., Borror, C. M., and Montgomery, D. C. (2009), “Response surface design evaluation and comparison”, *Journal of Statistical Planning and Inference*, Vol. 139, No. 2, pp. 629-641.

Androulakis, E., Angelopoulos, P., and Koukouvinos, C. (2013), “A comparison of three-level orthogonal arrays in the presence of different correlation structures in observations”, *Communications in Statistics - Simulation and Computation*, Vol. 42, No. 3, pp. 552-569.

Beattie, S. D., Fong, D. K. H., and Lin, D. K. J. (2002), “A two-stage Bayesian model selection strategy for supersaturated designs”, *Technometrics*, Vol. 44, No. 1, pp. 55-63.

Chatterjee, K., Koukouvinos, C., and Mylona, K. (2011), “A new lower bound to $A_2$-optimality
measure for multi-level and mixed level column balanced designs and its applications”, *Journal of Statistical Planning and Inference*, Vol. 141, No. 2, pp. 877-888.

Chen, J. and Liu, M. Q. (2008), “Optimal mixed-level supersaturated design with general number of runs”, *Statistics & Probability Letters*, Vol. 78, No. 15, pp. 2496-2502.

Cossari, A. (2008), “Applying Box-Meyer method for analyzing supersaturated designs”, *Quality Technology and Quantitative Management*, Vol. 5, No. 4, pp. 393-401.

Draper, N. R. and John, J. A. (1998), “Response surface designs where levels of some factors are difficult to change”, *Australian and New Zealand Journal of Statistics*, Vol. 40, No. 4, pp. 487-495.

Edwards, D. J. and Mee, R. W. (2011), “Supersaturated designs: are our results significant?”, *Computational Statistics & Data Analysis*, Vol. 55, No. 9, pp. 2652-2664.

Evangelaras, H., Georgiou, S., and Koukouvinos, C. (2003), “Inequivalent projections of Hadamard matrices of orders 16 and 20”, *Metrika*, Vol. 57, No. 1, pp. 29-35.

Evangelaras, H., Koukouvinos, C., and Lappas, E. (2007), “Further contributions to nonisomorphic two level orthogonal arrays”, *Journal of Statistical Planning and Inference*, Vol. 137, No. 6, pp. 2080-2086.

Georgiou, S. D. (2014), “Supersaturated designs: A review of their construction and analysis”, *Journal of Statistical Planning and Inference*, Vol. 144, pp. 92-109.

Georgiou, S., Koukouvinos, C., and Mantas, P. (2003), “Construction methods for three-level supersaturated designs based on weighing matrices”, *Statistics & Probability Letters*, Vol. 63, No. 4, pp. 339-352.

Gupta, V. K., Chatterjee, K., Das, A., and Kole, B. (2012), “Addition of runs to an s-level supersaturated design”, *Journal of Statistical Planning and Inference*, Vol. 142, No. 8, pp. 2402-2408.

Halawa, A. M., and Bassiouni, M. Y. E. (2000), “Tests of regression coefficients under ridge regression models,” *Journal of Statistical Computation and Simulation*, Vol. 65, No. 1-4, pp. 341-356.

Hoerl, A. E. and Kennard, R. W. (1976), “Ridge regression iterative estimation of the biasing parameter,” *Communications in Statistics - Theory and Methods*, Vol. 5, No. 1, pp. 77-88.

Hoerl, A. E., Kenward, R. W., and Baldwin, K. F. (1975), “Ridge regression: some simulations”, *Communications in Statistics*, Vol. 4, No. 2, pp. 105-123.

Ikeda, S., Matsuura, S., and Suzuki, H. (2014), “Two-step residual-based estimation of error variances for generalized least squares in split-plot experiments”, *Communications in Statistics - Simulation and Computation*, Vol. 43, No. 2, pp. 342-358.

Khuri, A. I. (1992), “Response surface models with random block effects”, *Technometrics*, Vol. 34, No. 1, pp. 26-37.

Kolaiti, E. and Koukouvinos, C. (2007), “A comparison of three-level orthogonal arrays in the presence of a possible correlation in observations”, *Journal of Applied Statistics*, Vol. 34, No. 2, pp. 167-175.

Koukouvinos, C., Mantas, P., and Mylona, K. (2007), “A general construction of E(fn4)-optimal mixed-level supersaturated designs”, *Sankhya*, Vol. 69, No. 2, pp. 358-372.

Koukouvinos, C., Massou, E., Mylona, K., and Parpoula, C. (2011), “Analyzing supersaturated designs with entropic measures”, *Journal of Statistical Planning and Inference*, Vol. 141, No. 3, pp. 1307-1312.

Letsinger, J. D., Myers, R. H., and Lentner, M. (1996), “Response surface methods for bi-randomization structures”, *Journal of Quality Technology*, Vol. 28, No. 4, pp. 381-397.

Li, P., Zao, S., and Zhang, Z. (2010), “A cluster analysis selection strategy for supersaturated designs”, *Computational Statistics & Data Analysis*, Vol. 54, No. 6, pp. 1605-1612.

Li, R. and Lin, D. K. J. (2003), “Analysis methods for supersaturated design: some comparisons”, *Journal of Data Science*, Vol. 1, No. 3, pp. 249-260.

Macharia, H. and Goos, P. (2010), “D-optimal and D-efficient equivalent-estimation second-order split-plot designs”, *Journal of Quality Technology*, Vol. 42, No. 4, pp. 358-372.

Marley, C. J. and Woods, D. C. (2010), “A comparison of design and model selection methods for supersaturated experiments”, *Computational Statistics & Data Analysis*, Vol. 54, No. 12, pp. 3158-3167.

Matsuura, S., Suzuki, H., Iida, T, Kure, H., and Mori, H. (2011), “Robust parameter design using a supersaturated design for a response surface model”, *Quality and Reliability Engineering International*, Vol. 27, No. 4, pp. 541-554.
Morris, M. (2000). “A class of three-level experimental designs for response surface modeling”, *Technometrics*, Vol. 42, No. 2, pp. 111-121.

Myers, R. H., Montgomery, D. C., and Anderson-Cook, C. M. (2009), *Response Surface Methodology: Process and Product Optimization Using Designed Experiments (3rd Edition)*, Wiley.

Myers, R. H., Montgomery, D. C., Vining, G. G., Borror, C. M., and Kowalski, S. M. (2004), “Response surface methodology: A retrospective and literature survey”, *Journal of Quality Technology*, Vol. 36, No. 1, pp. 53-77.

Niki, N., Iwata, M., Hashiguchi, H., and Yamada, S. (2011), “Optimal selection and ordering of columns in supersaturated designs”, *Journal of Statistical Planning and Inference*, Vol. 141, No. 7, pp. 2449-2462.

Parker, P. A., Kowalski, S. M., and Vining, G. G. (2007), “Unbalanced and minimal point equivalent estimation second-order split-plot designs”, *Journal of Quality Technology*, Vol. 39, No. 4, pp. 376-388.

Plackett, R. L. and Burman, J. P. (1946), “The design of optimum multifactorial experiments”, *Biometrika*, Vol. 33, No. 4, pp. 305-325.

Tibshirani, R. (1996), “Regression shrinkage and selection via the lasso,” *Journal of the Royal Statistical Society. Series B*, Vol. 58, No. 1, pp. 267-288.

Vining, G. G. and Kowalski, S. M. (2008), “Exact inference for response surface designs within a split-plot structure”, *Journal of Quality Technology*, Vol. 40, No. 4, pp. 394-406.

Wu, C. F. J. (1993), “Construction of supersaturated designs through partially aliased interactions,” *Biometrika*, Vol. 80, No. 3, pp. 661-669.

Wu, C. F. J. and Hamada, M. S. (2009), *Experiments: Planning, Analysis, and Optimization (2nd Edition)*, Wiley.

Yamada, S. (2004), “Selection of active factors by stepwise regression in the data analysis of supersaturated design”, *Quality Engineering*, Vol. 16, No. 4, pp. 501-513.

Yamada, S., Ikebe, Y., Hashiguchi, H., and Niki, N. (1999), “Construction of three-level supersaturated design”, *Journal of Statistical Planning and Inference*, Vol. 81, No. 1, pp. 183-193.

Yamada, S. and Lin, D. K. J. (2002), “Construction of mixed-level supersaturated designs”, *Metrika*, Vol. 56, No. 3, pp. 205-214.

Acknowledgements

Matsuura’s portion of this work was supported by JSPS KAKENHI Grant Number 15K15952. The authors are grateful to two anonymous reviewers for their insightful and constructive comments.

Author’s biographical notes

Dr. Shun Matsuura is an Assistant Professor at Faculty of Science and Technology, Keio University, Japan. He received his Ph.D. degree from Keio University in 2009. His research interests include applied statistics, statistical quality control, and multivariate analysis.

Mr. Shogo Tajima is an Engineer at Yokohama Plant, NISSAN MOTOR CO.,LTD., Japan. He received his Master degree from Keio University in 2013. His research interests include industrial engineering, statistical quality control, and quality management.

Dr. Hideo Suzuki is a Professor at Faculty of Science and Technology, Keio University, Japan. He received his Ph.D. degree from Tokyo Institute of Technology in 1996. His research interests include applied statistics, quality management, and marketing research.