Response Surface Methodology I

1. Introduction

Response surface methodology is a collection of experimental strategies, mathematical methods, and statistical inference which enable an experimenter to make efficient empirical exploration of the system of interest.

The work which initially generated interest in the package of techniques was a paper by Box and Wilson in 1951.

Many times these procedures are used to optimize a process. For example, we may wish to maximize yield of a chemical process by controlling temperature, pressure and amount of catalyst.

The basic strategy has four steps:

1. Procedures to move into the optimum region.
2. Behavior of the response in the optimum region.
3. Estimation of the optimum conditions.
4. Verification.

Now let’s set up the problem. We have \( p \) factors. Call them \( x_1, x_2, \ldots, x_p \). We have a response \( y \), and a function \( \phi \), such that

\[
E(y) = \phi(x_1, x_2, \ldots, x_p).
\]

Initially, \( \phi \) is usually approximated by a first order regression model over narrow regions of \( x \), that is, where there is little curvature. That is,

\[
E(y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p = \beta_0 + \sum_{i=1}^{p} \beta_i x_i.
\]

In regions of higher curvature, especially near the optimum, second order models are commonly used:

\[
E(y) = \beta_0 + \beta_1 x_1 + \cdots + \beta_p x_p \\
+ \beta_{11} x_1^2 + \cdots + \beta_{pp} x_p^2 \\
+ \beta_{12} x_1 x_2 + \cdots + \beta_{p-1,p} x_{p-1} x_p \\
= \beta_0 + \sum_{i=1}^{p} \beta_i x_i + \sum_{i=1}^{p} \beta_{ii} x_i^2 + \sum_{i=1}^{p} \sum_{j>i} \beta_{ij} x_i x_j.
\]
Therefore, the overall strategy is to use first order models to “climb” the response surface and then higher order models to explore the optimum region.

Let us now consider the first phase of experimentation. There are basically two issues that will be considered. First the types of experimental designs that are used and then procedures to determine where the next experimental design should be run. Remember we are climbing the response surface.

2. First order models

First we will consider designs for fitting first order models. In a regression problem, in matrix notation

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

where

\[
\begin{bmatrix}
  y_1 \\
  y_2 \\
  \vdots \\
  y_n
\end{bmatrix} =
\begin{bmatrix}
  \beta_0 \\
  \beta_1 \\
  \vdots \\
  \beta_p
\end{bmatrix} +
\begin{bmatrix}
  \epsilon_1 \\
  \epsilon_2 \\
  \vdots \\
  \epsilon_n
\end{bmatrix},
\]

and \( X \) is the design matrix

\[
X = 
\begin{bmatrix}
  1 & x_{11} & x_{12} & \cdots & x_{1p} \\
  1 & x_{21} & x_{22} & \cdots & x_{2p} \\
  \vdots & \vdots & \vdots & \ddots & \vdots \\
  1 & x_{n1} & x_{n2} & \cdots & x_{np}
\end{bmatrix},
\]

but we will code the data to center it at 0, and use ±1.

Using the above coding, one of the designs used with first order models is \( 2^p \) factorials. For example, suppose that \( p = 3 \) and we wish to center the experiment as follows:

\[
x_1 = 225 \\
x_2 = 4.25 \\
x_3 = 91.5
\]

Next we must decide how far to extend the design from the center. As a rough guideline:

1. Make them far enough apart to allow the effect of the factor to be seen.
2. Make them not so far apart as to feel the surface is curving appreciably.

For example,

\[
x_1 \pm 25 \\
x_2 \pm .25 \\
x_3 \pm 1.5
\]

This gives

\[
x_1 \begin{array}{cc}
  200 & 250 \\
  4.0 & 4.5 \\
  90 & 93
\end{array}
\]
Now let
\[ x'_1 = x_1 - \frac{225}{25}, \]
\[ x'_2 = x_2 - \frac{4.25}{0.25}, \]
\[ x'_3 = x_3 - \frac{91.5}{1.5}, \]
giving
\[ X = \begin{bmatrix}
1 & -1 & -1 & -1 \\
1 & +1 & -1 & -1 \\
1 & -1 & +1 & -1 \\
1 & +1 & +1 & -1 \\
1 & -1 & -1 & +1 \\
1 & +1 & -1 & +1 \\
1 & -1 & +1 & +1 \\
1 & +1 & +1 & +1
\end{bmatrix}. \]

From regression methods we know that
\[ \hat{\beta} = (X'X)^{-1}X'y \]
\[ \text{Cov}(\hat{\beta}) = \sigma^2(X'X)^{-1}. \]

Notice that for our example
\[ (X'X)^{-1} = \begin{bmatrix}
1/8 & 0 & 0 & 0 \\
0 & 1/8 & 0 & 0 \\
0 & 0 & 1/8 & 0 \\
0 & 0 & 0 & 1/8
\end{bmatrix}. \]

Since this is diagonal the estimates of the regression coefficients are independent. Of course we already knew that from the work done earlier in this semester. At any rate,
\[ \hat{\beta}_0 = \bar{y}, \]
\[ \hat{\beta}_i = \frac{\text{effect } i}{2} \quad i = 1, 2, 3 \]
\[ \text{Var}(\hat{\beta}_i) = \frac{\sigma^2}{8} \quad i = 0, 1, 2, 3 \]

In evaluating the designs one question we might ask concerns the problem that we might run into if a 2nd order model is required.
Of course, we know that the interactions are orthogonal to the main effects, so they will be okay. The quadratic forms, \( x_i^2 \), will give a column of 1’s, thus they will be confounded with \( \hat{\beta}_0 \).
As we have studied earlier we can fractionate the 2-level design. For example, we might have a $2^3 - 1$ design where we let $C = AB$. The design matrix will be

$$X = \begin{bmatrix} 1 & -1 & -1 & +1 \\ 1 & +1 & -1 & -1 \\ 1 & -1 & +1 & -1 \\ 1 & +1 & +1 & +1 \end{bmatrix}.$$ 

Again, what happens if we need a quadratic model?

$$x_0 x_1 x_2 x_3 x_1 x_2 x_1 x_3 x_2 x_3 x_1^2 x_2^2 x_3^2$$

<table>
<thead>
<tr>
<th>Source</th>
<th>df</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
<td>7</td>
</tr>
<tr>
<td>Regression</td>
<td>3</td>
</tr>
<tr>
<td>Residual</td>
<td>4</td>
</tr>
</tbody>
</table>

The “residual” is a composite of both lack of fit and experimental error.

To allow for an estimate of experimental error and a little information about quadratic terms, the $2^p$ or $2^p - q$ design can be supplemented by $n_c$ center points.

For this design the design matrix is

$$x_0 x_1 x_2 x_3 x_1 x_2 x_1 x_3 x_2 x_3 x_1^2 x_2^2 x_3^2$$

+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 0 0 0 0
+1 0 0 0 0 0 0 0 0 0
+1 0 0 0 0 0 0 0 0 0
+1 0 0 0 0 0 0 0 0 0

What we notice is that the quadratic terms can be estimated independently from $\beta_0$, although not from each other. Also, as we have run more than one center point, we can obtain an estimate of experimental error. The ANOVA for this design is
While other designs have been proposed, these are the most popular for first order models.

Other first order designs can be used. One type, which we have not seen before, is called a simplex design. This design will be used extensively later on in the semester when we discuss Mixture Experimentation. But we will briefly introduce the concept of a simplex design now.

The simplex designs that we will discuss today are orthogonal designs which have \( n = p + 1 \) points. Geometrically, the design points represent the vertices of a \( p \)-dimensional regular sided figure, or simplex. For example, if \( p = 2 \), the points form an equilateral triangle.

The design may be constructed with the following procedure. Construct the design matrix,

\[
X^{(p+1)\times(p+1)},
\]

by letting \( X = \sqrt{n}O \), where \( O \) is an orthogonal matrix (i.e., \( O^{-1} = O^T \)). For example, for \( p = 2 \), we may construct \( O \) as follows. First, find a \( (p+1) \times (p+1) \) matrix whose columns are independent and whose first column is 1. For example,
Now divide each element of a given column by $\sqrt{\sum c_i^2}$, that is, the length of the column considered as a vector. For our example

- Column 1: $\sqrt{\sum c_i^2} = \sqrt{3}$
- Column 2: $\sqrt{\sum c_i^2} = \sqrt{2}$
- Column 3: $\sqrt{\sum c_i^2} = \sqrt{6}$

giving

$$O = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix}.$$ 

Note that $O'O = I$. Thus

$$X = \sqrt{n}O = \sqrt{3}O = \begin{bmatrix} 1/\sqrt{3} & 1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & -1/\sqrt{2} & -1/\sqrt{6} \\ 1/\sqrt{3} & 0 & 2/\sqrt{6} \end{bmatrix}.$$ 

Note that

$$X'X = \begin{bmatrix} 3 & 0 & 0 \\ 0 & 3 & 0 \\ 0 & 0 & 3 \end{bmatrix},$$

indicating that the design is orthogonal.

If $p = 3$, starting with $2^{3-1}$,

$$O = \begin{bmatrix} 1/2 & 1/2 & -1/2 & -1/2 \\ 1/2 & -1/2 & -1/2 & 1/2 \\ 1/2 & 1/2 & 1/2 & 1/2 \\ 1/2 & -1/2 & 1/2 & -1/2 \end{bmatrix}.$$ 

Now, $\sqrt{n}O = 2O$, so

$$X = \begin{bmatrix} 1 & 1 & -1 & -1 \\ 1 & -1 & -1 & 1 \\ 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & -1 \end{bmatrix}.$$ 

Notice that this is in fact a $2^{3-1}$ factorial.

*Draw picture.*

One final note is that this procedure will not necessarily generate a unique design; this is because different $O$’s can be generated. For example, with $p = 3$,

$$O = \begin{bmatrix} 1/2 & 0 & 1/\sqrt{2} & -1/2 \\ 1/2 & -1/\sqrt{2} & 0 & 1/2 \\ 1/2 & 0 & -1/\sqrt{2} & -1/2 \\ 1/2 & 1/\sqrt{2} & 0 & 1/2 \end{bmatrix}.$$
giving

\[
X = \begin{bmatrix}
1 & 0 & \sqrt{2} & -1 \\
1 & -\sqrt{2} & 0 & 1 \\
1 & 0 & -\sqrt{2} & -1 \\
1 & \sqrt{2} & 0 & 1 \\
\end{bmatrix}.
\]

While both designs are simplex designs one may be preferable to another regarding the biases of the regression coefficients against second order coefficients. As we have seen, with \(2^3\)–1

\[\hat{\beta}_0\] is confounded with \(\hat{\beta}_{11}, \hat{\beta}_{22}, \hat{\beta}_{33}\)
\[\hat{\beta}_1\] is confounded with \(\hat{\beta}_{23}\)
\[\hat{\beta}_2\] is confounded with \(\hat{\beta}_{13}\)
\[\hat{\beta}_3\] is confounded with \(\hat{\beta}_{12}\)

In the second design,

\[
\begin{array}{cccccccc}
x_0 & x_1 & x_2 & x_3 & x_1^2 & x_2^2 & x_3^2 & x_1x_2 & x_1x_3 & x_2x_3 \\
1 & 0 & \sqrt{2} & -1 & 0 & 2 & 1 & 0 & 0 & -\sqrt{2} \\
1 & -\sqrt{2} & 0 & 1 & 2 & 0 & 1 & 0 & -\sqrt{2} & 0 \\
1 & 0 & -\sqrt{2} & -1 & 0 & 2 & 1 & 0 & 0 & \sqrt{2} \\
1 & \sqrt{2} & 0 & 1 & 2 & 0 & 1 & 0 & \sqrt{2} & 0 \\
\end{array}
\]

This indicates that \(x_1^2 + x_2^2 = x_0\) and \(x_3^2 = x_0\). Therefore,

\[\hat{\beta}_0\] is confounded with \(\hat{\beta}_{11}, \hat{\beta}_{22}, \hat{\beta}_{33}\)
\[\hat{\beta}_1\] is confounded with \(\hat{\beta}_{13}\)
\[\hat{\beta}_2\] is confounded with \(\hat{\beta}_{23}\)
\[\hat{\beta}_3\] is confounded with \(\hat{\beta}_{11}\) and \(\hat{\beta}_{22}\)

Now let’s briefly discuss how the computations can be carried out for these simplex designs. Continuing with the example suppose we used \(2^3\)–1

\[
(1)x_1 & x_2 & x_3 = x_1x_2 & y \\
+1 & -1 & -1 & +1 & 51.6 \\
+1 & +1 & -1 & -1 & 54.1 \\
+1 & -1 & +1 & -1 & 31.2 \\
+1 & +1 & +1 & +1 & 51.6 \\
\]

You can see that \(X'X = \text{diag}(4, 4, 4, 4)\), and since

\[\hat{\beta} = (X'X)^{-1}X'y,\]

we have

\[\hat{\beta}_0 = \bar{y} = \frac{195.5}{4} = 48.9\]
\[\hat{\beta}_1 = \frac{112.7 - 82.8}{4} = 7.5\]
\[\hat{\beta}_2 = \frac{89.8 - 105.7}{4} = -4.0\]
\[\hat{\beta}_3 = \frac{110.2 - 85.3}{4} = 6.2\]
This gives us the estimated regression model
\[ \hat{y} = 48.9 + 7.5x_1 - 4.0x_2 + 6.2x_3. \]

The next issue is one of obtaining a graphical representation of the model. A procedure that has been found to be successful is one of using response contours. For two variables,

This may be obtained from the estimated regression model
\[ \hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2, \]
by solving for, say, \( x_2 \),
\[ x_2 = \frac{\hat{y} - \hat{\beta}_0}{\hat{\beta}_2} - \frac{\hat{\beta}_1}{\hat{\beta}_2} x_1. \]
Now all we need to do is select a value of \( \hat{y} \), say \( \hat{y}_0 \), and we have a line
\[ x_2 = \frac{\hat{y}_0 - \hat{\beta}_0}{\hat{\beta}_2} - \frac{\hat{\beta}_1}{\hat{\beta}_2} x_1. \]
This can be repeated for several values of \( \hat{y} \) giving the contour plot.