RESPONSE SURFACE CONTOUR PLOITING IN SAS

Daniel M. Chilko
West Virginia University
Morgantown, West Virginia

ABSTRACT


Response surface equations are frequently used to analyze data consisting of a dependent variable and a set of independent variables, each measured on a continuous scale. The SAS procedure GLM facilitates the development of such equations. A SAS procedure has been developed to facilitate the visual exploration of response surfaces by contour plotting.

INTRODUCTION

Research into scientific or engineering processes often results in data consisting of a dependent variable and a set of independent variables, each measured on a continuous scale. The dependent variable, or response, measures some important feature of the process, such as yield or cost. The independent variables, or input variables, are subject to the control of the experimenter and have some effect on the response. For example, the yield of some chemical process might depend on reaction temperature, reaction pressure, and concentration of reactant.

The collection of procedures, involving experimental strategy, mathematical methods, and statistical inference, which enable the experimenter to make an efficient empirical exploration of his process is referred to as response surface methodology. The literature of response surface methodology (RSM) is reviewed by Hill and Hunter (1966).

In general, we have

\[ y = f(x_1, x_2, ..., x_k) \]

where the form of \( f \) is unknown and perhaps extremely complicated. The success of RSM results from the fact that the function \( f \) can be approximated closely by a low-order polynomial in the independent variables, at least over some region of interest. This polynomial function is referred to as a response surface equation and usually is a first order equation

\[ y = b_0 + b_1 x_1 + b_2 x_2 \]

or a second order equation

\[ y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1^2 + b_4 x_2^2 \]

or a third order equation

\[ y = b_0 + b_1 x_1 + b_2 x_2 + b_3 x_1^2 + b_4 x_2^2 + b_5 x_1 x_2 + b_6 x_1^2 x_2 + b_7 x_1 x_2^2 + b_8 x_1^2 x_2^2 \]

The SAS procedure GLM facilitates estimation and testing of such equations by allowing the user to specify, for example:

PROC GLM;
MODEL Z = X X Y Y Y X Y X Y;

Using the estimated coefficients, it is possible to construct a response surface of predicted values. Frequently the experimenter is interested in determining what values of the independent variables are optimum as far as the predicted response is concerned (Myers, 1971). The response surface, in the form of contour plots, can be studied visually to gain insight into the relationships between the response and the input variables.

CONTOUR PLOITING

Contour plots are an attempt to represent a threedimensional surface in a two-dimensional medium. For example, the equation

\[ y = x_1^2 + x_2^2 \]

defines a surface as \( x_1 \) and \( x_2 \) vary from, say, -1 to +1. Contour plots attempt to depict this surface in the following way:

Imagine the plane formed by the \( x_1, x_2 \) coordinates. The response surface equation gives the height of the surface above (or below) this plane for any given point. If one constructs another plane parallel to the base plane, but at a fixed distance, say 0.2 units above it, then this plane intersects the response surface. If the intersection is projected back into the base plane, the projection consists of those points which satisfy the equation

\[ 0.2 = x_1^2 + x_2^2 \]

That is, the intersection of the surface and a plane at height 0.2 is a circle with radius equal to the square root of 0.2.

In general, any plane that intersects this surface at height \( h \) produces an intersection whose projection is a circle with radius equal to the square root of \( h \). By varying \( h \) over a set of values in equal steps, by projecting the intersections into the base plane, and by identifying different projections with different symbols, such as letters of the alphabet, a contour plot results (Figure 1).
RSP PROCEDURE

To facilitate the production of contour plots for response equations, we wrote a SAS procedure RSP. An example of its use is:

```sas
PROC RSP;
MODEL Z = 5.2 + .09*X - .005*X**2 + .1*Y - .0005*Y**2 + .0004*X*Y;
RANGES X = 0 TO 400 Y = 0 TO 300 Z = 3 TO 18 BY 5;
PLOT X*Y;
```

The output appears as Figure 2.

The `MODEL` statement specifies the response surface equation to be plotted by PROC RSP. The equation may be up to third order.

The `RANGES` statement specifies the values over which the variables used in the `MODEL` statement are to vary in producing the contour plots. A specification must appear in the `RANGES` statement for each variable, including the dependent variable, used in the `MODEL` statement.

The `PLOT` statement specifies the axes to be used for each contour plot. Any number of specifications can appear in the `PLOT` statement.

If the equation specified in the `MODEL` statement includes just two independent variables, then a plot specification will produce a single contour plot. If the equation involves more than two independent variables, then a series of contour plots are produced. First, each variable specified in the `MODEL` statement, but not used in that plot specification, is assigned a low value, as specified in the `RANGES` statement, and a contour plot produced. Then these variables are "stepped" through all combinations of values specified in the `RANGES` statement. At each step, a contour plot is produced.

REFERENCES


Myers, R.H., Response Surface Methodology, Allyn and Bacon, Boston, 246 pp., 1971.
Y SYMBOLS AND VALUES
A=0.0  B=0.2  C=0.4  D=0.6  E=0.8
F=1.0

Figure 1.
## STATISTICAL ANALYSIS SYSTEM

<table>
<thead>
<tr>
<th>Symbol</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>3</td>
</tr>
<tr>
<td>B</td>
<td>8</td>
</tr>
<tr>
<td>C</td>
<td>13</td>
</tr>
<tr>
<td>D</td>
<td>18</td>
</tr>
</tbody>
</table>

Figure 2.