LECTURE 7

Response Surface Methodology III

1. Canonical Form of Response Surface Models

To examine the estimated regression model we have several choices. First, we could plot response contours. Remember that we set $\hat{y}$ to some specified value, y_0, and trace out contours relating x_1, x_2, and x_3.

An alternative is to reduce the equation to its "canonical form." That is, we form an equation of the form

$$y - y_s = \lambda_1 w_1^2 + \lambda_2 w_2^2 + \lambda_3 w_3^2,$$

where y_s is the center of the contours (that is, the stationary point) and w_1, w_2, and w_3 are a new set of axes called the principal axes. The coefficients λ_1, λ_2, and λ_3 give the shape of the surface (they are the eigenvalues of a matrix to be defined shortly).

Using matrix notation for a bit, we could write the model as

$$\hat{y} = \hat{\beta}_0 + X'\hat{\beta} + X'\hat{B}X,$$

where here X is just the linear part of the design matrix, β is the vector of linear coefficients, and

$$B = \begin{bmatrix} \hat{\beta}_{11} & \hat{\beta}_{12}/2 & \dots & \hat{\beta}_{1k}/2 \\ \vdots & \hat{\beta}_{22} & \dots & \hat{\beta}_{2k}/2 \\ & & \ddots & \vdots \\ \hat{\beta}_{1k}/2 & \dots & \hat{\beta}_{kk} & \end{bmatrix}.$$

To find the stationary point X_s, differentiate $\hat{y}$ to find

$$\frac{\partial y}{\partial X} = \hat{\beta} + 2\hat{B}X.$$

Setting this to zero, we find

$$X_s = -\hat{B}^{-1}\hat{\beta}/2.$$

As to finding the right form for the λ's and the z's, you might recall that if you form a matrix M with columns equal to the normalized eigenvectors of $\hat{B}$, then

$$M'\hat{B}M = \Lambda,$$

where Λ is a diagonal matrix with diagonal elements equal to the eigenvalues of $\hat{B}$. Now write

$$\begin{aligned} Z &= X - X_s \\ W &= M'Z, \end{aligned}$$

we have

$$\begin{aligned} \hat{y} &= \hat{\beta}_0 + X'\hat{\beta} + X'\hat{B}X \\ &= \hat{\beta}_0 + (Z + X_s)'\hat{\beta} + (Z + X_s)'\hat{B}(Z + X_s) \\ &= [\hat{\beta}_0 + X_s'\hat{\beta} + X_s'\hat{B}X_s] + Z'\hat{\beta} + Z'\hat{B}Z + 2X_s'\hat{B}Z \\ &= \hat{y}_s + Z'\hat{B}Z, \end{aligned}$$

because $2X_s'\hat{B}Z = -Z'\hat{\beta}$ from the definition of X_s. Rotating the coordinate system, we have

$$\begin{aligned} \hat{y} &= \hat{y}_s + Z'\hat{B}Z \\ &= \hat{y}_x + W'M'\hat{B}MW \\ &= \hat{y}_s + M'\Lambda M, \end{aligned}$$

which is what we want.

The reason to do this is because the eigenvalues, the diagonal values of Λ, can tell a great deal about the stable point.

- If the eigenvalues are all negative, the stable point is a maximum.
- If the eigenvalues are all positive, the stable point is a minimum.
- If the eigenvalues are of mixed sign, the stable point is a saddle.

That's not all. The relative sizes of the eigenvalues also tell a great deal. For example, if most of the eigenvalues are large positive numbers but a few are near zero, then there is a ridge in the graph of the response function. Moving along that ridge will make little difference in the value of the response (but might make a big difference in some other aspect of the system, like cost, for example).

Let's illustrate this with an example. Say

$$\hat{y} = 81.22 + 1.97x_1 + 0.22x_2 - 3.93x_1^2 - 1.38x_2^2 - 2.22x_1x_2.$$

To find y_s, we must find the values of x_1 and x_2 which represent the stationary point.

$$\begin{aligned} \frac{\partial \hat{y}}{\partial x_1} &= 1.97 - 7.86x_1 - 2.22x_2 = 0 \\ \frac{\partial \hat{y}}{\partial x_2} &= 0.22 - 2.76x_2 - 2.22x_1 = 0 \end{aligned}$$

Solving this system of equations gives

$$\begin{aligned} x_{1,s} &= 0.30 \\ x_{2,s} &= -0.16, \end{aligned}$$

which in turn gives

$$\hat{y}_s = 81.49.$$

To find the eigenvalues of B, we solve

$$\begin{aligned}
0 &= |B - \lambda I| \\
&= \begin{vmatrix} \hat{\beta}_{11} - \lambda & \hat{\beta}_{12}/2 \\ \hat{\beta}_{12}/2 & \hat{\beta}_{22} - \lambda \end{vmatrix} \\
&= \begin{vmatrix} -3.93 - \lambda & -1.11 \\ -1.11 & -1.38 - \lambda \end{vmatrix} \\
&= (-3.93 - \lambda)(-1.3\lambda) - (-1.11)(-1.11) \\
&= \lambda^2 + 5.31\lambda + 4.19
\end{aligned}$$

This yields

$$\lambda = \frac{-5.31 \pm \sqrt{5.31^2 - 4(1)(4.19)}}{2} = \frac{-5.31 \pm 3.38}{2}$$

or

$$\begin{aligned}
\lambda_1 &= -4.35 \\
\lambda_2 &= -0.96.
\end{aligned}$$

Note that the choice of λ_1 or λ_2 is not important.

The next step, following the determination the eigenvalues, is finding the eigenvectors.

Graphically, what we are doing is (in two dimensions)

Draw picture.

w_1 and w_2 are the major and minor axes. The eigenvectors corresponding to each λ_i can give

$$z_i = m_{i1}(x_1 - x_{1,s}) + m_{i2}(x_2 - x_{2,s}),$$

which shows how the axes are translated. If we let

$$M = \begin{bmatrix} m_{11} & m_{12} \\ m_{21} & m_{22} \end{bmatrix} = \begin{bmatrix} M_1 & M_2 \end{bmatrix},$$

then the M_i are the eigenvectors determined by

$$[B - \lambda I]M_i = 0.$$

For example, for $\lambda_1 = -4.35$,

$$\begin{bmatrix} -3.93 + 4.35 & -1.11 \\ -1.11 & -1.38 + 4.35 \end{bmatrix} \begin{bmatrix} m_{11} \\ m_{21} \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix},$$

or

$$\begin{aligned}
0.42m_{11} - 1.11m_{21} &= 0 \\
-1.11m_{11} + 2.97m_{21} &= 0.
\end{aligned}$$

Note that there is no unique solution to this system of equations, because λ_1 was chosen to make the matrix of coefficients singular.

What we want is to find a proportional relationship, subject to the constraint that $m_{11}^2 + m_{21}^2 = 1$. For example, setting $m'_{11} = 1$, we find $m'_{21} = .42/1.11 = 1.11/2.97 = .3784$. Since $\sqrt{1^2 + .3784^2} = 1.07$, we have $m_{11} = 1/1.07 = 0.94$ and $m_{21} = .3784/1.07 = 0.35$. Similarly, we could find that $m_{21} = 0.35$ and $m_{22} = -0.94$, so

$$M = \begin{bmatrix} .94 & .35 \\ .35 & -.94 \end{bmatrix}.$$

Note that the eigenvectors are orthogonal.

At any rate, we now know that

$$\begin{aligned} w_1 &= .94(x_1 - .30) + .35(x_2 + .16) \\ w_2 &= .35(x_1 - .30) - .94(x_2 + .16). \end{aligned}$$

Draw picture

As mentioned earlier, when all eigenvalues are negative the stationary point is a (local) maximum. When they are not all equal, the function looks like an ellipsoid near the maximum. In our case, the ellipses of the contour plots are elongated along the w_2 axis. This means a small value in w_1 corresponds to a larger value for w_2 as far as giving the same value of $\hat{y} - y_s$.

Having found the canonical form we can now find how it can be used to describe the surface without having to plot the response contours. First, $(x_{1,s}, x_{2,s})$ is the stationary point with y_s being the response at the stationary point. The eigenvalues λ_1, λ_2 give the behavior of the response as we move away from the stationary point.

In our example, $\lambda_1 = -4.35$ and $\lambda_2 = -.96$, which means that the response decreases as we move away from the stationary point. Furthermore, since

$$|\lambda_1| > |\lambda_2|$$

we know that the contours are elongated along the w_2 axis.

If $\lambda_1 = \lambda_2$ then the ellipses are actually into circles.

If λ_1 and λ_2 are both positive then the response is minimized at the stationary point.

If λ_2 is close to zero then we have a stationary ridge. This means that we have a variety of x_1 and x_2 which will result in the maximum value.

In the above situations, the optimum conditions have been reached. All that remains is that we run a experiment in the region the verify the model.

If $\lambda_1 < 0$ and $\lambda_2 > 0$ then we have found a stationary point that is a saddle point. For example

Draw picture

. We have reached a maximum on one axis and a minimum along the other axis. If our objective was to maximize the response then our next experiment should be in the direction of w_2.

If λ_1 and λ_2 are both negative but the stationary point lies well outside the region of the design

Draw picture

This is called a rising ridge. It means we have not yet found the proper region of the maximum, a common occurance. Again the next experiment needs to be run up the ridge.

2. Blocking in Response Surface Designs

The last issue that we will discuss is that of blocking response surface designs. We have already discussed blocking of two-level designs, which can be used in first order designs.

Now we will discuss blocking of second order models. The important issue is to determine how, if at all, one can assign treatments to blocks so that the block effects will be orthogonal to the model coefficients.

First we must consider the conditions necessary for the parameter estimates to be orthogonal to blocks. Let z_{mu} be a dummy variable which takes value $+1$ if the u-th point is in the m-the block and zero otherwise. The model is then

$$Y_u = \beta_0 + \sum_{i=1}^{k} \beta_i x_{iu} + \sum_{i=1}^{k} \beta_{ii} x_{iu}^2 + \sum_{i=1}^{k} \sum_{j>i} \beta_{ij} x_{iu} x_{ju} + \sum_{m=1}^{b} \delta_m (z_{mu} - \bar{z}_m) + \epsilon_u.$$

To have orthogonality of parameter estimates and block effects, we need

$$\sum_{u=1}^{N} x_{iu}(z_{mu} - \bar{z}_m) = 0 \qquad \text{for } i \le k \text{ and all } m$$

$$\sum_{u=1}^{N} x_{iu}^2(z_{mu} - \bar{z}_m) = 0 \qquad \text{for } i \le k \text{ and all } m$$

$$\sum_{u=1}^{N} x_{iu} x_{ju}(z_{mu} - \bar{z}_m) = 0 \qquad \text{for } i \ne j \le k \text{ and all } m$$

With the usual coding for x's so that $\sum x_{iu} = 0$ and $\sum x_{iu} x_{ju} = 0$ for $i \ne j \le k$, these conditions are equivalent to

$$\sum_{i=1}^{N} x_{iu} z_{mu} = 0 \qquad \text{for } i \le k \text{ and all } m$$

$$\sum_{u=1}^{N} x_{iu}^2 z_{mu} = \sum_{u=1}^{N} x_{iu}^2 \bar{z}_m \qquad \text{for } i \le k \text{ and all } m$$

$$\sum_{u=1}^{N} x_{iu} x_{ju} z_{mu} \qquad \text{for } i \ne j \le k \text{ and all } m.$$

The first equation implies that the sum of the observations in the mth block, for variable x_i, is zero. The third equation implies that the cross-product of x_i and x_j sums to 0 in the mth block.

In the second equation, $\bar{z}_m$ is the proportion of observations that occur in block m. Therefore, this equation specifies that the contributions from block m to the total SS for each variable x_i is proportional to the number of runs in the block.

For example, consider the following design.

x_1	x_2	x_3	
1	1	1	
1	-1	-1	
-1	1	-1	
-1	-1	1	Block 1
0	0	0	
0	0	0	
1	1	-1	
1	-1	1	
-1	1	1	Block 2
-1	-1	-1	
0	0	0	
0	0	0	
-1.633	0	0	
1.633	0	0	
0	-1.633	0	
0	1.633	0	Block 3
0	0	-1.633	
0	0	1.633	
0	0	0	
0	0	0	

As for the conditions

- The first equation: $\sum x_{iu} = 0$ within each block.
- The third equation: $\sum x_{iu}x_{ju} = 0$ within each block.
- The second equation. Now $\sum_{u=1}^{N} x_{iu}^2 = 13.33$, the total SS for each variable. For block 1, the sum $\sum_{u=1}^{N} x_{iu}^2 z_{mu} = 4$. For block 2, the sum is 4. For block 3, the sum is 5.33. On the other hand, for block 1, $\sum x_{iu}^2 \bar{z}_m = 4$. For blocks 2 and 3, the sums are 4 and 5.33

Since the conditions are met, the blocking in this design will be orthogonal to the parameter estimates.

In general, for a central composite design, the question that remains is what value of α do we select in order to make the blocks orthogonal to the regression parameters.

First, let's say we want two blocks. In this case we will place the 2^p points in one block and the axial points in the second block. In addition, we'll place c_F center points in the factorial block and C_A center points in the axial block.

From the second condition,

$$\frac{N_A}{N} SS_{TOT} = SS_A, \qquad \frac{N_F}{N} SS_{TOT} = SS_F,$$

which implies that

$$\frac{N_A}{N_F} = \frac{SS_A}{SS_F}.$$

In turn, this means

$$\frac{\sum_{\text{axial block}} x_{iu}^2}{\sum_{\text{factorial block}} x_{iu}^2} = \frac{\text{number of points in axial block}}{\text{number of points in factorial block}} = \frac{2k + C_A}{2^k + C_F}.$$

The left hand side of this equation is simply $2\alpha^2/2^k$, so

$$\frac{2\alpha^2}{2^k} = \frac{2k + C_A}{2^k + C_F},$$

yielding

$$\alpha = \sqrt{\frac{2^k(2k + C_A)}{2(2^k + C_F)}}.$$

Thus is the experimenter requires two blocks, the value of α given by this equation, for the specified values of C_A and C_F, gives a Central Composite Design that blocks orthogonally. For example, if $k = 3$ and $C_A = C_F = 2$, then

$$\alpha = \sqrt{\frac{2^3(2 \cdot 3 + 2)}{2(2^3 + 2)}} \approx 1.7889.$$

What about rotatability? We know that that is also a desirable criterion. We have seen that for rotatability,

$$\alpha = 2^{k/4}.$$

This implies that, for the design to be both rotatable and blocked orthogonally,

$$2^{k/4} = \sqrt{\frac{2^k(2k + C_A)}{2(2^k + C_F)}}$$

or

$$2^{k/2} = \frac{2^k(2k + C_A)}{2(2^k + C_F)}.$$

Thus, for a given value of k, the question remains of finding an appropriate value for C_A and C_F to make the design rotatable. For example, if $k = 2$, we find

$$2 = \frac{4(4 + C_A)}{2(4 + C_F)},$$

or $C_F = C_A$.

If $k = 3$, we get $1.5 + C_A = .5625C_F$, making no solution possible. That doesn't mean, however, that you might not pick a design with orthogonal blocking and near-rotatability. For example, with $C_A = 2$ and $C_F = 3$, for orthogonal blocking we need $\alpha = 1.7056$. For rotatability, we would want $\alpha = 1.6818$. By going with $C_A = 2, C_F = 3$, and $\alpha = 1.7056$, we will get orthogonal blocking and near-rotatability.

For $k = 4$, we get $C_F = 2C_A$.

Some experimental situations dictate the need for more than 2 blocks. To achieve this we fractionate the 2-level portion of the design. We have discussed how we can fractionate this portion of the design earlier. Again we need to determine the appropriate value for α.

$$\alpha = \sqrt{\frac{2^{k-p}(2k + C_A)}{2(2^{k-p} + C_F)}}.$$

For example, if $C_A = C_F = 2, k = 3$, and $p = 1$, we have

$$\alpha = \sqrt{\frac{2^{3-1}(2 \cdot 3 + 2)}{2(2^{3-1} + 2)}} \approx 1.7889,$$

the design that we saw earlier.