Response Surface Methodology

Design of Experiments - Montgomery
Chapter 11

Response Surface Methodology

- Response y and factors $\mathbf{x}$
- Factors influence response in unknown way
- Describe influence using model $f(\mathbf{x})$
- Objective is to find levels which maximize response

$$y = f(x_1, x_2, \ldots, x_k) + \epsilon$$

- ϵ represents noise or error in response
- Call $\eta = f(x_1, x_2, \ldots, x_k)$ the response surface
- Maximize response by maximizing response surface
- Contours - values of $\mathbf{x}$ such that η is constant

1st and 2nd Order Approximations

- Use suitable approximation of f to maximize
  - First order - Linear function of factors

$$y = \beta_0 + \beta_1 x_1 + \ldots + \beta_k x_k + \epsilon$$

  - Second order - Quadratic function of factors

$$y = \beta_0 + \sum \beta_i x_i + \sum \beta_{ii} x_i^2 + \sum \sum \beta_{ij} x_i x_j + \epsilon$$

- While two approximations unrealistic in general
- Often quite realistic in small region of surface
- Use sequential approach to find optimum

1 Response surface design determines $\{x_i\}$

2 Least squares to estimate parameters

3 Use contours to move in optimal direction

The Method of Steepest Ascent

- Move rapidly to general vicinity of optimum
- Use approximate model to move in proper direction
- Consider linear model

$$\widehat{\eta} = \widehat{\beta}_0 + \sum \widehat{\beta}_i x_i$$

- Contours of x_i and x_j are series of parallel lines
- Move in direction which increase $\widehat{\eta}$ the quickest

  Move perpendicular to contour lines
- Direction based on slope estimates $\widehat{\beta}_i$ and $\widehat{\beta}_j$
- Often center points to make determination easier
- Choose step size for one of the variables (Δx_j)
- Move others accordingly (Table 11-3)

$$\Delta x_i = \frac{\widehat{\beta}_i}{\widehat{\beta}_j / \Delta x_j}$$

Example

Consider Problem 11.1 in Montgomery. A chemical plant produces oxygen by liquefying air and separating it into its component gases by fractional distillation. Current operating conditions are Temp = -220°C and pressure ratio of 1.2. Interested in maximizing the purity of oxygen.

$$\begin{aligned} \text{temp} &= \frac{\text{Temp} + 220}{5} \\ \text{pres} &= \frac{\text{Ratio} - 1.2}{.1} \end{aligned}$$

```
options nocenter ls=75;

data purity;
input temp pres pure;
x1x1 = temp*temp;
x1x2 = temp*pres;
cards;
-1 -1 82.8
-1  1 83.5
 1 -1 84.7
 1  1 85.0
 0  0 84.1
 0  0 84.5
 0  0 83.9
 0  0 84.3
;

proc reg;
model pure = temp pres x1x1 x1x2;
```



```
Dependent Variable: PURE
                          Sum of         Mean
Source           DF      Squares       Square      F Value       Prob>F
Model             4      3.26000      0.81500       12.225       0.0335
Error             3      0.20000      0.06667
C Total           7      3.46000

Parameter Estimates

                 Parameter      Standard    T for H0:
Variable  DF      Estimate         Error   Parameter=0     Prob > |T|
INTERCEP   1     84.200000    0.12909944       652.210         0.0001
TEMP       1      0.850000    0.12909944         6.584         0.0071
PRES       1      0.250000    0.12909944         1.936         0.1482
X1X1       1     -0.200000    0.18257419        -1.095         0.3534
X1X2       1     -0.100000    0.12909944        -0.775         0.4950
```

A one degree change in temp is equivalent to a 1/5=.2 temp step once standardized. Thus means we increase pressure by (.2)(.25/.85) = .059. The following table summarizes moving in this direction and the observed purity (**simulated results**). It appears that a maximum is reached around ten steps. Another linear approximation should be made centered now at (-210,1.2590).

$$\widehat{y}_{ij} = 84.10 + .85x_T + .25x_P$$

	Coded Variables		Natural Variables		
Steps	x_1	x_2	Temp	Pres	Response
0	0.0	0.000	-220	1.2	84.2
1	0.2	0.059	-219	1.2059	84.7
5	1.0	0.295	-215	1.2295	85.2
10	2.0	0.590	-210	1.2590	85.3
15	3.0	0.885	-205	1.2885	85.1
20	4.0	1.180	-200	1.3180	84.7

Checking Linear Approximation

- Do Lack of Fit Test

  Use the center points to estimate pure error. Use $SS_{\text{Interaction}}$ or $\widehat{\beta}_{12}$ (created by adding x_1x_2 term in the model) to compute lack of fit error.

  $$\widehat{\sigma}^2 = \frac{84.1^2 + \cdots + 84.3^2 - 336.8^2/4}{3} = \frac{.2}{3} = .0666$$

  If linear, there will be no interaction effect. The estimate $\widehat{\beta}_{12}$ is simply 1/2 the estimated effect which is

  $$2\widehat{\beta}_{12} = .5(82.8 + 85.0 - 83.5 - 84.7) = -.2$$

  The SS_{12}=$(-.4)^2/4$ = .04. This is also the lack-of-fit SS. The F test is .04/.0666 = 0.6 and has a P-value of around .5.

- Compare average of center points to design points

  If there is no curvature, the average of the four design points should be equal to the average of the center points. The difference between these two averages is an estimate of the pure quadratic term $\beta_{11} + \beta_{22}$. In this example, it is

  $$336/4 - 336.8/4 = -.2$$

  The SS is $(4)(4)(-.2)^2/(4+4)$ = .08. The F test is .08/.06666 = 1.33 and has a P-value of around .3. This also suggests that the linear model is appropriate.

Analysis of Second Order Model

- Due to curvature, determine **stationary point**

  $$\frac{\partial \hat{\eta}}{\partial x_i} = 0$$

- Stationary point may be minimum, maximum, or saddle point
- Use contours or canonical analysis to determine behavior
- Can write quadratic approximation as (page 440)

  $$y = \beta_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x} + \epsilon$$

- Solution is $\mathbf{x}_s = -.5\mathbf{B}^{-1}\mathbf{b}$
- Canonical analysis looks at eigenvalues and eigenvectors of $\mathbf{B}$

Example

Consider Problem 11.8 in Montgomery. An experimenter wants to optimize crystal growth as a function of three variables x_1, x_2, and x_3. The design used is a factorial with six center points and six axial points (see Figure 11-20). We will use Proc RSREG which does a quadratic response surface analysis.

```
options nocenter ls=75;

data purity;
input x1 x2 x3 resp @@;

cards;
-1 -1 -1  66 -1 -1  1  70 -1  1 -1  78 -1 1 1 60
 1 -1 -1  80  1 -1  1  70  1  1 -1 100  1 1 1 75
-1.682 0 0 100 1.682 0 0 80 0 -1.682 0 68 0  1.682 0 63
0 0 -1.682 65 0 0  1.682 82 0 0 0 113
0 0 0 100 0 0 0 118 0 0 0 88 0 0 0 100 0 0 0 85
;

proc rsreg;
model resp=x1 x2 x3 / lackfit;
```


Breaks down regression SS into linear, quadratic terms

Regression	Degrees of Freedom	Type I Sum of Squares	R-Square	F-Ratio	Prob > F
Linear	3	77.854973	0.0141	0.139	0.9341
Quadratic	3	3291.741253	0.5960	5.896	0.0139
Crossproduct	3	292.375000	0.0529	0.524	0.6757
Total Regress	9	3661.971227	0.6630	2.186	0.1194

```
**Appears to be a quadratic component but very little crossproduct
**Overall regression not significant but that's not concern
**Would expect contours to be fairly circular
```

Does lack of fit test

Residual	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob > F
Lack of Fit	5	1001.645440	200.329088	1.166	0.4353
Pure Error	5	859.333333	171.866667		
Total Error	10	1860.978773	186.097877		

```
**No apparent lack of fit to quadratic model
```


Gives the parameter estimates and standard errors

Parameter	Deg of Freedom	Parameter Estimate	Standard Error	T for H0: Parameter=0
INTERCEPT	1	100.666301	5.563818	18.093
X1	1	1.271027	3.691248	0.344
X2	1	1.361082	3.691248	0.369
X3	1	-1.494042	3.691248	-0.405
X1*X1	1	-3.767908	3.592852	-1.049
X2*X1	1	2.875000	4.823094	0.596
X2*X2	1	-12.427833	3.592852	-3.459
X3*X1	1	-2.625000	4.823094	-0.544
X3*X2	1	-4.625000	4.823094	-0.959
X3*X3	1	-9.600102	3.592852	-2.672

Performs canonical analysis

Factor	Degrees of Freedom	Sum of Squares	Mean Square	F-Ratio	Prob > F
X1	4	347.989342	86.997336	0.467	0.7587
X2	4	2489.210554	622.302639	3.344	0.0553
X3	4	1585.399212	396.349803	2.130	0.1515

Canonical Analysis of Response Surface (based on coded data)

Factor	Critical Value Coded	Uncoded
X1	0.154420	0.259735
X2	0.065908	0.110858
X3	-0.083251	-0.140028

Predicted value at stationary point   101.011413

Eigenvalues	Eigenvectors X1	X2	X3
-8.711273	0.941384	0.210046	-0.263964
-25.327160	0.330987	-0.424011	0.843008
-38.941204	-0.065147	0.880963	0.468679

Stationary point is a maximum.

Response Surface Designs

- Want points sparse but give info over entire region
- Blocking is possible
- Designs can be built sequentially
- Provides internal estimate of error/lack of fit
- Orthogonal first order designs
  - Minimize the variance of regression coefficients
  - 2^k Factorial (with center points)
  - Fractional factorial of resolution III or higher
- Central composite designs (2nd order)
  - 2^k factorial
  - Fractional factorial of resolution V or higher
  - Add center points and $2k$ axial runs