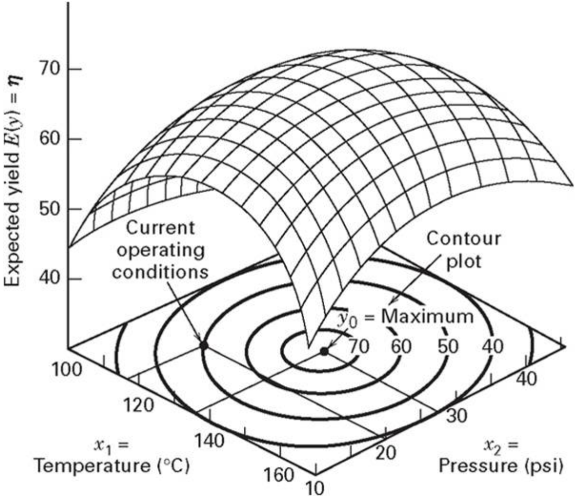
Lec 11: Response Surface Methodology

Ying Li

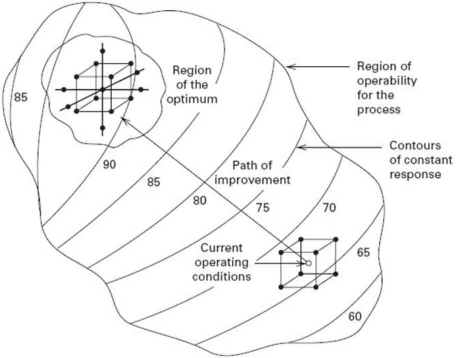
December 6, 2011

Response Surface Methodology (RSM)

- A collection of mathematical and statistical techniques;
- Model and analysis of problems in which a response of interest;
- The objective is to optimize the response.


Expected yield $E(y) = \eta$
70
60
50
40
Current operating conditions
Contour plot
y_0 = Maximum
70
60
50
40
100
120
140
160
x_1 =
Temperature (°C)
10
20
30
40
x_2 =
Pressure (psi)

RSM is a sequential procedure



■ FIGURE 11.3 The sequential nature of RSM

- Factor screening
- Finding the region of the optimum
- Modeling & Optimization of the response

Response Surface Models

- Screening:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$$

- Steepest ascent:

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$$

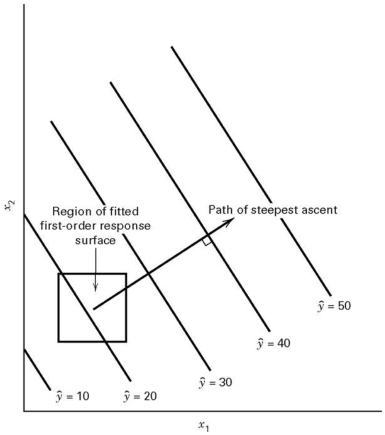
- Optimization

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \varepsilon$$

Steepest Ascent

- A procedure for moving sequentially from an initial "guess" towards to region of the optimum.
- Based on the fitted first-order model

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x_1 + \hat{\beta}_2 x_2$$



An Example of Steepest Ascent

A chemical engineer is interested in determining the operating conditions that maximize the yield of a process. Two controllable variables: reaction time (ξ_1) and reaction temperature (ξ_2).

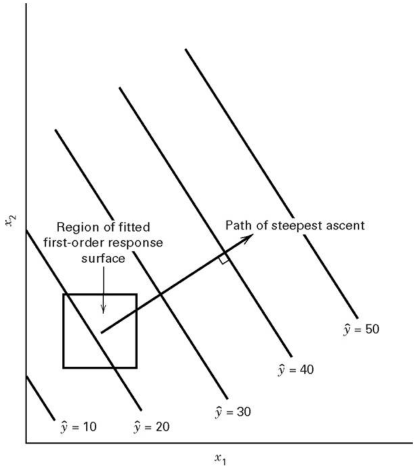
■ **TABLE 11.1**

Process Data for Fitting the First-Order Model

Natural Variables		Coded Variables		Response
ξ_1	ξ_2	x_1	x_2	y
30	150	−1	−1	39.3
30	160	−1	1	40.0
40	150	1	−1	40.9
40	160	1	1	41.5
35	155	0	0	40.3
35	155	0	0	40.5
35	155	0	0	40.7
35	155	0	0	40.2
35	155	0	0	40.6

A first-order model may be fit to these data by least squares.

$$\hat{y} = 40.44 + 0.775x_1 + 0.325x_2$$



■ **TABLE 11.3**

Steepest Ascent Experiment for Example 11.1

Steps	Coded Variables x_1	Coded Variables x_2	Natural Variables ξ_1	Natural Variables ξ_2	Response y
Origin	0	0	35	155	
Δ	1.00	0.42	5	2	
Origin + Δ	1.00	0.42	40	157	41.0
Origin + 2Δ	2.00	0.84	45	159	42.9
Origin + 3Δ	3.00	1.26	50	161	47.1
Origin + 4Δ	4.00	1.68	55	163	49.7
Origin + 5Δ	5.00	2.10	60	165	53.8
Origin + 6Δ	6.00	2.52	65	167	59.9
Origin + 7Δ	7.00	2.94	70	169	65.0
Origin + 8Δ	8.00	3.36	75	171	70.4
Origin + 9Δ	9.00	3.78	80	173	77.6
Origin + 10Δ	10.00	4.20	85	175	80.3
Origin + 11Δ	11.00	4.62	90	179	76.2
Origin + 12Δ	12.00	5.04	95	181	75.1

■ TABLE 11.4

Data for Second First-Order Model

Natural Variables		Coded Variables		Response
ξ_1	ξ_2	x_1	x_2	y
80	170	−1	−1	76.5
80	180	−1	1	77.0
90	170	1	−1	78.0
90	180	1	1	79.5
85	175	0	0	79.9
85	175	0	0	80.3
85	175	0	0	80.0
85	175	0	0	79.7
85	175	0	0	79.8

Summary of steepest ascent

- Points on the path of steepest ascent are proportional to the magnitudes of the model regression coefficients.
- The direction depends on the sign of the regression coefficient.

Second-order model

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum\sum_{i<j} \beta_{ij} x_i x_j + \epsilon \tag{11.4}$$

- These models are used widely in practice.
- Fitting the model is easy, some nice designs are available.
- Optimization is easy.
- There is a lot of empirical evidence that they work very well.

General solution for second-order model

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<j}\sum \beta_{ij} x_i x_j + \epsilon \qquad \textbf{(11.4)}$$

$$\hat{y} = \hat{\beta}_0 + \mathbf{x}'\mathbf{b} + \mathbf{x}'\mathbf{B}\mathbf{x}$$

$$\mathbf{x} = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_k \end{bmatrix} \quad \mathbf{b} = \begin{bmatrix} \hat{\beta}_1 \\ \hat{\beta}_2 \\ \vdots \\ \hat{\beta}_k \end{bmatrix} \quad \text{and} \quad \mathbf{B} = \begin{bmatrix} \hat{\beta}_{11}, \hat{\beta}_{12}/2, \ldots, \hat{\beta}_{1k}/2 \\ \hat{\beta}_{22}, \ldots, \hat{\beta}_{2k}/2 \\ \ddots \\ \text{sym.} \qquad \hat{\beta}_{kk} \end{bmatrix}$$

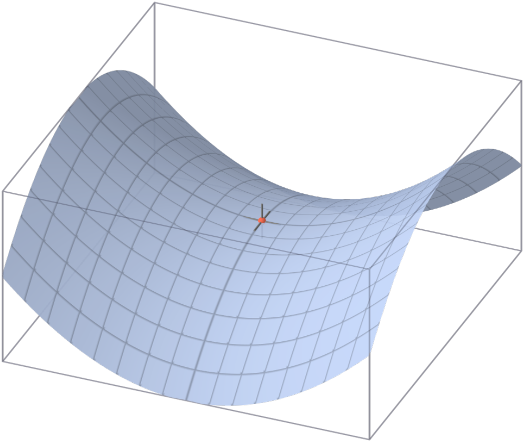
The stationary point

$$\frac{\partial \hat{y}}{\partial \mathbf{x}} = \mathbf{b} + \mathbf{2Bx} = 0$$

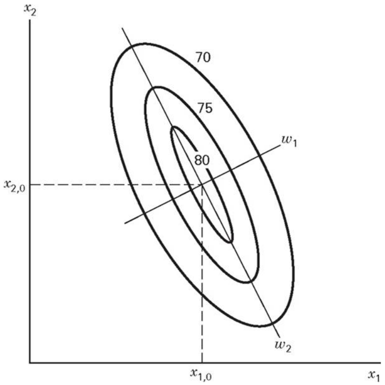
$$\mathbf{x}_s = -\frac{1}{2}\mathbf{B}^{-1}\mathbf{b} \Leftarrow \text{the stationary point}$$

$$\hat{y}_s = \hat{\beta}_0 + \frac{1}{2}\mathbf{x}_s'\mathbf{b}$$

The stationary point could represent a point of **maximum response**, a point of **minimum response**, or a **saddle point**.



Canonical Analysis



$$\mathbf{x}_s = -\frac{1}{2}\mathbf{B}^{-1}\mathbf{b} \quad \hat{y}_s = \hat{\beta}_0 + \frac{1}{2}\mathbf{x}_s'\mathbf{b}$$

The canonical form:

$$\hat{y} = \hat{y}_s + \lambda_1\omega_1^2 + \lambda_2\omega_2^2 + \cdot + \lambda_k\omega_k^2$$

the $\{\lambda_i\}$ are the eigenvalues or characteristic roots of the matrix $\mathbf{B}$

Canonical Analysis

The nature of the response surface can be determined from the stationary point & the signs and magnitudes of the $\{\lambda_i\}$.

- all positive: a minimum is found
- all negative: a maximum is found
- mixed: a saddle point is found

The response surface is steepest in the direction (canonical) corresponding to the largest absolute eigenvalue

Designs for the first-order model

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon$$

orthogonal first-order designs

A first-order design is orthogonal is the off-diagonal elements of the $\mathbf{XX}'$ matrix are all 0.

Designs for the first-order model

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \varepsilon$$

orthogonal first-order designs

A first-order design is orthogonal is the off-diagonal elements of the $\mathbf{XX}'$ matrix are all 0.

- 2^k experiment
- 2^k experiment with center runs ($x_i = 0$).
  - makes it possible to estimate the variance.
  - does not influence the estimates of β_i.
  - the estimate of β_0 is the ground mean.
- Simplex

Simplex

Simplex: a regularly sided figure with $k+1$ vertices in k dimensions.

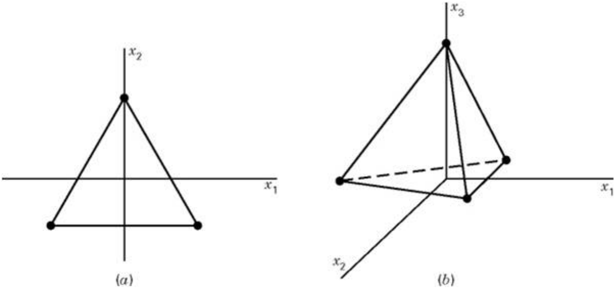


Figure: The simplex design for $k = 2$ and $k = 3$ variables

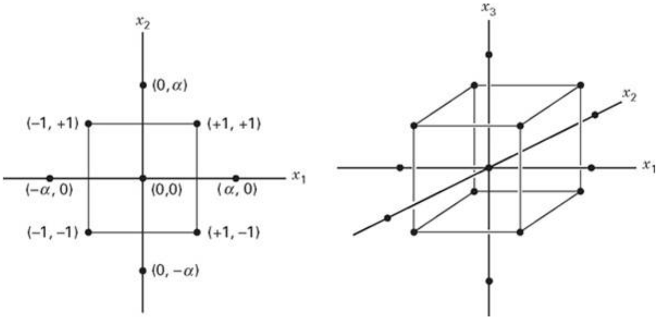
Designs for the second order model

$$y = \beta_0 + \sum_{i=1}^{k} \beta_i x_i + \sum_{i=1}^{k} \beta_{ii} x_i^2 + \sum_{i<j}\sum \beta_{ij} x_i x_j + \epsilon \qquad \textbf{(11.4)}$$

- **Central composite design**
- Box-Behnken design
- Face-centered design
- Equiradial design
- small composite design & hybrid design

Central composite design

The CCD consists of a 2^k factorial with n_F factorial runs, $2k$ axial or star runs, and n_c center runs.



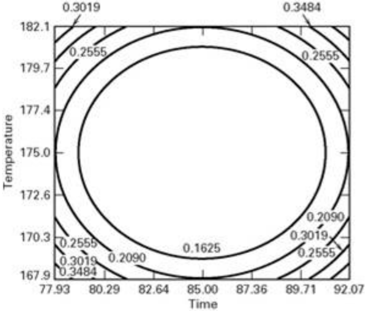
Two parameters in such design: the distance α of the axial runs from the center; the number of the center points n_c

Rotatability

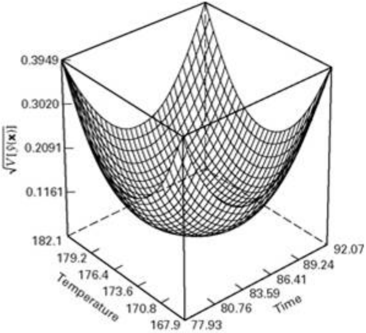
Rotatable

A second-order response surface design is rotatable if the variance of the predicted response $V[\hat{y}(x)]$ is the same at all the points of $\mathbf{x}$ that are at the same distance form the design center: $\alpha = (n_F)^{1/4}$

Rotatability

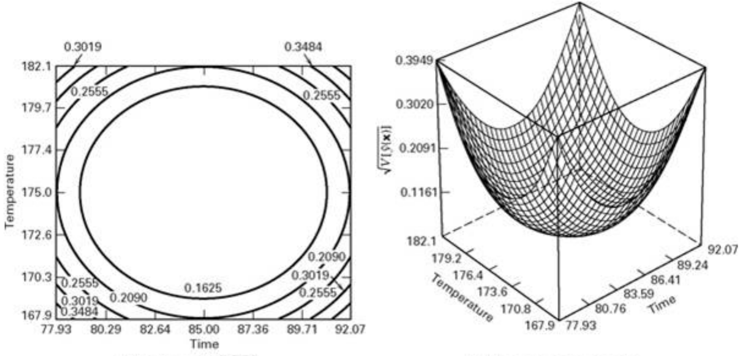


(a) Contours of $\sqrt{V[\hat{y}(\mathbf{x})]}$



(b) The response surface plot

Rotatability



(a) Contours of $\sqrt{V[\hat{y}(\mathbf{x})]}$ (b) The response surface plot

A rotatable design gives the same prediction in all directions.

Spherical CCD

All factorial runs and axial runs have the same distance to the center: $\alpha = (k)^{1/2}$

Spherical CCD

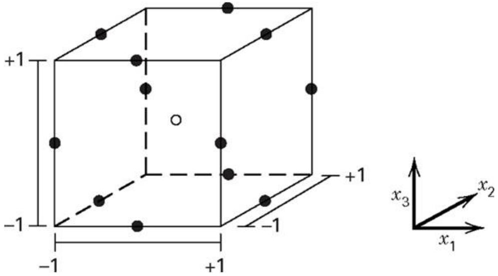
All factorial runs and axial runs have the same distance to the center: $\alpha = (k)^{1/2}$

k	$n_F = 2^k$	Rotatable α	Spherical α
2	4	1.41	1.41
3	8	1.68	1.73
4	16	2.00	2.00
5	32	2.38	2.24
6	64	2.83	2.45
7	128	3.36	2.65
8	256	4.00	2.83

Figure: The distance α between the axial points and the center

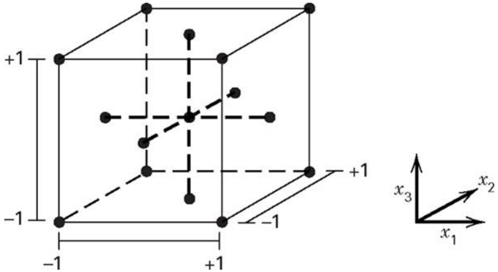
Box-Behnken Design

- Three-level designs for small number of runs.
- All points are on the distance $2^{1/2} = 0.707$ from the center.
- No points at the vertices.
- Either rotatable or nearly rotatable.



Face-centered cube

- The axial points are on the surface of the cube ($\alpha = 1$).
- Conveniently, because it only use three levels per factor.
- Not rotatable.



Face-centered cube

The face-centered cube is not rotatable.

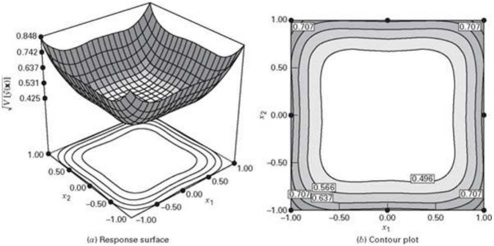
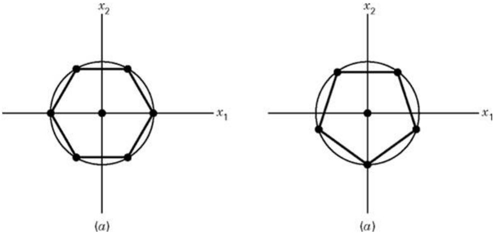


Figure: The standard deviation in the predicted values $k = 3, n_c = 3$

Equiradial design

- Design for two variables
- Rotatable



Graphical Evaluation of Response Surface Designs

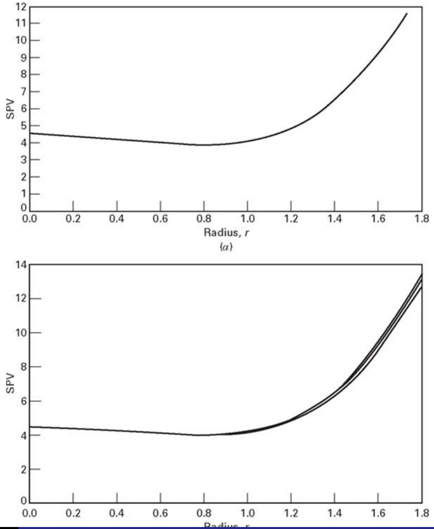
Scaled prediction variance (SPV):

$$\frac{NV[\hat{y}(x)]}{\sigma^2} = Nx'(X'X)^{-1}x$$

The **variance dispersion graph (VDG)** plot the maximum, the minimum and the average scaled prediction variance against the distance to the center.

Variance dispersion graph

a. Rotatable CCD for $k = 3$ ($n_c = 4$, $\alpha = 1.68$)
b. Spherical CCD for $k = 3$ ($n_c = 4$, $\alpha = 1.73$)



Variance dispersion graph

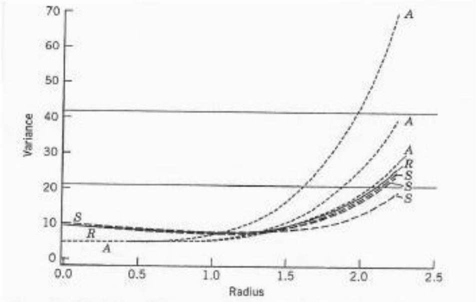


Figure: VDG for three CCDs for $k = 5$. Design A contains $\alpha = 1.5$. Design R contains $\alpha = 2$. Design S contains $\alpha = \sqrt{5}$. Each contains a 2^{5-1} and three center runs.

Variance dispersion graph

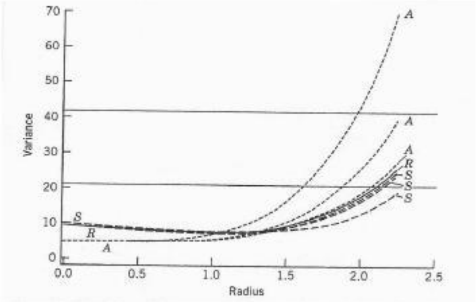


Figure: VDG for three CCDs for $k = 5$. Design A contains $\alpha = 1.5$. Design R contains $\alpha = 2$. Design S contains $\alpha = \sqrt{5}$. Each contains a 2^{5-1} and three center runs.

Variance dispersion graph

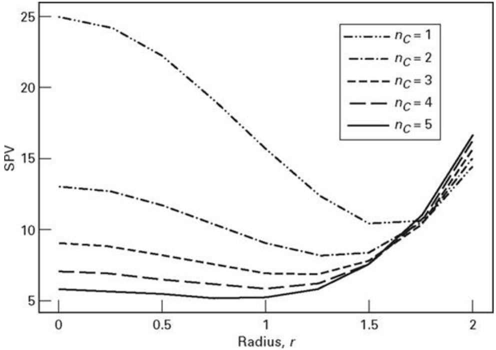


Figure: for central composite design for $k = 4$ and $\alpha = 2$. Four or five central point ($n_c = 4or5$) gives a stable variance in the predicted values.

Design criteria

- **D-optimal design** for minimizing the variance of the model regression coefficients. $\min |(X'X)^{-1}|$
- **G-optimal design** for minimizing the maximum prediction variance. $\min \max Var(\hat{y})$
- **I-optimal design** for its smallest possible value of the average prediction variance. $\min average Var(\hat{y})$

Which Criterion Should I Use?

- For fitting a first-order model, D is a good choice
  - Focus on estimating parameters
  - Useful in screening
- For fitting a second-order model, I is a good choice
  - Focus on response prediction
  - Appropriate for optimization