

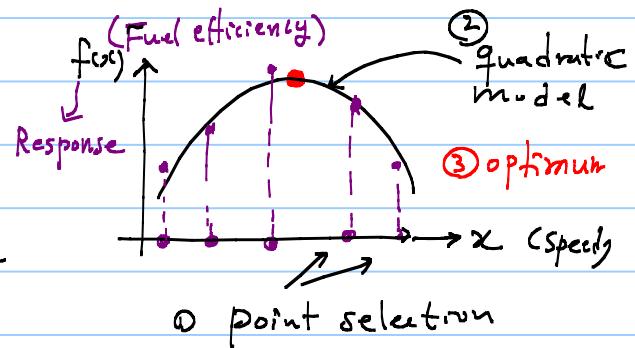
Lect 5-1 Response Surface Methodology : part 1

노트 제목

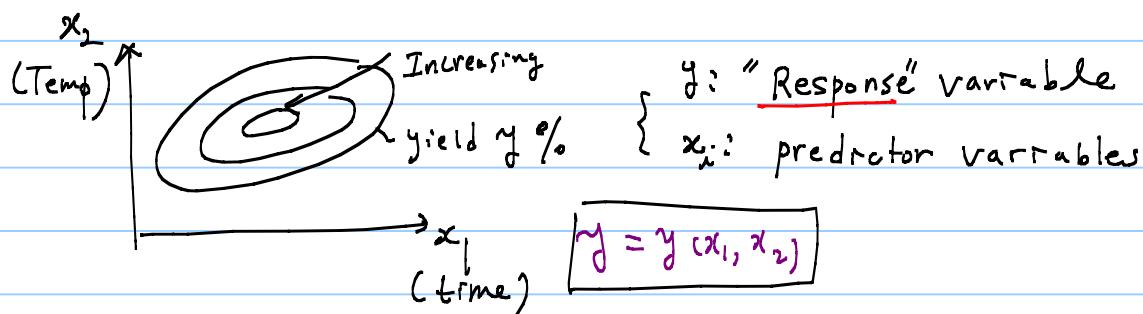
Reference: R.H. Myers and D.C. Montgomery
"Response surface Methodology"
2nd Ed., (2002), Wiley & Sons

RSM?

- { ① Experimental Design
- ② Regression Modeling
(such as least square method)
- ③ Elementary optimization



Basic Concept & Terminology



Mechanical Model

$$y = \underbrace{g(x_1, x_2, \dots, x_k)}_{\text{exact}} + \epsilon$$

↑ error in systems

Empirical (or Response Surface) Model

$$y = \underbrace{f(x_1, x_2, \dots, x_k)}_{\text{approximate}} + \epsilon$$

typical f :

(2)

Approximation of Response Functions

$$y = f(\xi_1, \xi_2, \dots, \xi_k) + \epsilon$$

↑
 natural variables
 (variables before
 being coded) such as temperature, pressure, time

ε some source of
 variability not
 accounted for in f

cf: Coded variables x_1, \dots, x_k
 (usually dimensionless
 with mean zero and the same
 spread or standard deviation)

* Because true response function f is either very complicated or unknown, an approximation of f is necessary.

(3)

"X" ε contains

- 1) Measurement error
- 2) Other sources of variant that are inherent in the process or system (background noise, etc.)

*** 3) thus ε will be treated as statistical error,

{ often assumed to have a normal distribution

$$\varepsilon \sim N(0, \sigma^2)$$

mean ↑ ↑ variance

More on " ε "

If mean of $\varepsilon = 0$, then

$$E(y) \stackrel{\triangle}{=} \gamma = E[f(\xi_1, \xi_2, \dots, \xi_K)] + E(\varepsilon)$$

↑
Expected value of y = $f(\xi_1, \dots, \xi_K)$

- Better to work with coded variables

$$\gamma = f(x_1, x_2, \dots, x_m)$$

↑
 $E(y)$

(4)

■ Typical response function:

Low-order polynomials are appropriate
in some relatively small region of
independent variable space

i) first-order model

(Main-effect Model)

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

(no interaction between x_1 and x_2)

ii) first-order model with interaction

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

iii) 2nd-order model

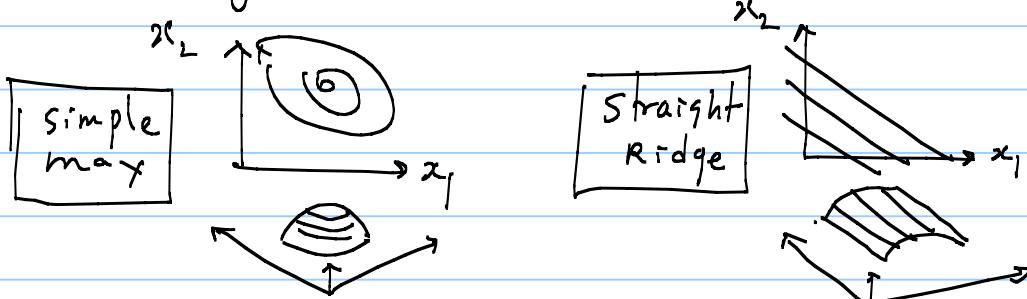
$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{11} x_1^2 + \beta_{22} x_2^2 + \beta_{12} x_1 x_2$$

< very popular & useful >

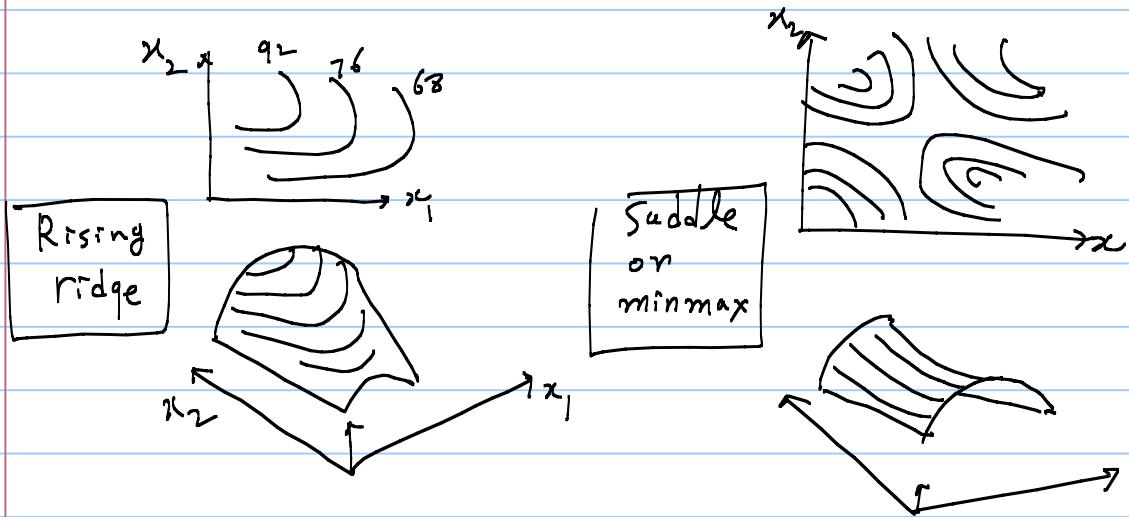
?

■ Why 2nd-order model so useful

① Very flexible



(5)



② Estimation of β 's is easy
(say, can use the least square method)

③ considerable practical experiences indicate
that second-order models work well in
solving real response surface problems

(6)

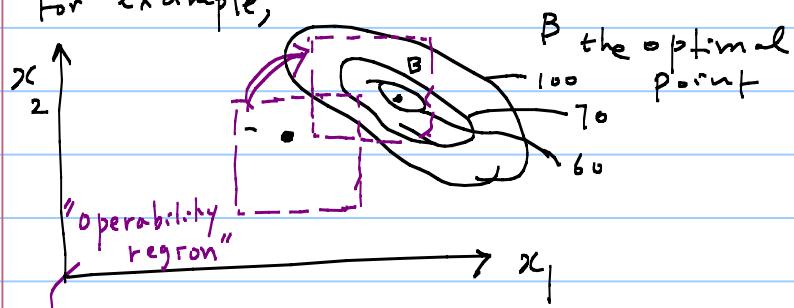
Sequential Nature of RSM

Phase 0 : Screening Experiments

(experiments designed towards
elimination of unimportant
variables or factor)

Phase 1 : determine if the current level or setting of the independent variables results in a value of the response that is near optima.

For example,



If current process pt is A,
then, we must move
near B.

region of independent variable space \Rightarrow May use "Method of steepest Ascent (or descent) to move from A near a point B. (i.e., need an optimization technique)

(\because The first-order model may be preferred if A is far away from B.)

7)

Phase 2 : Analysis of the Response Model
(typically using 2nd-order or higher-order model because the process is near the optimal)
and determination of optimum condition

Remarks on Robust Design

- Design of products or processes that are robust to environment conditions
- Variability minimization in the output response of a product around a target value
- Taguchi used "robust parameter design" to describe his approach to this issue.
→ Choose the level of controllable factors or parameters that make the system insensitive (or robust) to changes in a set of uncontrollable factors such as noise factors.

④

Building Empirical Models

< Linear Regression Model >

$$y = \beta_0 + \beta_1 x_1 + \dots + \beta_k x_k + \varepsilon$$

Multiple linear regression model

(β_k : regression coefficients)

example

① given: $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \varepsilon$

linear regression model

$$\rightarrow y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \varepsilon$$

β_{12} \nwarrow $x_1 x_2$

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

linear regression model

$$\rightarrow y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_4 + \beta_5 x_5 + \varepsilon$$

Thus,

Estimation of β_k' (or Model fitting) is an issue.

↳ will use [the least square method]

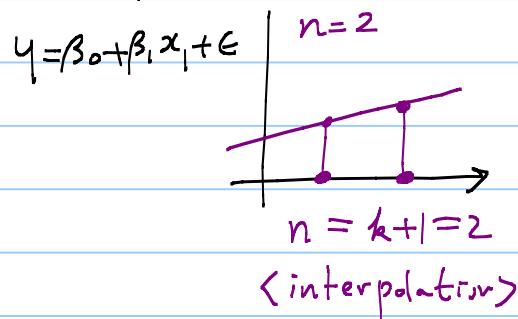
(9)

Given Data for Multiple Linear Regression

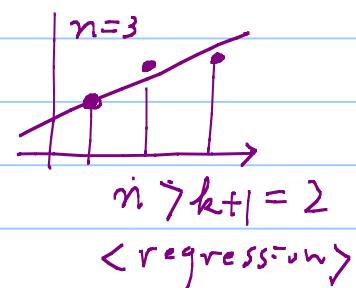
$$\text{Model: } y = \beta_0 + \beta_1 x_1 + \cdots + \beta_k x_k + \varepsilon$$

Observed Response y_i : $(x_{i1}, x_{i2}, \dots, x_{ik})$
 with observation or level of variable (x_1, \dots, x_k)
 $(i=1, \dots, n); \underline{n > k}$

For $k=1$.



? why = ?



Assumptions for Collected Data

$$\textcircled{1} \quad E(\varepsilon_i) = 0 \quad \text{Var}(\varepsilon_i) = \sigma^2$$

$$\textcircled{2} \quad \{\varepsilon_i\} = \text{uncorrelated random variables}$$

$\Leftrightarrow E(\varepsilon_i \varepsilon_j) = 0 \text{ if } i \neq j$

(1b)

Estimation of β_j ($j=0, \dots, k$) from measurement

$$\cdot \underline{y}_i = \beta_0 + \sum_{j=1}^k \beta_j \underline{x}_{ij} + \varepsilon_i \quad (i=1, \dots, n)$$

Easier to use Matrix Form

$$\boxed{\underline{y} = \underline{X} \underline{\beta} + \underline{\varepsilon}} \quad \rightarrow (1)$$

where $\underline{y} = \begin{Bmatrix} y_1 \\ \vdots \\ y_n \end{Bmatrix}_{n \times 1}$ $\underline{X} = \begin{bmatrix} 1 & x_{11} & \cdots & x_{1k} \\ 1 & x_{21} & \cdots & x_{2k} \\ \vdots & \vdots & & \vdots \\ 1 & x_{n1} & \cdots & x_{nk} \end{bmatrix}_{n \times (k+1)}$

$$n \times (k+1) \\ \equiv n \times p$$

$$\underline{\beta} = \begin{Bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{Bmatrix}_{p \times 1} \quad \underline{\varepsilon} = \begin{Bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{Bmatrix}_{n \times 1}$$

$$\boxed{\underline{b} \triangleq \underline{b}_{\text{LS}}}$$

< least square estimators of β : \underline{b} >

$$\begin{aligned} L &= \sum_{i=1}^n \varepsilon_i^2 = \underline{\varepsilon}^T \underline{\varepsilon} \\ &= (\underline{y} - \underline{X} \underline{\beta})^T (\underline{y} - \underline{X} \underline{\beta}) \\ &\quad \left(\beta^T \underline{X}^T \underline{y} = \underline{y}^T \underline{\beta} \right) \\ &= \underline{y}^T \underline{y} - 2 \underline{\beta}^T \underline{X}^T \underline{y} + \underline{\beta}^T \underline{X}^T \underline{X} \underline{\beta} \end{aligned}$$

(1)

To minimize error L

$$\left. \frac{\partial L}{\partial \beta_j} \right| = 0 \rightarrow -2 \underline{x}^T \underline{y} + 2 \underline{x}^T \underline{x} \underline{\beta} = 0$$

$\underline{\beta} = \underline{b}$

$$\Rightarrow (\underline{x}^T \underline{x}) \underline{\underline{b}} = \underline{x}^T \underline{y} \quad (2a)$$

$$\begin{bmatrix} n & \sum_{i=1}^n x_{i1} & \dots & \sum_{i=1}^n x_{ik} \\ \sum_{i=1}^n x_{i1} & \sum_{i=1}^n x_{i1}^2 & & \sum_{i=1}^n x_{i1} x_{ik} \\ \vdots & \vdots & & \vdots \\ \sum_{i=1}^n x_{ik} & \sum_{i=1}^n x_{ik} x_{i1} & \sum_{i=1}^n x_{ik}^2 & \end{bmatrix} \begin{Bmatrix} b_0 \\ b_1 \\ \vdots \\ b_k \end{Bmatrix} = \begin{Bmatrix} \sum_{i=1}^n y_i \\ \sum_{i=1}^n x_{i1} y_i \\ \vdots \\ \sum_{i=1}^n x_{ik} y_i \end{Bmatrix} \quad (2b)$$

Solving Eq-(1),

$$\boxed{\underline{\underline{b}} = (\underline{x}^T \underline{x})^{-1} \underline{x}^T \underline{y}} \quad (3)$$

\therefore The fitted regression model is

$$\begin{aligned} \hat{\underline{y}} &= \underline{x} \underline{\underline{b}} \\ (\hat{y}_i &= b_0 + \sum_{j=1}^k b_j x_{ij}, \quad i=1, \dots, n) \end{aligned}$$

$$\text{Residual: } \underline{\underline{e}} = \underline{\underline{y}} - \hat{\underline{\underline{y}}} \quad (e_i = y_i - \hat{y}_i; i=1, \dots, n)$$

(\checkmark See Text example 2-1; the transistor Gain Data)

(12)

Important facts regarding Least Square Method

$$\textcircled{1} \quad E(\hat{\beta}) = \beta$$

→ meaning: the least square method produces an unbiased estimator of β in the multiple linear regression model

$$\textcircled{2} \quad \text{cov}(\hat{\beta}) = \sigma^2 (\underline{X^T X}^{-1})$$

↑ choice of X affects covariances of $\hat{\beta}$

Covariance matrix

$p \times p$

$$= (k+1) \times (k+1)$$

$(\text{cov}(\hat{\beta})_{ij} = \text{covariance between } \hat{\beta}_i \text{ and } \hat{\beta}_j)$
 $\text{cov}(\hat{\beta})_{ii} = \text{variance of } \hat{\beta}_i$

< Detailed analysis >

$$\textcircled{1} \quad \text{To show } E(\hat{\beta}) = \beta$$

$$\begin{aligned} \hat{\beta} &\stackrel{(3)}{=} (X^T X)^{-1} X^T \underline{y} \\ &= (X^T X)^{-1} X^T (\underline{X} \beta + \underline{\varepsilon}) \end{aligned}$$

$$\begin{aligned} &= (X^T X)^{-1} (X^T \beta) + (X^T X)^{-1} X^T \underline{\varepsilon} \\ &= \beta + (X^T X)^{-1} \underline{\varepsilon} \end{aligned} \quad (4)$$

(13)

$$\begin{aligned}
 \therefore E(\hat{\beta}) &= E\left[\beta + (X^T X)^{-1} X^T \underbrace{\varepsilon}_{\text{random variable}}\right] \\
 &= \beta + (X^T X)^{-1} X^T E(\varepsilon) \\
 &= \beta
 \end{aligned}
 \quad \text{--- (5)}$$

(2) To show $\text{Cov}(\hat{\beta}) = \sigma^2 (X^T X)^{-1}$

$$\begin{aligned}
 \text{Cov}(\hat{\beta}) &\triangleq E\left\{ (\hat{\beta} - E(\hat{\beta})) (\hat{\beta} - E(\hat{\beta}))^T \right\} \\
 &= E\left\{ \underbrace{(\beta + (X^T X)^{-1} X^T \varepsilon - \beta)}_{\text{by (4)}} \underbrace{(\beta + (X^T X)^{-1} X^T \varepsilon - \beta)}_{\text{by (5)}}^T \right\} \\
 &= E\left\{ (X^T X)^{-1} X^T \underbrace{\varepsilon \varepsilon^T}_{\substack{\text{random} \\ \text{variables}}} X (X^T X)^{-1} \right\} \\
 &= (X^T X)^{-1} X^T E(\varepsilon \varepsilon^T) X (X^T X)^{-1}
 \end{aligned}$$

Recall $E(\varepsilon \varepsilon^T) = E\begin{bmatrix} \varepsilon_1^2 & \varepsilon_2 \varepsilon_1 & \cdots & \varepsilon_n \varepsilon_1 \\ \varepsilon_1 \varepsilon_2 & \varepsilon_2^2 & & \vdots \\ \vdots & & \ddots & \varepsilon_n^2 \end{bmatrix}$

(14)

$$= \begin{bmatrix} E(\epsilon_1^2) & E(\epsilon_2\epsilon_1) & \dots & E(\epsilon_n\epsilon_1) \\ & E(\epsilon_2^2) & & \\ \text{Sym} & & \ddots & \\ & & & E(\epsilon_n^2) \end{bmatrix}$$

$$= \sigma^2 \begin{bmatrix} 1 & 0 & & \\ 0 & 1 & \ddots & \\ & & \ddots & 1 \end{bmatrix} = \sigma^2 I$$

$$\begin{aligned} &= \sigma^2 (X^T X)^{-1} (X^T X) (X^T X)^{-1} \\ &= \sigma^2 (X^T X)^{-1} \end{aligned}$$

→ may select measurement point $x_{(j)}$ to make $(X^T X)^{-1}$ diagonal; orthogonal Design.