A Study on the Confidence Region of the Stationary Point in a Second Order Response Surface

Hong S. Jorn*

I. Introduction

When a response surface is fitted by a second order polynomial regression model, the stationary point is obtained by solving simultaneous linear equations. But the point is a function of random variables. We can find a confidence region for this point as Box and Hunter provided [1]. However, the confidence region is often too large to be useful for the experiments, and it is necessary to augment additional design points in order to obtain a satisfactory confidence region for the stationary point.

In this note, the author suggests a method how to augment design points "efficiently", and shows the change of the confidence region of the estimated stationary point in a response surface.

II. Simple Case

If we fit a respense surface $y(x)$ by second order of one independent variable x as

$$y(x)=\beta_0+\beta_1 x+\beta_{11}x^2+e \tag{1}$$

and the observations are taken at n selected points, then the coefficients are estimated by the least squares method

* Department of Computer Science and Statistics, Seoul National University. The author would like to express his grateful thanks to Prof. Sung H. Park for his helpful comments and suggestions. This work is partially supported by Korea Science & Engineering Foundation.

$$b=(X'X)^{-1}X'Y$$

where

$$X=\begin{pmatrix} 1, & x_1, & x_1^2 \\ 1, & x_2, & x_2^2 \\ \cdot & \cdot & \cdot \\ \cdot & \cdot & \cdot \\ 1, & x_n, & x_n^2 \end{pmatrix} \quad Y=\begin{pmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ y_n \end{pmatrix} \quad b=\begin{pmatrix} b_0 \\ b_1 \\ b_{11} \end{pmatrix}$$

under the assumption that e is normally and independently distributed with mean 0 and common variance σ^2, i,e., $e\sim N(0,\sigma^2)$.

Hence (1) can be rewritten as:

$$\hat{y}(x)=b_0+b_1x+b_{11}x^2 \tag{2}$$

The stationary point can be obtained by solving

$$\frac{\partial\hat{y}}{\partial x}=b_1+2b_{11}x=0 \tag{3}$$

If ρ is the true stationary point, then

$$b_1+2b_{11}\rho=\delta$$

may not be zero. But

$$E(\delta)=\beta_1+2\beta_{11}\rho=0$$

$$V(\delta)=V(b_1)+4\rho^2V(b_{11})+4\rho\mathrm{Cov}(b_1,b_{11})$$

$$=\sigma^2(c_{11}+4\rho^2c_{22}+4\rho c_{12})$$

where

$$\mathrm{Cov}(b_1,b_{11})=\begin{pmatrix} c_{11} & c_{12} \\ c_{12} & c_{22} \end{pmatrix}\sigma^2.$$

If σ^2 is known

$$\frac{\delta^2}{V(\delta)}=\frac{(b_1+2b_{11}\rho)^2}{\sigma^2(c_{11}+4\rho^2c_{22}+4\rho c_{12})}\sim\chi^2(1) \tag{4}$$

If σ^2 is unknown

$$\frac{\delta^2}{\hat{V}(\delta)}=\frac{(b_1+2b_{11}\rho)^2}{\hat{\sigma}^2(c_{11}+4\rho^2c^{22}+4\rho c_{12})}\sim F(1,n\text{-}3;\alpha) \tag{5}$$

Hence generally (σ^2 is unknown), the $(1\text{-}\alpha)\times100\%$ confidence region of ρ is

$$R^c=\{\rho\mid\frac{(b_1+2b_{11}\rho)^2}{\hat{\sigma}^2(c_{11}+4\rho^2c_{22}+4\rho c_{12})}\leq F(1,n\text{-}3;\alpha)\} \tag{6}$$

Actually the inequality

$$(b_1+2b_{11}\rho)^2\leq\hat{\sigma}^2(c_{11}+4\rho^2c_{22}+4\rho c_{12})F(1,n\text{-}3;\alpha)$$

can be written as

$$\rho^2(4b_{11}^2-4c_{22}F\hat{\sigma}^2)+\rho(4b_1b_{11}-4c_{12}F\hat{\sigma}^2)+(b_1^2-\hat{\sigma}^2c_{11}F)\leq 0 \qquad (7)$$

Example 1

In a laboratory experiment which is performed to fit a second order model relating the growth (y) of a particular organism to the percentage of glucose (x), the researcher fitted the model as

$$y=\beta_0+\beta_1x+\beta_{11}x^2+e.$$

The actual levels of variable were coded for the computations with the formular used in the coding being given by

$$x=\frac{\%\text{ glucose}-3.0}{1.0}.$$

Run	%Glucose(x)	Growth(y)
1	1.	76
2	0.5	81
3	0.	82
4	−0.5	77
5	−1	69

Table 1.

The observation values are listed in Table 1. The second order fitted response model can be obtained by the least squares method as

$$\hat{y}=b_0+b_1x+b_{11}x^2$$
$$=81.5714+3.6x-9.1427x^2.$$

The variance-covariance matrix for the coefficients is

$$\text{Cov }(b_0, b_1, b_{11})=(X'X)^{-1}\sigma^2=\begin{pmatrix} 0.485714 & 0. & -0.571428 \\ & 0.4, & 0. \\ (\text{sym.}) & & 1.142857 \end{pmatrix}\sigma^2 \qquad (8)$$

The analysis of variance is given in Table 2.

Source	S.S.	d.f.	M.S.	F
Regression	105.5768	2	52.7884	249.4726
Error	0.4232	2	0.2116	
Total	106.	4		

Table 2

The residual sum of squares, based on 2 degrees of freedom was 0.4232 giving an estimate of $\hat{\sigma}^2=0.2116$.

From Eq. (7), (8) and

$\hat{\delta}=3.6+18.2853x,\ \hat{\sigma}=0.2116,\ F(1,2;0.1)=8.53,$

we can find a confidence region for

$326.1047\rho^2-131.6549\rho+12.2380\leq 0$

i.e., $0.1451\leq\rho\leq 0.2586$ at the level of 90%.

III. General Case

If we fit a response surface $y(x)$ by second order polynomial of k independent variables

$$y(x)=\beta_0+\sum_{i=1}^{k}\beta_i x_i+\sum_{i=1}^{k}\beta_j x_i^2+\sum_{j>}^{k}\sum_{i=1}^{k}\beta_{ij}x_i x_j+e \tag{9}$$

and the observations, are taken at n selected comibinations of the x variables, then the coefficients are estimated by the least squares method,

$$b=(X'X)^{-1}X'Y \tag{10}$$

in which

$$X=\begin{pmatrix} 1, & x_{11},x_{21},\cdots,x_{k1},x_{11}^2,\cdots,x_{k1}^2,x_{11}x_{21},\cdots,x_{k-1,1},x_{k1} \\ 1, & x_{12},x_{22},\cdots,x_{k2},x_{12}^2,\cdots,x_{k2}^2,x_{12}x_{22}\cdots,x_{k-1,2}x_{k2} \\ \cdot & \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \quad \cdot \\ 1, & x_{1n},x_{2n},\cdots,x_{kn},x_{1n}^2,\cdots,x_{kn}^2,x_{1n}x_{2n}\cdots,x_{k-1,n}x_{kn} \end{pmatrix}$$

$$Y'=(y_1,y_2\cdots,y_n)$$

$$b=(b_0,b_1,\cdots,b_k,b_{11},\cdots,b_{kk},b_{12},\cdots,\ b_{k-1,k})$$

under the assumption that e is normally and independently distributed with mean 0 and common variance σ^2 *i.e.*, $e\sim N(0,\sigma^2)$.

$E(\hat{y})=\beta,\ V(b)=\sigma^2(X'X)^{-1}$ is a well known result.

The stationary point for this surface is obtained by solving following simultaneous linear equations.

$$\frac{\partial\hat{y}}{\partial x}=a+2Ax=0 \tag{11}$$

where $a=(b_1, b_2, \cdots, b_k)', x'=(x_1, x_2, \cdots. x_k)$

$$A=\begin{pmatrix} b_{11}, b_{12}/2, b_{13}/2, \cdots, & b_{1k}/2 \\ b_{22}, b_{23}/2, \cdots, & b_{2k}/2 \\ b_{33}, \cdots, & b_{3k}/2 \\ \text{(sym.)} \quad \cdots & \cdots \\ & b_{kk} \end{pmatrix}$$

If the true stationary point is $x=\rho$,

$$a+2A\rho=\delta \tag{12}$$

may not be $(0, 0, \cdots, 0)'$, where δ is a $k\times 1$ vector and

$$E(\delta)=0$$

$$\mathrm{Cov}(\delta)=E(\delta\delta')=E[(a+2A\rho)(a+2A\rho)']=\sigma_2 V.$$

When σ^2 is unknown, and the estimate of σ^2 is $\hat{\sigma}^2$ of ϕ d.f.,

$\dfrac{\delta' V^{-1}\delta}{k\hat{\delta}^2}$ is distributed as $F(k, \phi)$

Hence $(1-\alpha)\times 100\%$ confidence region is denoted by following inequality

$$\delta' V^{-1}\delta \leq k\hat{\sigma}F_\alpha(k,\phi) \tag{13}$$

Example 2

A 3^2 factorial experiment was performed, from which the researcher attempts to gain an insight into the influence of sealing temperature (x_1), cooling bar temperature (x_2) on the seal strength (y) in grams per inch of a breadwrapper stock. The actual levels of the variables were coded for the computations with the formulars used in the coding given by

$$x_1=\frac{\text{seal temp.}-255}{30}, \quad x_2=\frac{\text{cooling temp.}-55}{9}.$$

The observation values are listed in Table 3.

From Table 3 the second order surface fitted by the least squares methed to these points was

$$\hat{y}=89.7108+14.333x_1+6.4x_2-17.2666x_1^2$$
$$-5.2665x_2^2+4.525x_1x_2 \tag{14}$$

with stationary point (0.5242, 0.8328) which is indicated by a cross in Fig. 1. The analysis of variance is given in Table 4.

Run	x_1	x_2	y
1	1.	1.	93.6
2	1.	0.	87.7
3	1.	−1.	68.5
4	0.	1.	88.7
5	0.	0.	87.5
6	0.	−1.	82.4
7	−1.	1.	55.7
8	−1.	0.	59.4
9	−1.	−1.	48.7

Table 3

Sourse	S.S.	d.f.	M.S.	F
Regression	2211.9258	5	442.3852	30.9174
Error	42.9258	3	14.3086	
Total	2254.8516	8		

Table 4

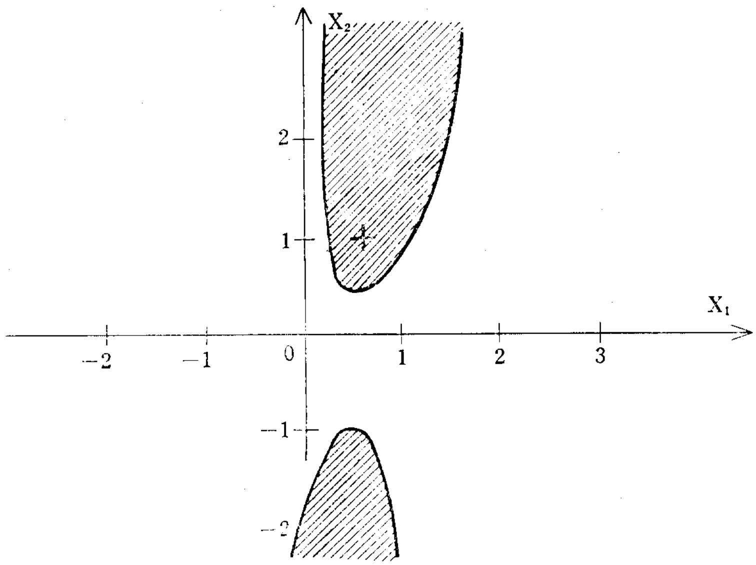
The reseidual sum of squares, based on 3 degree of freedom was 41.9258 giving an etimate of $\hat{\sigma}^2=14.3086$. The covariance matrix of the coefficients in Eq. (14) was

$$(X'X)^{-1}\sigma^2=\begin{pmatrix} 5/9, & 0, & 0, & -1/3, & -1/3, & 0 \\ & 1/6, & 0, & 0, & 0, & 0, \\ & & 1/6, & 0, & 0, & 0, \\ & & & 1/2, & 0, & 0, \\ & \text{(sym.)} & & & 1/2, & 0, \\ & & & & & 1/4, \end{pmatrix}\sigma^2 \quad (15)$$

and

$$\begin{aligned}\delta_1&=14.333-34.533x_1+4.525x_2\\ \delta_2&=6.4\quad-10.533x_2+4.525x_1\end{aligned}, \quad (16)$$

substituting the values of $\hat{\sigma}^2=14.3086$ and $F(2,3;01)=5.46$ and $k=2$ in Eq. (13) and from (15), (16), we obtain the 90% confidence region, the shaded portion indicated in Fig. 1.



IV. The Augmentation of Design Points

In a regression model

$$Y = X\beta + e$$

the vactor of estimates of β is given by

$$b = (X'X)^{-1}X'Y$$

where Y is a vector of n responses, X is the $n \times p$ matrix of independent variables having 1's in the first column, β is the vector of regression coefficients, b is the least squares estimate of β and e is the vector of errors, under the assumption that e is distributed as $N \sim (0, \sigma^2 I)$.

If a_{ii} is the i-th diagonal element of the matrix $A = X'X$ and c_{ii} is the i-th diagonal element of $C = (X'X)^{-1}$, then the squared multiple correlation of variable x_i with the other x variables, denoted by R_i^2, is

$$R_i^2 = 1 - \frac{1}{a_{ii}c_{ii}}.$$

If this value of R_i^2 is near 1, then the variable x_i is virtually, if not completely, useless as a predictor. The variance of each regression coefficient β_i is

$$\text{Var}\ (b_i) = c_{ii}\sigma^2 = \sigma^2/a_{ii}(1-R_i^2)$$

$\text{Var}(b_i)$ approaches infinity as R_i^2 approaches 1. One criterion mentioned by Kiefer is minimizing the generalzed variance, which is equivalent to maximizing $|X'X|$. [2] By maximizing $|X'X|$ produces designs which

1. Have a confidence region for the parameters of smallest (hyper) volume in the parameter spaces,
2. Minimize the generalized variance of the parameter estimates,
3. Are invariant to linear changes of scale of the parameters,
4. Minimize the maximum variance of any predicted value over the experimental space. [3]

Hence it is reasonable in the view of variance criteria that the augmentation of design points for confidence region should follow the way of maximizing $|X'X|$.

Let $\tilde{X}' = (X', F')$ be $p \times (n+m)$ matrix, where X is $n \times p$ for existing data, and F is $m \times p$ for additional m design points.

We use a result relating to the inverse of matrices, namely,

$$(\tilde{X}'\tilde{X})^{-1} = (A+F'F)^{-1} = A^{-1} - A^{-1}F' \ (I+FA^{-1}F')^{-1}FA^{-1}$$

which is proved by Plackett [3], where $A = X'X$.

Let

$$H = \begin{bmatrix} A^{-1}, & A^{-1}F' \\ FA^{-1}, & I+FA^{-1}F' \end{bmatrix}$$

then

$$\begin{aligned} |H| &= |I+FA^{-1}F'|\,|A^{-1} - A^{-1}F'(I+FA^{-1}F')^{-1}FA^{-1}| \\ &= |A^{-1}|\,|I+FA^{-1}F' - FA^{-1}AA^{-1}F'| \\ &= |A^{-1}|\,|I| = |/|A|. \end{aligned}$$

$$|I+FA^{-1}F'|\,|(\tilde{X}'\tilde{X})^{-1}| = 1/|A|.$$

Hence we obtain

$$|\tilde{X}'\tilde{X}| = |A| \cdot |I+FA^{-1}F'|.$$

We see that $|X'X|$ is maximized by taking as the next m runs the points in the design space where $|I+FA^{-1}F'|$ is greatest.

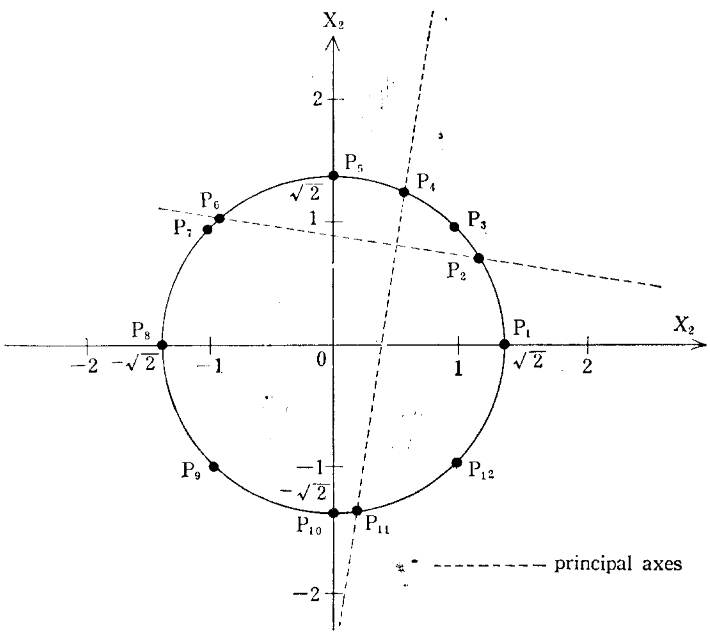
If F is $1 \times p$, *i.e.*, only one additional point is added, $|I+FA^{-1}\ F'|$ is the same as that of Dykstra, $1+x_0'(X'X)^{-1}x_0$. [2]

Example 3

The confidence region obtained in Example 2 is unsatisfactory. Since the region of interest is $(-1\leq x_1\leq 1,\ -1\leq x_2\leq 1)$ with center point (0, 0) and the region of experiment is $\{(x_1, x_2)\mid x_1^2+x_1^2\leq 2\}$, the candidates for the next runs are the corner points and the axial points which are obtained by

$$x_1^2+x_2^2=2,$$

and the points which are in the principal axes (obtained by cannonical analysis [4]). The candidates are indicated in Fig. 2.



$P_1(\sqrt{2},\ 0)$
$P_2(1.22, 0.71)$
$P_3(1,\ 1)$
$P_4(0.61, 1.28)$
$P_5(0,\ \sqrt{2})$
$P_6(-0.89, 1.09)$
$P_7(-1, 1)$
$P_8(-\sqrt{2}, 0)$
$P_9(-1, -1)$
$P_{10}(0, -\sqrt{2})$
$P_{11}(0.11, -1.41)$
$P_{12}(0.52, 0.83)$

Fig. 2

The maximum value of $|I+FA^{-1}F|$ for the next runs and additional points are listed in Table 4.

Five further points were added as shown in Fig. 3, the values which might have been obtained at these points using Eq. (14) and adding random normal deviates, $\sigma=3.78$. Using all 14 points the equation was refitted giving

$$\hat{y}=89.9272+14.317x_1+6.2269x_2-16.215x_1^2-6.9685x_2^2+4.9129x_1x_2$$

No. of add. points	points	$\lvert I+FA^{-1}F' \rvert$
1	$P1$ or $P5$ or $P8$ or $P10$	2.5555
2	$P1$, $P11$	5.9365
3	$P1$, $P5$, $P11$	12.6786
4	$P1$, $P5$, $P8$, $P10$	26.2345
5	$P1$, $P4$, $P6$, $P8$, $P10$	41.7272

Table 5

The analysis of variance is given in Table 5.

Source	S.S.	d.f.	M.S.	F
Regression	4263.8225	5	852.7645	116.6174
Error	58.5	8	7.3125	
Total	4205.3225	13		

Table 6

The newly obtained confidence region was the shaded portion indicated in Fig. 3.

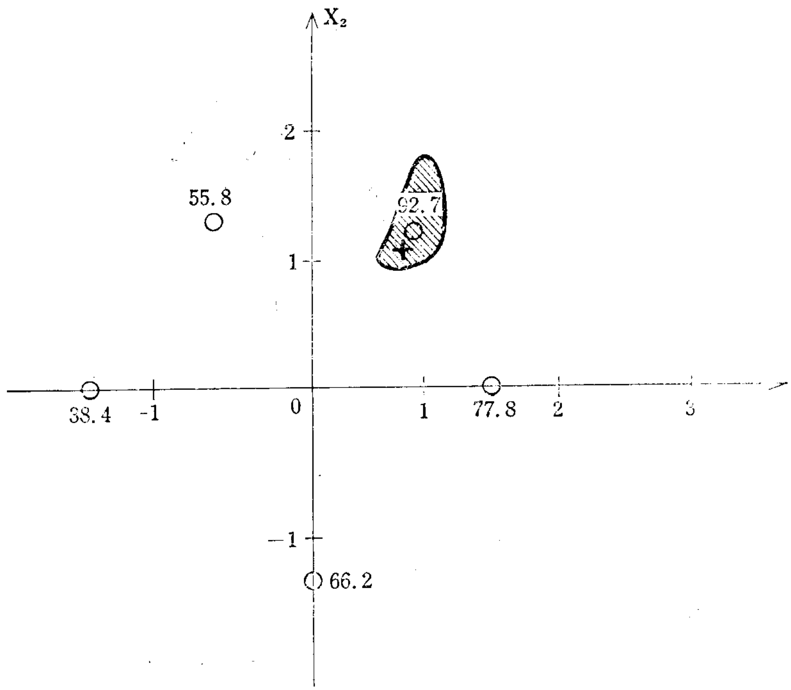


Fig. 3

REFERENCES

[1] Box G. E. P. and J. S. Hunter (1954), "A confidence region for the solution of a set of simultaneous equations with an applications to experimental design," *Biometrika*, Vol. 41 pp, 190—199

[2] Dykstra O. Jr. (1971) "The augmentation of experimental data to maximize $|X'X|$," *Technometrics*, Vol. 13, No. 3, pp. 682—688

[3] Evans J.W. (1974), "Computer generated D-optimal, V-optimal and trace optimal maximum bias experimental designs." Institute of Statistics, North Carolina State Univ. Mimeograph Series No. 936 July, 1974

[4] Myers R.H. 1971, *Response Surface Methodology* Allyn and Bacon, Inc. Boston.