ertrude M. Cox

A CONFIDENCE REGION FOR THE SOLUTION OF A SET OF SIMULTANEOUS EQUATIONS WITH AN APPLICATION TO EXPERIMENTAL DESIGN

Prepared Under Office of Ordnance Research

Contract No. DA-36-034-ORD-1177 (RD)

by

G. E. P. Box
J. S. Hunter

Institute of Statistics
Mimeo Series No. 72
June, 1953

TECHNICAL REPORT NO. 3

A CONFIDENCE REGION FOR THE SOLUTION OF A SET OF SIMULTANEOUS EQUATIONS WITH AN APPLICATION TO EXPERIMENTAL DESIGN.

Prepared Under Contract No. DA-36-034-ORD-1177 (RD)
(EXPERIMENTAL DESIGNS FOR INDUSTRIAL RESEARCH)

Philadelphia Ordnance District

Department of The Army, Department of Defense

with

Institute of Statistics

North Carolina State College of
The University of North Carolina
Raleigh, North Carolina

Technical Supervisor
Ballistics Research Laboratories
Aberdeen Proving Ground
Aberdeen, Maryland

G. E. P. Box
J. S. Hunter
Authors of Report

ABSTRACT

The problem of finding a confidence region for an experimentally determined maximum or stationary point involves the more general one of finding a confidence region for the solution of a set of linear equation of the form

$$
\begin{array}{ccccccccccc}
a_{10} & + & a_{11}x_1 & + & a_{12}x_2 & + & \cdots & + & a_{1k}x_k & = & \delta_1 \\
a_{20} & + & a_{21}x_1 & + & a_{22}x_2 & + & \cdots & + & a_{2k}x_k & = & \delta_2 \\
\vdots & & \vdots & & \vdots & & & & \vdots & & \vdots \\
a_{k0} & + & a_{k1}x_1 & + & a_{k2}x_2 & + & \cdots & + & a_{kk}x_k & = & \delta_k
\end{array}
$$

where the coefficients $a_{10}, a_{11}, \dots, a_{1k}, \dots, a_{kk}$ are subject to random error. This problem is discussed and a method is described which yields the exact confidence region about the solution. In addition, the manner in which both the conditioning of the equations and the errors in the coefficients affect the shape, orientation and size of the confidence region is discussed.

1. INTRODUCTION

The problem of finding limits of error for the solutions of a set of k linear equations obtained by equating each of the quantities

$$
(1.0.1)\qquad
\begin{aligned}
\delta_1 &= a_{10} + a_{11}x_1 + a_{12}x_2 + \dots + a_{1k}x_k \\
\delta_2 &= a_{20} + a_{21}x_1 + a_{22}x_2 + \dots + a_{2k}x_k \\
&\vdots \\
\delta_k &= a_{k0} + a_{k1}x_1 + a_{k2}x_2 + \dots + a_{kk}x_k
\end{aligned}
$$

to zero, where $a_{10}, a_{11}, \dots, a_{1k}, \dots, a_{kk}$ are subject to error, was considered by Lonseth (1942) who gives a number of references to earlier writers. He obtained a series for the error of any unknown, considered as a function of the $k(k+1)$ errors in the coefficients, and also a criterion for the convergence of this series which depended on the conditioning of the equations and the smallness of the errors relative to their coefficients. Lonseth and more recent writers such as Hotelling (1943) and Turing (1948) have been chiefly concerned with the effect of rounding errors. Not infrequently however we are faced with the problem where large observational errors occur in the coefficients and the equations may not be well conditioned. It then seems essential to consider the errors in the solution not individually but jointly, in fact to determine a confidence region for the possible solution in the space of $x_1, x_2, \dots, x_k$. One example of this circumstance which has attracted the attention of the authors to the problem occurs when attempts are made to attach limits of error to the position of a maximum in experiments of the type discussed by Box and Wilson (1951). Further reference will be made to this problem later.

2. AN EXACT CONFIDENCE REGION

We assume that the errors in the coefficients a_{10}, a_{11}, ..., a_{kk} are distributed multinormally with a k(k+1) by k(k+1) variance-covariance matrix $\underline{\Omega}\,\sigma^2$, known apart from the factor σ^2, and that an estimate s^2 of σ^2 is avilable based on φ degrees of freedom and distributed as $\chi^2_{(\varphi)}/\varphi$ independently of the errors in the coefficients. We use the notation $\chi^2_{(\varphi)}$ to denote a quantity distributed as the χ^2 distribution with φ degrees of freedom, similarly $F(\varphi_1, \varphi_2)$ denotes a quantity distributed in the Fisher-Snedecor F distribution with φ_1 and φ_2 degrees of freedom. A confidence region may now be found by an extension of an argument given by Feiller (1940).

Consider the expression δ_1, δ_2, ..., δ_k of equation (1.0:1) for a fixed set of values x_1^o, x_2^o, ..., x_k^o. Knowing $\underline{\Omega}$ we may readily calculate the variance-covariance matrix $\mathcal{E}(\underline{\delta}'\underline{\delta}) = \underline{V}\sigma^2$ where $\underline{\delta}$ denotes the k x 1 column vector of the δ's. Since $\dfrac{\underline{\delta}'\underline{V}^{-1}\underline{\delta}}{ks^2}$ is distributed as $F(k, \varphi)$, the probability is $1 - \alpha$ that the inequality

$$(2.0:1) \qquad \frac{\underline{\delta}'\underline{V}^{-1}\underline{\delta}}{ks^2} < F_\alpha(k, \varphi)$$

is true, where F_α denotes the α probability point of the F distribution.

Now consider the space of the $k(k+2) + 1$ quantities a_{10}, a_{11}, ..., a_{kk}, s^2, x_1, x_2, ..., x_k.

The expression,

$$\frac{\underline{\delta}'\underline{V}^{-1}\,\underline{\delta}}{ks^2},$$

is a function of these quantities and equation(2.0:1) defines a region $R(x_1^o, x_2^o, \ldots, x_k^o)$ on the hyperplane $x_1 = x_1^o$, $x_2 = x_2^o$, $\ldots$, $x_k = x_k^o$ such that the probability that the point $a_{10}, a_{11}, \ldots, a_{1k}, s^2, x_1^o, x_2^o, \ldots, x_k^o$ falls within this region is $1 - \alpha$. Now if we combine such regions for all values of $x_1, x_2, \ldots, x_k$ we obtain a region R in the whole $k \cdot (k+2)+1$ dimensional space which is such that the chance that a point $a_{10}, a_{11}, \ldots, a_{1k}, s^2, x_1, x_2, \ldots, x_k$ falls within R is $1 - \alpha$. It follows that, given a set of observed values $a_{10}^o, a_{11}^o, \ldots, a_{kk}^o, (s^o)^2$, the region $R\ (a_{10}, a_{11}, \ldots, a_{1k}, (s^o)^2\ )$ on the hyperplane $a_{10} = a_{10}^o$, $a_{11} = a_{11}^o$, $\ldots$, $a_{kk} = a_{kk}^o$, $s^2 = (s^o)^2$ supplies the $1 - \alpha$ confidence region in the usual sense for $x_1, x_2, \ldots, x_k$.*

The limits of the $1 - \alpha$ confidence region are given therefore by

$$(2.0{:}2) \qquad \underline{\delta}'\ \underline{V}^{-1}\ \underline{\delta} = s^2 k\, F_\alpha\ (k, \quad)$$

Now the original quadratic form may be written as a ratio of two determinants

$$(2.0{:}3) \qquad \underline{\delta}'\underline{V}^{-1}\underline{\delta} = -\frac{\begin{vmatrix} 0 & \delta_1 & \delta_2 & \cdots & \delta_k \\ \delta_1 & & & & \\ \delta_2 & & V & & \\ \vdots & & & & \\ \delta_k & & & & \end{vmatrix}}{|V|}$$

---

* We are indebted to Drs. R. C. Bose and S. N. Roy for the above argument which indicates that the region defined by (2.0:2) is a confidence region in the usual sense.

Consequently, in general the boundary of the confidence region is defined by those values of $x_1, x_2, \ldots, x_k$ which cause the equation

$$(2.0{:}4)\qquad \begin{vmatrix} s^2kF_\alpha & \delta_1 & \delta_2 & \cdots & \delta_k \\ \delta_1 & & & & \\ \delta_2 & & V & & \\ \vdots & & & & \\ \delta_k & & & & \end{vmatrix} = 0$$

to be satisfied. If σ^2 is actually known, s^2kF_α will be replaced by $\chi^2\sigma^2$ in (2.0:4.).

In the important special case in which the estimates $a_{10}, a_{11}, \ldots, a_{kk}$ are uncorrelated (2.0:2) gives for the boundary of the confidence region

$$(2.0{:}5)\qquad \sum_{i=1}^{k}\left\{(\delta_i^2)/V(\delta_i)\right\} = \sum_{i=1}^{k}\left\{\left(\sum_{j=0}^{k}a_{ij}x_j\right)^2\Big/\sum_{j=0}^{k}V(a_{ij})x_j^2\right\} = s^2kF_\alpha$$

where $x_0 = 1$. If, in addition, the variances for all the coefficients are equal ($V(a_{ij}) = V(a)$; $i,j = 1,2,\ldots,k$) the boundary of the confidence region is given simply by the equation

$$(2.0{:}6)\qquad \sum\delta_i^2 = \sum_{i=1}^{k}\left(\sum_{j=0}^{k}a_{ij}x_j\right)^2 = \sum_{j=0}^{k} x_j^2\, V(a)\, s^2kF_\alpha(k,\varphi)$$

3. CONDITIONING OF THE EQUATIONS

A circumstance which profoundly affects the nature of the confidence region is whether the equations are well or poorly "conditioned". We first explain what we mean by this term.

In an obvious matrix notation (1.0:1) becomes

(3.0:1) $$\underline{\delta} = \underline{a}_o + \underline{A}\ \underline{x}.$$

Consider the quadric defined by

(3.0:2) $$\delta'\delta = (\underline{A}\ \underline{x} + \underline{a}_o)'\ (\underline{A}\ \underline{x} + \underline{a}_o)$$

For any fixed value of z

(3.0:3) $$\underline{\delta}'\underline{\delta} = \sum_{i=1}^{k} \delta_i^2 = z$$

defines a surface which we call the conditioning surface in the k + 1 dimensional space of $z, x_1, x_2, \ldots, x_k$. In the k space of $x_1, x_2, \ldots, x_k$ the contours of (3.0:3) for fixed values of z are ellipsoids. Referring these contours to their center as the origin and using their principle axes as new co-ordinates $x_1, x_2, \ldots, x_k$, equation (3.0:3) may be written

(3.0:4) $$\sum_{i=1}^{k} \delta_i^2 = \sum_{i=1}^{k} \lambda_i X_i^2 = z$$

where $X_i = \underline{u}_i'\ \underline{x}$, $u_i (i = 1,2, \ldots, k)$ being a latent vector of $\underline{A}'\underline{A}$ and λ_i the corresponding latent root which is, of course, essentially positive. Clearly, if one or more of the latent roots are small compared with the remainder, the conditioning surface is attenuated in the direction of these axes. Thus, suppose $\lambda_1, \lambda_2, \ldots, \lambda_r$ are small compared with the remainder, then values of $x_1, x_2, \ldots, x_k$ differing greatly from the correct solution but corresponding to points near the hyperplane h_r passing through the axes of $X_1, X_2, \ldots, X_r$

produce only small values of $\sum_{i=1}^{k} \delta_i^2$ and therefore only small values of $\delta_1, \delta_2, \ldots, \delta_k$ in (1.0:1). Thus a wide variety of "nearly correct" solutions of the equations $\underset{\sim}{a}_0 + \underset{\sim}{A}\,\underset{\sim}{x} = 0$ exist leading to difficulties in the numerical solution, and the equations are said to be "poorly conditoned". An interesting example of such equations is given by Morris (1946). In an extreme case in which $\lambda_1, \lambda_2, \ldots, \lambda_r$ are exactly zero, any point in the hyperplane h_r satisfies the equations, $\underset{\sim}{A}$ being now of rank $k - r$.

Turing (1948) proposed as a measure of ill conditioning the quantity

$$c = k^{-1}\, N(\underset{\sim}{A})\, N(\underset{\sim}{A}^{-1}) \tag{3.0:5}$$

where $N(A) = (\text{trace } \underset{\sim}{A}'\underset{\sim}{A})^{\frac{1}{2}} = (\sum_{ij} a_{ij}^2)^{\frac{1}{2}} = (\sum_i \lambda_i)^{\frac{1}{2}}$ is the norm of the matrix $\underset{\sim}{A}$. The criterion is in line with the discussion above for we see that

$$c = k^{-1} \left(\sum_i \lambda_i \sum_i \lambda_i^{-1} \right)^{1/2} \tag{3.0:6}$$

is a homogeneous function of degree sero in the λ's and is thus dependent only on their relative magnitudes. It takes the values of unity if all the λ's are equal (so that $\underset{\sim}{A}$ is orthogonal), and is greater than unity otherwise. Its value will be large if any of the λ's are small compared with the others.

4. EXAMPLES OF CONFIDENCE REGIONS

The confidence region for the solution of a set of linear equations would be expected to depend on 1) the magnitude of the errors in the coefficients and 2) the state of the conditioning of the equations. The separate contribution of these two influences can be seen particularly readily when the coefficients have equal variance and are uncorrelated so that equation (2.0:6) may be used to obtain the confidence interval.

For example, consider the pair of equations

(4.0:1) $$3x_1 - 5x_2 - 1 = \delta_1$$

(4.0:2) $$3x_1 + 3x_2 - 9 = \delta_2$$

when $\delta_1 = \delta_2 = 0$. These have the solution $x_1 = 2$, $x_2 = 1$ indicated by the point O in Figure 1 a . If the coefficients are subject to error, each with a known standard deviation equal to one, and they are uncorrelated, then the 95% confidence region obtained by setting $V(a)s^2\ kF(k, \varphi)$ equal to $1 \times \chi^2_{.05}(2) = 5.99$ in equation (2.0:6) is shown by the bold line in Figure 1 a . We note from equation (2.0:6) that the boundary of the confidence region is the intersection of the surface

(4.0:3) $$\chi^2_\alpha\ \sigma^2\ V(a)(1 + x_1^2 + x_2^2) = z$$

with the surface

(4.0:4) $$\delta_1^2 + \delta_2^2 = z$$

Contours of these two surfaces for $z = 25$ and $z = 120$ are shown by the dotted lines in the figure which intersect on the confidence region. In separating equation (2.0:6) into the two parts we separate the two features which decide the confidence region. Equation (4.0:3) defines a saucer-like surface represented by the circular contours. The steepness of the surface depends on V(a), the variance of the coefficients. Equation (4.0.4) on the other hand defines the conditioning surface as an elliptical valley whose contours are concentric ellipses and whose attenuation is a measure of the conditioning of the equations. If V(a) is small then the surface described by (4.0:3) will be shallow and will cut (4.0:4) to form a small region whose shape is very like that of the contours of the conditioning surface. If V(a) is large however, the surface (4.0:3) will rise more steeply and the confidence region will be larger and its shape will, to some extent, be distorted away from the shape of the contours of (4.0:4). In particular, the region will tend to extend farther from the solution point on the side remote from the origin.

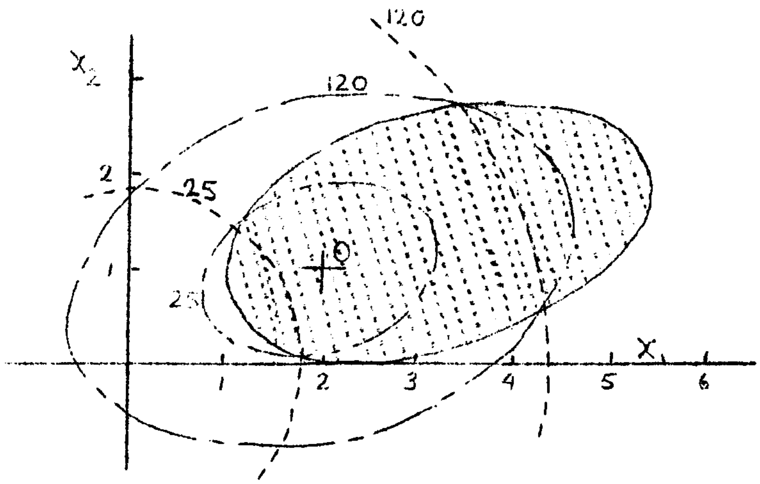
Equations (4.0:1) and (4.0:2) are a well conditioned pair ($\lambda_1 = 16$, $\lambda_2 = 36$, $c = 1.08$). A somewhat less conditioned pair is obtained if one-half of equation (4.0:1) is added to equation (4.0:2) to give a new equation to replace (4.0:1) so that we have

(4.0:5) $$4.5\, x_1 + 0.5\, x_2 - 9.5 = 0$$

(4.0:6) $$3.0\, x_1 + 3.0\, x_2 - 9.0 = 0$$

having the same solutions as before, but now $\lambda_1 = 4.2$, $\lambda_2 = 34.36$, $c = 1.55$. Then if we again assume that the coefficients have standard deviations $\sigma = 1$ and are uncorrelated we obtain the open confidence interval shown by the bold line in Figure 1b. .

Figure 1a

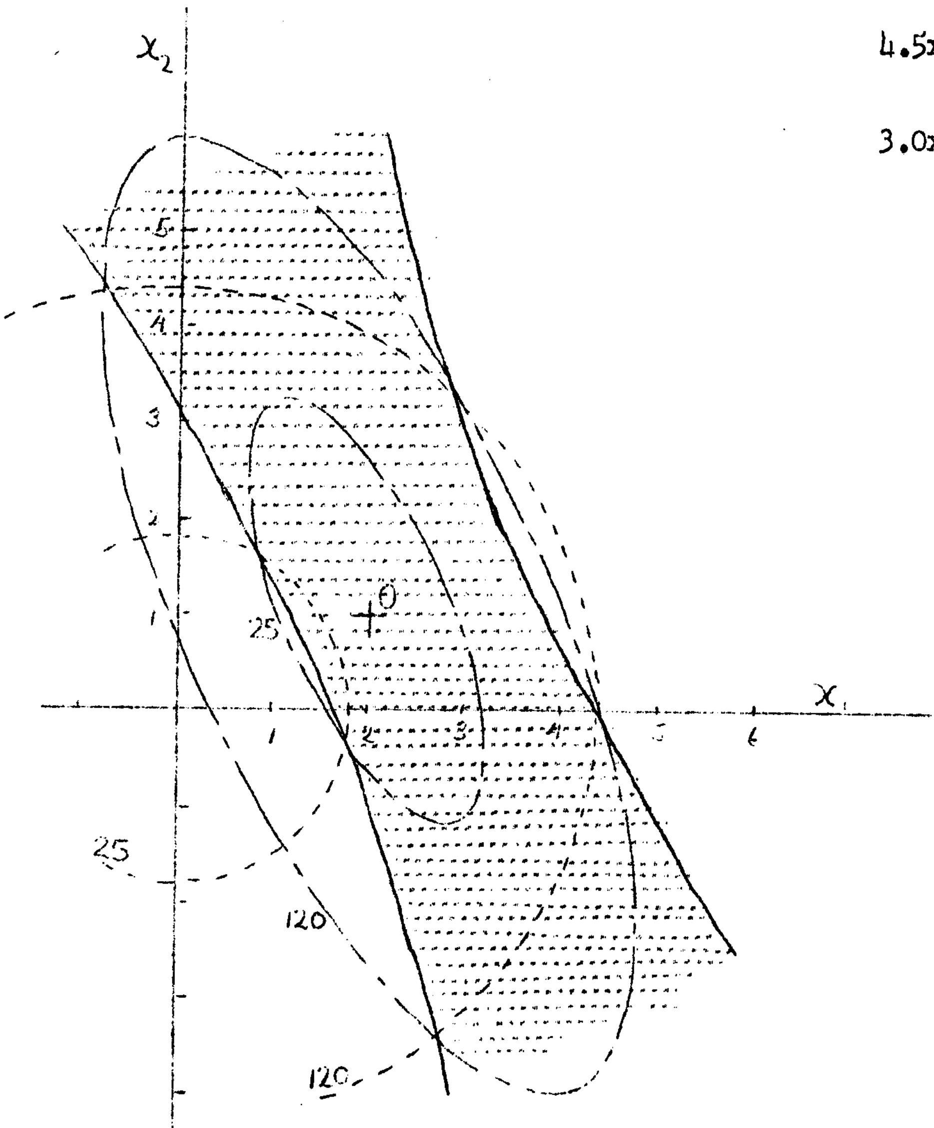


95% Confidence Region for the solution of the equations

$$3x_1 - 5x_2 - 1 = \delta_1$$

$$3x_1 + 3x_2 - 9 = \delta_2$$

Figure 1b



95% Confidence Region for the solution of the equations

$$4.5x_1 + 0.5x_2 - 9.5 = \delta_1$$

$$3.0x_1 + 3.0x_2 - 9.0 = \delta_2$$

The great difference in the shape of the confidence regions found with the two different pairs of equations is seen to be due to differences in conditioning as typified by the conditioning surfaces. Although the second set of equations is not seriously ill-conditioned, its conditioning surface is considerably more attenuated than that of the first set. Thus, when the surface corresponding to (4.0:3) is cut by that of (4.0:4) the open ended hyperbolic confidence region shown in Figure 1 b is obtained.

The effect of ill-conditioning is, in general, to cause the confidence region to spread out in the direction of the axes of the conditioning surface corresponding to the small latent root of $\underline{A}'\underline{A}$.

5. THE CONFIDENCE REGION FOR THE STATIONARY POINT ON A FITTED SECOND DEGREE SURFACE

In a paper concerning experimental methods for attaining optimum conditions, Box and Wilson (1951) discussed the fitting by least squares of a polynomial equation to a set of experimental points. Suggested experimental designs were discussed from the point of view of minimizing both the random errors in the estimates of the coefficients, and of the biases contributed by possible higher order terms which had been ignored in the response function.

Suppose that a polynomial equation of the second degree is fitted and for simplicity assume that there are only two independent variables x_1 and x_2. If the equation of the fitted surface is

$$y = a_{00} + a_{10}x_1 + a_{20}x_2 + \tfrac{1}{2}a_{11}x_1^2 + \tfrac{1}{2}a_{22}x_2^2 + a_{12}x_1x_2 \qquad (5.0:1)$$

the center of the fitted system is at a point corresponding to the solution of the equations

$$\frac{\partial y}{\partial x_1} = a_{10} + a_{11}x_1 + a_{12}x_2 = \delta_1 \tag{5.0:2}$$

$$\frac{\partial y}{\partial x_2} = a_{20} + a_{12}x_1 + a_{22}x_2 = \delta_2 \tag{5.0:3}$$

which are of the same form as equations (1.0:1) and (3.0:1) in the special case in which the matrix $\underline{A}$ is symmetric. Furthermore, the variance-covariance matrix of the least squares estimates which correspond to $\Omega\sigma^2$ is known apart from σ^2 and is dependent only on the experimental design used, and an estimate s^2 of σ^2 is usually available either from the residual sum of squares (assuming an adequate model) or from some independent estimate obtained by replication. Thus to obtain a confidence region for the center of the system we can immediately apply equation (2.0:4).

It should of course be remembered that in practice an important source of error not taken account of in the above confidence region arises due to the possible lack of fit of the second degree equation. Such lack of fit introduces errors not only directly in the sense that there is no second degree equation which can adequately represent the surface, but also because the omission of higher order terms necessary to give an exact fit may cause the least squares estimates of the second degree equation to be biased. However, provided this limitation is borne in mind it is instructive to consider the size and type of confidence region that arises due to sampling errors in the coefficients alone.

We may first note the relationship that exists between the fitted response surface and the conditioning surface discussed in Section 3. The equation of the fitted surface written in matrix notation is

$$(5.0:4) \qquad y - a_{00} = \underline{x}'\underline{a}_0 + \frac{1}{2}\underline{x}'\underline{A}\underline{x}$$

and that of the conditioning surface is given by equation (3.0:2) where $\underline{A}$ is necessarily symmetric. On differentiation we see that if $\underline{A}$ is non-singular both systems have the same center. Furthermore the latent vectors of $\underline{A}'\underline{A} = \underline{A}^2$ are the same as those of $\underline{A}$ and the latent roots of $\underline{A}^2$ are the squares of those of $\underline{A}$. For example, if the canonical form of the yield surface is

$$(5.0:5) \qquad z = \pm \lambda_1 X_1^2 \pm \lambda_2 X_2^2$$

that of the conditioning surface is

$$(5.0:6) \qquad w = + \lambda_1^2 X_1^2 + \lambda_2^2 X_2^2.$$

We see therefore that the center and the principal axes of the conditioning surface are the same as those of the yield surface, but attenuation of the yield surface is accompanied by even greater attentuation of the conditioning surface and the contours of the conditioning surface are always ellipses. We see therefore, that to the extent to which the confidence region reflects the influence of the conditioning surface, it will (to some lesser extent) reflect the characteristics of the yield surface itself.

It should be remembered that the method described provides a confidence region for a stationary point, which has as its co-ordinates the solution to the equations (5.0:2) and (5.0:3). If the coefficients in (5.0:2) and (5.0:3) are such as to indicate that the fitted surface possesses a maximum, <u>and</u> the confidence region is closed, we may assert that this is a confidence region for a true maximum, (that is to say not merely a confidence region for a stationary point.) This is so because, assuming that the surface can be represented by an equation of second degree, if the plausible range of variation in the a's includes surfaces in which one or more of the λ's could change in sign, then points lying along the axes corresponding to these roots would be acceptable as solutions of equations (5.0:2) and (5.0:3), and open contours would necessarily result.

6. EXAMPLE OF A CONFIDENCE REGION OBTAINED USING AN EXPERIMENTAL DESIGN.

To give some appreciation of the type of confidence regions that might be met in practice, a statistical experiment was carried out as follows. It was assumed that the true equation of a surface was

$$(6.0:1) \qquad \eta = -3.333x_1^2 - 3.333x_2^2 + 4.000x_1x_2 + 4.533x_1 - 1.867x_2 + 78.373$$

and that a 3^2 factorial experiment was conducted at the levels -1,0,1 for the independent variables x_1 and x_2. Random normal deviates with $\sigma = 1$, $\mu = 0$ were added to the nine values of η found at the nine points in the design and these values are recorded in Figure 2. The second degree equation fitted by least squares to these points was

$$(6.0:2) \qquad Y = -2.83x_1^2 - 2.58x_2^2 + 4.02x_1x_2 + 3.70x_1 - 1.62x_2 + 77.82$$

whose contours are indicated by the dotted lines in Figure 2 and whose center, $x_1^o = 0.96$, $x_2^o = 0.44$, is indicated by a cross. The residual sum of squares, (assuming, as is customarily the case, that the model is adequate) based on

3 degrees of freedom was 7.20 giving an estimate of $s^2 = 2.40$. The variance-covariance matrix $\Omega\sigma^2$ for the coefficients of (6.0:2) as estimated from a 3^2 factorial is

$$(6.0{:}3) \qquad C^{-1}\sigma^2 = \begin{bmatrix} \frac{5}{9} & 0 & 0 & -\frac{1}{3} & -\frac{1}{3} & 0 \\ & \frac{1}{6} & 0 & 0 & 0 & 0 \\ & & \frac{1}{6} & 0 & 0 & 0 \\ & & & \frac{1}{2} & 0 & 0 \\ & & & & \frac{1}{2} & 0 \\ & & & & & \frac{1}{4} \end{bmatrix} \sigma^2$$

Substituting the values of $s^2 = 2.40$, and the value of $F_{.05}(2,3) = 9.55$ in equation (2.0:4), we obtain for the 95% confidence region the shaded portion indicated in Figure 2.

In practice, having fitted the second degree surface the experimenter would perform confirmatory experiments. A natural location for these additional experiments would be along the principle axes of the fitted conic. We have assumed that six further points were added as shown in Figure 3. The values which might have been obtained at these points, indicated in Figure 3, were secured by calculating the values at the points using equation (6.0:1) and adding random normal deviates as before. Using all 15 points the equation was refitted giving

$$(6.0{:}4) \qquad Y = -2.87x_1^2 - 2.64x_2^2 + 3.84x_1x_2 + 3.76x_1 - 1.57x_2 + 77.95$$

The variance co-variance matrix for the new estimates was

$$(6.0{:}5)\quad C_0^{-1}\sigma^2 = \begin{bmatrix} .18684 & -.00296 & .00921 & -.05141 & -.11581 & .08576 \\ & .13091 & -.01361 & -.06170 & .02881 & -.01263 \\ & & .13547 & .00592 & .04130 & -.02844 \\ & & & .10962 & -.03792 & -.04748 \\ & & & & .23883 & -.09434 \\ & & & & & .14688 \end{bmatrix} \sigma^2$$

The residual sum of squares was 9.60 and hence the residual error, $s^2 = 1.07$, now based on 9 degrees of freedom. The center of the newly fitted equation was at the point $x_1^0 = 0.89$, $x_2^0 = 0.37$ indicated by the cross in Figure 3.

The confidence region would now be expected to be considerably smaller (i), because of the influence of the extra points, (ii), because of the larger number of degrees of freedom upon which F is based (the critical value of F is reduced from 9.55 to 4.26), and (iii), the first estimate of σ^2 ($s^2 = 2.4$) happens to be considerably greater than the second estimate ($s^2 = 1.07$). The new confidence region, shown as the shaded area in Figure 3, is closed but attenuated.

The size and attenuation of the regions found emphasize the considerable uncertainty that would exist in the location of a maximum, even when the profound effect of lack of fit of the second degree equation is ignored.

CONCLUSION

In this paper the discussion has been confined to obtaining confidence regions for the solution of sets of linear equations. However, it is worth noting that this general method may be used for any set of simultaneous equations which

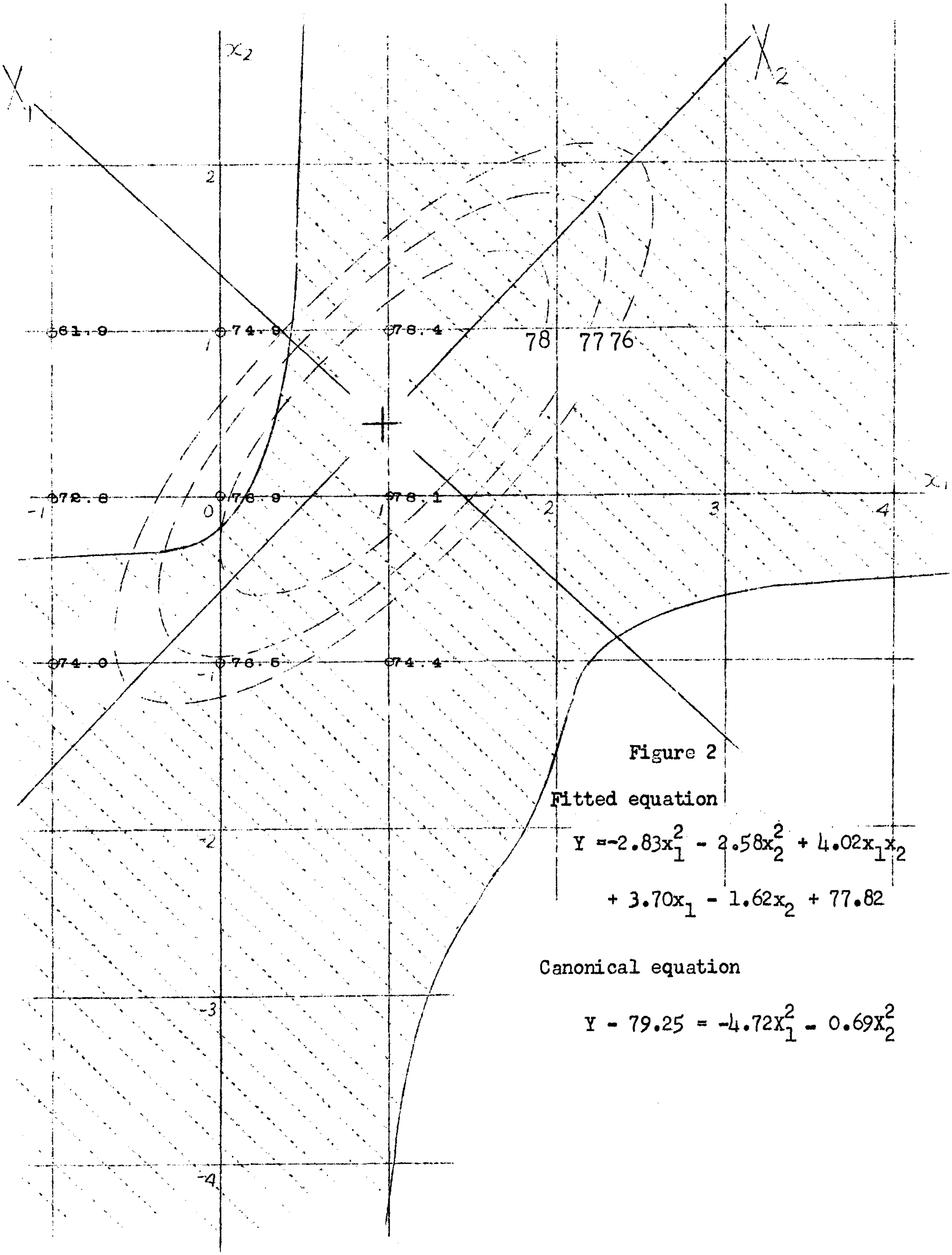


Figure 2

Fitted equation

$$Y = -2.83x_1^2 - 2.58x_2^2 + 4.02x_1x_2 + 3.70x_1 - 1.62x_2 + 77.82$$

Canonical equation

$$Y - 79.25 = -4.72X_1^2 - 0.69X_2^2$$

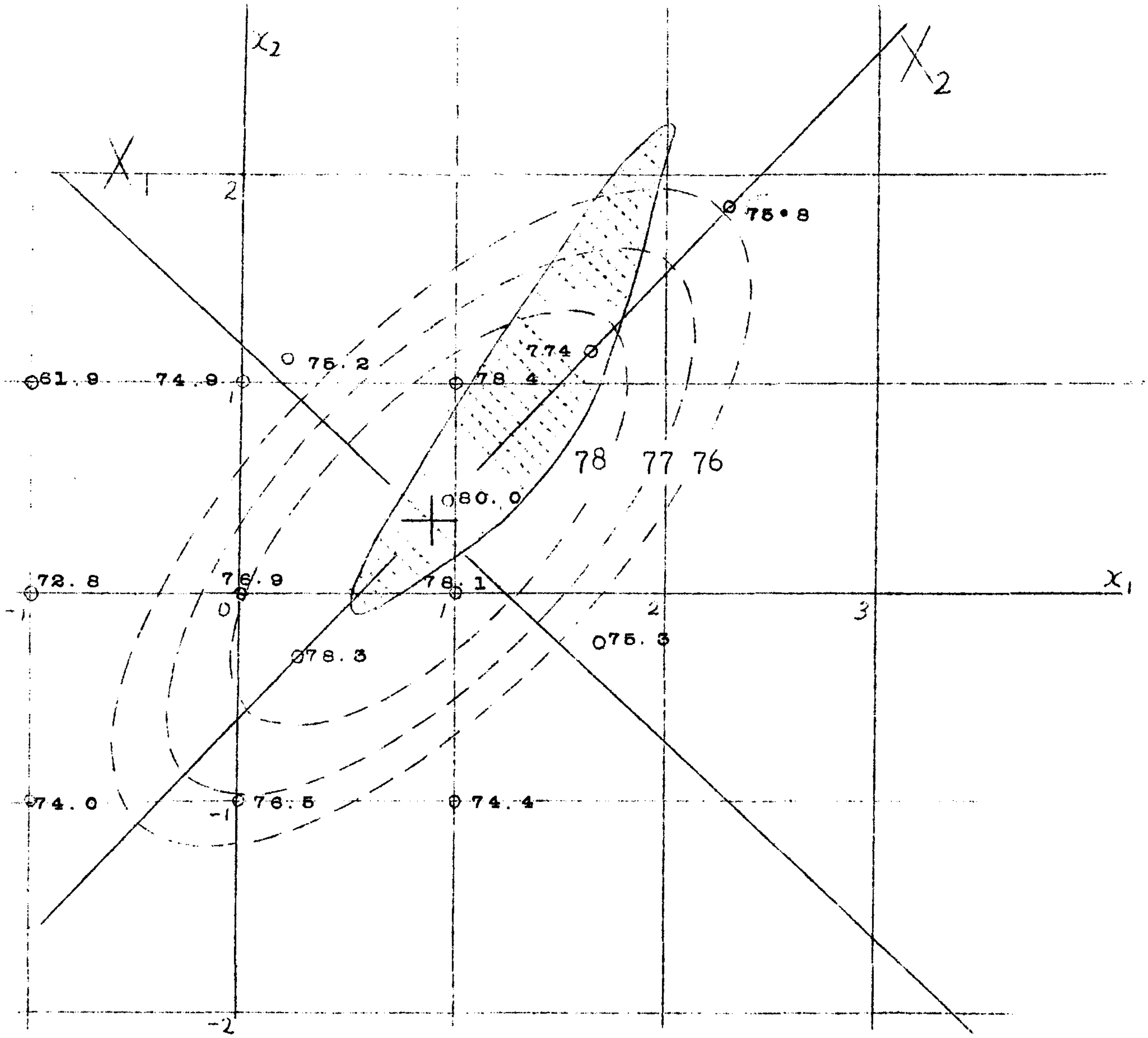


FIGURE 3

Fitted equation $\quad Y = -2.87x_1^2 - 2.64x_2^2 + 3.84x_1x_2 + 3.76x_1 - 1.57x_2 + 77.95$

Canonical equation $\quad Y - 79.35 = -4.68X_1^2 - 0.83X_2^2$

are linear in the coefficients. Thus in principal, we could use the method to find a confidence interval for the solution of the k equations

$$a_{io} + \sum_{j=1}^{k} a_{ij}\, f(x_1, x_2, \ldots, x_k) = 0 \quad (i = 1,2,\ldots,k)$$

Consequently, confidence intervals could be obtained for a stationary point on a surface represented by any equation linear in the coefficients, and not only for the quadric surface we have discussed.

BIBLIOGRAPHY

Box, G. E. P. and Wilson, K. B. (1952). On the Experimental Attainment of Optimum Conditions. J. R. Statist. Soc. B, 13, 1.

Feiller, E. C. (1940). The Biological Standardization of Insulin. J. R. Statist. Soc. Suppl., 7, 1.

Hotelling, H. (1943). Some New Methods in Matrix Calculations. Ann. Math. Statist., 14, 1.

Lonseth, A. T. (1942). Systems of Linear Equations with Coefficients Subject to Error. Ann. Math. Statist., 13, 332.

Morris, J. (1946). An Escalator Process for the Solution of Linear Simultaneous Equations. Phil. Mag., 7, 106.

Turing, A. M. (1947). Rounding Off Errors in Matrix Processes. Quart. J. Mech. and Applied Math., 1, 287.