THE N RESPONSE PROBLEM

By
CHARLES EDWARD ANTLE

Bachelor of Science
Eastern Kentucky State College
Richmond, Kentucky
1954

Master of Arts
Eastern Kentucky State College
Richmond, Kentucky
1955

Submitted to the Faculty of the Graduate School of
the Oklahoma State University
in partial fulfillment of the requirements
for the degree of
DOCTOR OF PHILOSOPHY
August, 1962

THE N RESPONSE PROBLEM

Thesis Approved:

Leroy Folker
Thesis Adviser

Earl E. Marshall

Robert D Morrison

David L. Huber

John B West

Louis E. Hoffman

Dean of the Graduate School

ACKNOWLEDGEMENTS

The author is indebted to Dr. Leroy Folks for suggesting the problem and directing the research throughout the preparation of this thesis.

The author is also indebted to the National Science Foundation and the Department of Health, Welfare, and Education for the financial support they provided.

The author's interest in mathematics was first kindled in Dr. Smith Park's calculus class at Eastern Kentucky State College. Interest in statistics was first acquired while attending a Summer Institute in Statistics directed by Dr. Carl E. Marshall and taught by Dr. Franklin A. Graybill.

Gratitude is extended to Professors Carl E. Marshall, Robert D. Morrison, David L. Weeks, John E. Hoffman, John B. West, and Roy B. Gladstone for serving on the author's advisory committee.

Thanks are also extended to Mrs. Virginia Schenandoah for having typed this thesis.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

The important problem of maximizing a single response by the choice of values for certain variables $x_1, x_2, \ldots, x_p$ has received and is continuing to receive considerable attention. The methods used in seeking these optimum values of the x_i's may be classified as sequential or non-sequential. The sequential methods receiving the most attention are the method of steepest ascent, the method of parallel tangents, and the one factor at a time method. No comparative study of these three has been published.

The method of steepest ascent was proposed by G. E. P. Box and K. B. Wilson (1)[1] in 1951. The method consists of fitting a linear response function by means of a first order design, and then proceeding to experiment in the direction of the gradient. The path thus determined is not invariant under scale transformations of the x_i's. Nevertheless, it has been used extensively with satisfactory results.

O. Kempthorne (2) has recently introduced the method of parallel tangents for seeking the region of maximum response. This method is illustrated in Figure 1 and consists of the following steps:

---

[1]Note: ( ) refers to Selected Bibliography.

1. Center a first order design at x^o and from the results determine the direction of a line L_1 that is tangent to the response contour at x^o.

2. Experiment along a line L_2 that is parallel to L_1 until the point of maximum response along L_2 is determined. Let L_3 be the line connecting this point and x^o.

3. Experiment along L_3 until the point of maximum response along L_3 is determined. In the ideal situation shown in Figure 1, this will occur in the region for which the response is a maximum.

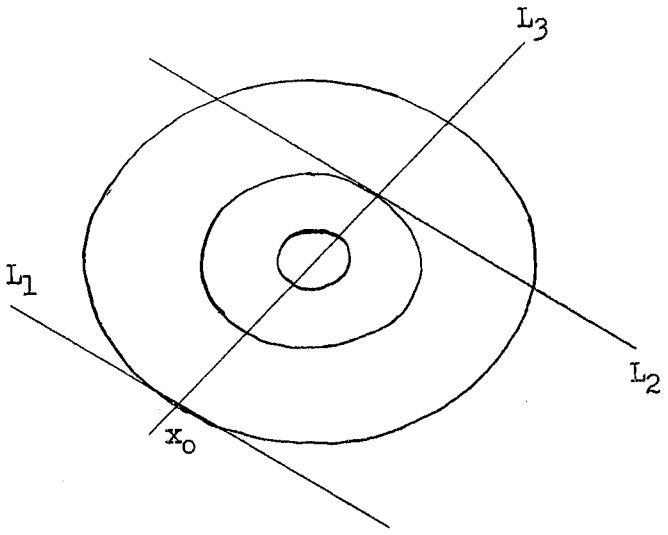

Figure 1. The Method of Parallel Tangents

The one factor at a time method is self-explanatory. Its primary use has been in the engineering fields. It is rather poor when there is considerable interaction among the effects of the variables.

The methods described above are often followed by the fitting of at least a quadratic surface so that the point for maximum response can be better estimated. This last step, though often in a sequence of experiments, is not considered as a sequential method. Another non-sequential method consists of a large experiment conducted at a random selection of points in the design space and simply selecting the combination of x_i's which produced the maximum response.

The problem of selecting optimum values for the control variables $x_1, x_2,\ldots,x_p$ when there are N responses of interest is the subject of this thesis. As an example, suppose an alloy is to be developed for use as an electrical transmission line. It would be desirable to produce one with maximum conductance, minimum weight, and maximum strength. We may choose the levels of certain control variables, such as the amount of copper, amount of steel, etc., in order to accomplish these desirable results. However, it is very unlikely that any choice of these levels will simultaneously maximize the conductance, minimize the weight, and maximize the strength. If the development of this alloy calls for an experimental program, then the question arises as to the object of the experimentation; that is, what is meant by optimum values of the control variables $x_1, x_2,\ldots,x_p$ when there are N responses of interest? We shall find that there will usually be a set of points such that each of them has some sort of optimum property associated with it.

Let us consider a simple example in which there are only five possible combinations of the control variables, say $x^{(1)}, x^{(2)},\ldots,x^{(5)}$, and there are two responses of interest. If the responses are as recorded in the table

	$x^{(1)}$	$x^{(2)}$	$x^{(3)}$	$x^{(4)}$	$x^{(5)}$
y_1	6	5	4	3	4
y_2	3	4	5	6	4

and high values of both responses are desired, then it is clear that $x^{(2)}$ and $x^{(3)}$ are better than $x^{(5)}$, but none of the other points give response vectors that can be compared. We shall see that $x^{(1)}$, $x^{(2)}$, $x^{(3)}$, and $x^{(4)}$ belong to a class of points called the complete set of efficient points. The description and determination of such sets constitutes the major result of this thesis.

Since $y(x)$ is minimized when $-y(x)$ is maximized, the problem is formulated in terms of maximizations only. The problem considered is that of selecting the values of the p control variables $x_1, x_2, \ldots, x_p$ so that the N responses $y_1(x), y_2(x), \ldots, y_N(x)$ will in some sense be jointly maximized.

CHAPTER II

EFFICIENT POINTS AND THE COMPLETE SET OF EFFICIENT POINTS

Let x be a p dimensional vector. Suppose there are N response functions $y_1(x)$, $y_2(x),\ldots,y_N(x)$, and it is desirable to have high values of all the responses.

Definition 1. A point x^o is better than the point x for the responses $y_1(x)$, $y_2(x),\ldots,y_N(x)$ if

i) $y_i(x^o) \geq y_i(x)$ for all i and

ii) $y_k(x^o) > y_k(x)$ for at least one k.

While better than certainly depends upon the set of responses, when there is no ambiguity the reference "for the responses $y_1(x)$, $y_2(x),\ldots,y_N(x)$" will be omitted.

Definition 2. The point x^o is an efficient point for $y_1(x)$, $y_2(x),\ldots,y_N(x)$ if there exists no x better than x^o.

Definition 3. The complete set of efficient points, if it exists, is the set of all points such that there are none better, and given any x not in the set there exist an x^o in the set that is better.

It is easily seen from the above definitions that when the complete set of efficient points exists, it is simply the set of all efficient points. It follows then that it is unique and may be referred to as the complete set of efficient points.

As an example suppose $p = 2$, $N = 2$ and the responses are

$$y_1(x) = 100 - (x_1-2)^2 - 4(x_2-3)^2$$

$$y_2(x) = 50 - 2(x_1-5)^2 - (x_2-2)^2.$$

Then the contours are as shown in Figure 2 and the complete set of efficient points is given by the section of a hyperbola connecting the two points of maximum responses. Note that at each efficient point the gradients of the responses are in opposite directions.

It is not difficult to construct a set of responses for which the complete set of efficient points will fail to exist. For example, if $y_1(x) = 1$ and $y_2(x) = |x|$, then the complete set of efficient points does not exist. Hence it is of interest to determine conditions sufficient for the existence of the complete set of efficient points. These conditions are set forth in Theorem I and it is seen that they are met by many of the response functions used for models.

Theorem I. If $y_i(x)$ is everywhere continuous for all i and at least one of the sets $S_i(C) = \{x | y_i(x) \geq C\}$ is bounded for all C, then the complete set of efficient points for $y_1(x)$, $y_2(x),\ldots,y_N(x)$ exists.

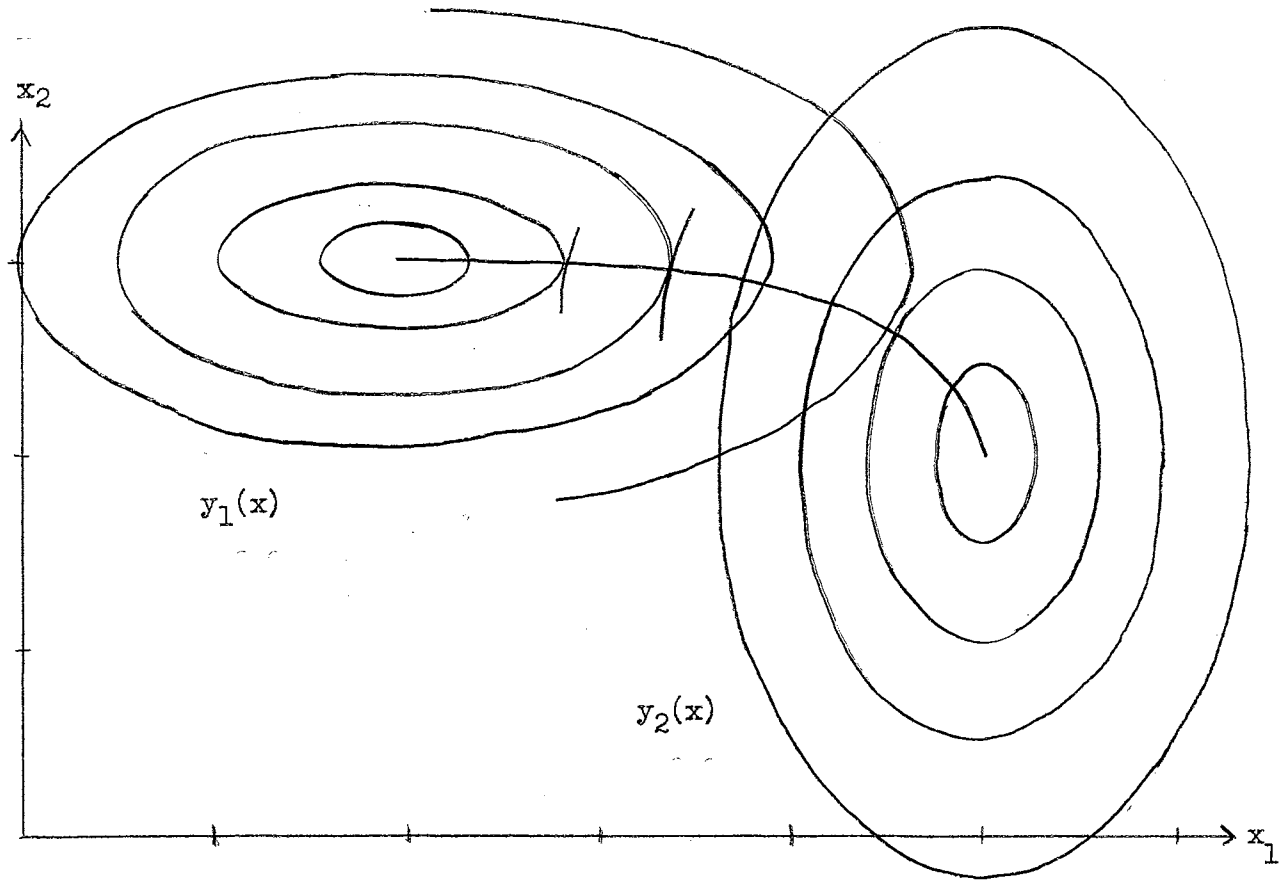

Figure 2. Contours and Efficient Points

Proof: It suffices to show that given any x^{φ}, there exists an $x^{(1)}$ such that

1) $$\begin{bmatrix} y_1(x^{(1)}) \\ \vdots \\ y_N(x^{(1)}) \end{bmatrix} \geq \begin{bmatrix} y_1(x^{\varphi}) \\ \vdots \\ y_N(x^{\varphi}) \end{bmatrix}$$

while there exists no x such that

2) $$\begin{bmatrix} y_1(x) \\ \vdots \\ y_N(x) \end{bmatrix} > \begin{bmatrix} y_1(x^{(1)}) \\ \vdots \\ y_N(x^{(1)}) \end{bmatrix}.$$

Let x^{φ} be any given point with responses $y_1(x^{\varphi}),\ldots,y_N(x^{\varphi})$. The set S given by

3) $$S = \left\{ x \mid y_i(x) \geq y_i(x^{\varphi}),\ i = 1,2,..,N \right\}$$

is a closed and bounded non-empty region. Since $y_N(x)$ is continuous, it follows that there exists at least one $x^{(N)}$ in the set S such that

4) $$y_N(x^{(N)}) = \max_{x \in S} y_N(x)$$

Let

5) $$P_i = \left\{ x \mid y_i(x) \geq y_i(x^{(i)}) \right\}$$

Now $S \cap P_N$ is a closed and bounded non-empty region over which $y_{N-1}(x)$ is continuous so that there exists at least one $x^{(N-1)}$ in the set $S \cap P_N$ such that

6) $$y_{N-1}(x^{(N-1)}) = \max_{x \in S \cap P_N} y_{N-1}(x).$$

We can repeat the above process until

7) $$y_1(x^{(1)}) = \max_{x \in S \cap P_N \cap P_{N-1} \cdots \cap P_2} y_1(x).$$

Clearly the method of selecting $x^{(1)}$ assures us that inequality 1 holds and there is no x such that inequality 2 holds. Thus, the proof is complete.

It follows from the definitions that if the complete set of efficient points exists and a convex combination of the responses has a maximum, then it must be attained at some efficient point. Thus, if a non-negative value is assigned to each of the responses, the sum S given by

$$S = \Sigma\, a_i y_i(x),\ a_i \geq 0 \text{ for all } i$$

will have its maximum at an efficient point, provided the complete set exists and S has a maximum value. The following theorems have been formulated in such a way as to make the identification of these efficient points particularly simple.

Theorem 2. If $\nabla y_1(x), \nabla y_2(x), \ldots, \nabla y_N(x)$ exist at a point x^0, then a necessary condition for x^0 to be an efficient point is that there exist a vector α such that

8) $$\Sigma\, \alpha_i \nabla y_i(x^0) = \emptyset,\ \alpha_i \geq 0 \text{ for all } i, \text{ and } \Sigma\, \alpha_i = 1.$$

Proof: Assume no such vector exists. Then none of the $\nabla y_i(x^o)$ are zero, and the convex hull of the tips of the vectors $\nabla y_i(x^o)$ does not contain x^o. Call this convex hull D. Since x^o and D are convex and disjoint, there exists a hyperplane that strictly separates them. Let the normal to this hyperplane that is directed toward D be V. All $\nabla y_i(x^o)$ have a positive component in the direction of V. Therefore, there exists a point x in the direction of V such that all responses are higher than they were at x^o, and thus x^o is not an efficient point.

At this point it should be noted that if the conditions of Theorems 1 and 2 are satisfied, then the only x which should be considered are those which are solutions to equation 8. Let the solution for a given α be written as $x(\alpha)$. Since there are only N-1 independent components of α, it follows that when $x(\alpha)$ is single valued the efficient points have been identified by an N-1 dimensional vector. Furthermore, we may without any loss at all reduce the domain of the $y_i(x)$ to that of the efficient points. Thus, it is clear that the points which should be considered and their responses are a function of the N-1 independent components for α. For example, if p = 5 and N = 2, then instead of considering the responses of $y_1(x)$ and $y_2(x)$ in the 5 dimensional x space, we can consider them in the 1 dimensional space of α.

It would be convenient if Theorem 2 contained sufficient conditions for x^o to be an efficient point. That it does not is apparent when we consider the response functions

$$y_1(x) = 100 - (x_1-1)^2 - (x_2-2)^2$$

$$y_2(x) = x_1^2 - x_2^2.$$

Note that the conditions of Theorem 1 are satisfied so that the complete set of efficient points exists. Also the conditions of Therorem 2 are satisfied everywhere, and it follows that the complete set of efficient points are among the set E where

$$E = \left\{ x \middle| \alpha \bigtriangledown y_1(x) + (1-\alpha) \bigtriangledown y_2(x) = \emptyset,\ 0 \le \alpha \le 1, \right\}.$$

The parametric equations of this set are

$$x_1 = \alpha(2\alpha-1)^{-1}, \quad x_2 = 2\alpha; \quad 0 \le \alpha \le 1.$$

The graph of these responses as a function of α is given in Figure 3. It is clear from the figure that the $x(\alpha)$ for which $\alpha < .5$ are not efficient points and the complete set of efficient points is given by the $x(\alpha)$ for which $\alpha \ge .5$.

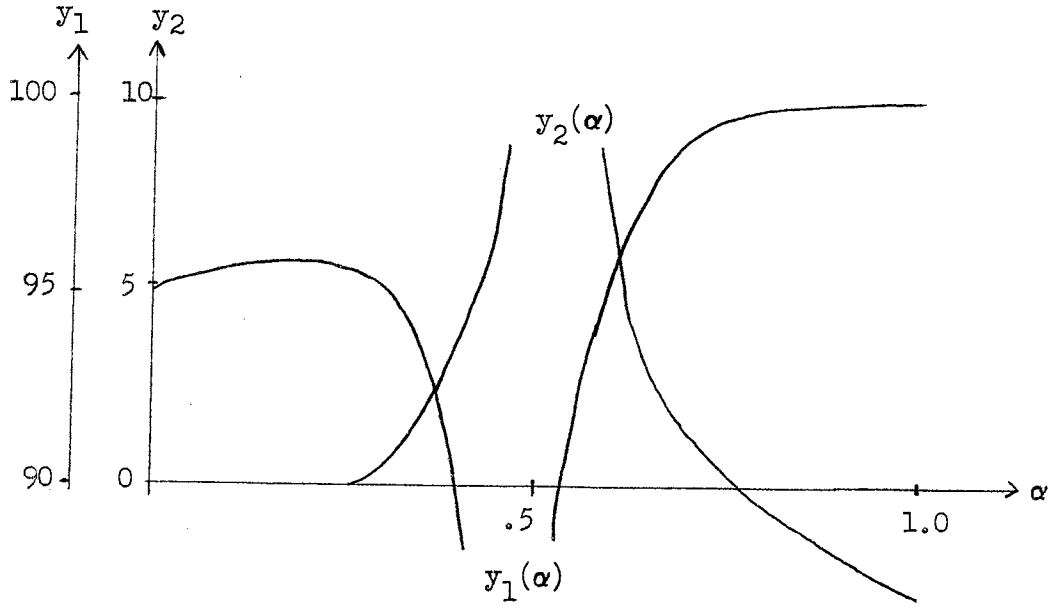

Figure 3. Responses at the Efficient Points

In general the values of α which correspond to efficient points depend upon the structure of the contours of the response functions. The following definitions will facilitate the study of these.

Definition 4. A function $y(x)$ is a type I function if the sets $S(C) = \{x | y(x) \geq C\}$ are convex for all C.

We might think of a type I function of two variables as one which has at most a single "mound" and no valleys. An example of a type I function is given by

$$y(x) = c + \exp - \left[(x_1-h)^2 + (x_2-k)^2\right].$$

Definition 5. A set of functions such that all convex combinations of the functions result in a type I function will be called a type I set of functions.

Definition 6. Let $y(x)$ be a type I function and $H(x^*)$ a hyperplane tangent to the set $S\left[y(x^*)\right] = \{x | y(x) \geq y(x^*)\}$ at x^*. Then $y(x)$ will be a type IA function if $H(x^*) \cap S\left[y(x^*)\right] = x^*$ for all x^*.

Note that if $y(x)$ is a type IA function, then every point on a contour is an extreme point for the set enclosed, that is there are no straight segments on any of the contours. The function $y(x) = c + \exp - \left[(x_1-h)^2 + (x_2-k)^2\right]$ is also a type IA function.

Theorem 3. If the functions $y_1(x), y_2(x),\ldots,y_N(x)$ form a type I set of functions and $\nabla y_1(x), \nabla y_2(x),\ldots,\nabla y_N(x)$ exist and are non-zero at x^o, then a necessary and sufficient condition for x^o to be an efficient point is that there exist a vector α such that

9) $\Sigma\, \alpha_i\ \nabla y_i(x^o) = \emptyset$ where $\alpha_i \geq 0$ for all i, and $\Sigma\, \alpha_i = 1$.

Proof: It follows from equation 9 that there exists a non-xero α_k such that

10) $$\alpha_k\ \nabla y_k(x^o) = -\sum_{i\neq k} \alpha_i \nabla y_i(x^o).$$

Let

11) $$y(x) = \sum_{i\neq k} \alpha_i y_i(x)$$

so that $y(x)$ is a type I function and

12) $$\alpha_k\ \nabla y_k(x^o) = -\ \nabla y(x^o).$$

Consider the hyperplane

13) $$(x-x^o)\cdot\ \nabla y(x^o) = 0\ .$$

It follows from the type I property of the functions $y_k(x)$ and $y(x)$ that

14) $y_k(x) > y_k(x^o) \Rightarrow (x-x^o)\cdot\ \nabla y_k(x^o) > 0 \Rightarrow (x-x^o)\cdot\ \nabla y(x^o) < 0$

15) $y_k(x) \geq y_k(x^o) \Rightarrow (x-x^o)\cdot\ \nabla y_k(x^o) \geq 0 \Rightarrow (x-x^o)\cdot\ \nabla y(x^o) \leq 0$

16) $y(x) > y(x^o) \Rightarrow (x-x^o)\cdot\ \nabla y(x^o) > 0$

17) $y(x) \geq y(x^o) \Rightarrow (x-x^o)\cdot\ \nabla y(x^o) \geq 0\ .$

Now consider the two ways in which x^o could fail to be an efficient point. First, it could be that there exists an x such that

18) $y_i(x) \geq y_i(x^o)$ for all i, and $y_k(x) > y_k(x^o)$.

If we note that this requires

19) $y(x) \geq y(x^o)$ and $y_k(x) > y_k(x^o)$

we see from inequalities 14 and 17 that this is not possible. The other way in which x^o could fail would occur if

20) $y_i(x) \geq y_i(x^o)$ for all i, and $y_j(x) > y_j(x^o)$ for some $j \neq k$.

This would require an x such that

21) $y_k(x) \geq y_k(x^o)$ and $y(x) > y(x^o)$

and it follows from inequalities 15 and 16 that there is no such x. Therefore, we must conclude that x^o is an efficient point. Since the necessity of equation 1 follows from Theorem 2, this completes the proof.

The conditions of the theorem may be weakened somewhat when N = 2. In the proof of the theorem the condition that the functions form a type I set of functions was used in order to make y(x) a type I function. However, when N = 2 and $y_1(x)$ and $y_2(x)$ are type I functions, it follows that y(x) is a type I function. Thus it is not necessary to require $y_1(x)$ and $y_2(x)$ to form a type I set of functions. We state this in the form of a corollary.

Corollary: If $y_1(x)$ and $y_2(x)$ are type I functions and $\nabla y_1(x)$ and $\nabla y_2(x)$ exist and are non-xero at x^o, then a necessary and sufficient condition for x^o to be an efficient point for $y_1(x)$ and $y_2(x)$ is that there exist an α such that

22) $\alpha \nabla y_1(x^o) + (1-\alpha) \nabla y_2(x^o) = \emptyset$, $0 \leq \alpha \leq 1$.

Theorem 4. If the functions $y_1(x), y_2(x),\ldots,y_N(x)$ are type IA functions and $\nabla y_1(x), \nabla y_2(x),\ldots, \nabla y_N(x)$ exist and are non-zero at x^o, then a necessary and sufficient condition for x^o to be an efficient point is that there exist a vector α such that

23) $\Sigma \alpha_i \nabla y_i(x^o) = \emptyset$, $\Sigma \alpha_i = 1$, and $\alpha_i \geq 0$ for all i.

Proof: The necessity follows immediately from Theorem 2. For the proof of the sufficiency, first note that since $y_i(x)$ is a type IA function it follows that for all i

24) $(x-x^o) \cdot \nabla y_i(x^o) > 0$ if $y_i(x) \geq y_i(x^o)$ and $x \neq x^o$.

Thus we have

25) $\Sigma \alpha_i (x-x^o) \cdot \nabla y_i(x^o) > 0$ if $y_i(x) \geq y_i(x^o)$ for all i, and $x \neq x^o$

which requires that

26) $(x-x^o) \cdot \Sigma \alpha_i \nabla y_i(x^o) > 0$ if $y_i(x) \geq y_i(x^o)$ for all i, and $x \neq x^o$.

Equations 23 and 26 combine to tell us that there is no $x \neq x^o$ such that $y_i(x) \geq y_i(x^o)$ for all i. Thus x^o is an efficient point.

We note that Theorem 4 places the requirements on the individual response functions while Theorem 3 requires something of the set of response functions. We shall see that there are times when each will be useful.

We recall that y(x) is a concave function if

$$y\left[(1-\alpha)x^{(1)} + \alpha x^{(2)}\right] \geq (1-\alpha)y(x^{(1)}) + \alpha y(x^{(2)})$$

for all $x^{(1)}$, $x^{(2)}$ and $0 \leq \alpha \leq 1$. Then it is seen that a concave function is also a type I function, that is its contours enclose convex sets. It is also seen that the sum of two concave functions is a concave function so that any set of concave functions is also a type I set of functions. A concave function will be a type IA function if it is a strictly concave function.

Theorem 5. All strictly concave functions are type IA functions.

Proof: It is necessary to show that x^o is the unique solution of the simultaneous equations

27) $$(x-x^o)\cdot \nabla y(x^o) = \emptyset$$

28) $$y(x) \geq y(x^o).$$

Suppose there is an $x^{(1)} \neq x^o$ such that the above equations are true. If we let $x^{(2)} = \alpha x^o + (1-\alpha)x^{(1)}$ with $0 < \alpha < 1$, then it follows from the strictly concave property of $y(x)$ that $y(x^{(2)}) > y(x^o)$. Clearly, $x^{(2)}$ is on the supporting hyperplane of the convex set $S\left[y(x^o)\right]$. Since a concave function is always continuous, it follows that there are points on both sides of the hyperplane for which $y(x) > y(x^o)$. Since this is not possible, we must conclude that x^o is the unique solution of equations 27 and 28, and therefore, $y(x)$ is a type IA function.

A further property of concave functions that will also be useful is given in the following theorem.

Theorem 6. If all $\nabla y_i(x)$ exist everywhere and all $y_i(x)$ are concave, then every efficient point maximizes some convex combination of the $y_i(x)$.

Proof: Assume x^o is an efficient point. Then from Theorem 2 it follows that there exists a vector α such that

29) $$\Sigma \alpha_i \nabla y_i(x^o) = \emptyset, \ \Sigma \alpha_i = 1, \text{ and } \alpha_i \geq 0 \text{ for all } i.$$

Now consider the function $y(x)$ given by

30) $$y(x) = \Sigma \alpha_i y_i(x).$$

Clearly $y(x)$ is a concave function with $\nabla y(x^o) = \emptyset$. It follows that $y(x)$ attains its maximum value at x^o.

Let us return to the problem of maximizing the sum $S(x)$ where

$$S(x) = \Sigma \alpha_i y_i(x), \ \alpha_i \geq 0 \text{ for all } i \text{ and } \sum_i \alpha_i = 1.$$

If the $y_i(x)$ satisfy the conditions of Theorem 3 or Theorem 4 and $S(x)$ has a maximum value, then it occurs at a point x^o for which

$$\nabla S(x^o) = \Sigma \alpha_i \nabla y_i(x^o) = 0.$$

It follows that the maximum occurs at the efficient point corresponding to the vector α, that is at $x(\alpha)$. Thus, if we have used Theorem 3 or Theorem 4 to find the complete set of efficient points, then we can immediately obtain the x that will maximize a convex combination of the responses.

QUADRATIC RESPONSE SURFACES

Quite often the response surfaces obtained from an experiment will be quadratic functions. The special properties of these will now be investigated. Let the quadratic response be written as

31) $$y(x) = a - x'Ax + B'x$$

where A is a pxp symetric matrix and B is a pxl vector.

Theorem 7. If y(x) is given by equation 31, then y(x) is a concave function if and only if A is either positive definite or positive semidefinite. If A is positive definite, then y(x) is strictly concave.

Proof: Let $x^0 = \alpha x^{(1)} + (1-\alpha)x^{(2)}$ where $0 < \alpha < 1$. Then it follows from equation 31 that

32) $$y(x^0) = a - \left[\alpha x^{(1)} + (1-\alpha)x^{(2)}\right]' A \left[\alpha x^{(1)} + (1-\alpha)x^{(2)}\right] + B' \left[\alpha x^{(1)} + (1-\alpha)x^{(2)}\right].$$

It follows from equation 32 that

33) $$y(x^0) = \alpha \left[a - x^{(1)'}Ax^{(1)} + B'x^{(1)}\right] + (1-\alpha)\left[a - x^{(2)'}Ax^{(2)} + B'x^{(2)}\right] - \alpha(1-\alpha)\left[x^{(1)'}Ax^{(2)} + x^{(2)'}Ax^{(1)} - x^{(1)'}Ax^{(1)} - x^{(2)'}Ax^{(2)}\right]$$

34) $$y(x^0) = \alpha y(x^{(1)}) + (1-\alpha)y(x^{(2)}) + \alpha(1-\alpha)\left[x^{(1)} - x^{(2)}\right]' A \left[x^{(1)} - x^{(2)}\right].$$

Thus, when A is positive definite or positive semidefinite

35) $$y(x^o) \geq \alpha y(x^{(1)}) + (1-\alpha)y(x^{(2)})$$

and if A is positive definite the strict inequality holds. This completes the proof.

Since all concave function are type I functions, and all strictly concave functions are IA functions, we have the following corollary.

Corollary 1. If the matrix A in equation 31 is positive definite or positive semidefinite, then y(x) is a type I function. If A is positive definite, then y(x) is a type IA function.

Since the sum of two concave functions is a concave function, we have the additional corollary.

Corollary 2. If $y_i(x) = a_i - x'A_ix + B_ix$ for $i = 1,2,\ldots,K$, and all the A_i are positive definite or positive semidefinite, then $\{y_1(x), y_2(x),\ldots,y_K(x)\}$ is a type I set of functions.

It follows from Theorem 3 and the above corollary that if the responses are as given in Corollary 2, then except for those points at which some $\nabla y_i(x) = \varphi$, the complete set of efficient points is given by the set

36) $$\left\{x \middle| x = .5\left[\Sigma \alpha_i A_i\right]^{-1} \Sigma \alpha_i B_i;\ 0 \leq \alpha_i < 1,\ \Sigma \alpha_i = 1\right\}.$$

In the event N = 2, this set may be written

37) $$\left\{x \middle| x = .5\left[\alpha A_1 + (1-\alpha)A_2\right]^{-1}\left[\alpha B_1 + (1-\alpha)B_2\right];\ 0 < \alpha < 1,\right\}.$$

Note that the points for which $\nabla y_i(x) = \phi$ are efficient points if and only if A_i is positive definite.

The simplicity with which the possible responses of interest may be observed in this case should be noted. Though the space spanned by the x vectors may be p dimensional, the choice of the x may be made by considering the responses as a function of an N - 1 dimensional vector. For example, let the response functions be

$$y_1(x) = 46-4x_1^2-4x_2^2-x_3^2-x_4^2+24x_1+16x_2+2x_3+2x_4$$

$$y_2(x) = 100-x_1^2-4x_2^2-4x_3^2-x_4^2.$$

Then the complete set of efficient points will be given by the set

$$\left\{x \middle| x_1 = 12\alpha(3\alpha+1)^{-1},\ x_2 = 2\alpha, x_3 = \alpha(4-3\alpha)^{-1},\ x_4 = \alpha;\ 0 \le \alpha \le 1\right\}.$$

The responses at the efficient points may then be plotted as a function of α as shown in Figure 4.

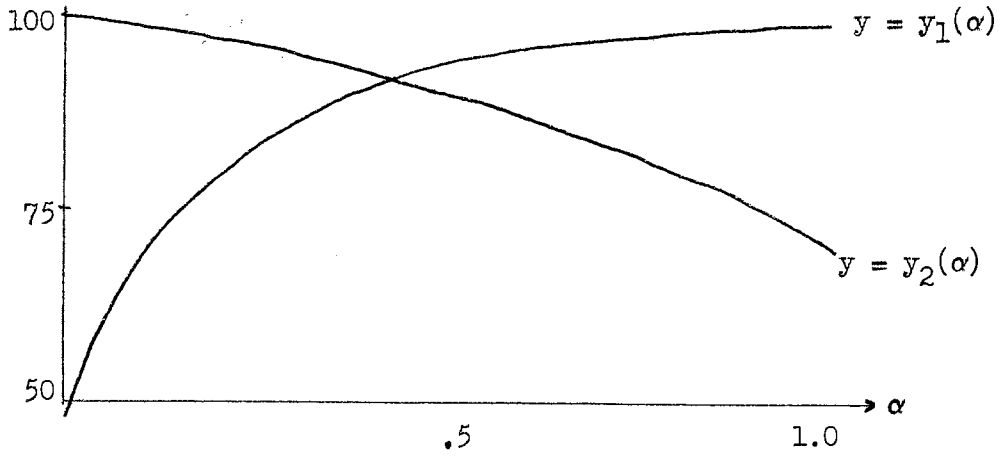

Figure 4. Responses at the Efficient Points

EFFICIENT POINTS SUBJECT TO RESTRICTIONS

We now consider the effect of restricting the domain for x to a set F.

Definition 4. The point x is a feasible point if it is contained in the set F.

Definition 5. The complete set of feasible efficient points is the complete set of efficient points when only feasible points are considered.

It follows from the definitions that the complete set of feasible efficient points includes all the feasible efficient points. Furthermore, if all the efficient points are feasible, then the complete set of feasible efficient points is simply the complete set of efficient points. However, if some efficient points are not feasible, then some new points may become members of the complete set of feasible efficient points. It is clear from the proof of Theorem 2 that any new efficient points must lie on the restricting boundary or be at interior points where $\Sigma\, \alpha_i \nabla y_i(x) = \emptyset$ with $\alpha_i \geq 0$ and $\Sigma\, \alpha_i = 1$, provided the conditions in Theorem 2 are satisfied.

The above considerations suggest the following procedure. First obtain the complete set of efficient points without considering the restrictions. Then if the efficient points are all feasible, the restrictions are of no concern. Furthermore, if from the complete

set a feasible efficient point is the choice for the operation, then do not consider the additional efficient points that may be introduced because some were not feasible. The additional efficient points introduced when some were not feasible will not be as good as the ones they replace. In the event the efficient point desired for the operation is not feasible, the new ones must be considered. The following theorem will be an aid in finding these new efficient points.

Theorem 8. Suppose all $\nabla y_i(x)$ exist everywhere and all $y_i(x)$ are concave. Let the set of feasible values of x be the set $\{x|Q(x) \geq 0\}$ where $Q(x)$ is a concave function and $\nabla Q(x)$ exists everywhere. Also let x^o be an efficient point introduced by the restrictions upon x. Then there exists a vector α with $\alpha_i \geq 0$ and $\Sigma \alpha_i = 1$ such that $\Sigma \alpha_i y_i(x^o)$ is the maximum of $\Sigma \alpha_i y_i(x)$ over the feasible values of x. Furthermore, x^o is on the restricting boundary.

Proof: The conditions imposed upon the $y_i(x)$ are sufficient for Theorem 3. Thus no points such that $\Sigma \alpha_i \nabla y_i(x) = \emptyset$ are introduced as efficient points because they are already efficient points. It follows from the considerations preceding the theorem that x^o must be on the boundary. Thus $Q(x^o) = 0$.

Now suppose $Q(x)$ is considered as an $N + 1$ response. Let us show that x^o is also an efficient point for $y_1(x)$, $y_2(x)$,..., $y_N(x)$, $Q(x)$. Assume x^o is not an efficient point for the $N + 1$ responses. Then there is an x^φ that is better than x^o for $y_1(x)$,

$y_2(x),\ldots,y_N(x)$, $Q(x)$. However, x^{φ} cannot be better than x^o for $y_1(x)$, $y_2(x),\ldots,y_N(x)$, so $Q(x^{\varphi})$ must be greater than $Q(x^o)$. Thus x^{φ} is not on the boundary and therefore was not introduced by the restriction as an efficient point for $y_1(x)$, $y_2(x),\ldots,y_N(x)$. However, x^{φ} is a feasible efficient point for the N responses. Thus it was an efficient point for the N responses before the restriction of the x. It would then follow that x^o was also an efficient point for $y_1(x)$, $y_2(x),\ldots,y_N(x)$ before the restriction of the x. Since this contradicts the definition of x^o it follows that the assumption is not correct and x^o is an efficient point for the responses $y_1(x)$, $y_2(x),\ldots,y_N(x)$, $Q(x)$.

Theorem 6 assures us that there is a β such that

$$\sum^{N} \beta_i y_i(x^o) + \beta_{N+1} Q(x^o) \geq \sum \beta_i y_i(x) + \beta_{N+1}\, Q(x) \text{ for all } x$$

where

$$\sum^{N+1} \beta_i = 1, \text{ and } \beta_i \geq 0 \text{ for all } i.$$

Since $Q(x^o) = 0$ it follows that

$$\sum^{N} \beta_i y_i(x^o) \geq \sum^{N} \beta_i y_i(x) \text{ for all } x \text{ such that } Q(x) \geq 0.$$

Thus $\sum^{N} \beta_i y_i(x^o)$ is the maximum value of $\sum^{N} \beta_i y_i(x)$ when only feasible values of x are considered. The proof is completed by letting α be a vector with

$$\alpha_i = \beta_i \left[\sum^{N} \beta_i\right]^{-1}, \quad i = 1,2,\ldots,N.$$

It should be emphasized that if $\Sigma \alpha_i y_i(x)$ has a maximum value, then it occurs at an efficient point. However, it is not true that all efficient points maximize some such function. For example if x is a one dimensional vector and $y_1(x) = \left[4\right]^{-|x|}$, $y_2(x) = \left[4\right]^{-|x-1|}$, then the complete set of efficient points is given by the set $\left\{x|0 \leq x \leq 1\right\}$, but the point $x = .5$ does not maximize any convex combination of $y_1(x)$ and $y_2(x)$. In fact, the only efficient points which maximize some convex combination of the responses are the points $x = 0$ and $x = 1$. However the conditions of Theorem 3 are satisfied and the theorem provides the efficient points except for $x = 0$ and $x = 1$ at which the gradients fail to exist.

When the conditions of Theorem 8 are satisfied, we can employ any of the standard methods for maximizing a function subject to certain restrictions and be assured that all of the efficient points can be obtained by this method. If the conditions of the theorem are not satisfied, as in the above example, there may be efficient points that cannot be obtained through the maximization of convex combinations of the responses.

CHAPTER III

EXPERIMENTAL DETERMINATION OF THE EFFICIENT POINTS

In the previous chapter it was assumed that the responses were certain known functions and from these the complete set of efficient points was obtained. However, the N response problems we are most likely to encounter will be those in which the response functions are not known. In this event it will be necessary to estimate the response functions, or at least their gradients, in order to apply the theory developed in Chapter II. This will be accomplished by means of an experimental program.

Now that the response functions are to be obtained from an experiment, we must recognize that the response for a given x will be a random variable. The response functions for which we shall seek efficient points will be either the surfaces which represent the expected values of the responses or the medians of the responses. For example, if we assume the model

$$y(x) = \beta_o + \Sigma\, \beta_i x_i + \sum_{i \le j} \sum \beta_{ij}\, x_i x_j + e,\ e \sim N(o,\sigma^2)$$

then the corresponding response function will be $\mathcal{E}\left[y(x)\right]$. Hereafter when we refer to an efficient point for a set of responses, we shall mean an efficient point for either their expected or median responses.

Three methods for estimation of the efficient points are presented. Each is in some way suggested by the results of Chapter II. Methods 2 and 3 are sequential methods.

Method 1. Fitted response functions are obtained from an experiment and the efficient points for the fitted responses are then used as estimates of the efficient points for the expected responses. It is clear from the results of Chapter II that fitted quadratic functions would be desirable for this purpose. Another possible surface that may be useful is given by

$$y(x) = k + \exp\left[\beta_0 + \Sigma\, \beta_i x_i + \sum_{i \le j} \beta_{ij} x_i x_j\right].$$

It seems that some responses could be better represented over a large region by this surface than by a quadratic surface, especially if the responses are all non-negative. If we assume the model

$$y(x) = k + \exp\left[\beta_0 + \Sigma\, \beta_i x_i + \sum_{i \le j} \beta_{ij} x_i x_j + \epsilon\right]$$

with $\epsilon \sim N(0, \sigma^2)$ and k known, then we can obtain the minimum variance unbiased estimates of the β's by considering $z = \log\left[y(x)-k\right]$ as the response. Furthermore, since z is a strictly increasing function of y(x) we obtain the complete set of efficient points for $y_1(x), y_2(x), \ldots, y_N(x)$ when we obtain the complete set of efficient points for $z, y_2(x), \ldots, y_N(x)$. Since z is a quadratic function, it follows that the special methods developed in Chapter II for quadratic response functions may also be used with these response functions. This model will be used in a later example.

It is important to obtain good estimates of the gradients in the region of the efficient points of interest. If this region of interest is large, as it may well be when we seek the complete set of efficient points, then lack of fit of the model may be a serious problem. In this event we must either choose a better model, or partition the region and fit a response surface in each of the sub-regions. The efficient points thus determined probably will not be connected, but this is of no great concern to us. It would provide some indication of the variability of the estimates. An example illustrating Method 1 is given in the Appendix.

Method 2. Suppose the unknown reponse functions are type IA functions. Then according to Theorem 3 when $\nabla y_1(x)$, $\nabla y_2(x),\ldots,$ $\nabla y_N(x)$ exist and are non- zero at x^o, a necessary and sufficient condition for x^o to be an efficient point is that there exists a vector α such that

38) $$\Sigma \alpha_i \nabla y_i(x^o) = 0,\ \alpha_i \geq 0 \text{ for all } i, \text{ and } \Sigma \alpha_i = 1.$$

This suggests that response surfaces be fitted by means of a first order design and then the resulting $\nabla \hat{y}_i(x)$ examined to see if there is a vector α such that equation 38 is approximately true for the $\nabla \hat{y}_i(x)$. If such a α is obtained and the design is centered at x^o, then x^o is a reasonable estimate of an efficient point for the expected responses. If there is no such α, then we should choose a vector β such that $\nabla \hat{y}_i(x) \cdot \Sigma \beta_i \frac{\nabla \hat{y}_i(x^o)}{\nabla \hat{y}_i(x^o)} \geq 0$ for all i, and then

proceed to experiment at the points

$$39) \qquad x^{(k)} = x^o + k \sum \beta_i \frac{\nabla \hat{y}_i(x^o)}{|\nabla \hat{y}_i(x^o)|}$$

until at least one response decreases.

The choice of β is rather arbitrary and should be determined by considering the relative importance of increasing the various responses. It may be changed at any step in the experiment. This is an advantage of this method in that it permits us to work toward an efficient point of our choice. When $N = 2$ we could choose all components of β to be positive and equal. It follows that

$$\nabla \hat{y}_i(x^o) \cdot \left[\sum \beta_i \frac{\nabla \hat{y}_i(x^o)}{|\nabla \hat{y}_i(x^o)|}\right] \geq 0 \text{ for } i = 1,2$$

because $x^{(k)}$ as given by equation 39 is then along the angle bisector of the angle between $\nabla \hat{y}_1(x)$ and $\nabla \hat{y}_2(x)$. This method is illustrated in Figure 5 where β is taken to have equal positive components. Note that in some cases we can expect to find an efficient point very quickly by this method, and furthermore, we can exercise a good deal of control over the choice of the efficient point.

When at least one response decreases, the experimenter must decide if he wishes to use another first order design and repeat the above process, or if he wishes to perform a larger experiment. Even though it appears that an efficient point has been obtained, the experiment probably should not be terminated. It would be desirable to estimate the complete set of efficient points for the expected responses, or at least a subset of the complete set in the region of interest.

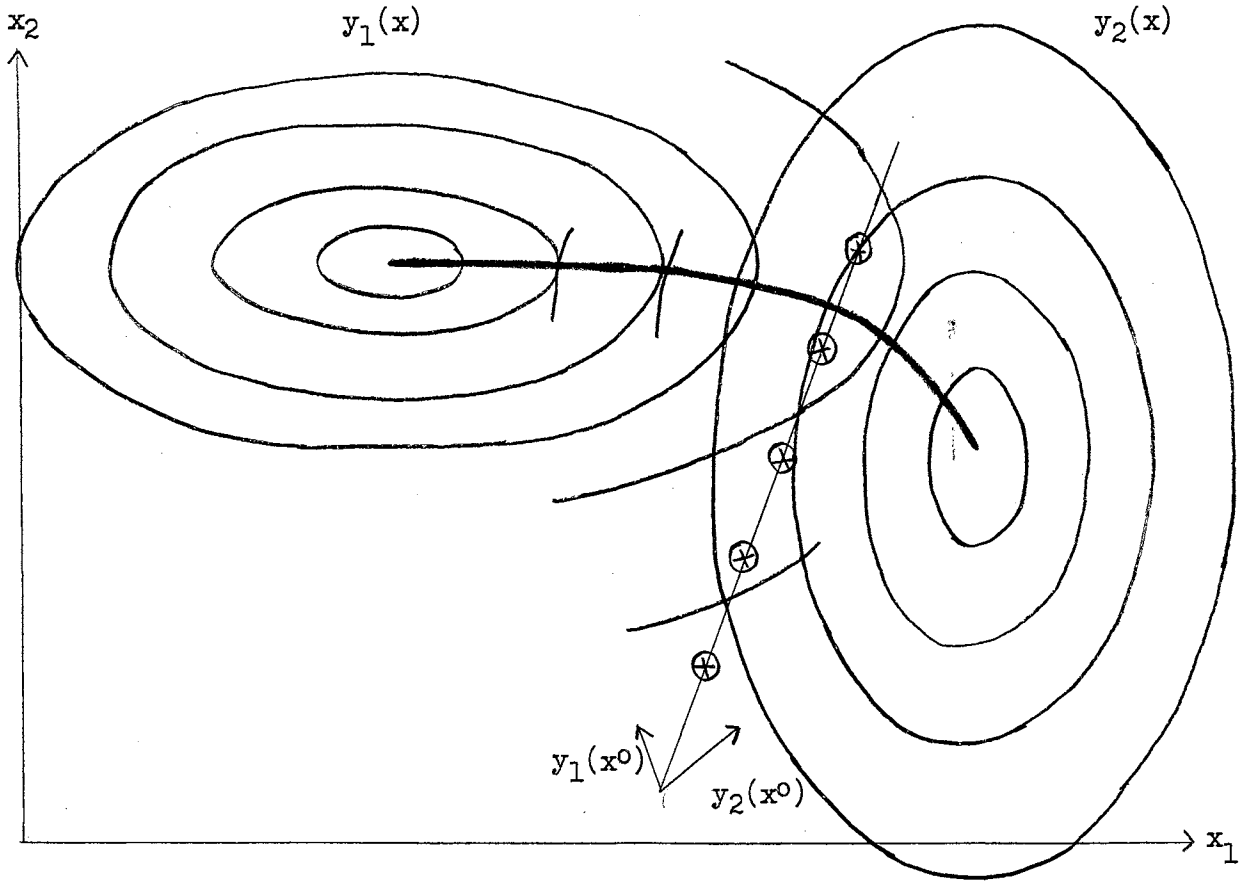

Figure 5. Method 2 for Seeking an Efficient Point

Method 3. Let

$$y(x) = \Sigma\, \beta_i y_i(x),\ \beta_i \geq 0,\ \Sigma\, \beta_i = 1$$

and find all x such that y(x) is maximized for some fixed β. If there are any such x, then there is at least one efficient point among them. This is true because if the complete set of efficient points exists and a convex combination of the $y_i(x)$ has a maximum value, then it occurs at an efficient point. Thus the method of

of steepest ascent, the method of parallel tangents, or any other method for experimental determination of maximum responses may be used to locate the corresponding efficient point.

In a sense this method is independent of the nature of the response surfaces. However, there exist response functions such that this method cannot lead to some of the efficient points. Such a set is given as an example on page 24 of Chapter II. If the response surfaces are all concave, then we do obtain all of the efficient points by this method.

The choice between Methods 2 and 3 depends largely upon the use anticipated for the efficient point. If we desire the efficient point corresponding to a particular β, then we should use Method 3. However, if convex combinations of the responses have no particular meaning, and it is the individual responses that are important, then we should use Method 2. For example, if the responses are the amounts of A, B, and C produced by a given process and the values of each per unit response are a, b, and c, then we would desire the efficient point that maximizes $\left[a + b + c\right]^{-1}\left[ay_1(x) + by_2(x) + cy_3(x)\right]$ and thus we would choose Method 3. On the other hand, if the responses are certain current measurements on a transistor, then we would not be interested in maximizing a convex combination of the responses and we would employ Method 2 in an attempt to reach a useful efficient point.

Once the experimenter has estimated an efficient point of interest, he should be interested in a joint confidence region for the responses at this point. This is the subject of the next chapter.

CHAPTER IV

JOINT CONFIDENCE REGIONS ON THE N RESPONSES

In this chapter we shall obtain joint confidence regions for the means of k future observations of the N responses at any given choice of the control variables. Some special uses for these are:

1. When we let $k = 1$, we obtain a joint tolerance region for the responses at any chosen x. This is a region which, on the average, contains $(1-\alpha)$ of the population of responses at the chosen x.

2. When we let $k \to \infty$ we obtain a joint confidence region on the expected values of the future observations at any point x. This will be the smallest of the regions, and may be all that is needed for a decision.

3. When we use some k such that $2 \leq k < \infty$, we obtain a joint confidence region on the means of the k future observations at the given value of x. If we multiply by k we have a joint confidence region on the sum of k future observations. Whether this is of interest depends upon the nature of the responses. If they are the amount of chemicals A, B, and C produced, then we may wish to have a joint confidence region on the totals of each produced in the next month. On the other hand, if they are the various responses of a transistor,

then their sums may be of no interest at all. In this event, we would be interested only in the joint tolerance region obtained by letting k = 1.

Three statistical models are considered for the responses. They differ only in the covariances of the errors. The models are of the form

41) $$y_i(x) = q'\beta_i + e_i,\ i = 1,2,\ldots,N$$

where

β_i is a vector of unknown parameters

e_i is a random variable with normal distribution about zero

42) $q' = (1, x_1, x_2,\ldots,x_p)$ if the model is a linear function of the control variables

43) $q' = (1, x_1, x_2,\ldots,x_p, x_1^2,\ldots,x_p^2, x_1x_2,\ldots,x_{p-1}x_p)$ if the model is a quadratic function of the control variables, etc.

While we shall be primarily concerned with q as given by one of these equations, the resulting confidence regions are not restricted to these forms for the model.

The covariances of the errors will be given in terms of the matrix models for the observations. The structure of the n values of each response observed in an experiment is given by

44) $$\begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_N \end{bmatrix} = \begin{bmatrix} X\beta_1 \\ X\beta_2 \\ \vdots \\ X\beta_N \end{bmatrix} + \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix} \quad \text{where} \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_N \end{bmatrix} \sim MVN(\varphi, \Sigma)$$

where

Y_i is an nxl vector of the observed values of the $i^{\underline{th}}$ response

X is the nxr matrix of known constants and rank (X) = r

β_i is an rxl vector of unknown parameters.

Model 1. The observed responses are given by equation 44 with

$$\Sigma = \begin{pmatrix} \sigma_{11}I & \varphi & \cdots & \varphi \\ & \sigma_{22}I & \cdots & \varphi \\ \text{Sym.} & & \cdots & \\ & & & \sigma_{NN}I \end{pmatrix}.$$

This model calls for complete independence of the responses. This is a very strong assumption and should not be made unless substantial information concerning the responses indicates that it is reasonable. However, when the model can be used, the joint confidence regions are especially easy to obtain. We could obtain confidence regions of size $(1-\alpha)^{\frac{1}{N}}$ for each of the N responses and the collection would be a joint confidence region of size $(1-\alpha)$ for the N responses. This approach would provide a joint confidence region for the means of k future observations of the form

$$\left\{(y_1, y_2, \ldots, y_N): y_i \geq C_i,\ i = 1,2,\ldots,N\right\}.$$

This form of a confidence region seems particularily desirable when it is important to obtain large values of all responses.

The above approach will also provide a bounded confidence region if desired. Such a region is certainly easy to use, but contains

more volume than an elliptical confidence region with the same confidence. If the σ_{ii} are known we have for a joint confidence region of the size $(1-\alpha)$ for the means of k future observations of the N responses.

$$45)\quad \left\{(y_1, y_2,\ldots,y_N): \sum_i^N \frac{(y_i - q^{o\prime}\hat{\beta}_i)^2}{\sigma_{ii}} \leq \chi^2_\alpha(N)\left[\frac{1}{k} + q^{o\prime}(X'X)^{-1}q^o\right]\right\}$$

where

$$\hat{\beta}_i = (X'X)^{-1}X'Y_i$$

and q^o is the vector given by equation 42 and 43 with the co-ordinates of x^o for the x_i, and χ^2_α is such that

$$\int_{\chi^2_\alpha}^{\infty} f(\chi^2;N)d\chi^2 = \alpha.$$

The problem of obtaining an elliptical confidence region when the σ_{ii} are not known is more difficult. Such regions are obtained for Model 2 which follows.

Model 2. The observed responses are given by equation 44 with

$$\Sigma = \begin{bmatrix} \sigma_{11}I & \sigma_{12}I\ldots\sigma_{1N}I \\ & \sigma_{22}I\ldots\sigma_{2N}I \\ \text{Sym.} & \ldots \\ & \sigma_{NN}I \end{bmatrix}.$$

In general it seems that Model 2 should be the most useful. Note that the N responses for a given trial may be correlated, but from trial to trial the responses are independent. With this model the maximum likelihood estimates of the parameters are

46) $$\hat{\beta}_i = (X'X)^{-1}X'Y_i, \quad i = 1, 2, \ldots, N$$

47) $$\hat{\sigma}_{ij} = \frac{1}{n} Y'_i \left[I - X(X'X)^{-1}X'\right] Y_j, \quad i, j = 1, 2, \ldots, N.$$

If q corresponds to x as in equation 42 or 43 and $\bar{y}_i(x)$ is the mean of k future observations of the $i^{\underline{th}}$ response variable, then

$$\begin{bmatrix} \bar{y}_1(x) - q'\hat{\beta}_1 \\ \bar{y}_2(x) - q'\hat{\beta}_2 \\ \vdots \\ \bar{y}_N(x) - q'\hat{\beta}_N \end{bmatrix} \sim MVN(\varphi, V)$$

where

$$V = \left[q'(X'X)^{-1}q + \frac{1}{k}\right] \begin{bmatrix} \sigma_{11} & \sigma_{12} \cdots \sigma_{1N} \\ & \sigma_{22} \cdots \sigma_{2N} \\ \text{Sym.} & \cdots \\ & \sigma_{NN} \end{bmatrix}.$$

Thus

$$\begin{bmatrix} \bar{y}_1(x) - q'\hat{\beta}_1 \\ \bar{y}_2(x) - q'\hat{\beta}_2 \\ \vdots \\ \bar{y}_N(x) - q'\hat{\beta}_N \end{bmatrix}' \frac{\begin{bmatrix} \sigma_{11} & \sigma_{12} \cdots \sigma_{1N} \\ & \sigma_{22} \cdots \sigma_{2N} \\ \text{Sym.} & \cdots \\ & \sigma_{NN} \end{bmatrix}^{-1}}{q'(X'X)^{-1}q + k^{-1}} \begin{bmatrix} \bar{y}_1(x) - q'\hat{\beta}_1 \\ \bar{y}_2(x) - q'\hat{\beta}_2 \\ \vdots \\ \bar{y}_N(x) - q'\hat{\beta}_N \end{bmatrix}$$

is distributed as $x^2(N)$, and it follows that when all σ_{ij} are known the desired joint confidence region on the means of k future observations is given by

$$48)\quad \begin{bmatrix} \bar{y}_1(x) - q'\hat{\beta}_1 \\ \bar{y}_2(x) - q'\hat{\beta}_2 \\ \vdots \\ \bar{y}_N(x) - q'\hat{\beta}_N \end{bmatrix}' \frac{\begin{bmatrix} \sigma_{11}\ \sigma_{12}\cdots\sigma_{1N} \\ \sigma_{22}\cdots\sigma_{2N} \\ \text{Sym.}\quad \cdots \\ \sigma_{NN} \end{bmatrix}^{-1}}{q'(X'X)^{-1}q+k^{-1}} \begin{bmatrix} \bar{y}_1(x) - q'\hat{\beta}_1 \\ \bar{y}_2(x) - q'\hat{\beta}_2 \\ \vdots \\ \bar{y}_N(x) - q'\hat{\beta}_N \end{bmatrix} \leq x^2_\alpha(N).$$

If the σ_{ij} are not known, the above suggests that the distribution of

$$49)\quad U = \begin{bmatrix} \bar{y}_1(x) - q'\hat{\beta}_1 \\ \bar{y}_2(x) - q'\hat{\beta}_2 \\ \vdots \\ \bar{y}_N(x) - q'\hat{\beta}_N \end{bmatrix}' \frac{\begin{bmatrix} \hat{\sigma}_{11}\ \hat{\sigma}_{12}\cdots\hat{\sigma}_{1N} \\ \hat{\sigma}_{22}\cdots\hat{\sigma}_{2N} \\ \text{Sym.}\quad \cdots \\ \hat{\sigma}_{NN} \end{bmatrix}^{-1}}{q'(X'X)^{-1}q+k^{-1}} \begin{bmatrix} \bar{y}_1(x) - q'\hat{\beta}_1 \\ \bar{y}_2(x) - q'\hat{\beta}_2 \\ \vdots \\ \bar{y}_N(x) - q'\hat{\beta}_N \end{bmatrix}$$

be obtained. For convenience let

$$50)\quad S = \begin{bmatrix} \hat{\sigma}_{11}\ \hat{\sigma}_{12}\cdots\hat{\sigma}_{1N} \\ \hat{\sigma}_{22}\cdots\hat{\sigma}_{2N} \\ \text{Sym.}\quad \cdots \\ \hat{\sigma}_{NN} \end{bmatrix}.$$

Then

$$S = \frac{1}{n}\begin{bmatrix} e_1'\left[I-X(X'X)^{-1}X'\right]e_1 & e_1'\left[I-X(X'X)^{-1}X'\right]e_2 \cdots e_1'\left[I-X(X'X)^{-1}X'\right]e_N \\ & e_2'\left[I-X(X'X)^{-1}X'\right]e_2 \cdots e_2'\left[I-X(X'X)^{-1}X'\right]e_N \\ \text{Sym.} & \cdots \\ & e_N'\left[I-X(X'X)^{-1}X'\right]e_N \end{bmatrix}$$

which may be written as

$$S = \frac{1}{n}\begin{bmatrix} e_1' \\ e_2' \\ \cdot \\ \cdot \\ \cdot \\ e_N' \end{bmatrix}\left[I-X(X'X)^{-1}X'\right]\left[e_1, e_2, \ldots, e_N\right].$$

Let

51) $$z_i = P'e_i$$

where P is an orthogonal matrix such that

$$P'\left[I-X(X'X)^{-1}X'\right]P = \begin{bmatrix} I_{rxr} & \varphi \\ \varphi & \varphi \end{bmatrix} \qquad \text{rank } (X) = r,$$

Then

$$S = \frac{1}{n}\begin{bmatrix} z_1' \\ z_2' \\ \cdot \\ \cdot \\ \cdot \\ z_N' \end{bmatrix}\begin{bmatrix} I_{rxr} & \varphi \\ \varphi & \varphi \end{bmatrix}\left[z_1, z_2, \ldots, z_N\right]$$

52) $$S = \frac{1}{n}\begin{bmatrix} \sum_{n-r} z_{1i}^2 & \sum_{n-r} z_{1i}z_{2i} & \cdots & \sum_{n-r} z_{1i}z_{Ni} \\ & \sum_{n-r} z_{2i}^2 & \cdots & \sum_{n-r} z_{2i}z_{Ni} \\ \text{Sym.} & & \cdots & \\ & & & \sum_{n-r} z_{Ni}^2 \end{bmatrix}.$$

Let

53) $$Z_i = \begin{bmatrix} z_{1i} \\ z_{2i} \\ \cdot \\ \cdot \\ \cdot \\ z_{Ni} \end{bmatrix} = \sum_{j=1}^{n} p'_{ij} \begin{bmatrix} e_{1j} \\ e_{2j} \\ \cdot \\ \cdot \\ \cdot \\ e_{Nj} \end{bmatrix}$$

so that

$$S = \frac{1}{n} \sum_{n-r} Z_i Z_i'$$

and from page 51 of (3) it follows that

54) $$Z_i \sim \text{NID} \left\{ \varphi, \begin{bmatrix} \sigma_{11} & \sigma_{12} & \cdots & \sigma_{1N} \\ & \sigma_{22} & \cdots & \sigma_{2N} \\ \text{Sym.} & & \cdots & \\ & & & \sigma_{NN} \end{bmatrix} \right\}.$$

Thus

55) $$(n-r)\left[\frac{nS}{n-r}\right] \sim \sum_{i=1}^{n-r} Z_i Z_i'$$

and from Theorem 5.22 of (3) it follows that

56) $$\frac{n-r}{n} U \sim T^2(n-r)$$

and

57) $$\frac{U}{n} \frac{(n-r-N+1)}{N} \sim F(N, n-r-N+1).$$

Therefore a confidence region of size $1-\alpha$ on the means of k future observations is given by

58) $$U \leqslant \frac{Nn}{n-r-N+1} F_{\alpha}(N, n-r-N+1)$$

with U given by equation 49.

Model 3. The observed responses are given by equation 44 with

$$\Sigma = V\sigma^2$$

where V is the known matrix

$$V = \begin{bmatrix} v_{11}I & v_{12}I \ldots v_{1N}I \\ & v_{22}I \ldots v_{2N}I \\ \text{Sym.} & \ldots \\ & v_{NN}I \end{bmatrix},$$

$\{v_{ij}\}$ is a non-singular matrix, and I is an nxn matrix. Note that the structure of Σ is the same as for Model 2. However, it is convenient to consider this as a distinct model.

The maximum likelihood estimates of β_i and σ^2 are

59) $$\hat{\beta}_i = (X'X)^{-1}X'Y_i$$

$$60)\quad \hat{\sigma}^2 = \frac{1}{Nn} Y' \begin{bmatrix} v^{11}\left[I-X(X'X)^{-1}X'\right] & \cdots & v^{1N}\left[I-X(X'X)^{-1}X'\right] \\ \text{Sym.} & \cdots & \\ & & v^{NN}\left[I-X(X'X)^{-1}X'\right] \end{bmatrix} Y$$

where

$$\left\{v^{ij}\right\} = \left\{v_{ij}\right\}^{-1}.$$

If we let

$$\beta' = (\beta_1', \beta_2', \ldots, \beta_N') \text{ and } \hat{\beta}' = (\hat{\beta}_1', \hat{\beta}_2', \ldots, \hat{\beta}_N')$$

then we find

$$61)\qquad \hat{\beta} \sim \text{MVN}\left[\beta,\ (X'X)^{-1}\left\{v_{ij}\right\}\sigma^2\right]$$

and

$$62)\qquad \frac{Nn\hat{\sigma}^2}{\sigma^2} \sim \chi^2\left[N(n-r)\right],\ r = \text{rank}(X).$$

Furthermore $\hat{\beta}$ and $\hat{\sigma}^2$ are independent.

Let the mean of k future observations at x be given by the vector $\bar{y}(x)$ where

$$\bar{y}(x) = \begin{bmatrix} \bar{y}_1(x) \\ \bar{y}_2(x) \\ \vdots \\ \bar{y}_N(x) \end{bmatrix}.$$

Then according to the model

$$63)\qquad \bar{y}(x) \sim \text{MVN}\left[u'\beta,\ k^{-1}\sigma^2\left\{v_{ij}\right\}\right]$$

where

$$u = \begin{bmatrix} q\ \varphi \ldots \varphi \\ \varphi\ q \ldots \varphi \\ \cdots \\ \varphi\ \varphi \ldots q \end{bmatrix}$$

and q is the vector corresponding to x and given by equation 42 or 43. Since $\bar{y}(x)$, $\hat{\beta}$ and $\hat{\sigma}^2$ are independent, it follows that

64) $$\left[\bar{y}(x)-u'\hat{\beta}\right] \sim \text{MVN}\left\{\varphi, \sigma^2\left[k^{-1}+q'(X'X)^{-1}q\right]\left\{v_{ij}\right\}\right\}$$

and hence

65) $$\frac{\left[\bar{y}(x)-u'\hat{\beta}\right]'\left\{v^{ij}\right\}\left[\bar{y}(x)-u'\hat{\beta}\right]}{\sigma^2\left[k^{-1}+q'(X'X)^{-1}q\right]} \sim \chi^2(N).$$

Thus a joint confidence region of size $1-\alpha$ on the means of k future responses at x is given by

66) $$\left[\bar{y}(x)-u'\hat{\beta}\right]'\left\{v^{ij}\right\}\left[\bar{y}(x)-u'\hat{\beta}\right] \leq \left[k^{-1}+q'(X'X)^{-1}q\right]\frac{Nn\hat{\sigma}^2 F_{\alpha}}{n-r}\left[N, N(n-r)\right]$$

CHAPTER V

SUMMARY

In considering the N response problem in which it is desirable to have all responses as large as possible, it was first necessary to recognize that we probably cannot simultaneously maximize all N responses. This led to the definition of an efficient point as any x whose responses are not dominated by those for some other x. The set of all such x usually constitutes the complete set of efficient points. It was seen that this set may be a very small subset of the set of all possible x. This is an important property of the complete set of efficient points.

Means for obtaining the efficient points from known response functions were presented in Chapter II. Since quadratic response functions are frequently used for models, these were given special consideration. It was seen that the complete set of efficient points is readily obtained when the responses are quadratic functions, especially when they are positive definite. In this case a formula is obtained which provides the complete set of efficient points. The set is indexed by a vector α which contains N-1 independent components. When the response functions are restricted to the domain of the efficient points, they also become functions of the vector α. Since the efficient points are the only ones which should

be considered, no loss results from this restriction. When the dimension of x is larger than N-1, this permits us to consider the smaller problem of selecting the N-1 components of α instead of the p components of x.

The problem of locating the efficient points for unknown response functions by experimental means was considered in Chapter III. Sequential and non-sequential methods were presented. The choice of method was seen to depend upon the structure assumed for the response functions and the nature of the responses. In general it seems that Method 2 will be the better choice of the sequential methods as it assumes less regarding the structure of the responses and allows the experimenter considerable freedom in the selection of a particular efficient point.

Joint confidence regions for the future responses at a particular value of x (not necessarily an efficient point) are obtained in Chapter IV. Three statistical models for the responses are considered. The most general of these is Model 2, and it would seem to be the most useful for that reason. These regions provide, among other things, a β expectation joint tolerance region for the responses at a given value of x.

Areas for Future Research

The lack of fit of the quadratic statistical models may be a serious problem when the regions at which the N responses attain their maxima are widely separated. In this event it may be desirable

to use cubic or quartic response models. Although the results already obtained will apply to these models also, the work involved in applying them may be prohibitive. Other possible models should be investigated.

There are situations in which it may be desirable to choose two or more of the efficient points. For example, if the responses are simply the amounts of A and B produced and the responses are as given in Figure 6 , then we may wish to choose $x^{(1)}$ and $x^{(2)}$ equally often in order to produce the desired amounts of A and B most efficiently. In other words, the problem of selecting one or more points from the complete set of efficient points should receive further attention.

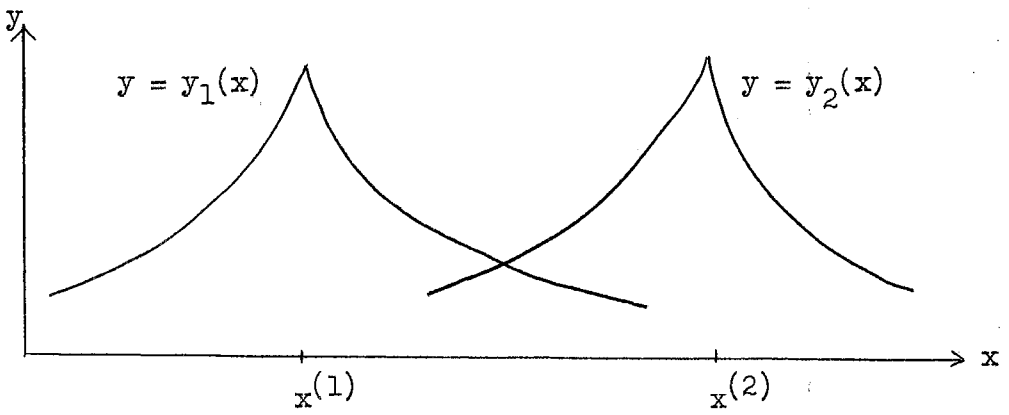

Figure 6. Amounts of A and B Produced

The joint tolerance regions developed were of the β expectation type. It would also be desirable to have joint γ probability of β content tolerance regions.

We would like to have a confidence region on the complete set of efficient points. If this is not possible, then we would like to have a confidence region for the particular x that maximizes a convex combination of the responses. By considering this convex combination as a function of x, we see that when the response functions are quadratic the problem is the same as the one considered by Box and Hunter (4) in which they attempt to obtain a confidence region for the x that maximizes a quadratic function of x. However, the confidence region they derive is really a confidence region on the expected value of the estimated x. Since the estimated x for maximum response is a biased estimate of the x that maximizes the desired function, it follows that the confidence regions are not really confidence regions on the x that maximizes the function. The bias in the estimated point for maximum response is a problem worthy of investigation.

There are situations in which it would be helpful to know if the complete set of efficient points is connected. No general results have been obtained for this problem.

SELECTED BIBLIOGRAPHY

(1) Box, G. E. P., and K. B. Wilson. "On the Experimental Attainment of Optimum Conditions." Journal of the Royal Statistical Society. Vol. 13, 1-45, 1951.

(2) Kempthorne, O., B. V. Shah, and R. J. Buehler. "The Method of Parallel Tangents (PARTAN) for finding an Optimum." The Annals of Mathematical Statistics. Vol. 32, 632, 1961.

(3) Anderson, T. W. Introduction to Multivariate Statistical Analysis. New York: John Wiley and Sons, 1958.

(4) Box, G. E. P., and J. S. Hunter. "A Confidence Region for the Solution of a Set of Simultaneous Equations with an Application to Experimental Design." Biometrika. Vol. 41, 190-199, 1954.

(5) Graybill, Franklin A. Introduction to Linear Statistical Models, Vol. 1. New York: McGraw-Hill Publishing Co., 1961.

APPENDIX

AN EXAMPLE

The following problem will illustrate the ideas presented in this thesis. Suppose there are two responses of interest denoted by $y_1(x)$ and $y_2(x)$. Let the model for these responses be

67) $$y_1(x) = a_1 - x'A_1x + \beta'_1x + e_1$$

68) $$y_2(x) = \exp(a_2 - x'A_2x + \beta'_2x + e_2)$$

where

$$\begin{bmatrix} e_1 \\ e_2 \end{bmatrix} \sim \text{NID} \left\{ \varphi, \begin{bmatrix} \sigma_{11} & \sigma_{12} \\ \sigma_{12} & \sigma_{22} \end{bmatrix} \right\}$$

when the associated $y_1(x)$ and $y_2(x)$ are observed simultaneously. If we let

$$z_2 = \ln y_2$$

it follows that

69) $$z_2(x) = a_2 - x'A_2x + \beta'_2x + e_2.$$

Now let us suppose the results of a 3^3 experiment are as given in Table 1. We shall use these experimental results to

1. Estimate the parameters in the model,

2. Estimate the efficient points for the responses $y_1(x)$ and $z_2(x)$,
3. Estimate the future responses for any given x, and
4. Provide a $(1-\alpha)$ joint confidence region on the means of k future observations at any given x.

TABLE I

THE RESPONSES FOR A 3^3 EXPERIMENT

x_1	x_2	x_3	$y_1(x)$	$z_2(x)$
1	1	1	64.84	4.62
1	1	0	71.09	5.14
1	1	-1	66.65	5.61
1	0	1	78.88	4.77
1	0	0	81.54	5.74
1	0	-1	78.25	5.93
1	-1	1	78.91	4.42
1	-1	0	87.70	5.28
1	-1	-1	82.82	5.56
0	1	1	75.82	3.85
0	1	0	81.55	4.70
0	1	-1	73.04	4.96
0	0	1	87.64	4.20
0	0	0	92.93	5.02
0	0	-1	86.77	5.32
0	-1	1	91.13	3.79
0	-1	0	98.63	4.72
0	-1	-1	91.23	4.86

TALBE I (Continued)

x_1	x_2	x_3	$y_1(x)$	$z_2(x)$
-1	1	1	78.05	2.15
-1	1	0	84.65	3.21
-1	1	-1	77.25	3.13
-1	0	1	89.77	2.49
-1	0	0	96.20	3.32
-1	0	-1	90.03	3.49
-1	-1	1	92.16	1.95
-1	-1	0	100.30	2.79
-1	-1	-1	93.78	3.43

Since the complete set of efficient points for $y_1(x)$ and $y_2(x)$ is also the complete set of efficient points for $y_1(x)$ and $z_2(x)$, we shall consider the estimate obtained for $y_1(x)$ and $z_2(x)$ as our estimate for the complete set of efficient points for $y_1(x)$ and $y_2(x)$.

The least squares estimates of the a_i, A_i, and β_i obtained in the usual manner are

$$\hat{a}_1 = 93.44 \qquad \hat{a}_2 = 5.04$$

$$\hat{A}_1 = \begin{bmatrix} 3.59 & .025 & .16 \\ & 4.13 & -.31 \\ \text{Sym.} & & 6.23 \end{bmatrix} \qquad \hat{A}_2 = \begin{bmatrix} .55 & .01 & .00 \\ & .36 & -.02 \\ \text{Sym.} & & .30 \end{bmatrix}$$

$$\hat{\beta}_1 = \begin{bmatrix} -6.2 \\ -7.98 \\ -.15 \end{bmatrix} \qquad \hat{\beta}_2 = \begin{bmatrix} 1.17 \\ .03 \\ -.56 \end{bmatrix}$$

Since $\hat{A}_1$ and $\hat{A}_2$ are positive definite, the complete set of efficient points for $\hat{y}_1(x)$ and $\hat{z}_2(x)$ is given by

$$\left\{x \mid x = .5\left[A_1 + (1-\alpha)A_2\right]^{-1}\left[\beta_1 + (1-\alpha)\beta_2\right] ;\ 0 \le \alpha \le 1\right\}.$$

When the necessary matrix operations have been carried out the resulting parametric equations of the complete set of efficient points are

$$x_1 = \frac{-155.948\alpha^3 + 2.908\alpha^2 + 3.07\alpha + .1259}{67.7\alpha^3 + 22.14\alpha^2 + 2.12\alpha + .06}$$

$$x_2 = \frac{-136.062\alpha^3 - 34.022\alpha^2 - 1.375\alpha + .005}{67.7\alpha^3 + 22.14\alpha^2 + 2.12 + .06}$$

$$x_3 = \frac{2.1\alpha^3 - 7.146\alpha^2 - 1.844\alpha - .11}{67.7\alpha^3 + 22.14\alpha^2 + 2.12\alpha + .06}$$

where

$$0 \le \alpha \le 1.$$

The most convenient presentation of the predicted responses is the one in which $\hat{y}_2(x)$ is plotted as a function of $\hat{y}_1(x)$ as in Figure 7. This is possible when we restrict x to the complete set of efficient points.

The graph of the predicted responses at the efficient points illustrates the situation when one attempts to simultaneously maximize two or more responses and the need for a compromise. Now suppose

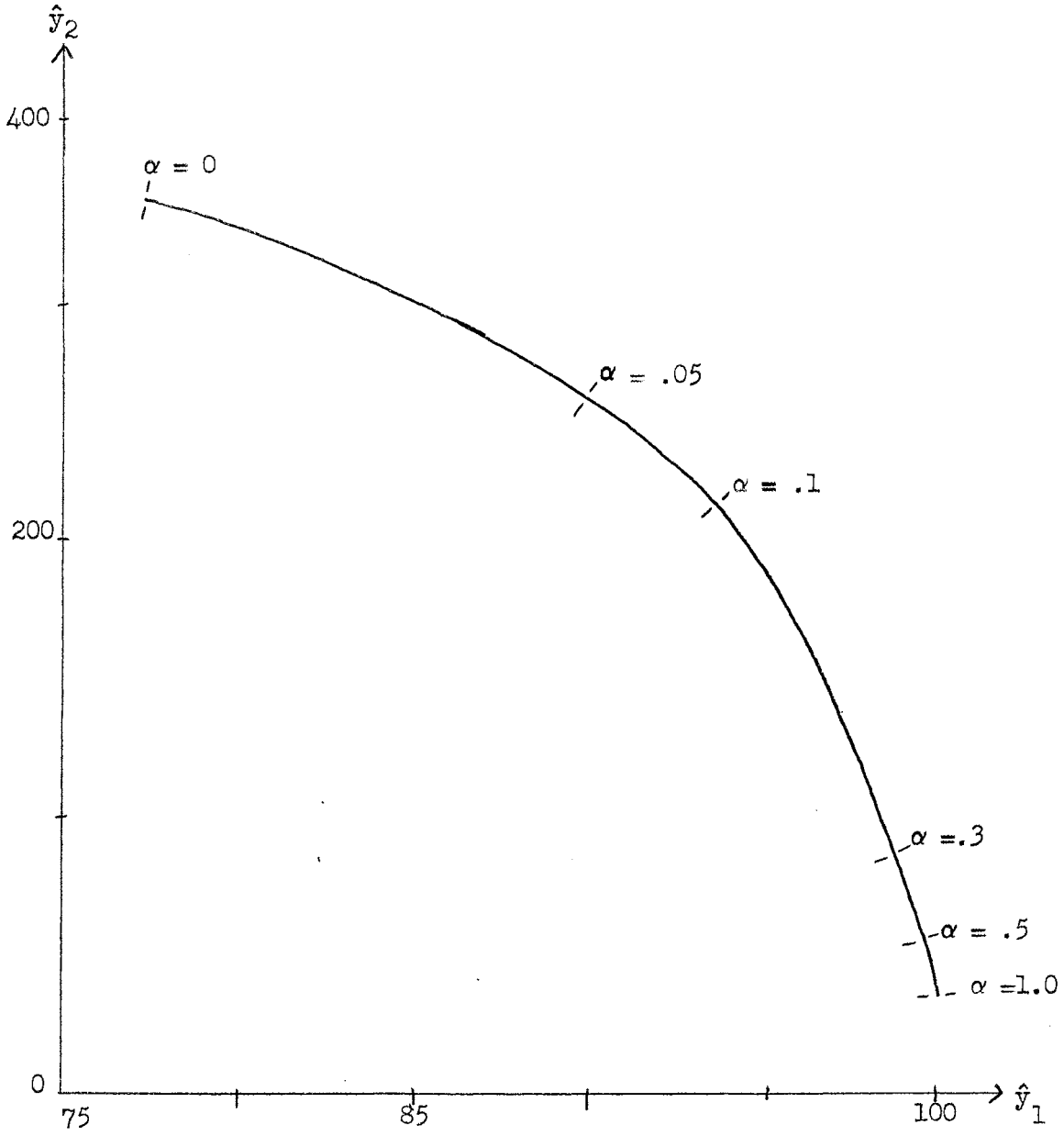

Figure 7. The Predicted Responses at the Efficient Points

the predicted responses at the efficient point for which $\alpha = .1$ are considered the most desirable. Then let us obtain a joint tolerance region on the future responses at the x associated with $\alpha = .1$.

It is seen that the responses $y_1(x)$ and $z_2(x)$ observed in the experiment satisfy the conditions for Model 2 as defined in Chapter IV. The quantities used with that model in calculating the joint tolerance region for $y_1(x)$ and $z_2(x)$ are $\hat{\sigma}_{11}$, $\hat{\sigma}_{12}$, $\hat{\sigma}_{22}$, $q'(X'X)^{-1}q$, $\hat{y}_1(x^o)$, and $\hat{z}_2(x^o)$.

The experimental data for the problem considered yield the following values:

$$\hat{\sigma}_{11} = \frac{1}{n} Y_1' \left[I - X(X'X)^{-1}X'\right] Y_1 = .808$$

$$\hat{\sigma}_{12} = \frac{1}{n} Y_1' \left[I - X(X'X)^{-1}X'\right] Z_2 = -.00263$$

$$\hat{\sigma}_{22} = \frac{1}{n} Z_2' \left[I - X(X'X)^{-1}X'\right] Z_2 = .00665$$

$$q' (X'X)^{-1} q = .297$$

$$\hat{y}_1(x^o) = 94.2$$

$$\hat{z}_2(x^o) = 5.35.$$

When Model 2 applies, the joint confidence region of size $(1-\alpha)$ for the means of k future responses at x^o is given by

70)
$$\begin{bmatrix} \bar{y}_1(x^o) - \hat{y}_1(x^o) \\ \bar{z}_2(x^o) - \hat{z}_2(x^o) \end{bmatrix}' \begin{bmatrix} \hat{\sigma}_{11} & \hat{\sigma}_{12} \\ \hat{\sigma}_{12} & \hat{\sigma}_{22} \end{bmatrix}^{-1} \begin{bmatrix} \bar{y}_1(x^o) - \hat{y}_1(x^o) \\ \bar{z}_2(x^o) - \hat{z}_2(x^o) \end{bmatrix}$$

$$\leq \frac{n - r - N + 1}{Nn} F_\alpha(N, n - r - N + 1).$$

In this problem N = 2, n = 27, r = 10, and we shall take $\alpha = .10$ and k = 1. Then the confidence region is a tolerance region for the future responses. The region is given by

$$\begin{bmatrix} y_1(x^0) - 94.2 \\ z_2(x^0) - 5.35 \end{bmatrix}' \begin{bmatrix} 1.239 & .49 \\ .49 & 150.57 \end{bmatrix} \begin{bmatrix} y_1(x^0) - 94.2 \\ z_2(x^0) - 5.35 \end{bmatrix} \leq 1.026,$$

This region is sketched in Figure 8.

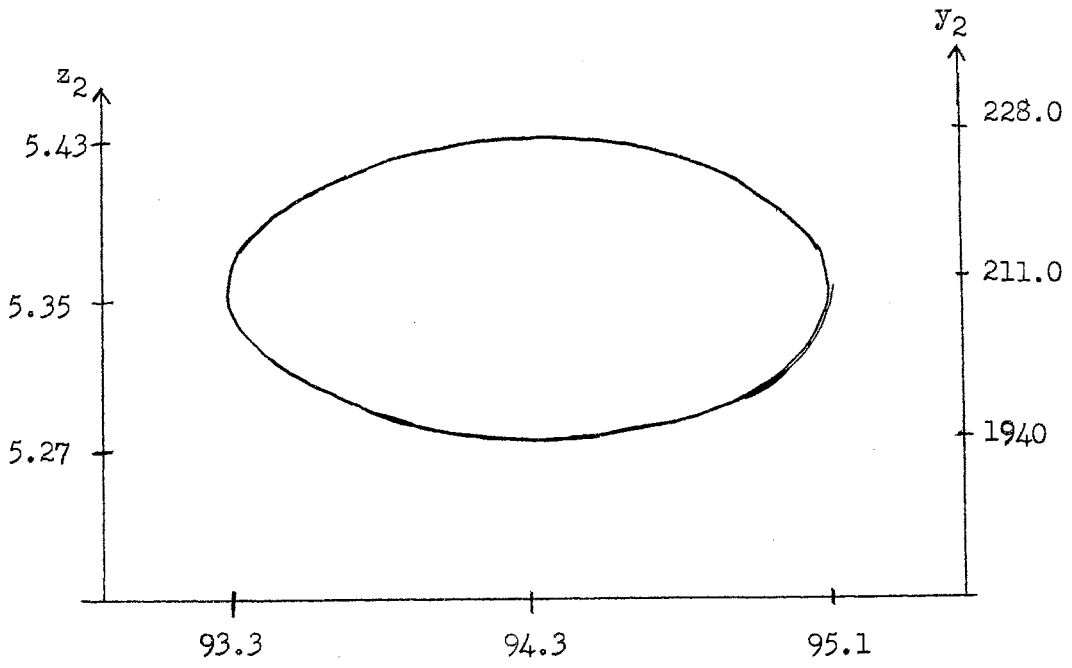

Figure 8. A β Expectation Tolerance Region for the Future Responses

The interpretation of this region as a tolerance region is that, on the average, a region obtained in this manner will contain .9 or more of the population of responses at the x associated with the efficient point for which $\alpha = .1$.

VITA

Charles Edward Antle

Candidate for the Degree of

Doctor of Philosophy

Thesis: THE N RESPONSE PROBLEM

Major Field: Mathematics

Biographical:

Personal Data: Born at East View, Kentucky, November 11, 1930, the son of Bayard and Mary Antle.

Education: Attended grade schools in Louisville, Mell, and Columbia, Kentucky; graduated from Columbia High School in 1948; received the Associate of Arts degree from Lindsey Wilson Junior College in June, 1950; received the Bachelor of Science degree from Eastern Kentucky State College, with majors in Mathematics and Physics, in June, 1954; received the Master of Arts degree from Eastern Kentucky State College in August, 1955; completed the requirements for the Doctor of Philosophy degree in August, 1962.

Professional Experience: Entered the United States Army in January, 1951; worked as an Aerophysics Engineer at Convair 1955-57; taught in the Evening College at Texas Christian University 1956-57; taught as an Instructor of Mathematics at the Missouri School of Mines and Metallurgy 1957-59; worked as a Senior Aerophysics Engineer at Convair during the summer of 1958; taught as an Assistant Professor of Mathematics at the Missouri School of Mines and Metallurgy 1959-60; worked as a Physicist with the U. S. Naval Weapons Evaluation Facilities in Albuquerque during the summer of 1961.