A COMPARISON OF TWO ALTERNATIVE UNCONSTRAINED NON-LINEAR OPTIMIZATION TECHNIQUES

John Anthony Murray

NAVAL POSTGRADUATE SCHOOL
Monterey, California

THESIS

A Comparison of Two
Alternative Unconstrained Non-Linear
Optimization Techniques

by

John Anthony Murray

Thesis Advisor: James K. Hartman

March 1972

Approved for public release; distribution unlimited.

A Comparison of Two Alternative Unconstrained Non-Linear Optimization Techniques

by

John Anthony Murray
Major, United States Marine Corps
B.S., Marquette University, 1962

Submitted in partial fulfillment of the requirements for the degree of

MASTER OF SCIENCE IN OPERATIONS RESEARCH

from the

NAVAL POSTGRADUATE SCHOOL
March 1972

ABSTRACT

Two alternative methods for optimizing an unconstrained non-linear function are investigated and compared. The investigations are made subject to a restriction as to the number of function evaluations available to conduct the optimization procedures. Powell's method of conjugate directions is employed as the direct search method and is considered the reference method. The alternate method is based on fitting a quadratic surface to the available function evaluations and optimizing over the resulting fitted surface. The test functions considered in the investigation were limited to unimodal functions.

TABLE OF CONTENTS

LIST OF TABLES

I. NATURE OF THE PROBLEM

A. INTRODUCTION

In Operations Research a major problem concerns the determination of the optimum response for an objective function. In many cases, the objective function may not lend itself to expression in closed analytical form. Information concerning the nature of the objective function may only be available through computer simulation, field testing or a combination of both. Furthermore, restrictions may exist with respect to the quantity of information available or obtainable due to the costs involved in securing the data. Thus, it is incumbent that the most efficient use be made of the limited information available in order to achieve the best possible results in determining the optimum response.

There exist several search techniques that could be adapted to solve the problem as stated above. Generally, these search techniques could be classified as direct search techniques or gradient search techniques. Direct search methods do not require the computation of partial derivatives of the objective function. Determination of the optimum response is based solely on values of the objective function itself. On the other hand, gradient search methods compute values of partial derivatives of the objective function. These resulting values are then used in selecting future search directions. Search techniques employing gradient methods require more information (i.e., additional function evaluations, computer simulations, etc.) than direct search procedures. Thus, with the presence of restrictions on the quantity of information available it would be more advantageous to use a direct search method.

Direct Search methods can be broadly divided into three classifications: tabulation methods, sequential methods and linear methods. Tabulation methods, as the name implies, result in a tabular listing of objective function responses for different values of the input variables. The "optimum" response is considered to be the "best" value found in the table. In the case of minimization problems, the "best" value would be the smallest value. An example of a tabulation method is random search.

Sequential methods employ the use of geometric designs in the input variable space. The objective function is evaluated at the vertices of the geometric design. The sequence is repeated until the desired accuracy is achieved or the restriction of the quantity of information obtainable precludes additional testing. Factorial designs and the simplex method are examples of sequential methods.

Linear methods employ the use of a set of direction vectors throughout the conduct of the search. The objective function is evaluated at different points along these search directions and conduct of the search is dependent upon the results obtained. The set of direction vectors may or may not be changed during the course of the search. Examples of linear methods are the alternating variable search and Powell's method of conjugate directions.

An alternative approach to employing search techniques to solve the problem would be to use curve-fitting techniques. Using this technique, test points would be established in the input variable space and the objective function would be evaluated at these test points. A least-square polynomial surface would then be fitted to these resulting objective function values. To complete the procedure the location of the "optimum" response of the fitted surface would be determined and the

resulting input variable values would be input to the true objective function to determine the "optimum" response.

B. STATEMENT OF THE PROBLEM

The specific problem that was the subject of investigation was to compare the relative efficiency of a curve-fitting technique as opposed to a search technique. The assumption was made that a restriction exists on the quantity of information available (i.e., there were a finite number of functional evaluation that could have been made). A direct result of the above assumption was that the investigation was limited to those techniques which do not require the computation of partial derivatives. Furthermore, the investigation was restricted to test objective functions of a unimodal nature insuring the existence of only one true optimum response.

II. EXPERIMENTAL PROCEDURES

A. SEARCH PROCEDURE

Powell's method of conjugate directions was chosen as the direct search technique to be used. Powell's method is considered to be one of the most efficient direct search techniques available [Ref.1]. Its efficiency is exceptionally good for functions that can be approximated by a quadratic in the region of the optimum.

Powell's method commences with a search along a set of linearly independent directions that span the input variable space starting from an arbitrarily chosen initial point. The initial set of directions chosen are the co-ordinate directions. At the completion of each iteration a new direction is defined and replaces one of the presently existing directions only if the new set of directions is at least as efficient as the present set. If this is not the case, an additional iteration is conducted retaining the present set of vectors. A complete description of Powell's method and a detailed discussion of its efficiency and convergence is contained in [Ref.2].

An existing version of Powell's method programmed in FORTRAN for use on the IBM 360 Computer was used as a subroutine in the overall computer program written to conduct the search technique optimization. The main program provided the calling argument required by the subroutine to conduct the optimization.

B. CURVE-FITTING TECHNIQUE

The curve-fitting technique used was based on standard multiple linear regression techniques [Refs. 3 and 4] modified as follows.

Because the test functions to be investigated were unimodal the decision was made to fit a quadratic surface to the experimental data. Thus, the order of the regression was double the number of input variables.

An early decision required to be made concerned exactly what procedure would be used to place the test points in the input variable space. Two possibilities were considered. A completely random selection of the test points was one alternative. A second approach was to place the test points throughout the space in accordance with an experimental design. The first experimental design considered was a composite factorial design [Ref. 5]. This approach was attractive for objective functions of few input variables. However, the requirement for an excessive number of test points and consequently an increased number of function evaluations precluded the use of this design for objective functions of more than four variables.

For functions of more than four variables it was decided to use an orthogonal design requiring 2n+1 test points, where n is the number of input variables. The initial test point was placed in the center of the input variable space and the remaining points were placed at a fixed distance (0.5) in both the positive and negative direction along the co-ordinate axes passing through the center of the design. In the investigations a third alternative was also considered. For those cases where the number of observations (function evaluations) available exceeded the number required for the experimental design, the remaining points were located randomly in the input variable space.

Once the test points were placed, the function was evaluated at these points. The value of input variables, the squares of the input variables and the value of the objective function for each test point were then

input to the regression subroutine to determine the regression coefficients of the fitted surface. The resulting fitted surface was a quadratic of the form:

$$f = \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n} b_i x_i^2 + \text{constant term}$$

Provided the quadratic was positive (negative) definite an explicit determination of the minimum (maximum) value of the quadratic could be computed. In that case, the minimum (maximum) value of the quadratic occurs at the point:

$$x_i = -\frac{1}{2} \frac{a_i}{b_i} \; ; \quad i = 1, \ldots , n$$

These values of x_i were computed and substituted into the true objective function to determine the "optimal" objective function response.

Preliminary results indicated that the curve-fitting technique was generating a large number of indefinite quadratic surfaces. This situation occurred more commonly when random choice was used to locate the test points. Any indefinite quadratics generated were not useful because it was not possible to find the location of the minimum response of the fitted surface. Thus, an "optimum" response could not be determined in those situations where the curve-fitting technique generated an indefinite quadratic.

In an effort to overcome this problem, a procedure was introduced into the curve-fitting technique to reduce the number of observations considered by the regression procedure. The basic idea behind the procedure was for the regression procedure to restrict its consideration to those test points that were close to the optimum. In that manner points at a distance from the optimum did not affect the fit. It was

felt that this procedure would insure a better fit in the area of the optimum and thereby, eliminate the number of indefinite quadratics generated.

A search was made of the objective function values for each test point. The worst 25 percent of the test points were then discarded reducing the number of observations considered by the regression procedure. Because the curve-fitting procedure was designed to find the minimum objective response, the worst 25 percent of the test points were considered to be the 25 percent with the highest values of the objective function. This procedure was then repeated eliminating the worst 50 percent of the test points.

This entire procedure was programmed in FORTRAN for use on the IBM 360 computer. The program as written is capable of handling up to 20 input variables and 50 observations of test points. A detailed description of the curve-fitting program is contained in Appendix A.

C. TEST FUNCTIONS

Two types of objective functions were considered in the investigation; quadratic functions, with cross product terms, and exponential functions. The decision to include cross product terms in the quadratic functions was made to preclude the possibility of achieving an exact fit when employing the curve-fitting technique. In the case of exponential functions, the exponent term was a quadratic function chosen so that the function itself was unimodal.

The test functions used were positive definite quadratic functions. Thus, the optimum response was the minimum objective function value. The test functions were constructed so that this minimum occurred at values of the input variables restricted to the domain zero to one.

This restriction limited the selection of initial starting points for Powell's method to points lying within this domain. Test points for the curve-fitting technique were also restricted to the interval zero to one for all input variables.

To insure that the quadratic functions of eight variables were in fact positive definite, the following procedure was used to generate eight by eight matrices of quadratic term coefficients. An eight by eight matrix was generated and pre-multiplied by its transpose. The resulting eight by eight matrix is assured to be positive definite.

An explicit expression of all test functions considered can be found in Appendix B.

D. NUMBER OF VARIABLES

Functions of two different levels of input variables were considered in the investigation. Initially, functions of three input variables were investigated. The investigation concluded with a study of functions of eight input variables. While these two cases are not all inclusive, it was felt they gave a comparative indication of how well the two techniques handled functions of low and high levels of input variables.

E. NUMBER OF OBSERVATIONS

Again two different levels of observations were considered. The low level was fixed at 2n+1 observations. This was the minimum number required for the orthogonal design used in the curve-fitting technique. Twice the number of observations in the low level was arbitrarily decided upon for the high level. An exception to this high level number was made in the case of functions of three input variables. An additional observation was added to the high level number to allow the use of a composite factorial design with the curve-fitting technique.

F. NUMBER OF TRIALS

Ten trials were conducted for each procedure that was non-deterministic. These procedures included all Powell Searches and those curve-fitting techniques where random choice was used to locate some or all of the points.

G. MEASURES OF EFFECTIVENESS

Two measures of effectiveness were employed to compare results:

1. the mean optimum response attained by optimization technique,
2. the mean Euclidean distance between the location of the true optimum and the location of the optimum response found by the optimization technique.

III. PRESENTATION OF RESULTS

Tables I through XI contain the results of the eleven test functions investigated. The following definition of terms is provided to assist in understanding the tabulated data:

TECHNIQUE - The name of the optimization technique used.

POWELL - direct search technique, Powell's method of conjugate directions.

C.F. DESIGN - curve-fit technique, employing experimental design to locate the test points.

C.F. RANDOM - curve-fit technique, employing random choice to locate the test points.

75 PERCENT C.F. DESIGN - curve-fit technique, employing experimental design to locate the test point, eliminating the worst 25 percent of the test points.

50 PERCENT C.F. DESIGN - curve-fit technique, employing experimental design to locate the test points, eliminating the worst 50 percent of the test points.

75 PERCENT C.F. RANDOM - curve-fit technique, employing random choice to locate the test points, eliminating the worst 25 percent of the test points.

50 PERCENT C.F. RANDOM - curve-fit technique, employing random choice to locate the test points, eliminating the worst 50 percent of the test points.

#OBS - The number of observations included in the investigation.

OPTIMUM RESPONSE - The mean and standard deviation, where applicable, of the optimum response found using the optimization technique.

The mean and standard deviation were computed considering only the results associated with positive definite quadratics generated by the curve-fitting technique. Results associated with indefinite quadratic fitted surfaces were meaningless.

DISTANCE - The mean and standard deviation, where applicable, of the Euclidean distance between the location of the true optimum and the location of the optimum response using the optimization technique. The mean and standard deviation were computed considering only the results associated with positive definite quadratics generated by the curve-fitting technique. Results associated with indefinite quadratic fitted surfaces were meaningless.

REMARKS - This column contains the number of indefinite quadratic surfaces generated using the different curve-fitting techniques. In addition, curve-fitting techniques that yielded deterministic results are noted by the statement, 1 trial, in this column.

TABLE I

3 VARIABLES

QUADRATIC I

TRUE OPTIMUM = -7.9875

TECHNIQUE	#OBS	OPTIMUM RESPONSE MEAN	OPTIMUM RESPONSE STD. DEV.	DISTANCE MEAN	DISTANCE STD. DEV.	REMARKS
POWELL	7	-7.3566	0.526	0.381	0.273	
	15	-7.5756	0.277	0.294	0.218	
C.F. DESIGN	7	-7.625	--	0.170	--	1 trial
	15	-7.625	--	0.170	--	1 trial
C.F. RANDOM	7	14.9451	31.03	1.424	1.369	7 indef quad
	15	1.6504	24.01	0.608	1.026	2 indef quad
75% C.F. DESIGN	15	--	--	--	--	1 trial indef quad
50% C.F. DESIGN	15	-7.5283	--	0.116	--	1 trial
75% C.F. RANDOM	15	-7.2643	0.409	0.297	0.074	7 indef quad
50% C.F. RANDOM	15	-5.8791	2.569	0.551	0.570	7 indef quad

NOTES:

1. C.F. design technique achieved the best result. Of interest is the fact that increasing the number of observations did not improve the results.

2. C.F. random technique was quite erratic. An improvement accompanied an increase in the number of observations. However, when eliminating the worst 25% (50%) of the test points an improvement in mean optimum response and distance was accompanied by an increase in the number of indefinite quadratics.

TABLE II

3 VARIABLES

QUADRATIC 2

TRUE OPTIMUM = -59.90

TECHNIQUE	#OBS	OPTIMUM RESPONSE MEAN	OPTIMUM RESPONSE STD DEV.	DISTANCE MEAN	DISTANCE STD. DEV.	REMARKS
POWELL	7	-58.52	1.02	0.485	0.167	
	15	-58.76	0.816	0.449	0.169	
C.F. DESIGN	7	-58.14	--	0.290	--	1 trial
	15	-58.14	--	0.290	--	1 trial
C.F. RANDOM	7	-23.15	35.73	0.896	0.365	8 indef quad
	15	-59.03	0.657	0.274	0.074	6 indef quad
75% C.F. DESIGN	15	7.77	--	0.919	--	1 trial
50% C.F. DESIGN	15	210.60	--	2.529	--	1 trial
75% C.F. RANDOM	15	-39.82	--	0.885	--	9 indef quad
50% C.F. RANDOM	15	161.83	219.68	1.619	1.235	8 indef quad

NOTES:

1. Powell's method achieved the best results at both 7 and 15 observations.

2. At 15 observations the C.F. random techniques achieved better results than Powell's method. However, 6 out of 10 quadratics generated using this technique were indefinite.

3. Eliminating the worst 25% (50%) of the test points did not reflect any improvement using the C.F. design technique. Although the mean optimum response and distance improved when eliminating the worst 25% of the test points using the C.F. random technique, the improvement was accomplished at the expense of 3 more indefinite quadratics being generated.

TABLE III

3 VARIABLES

EXPONENTIAL 1

TRUE OPTIMUM = 4.476×10^{-3}

TECHNIQUE	#OBS	OPTIMUM RESPONSE MEAN	OPTIMUM RESPONSE STD. DEV.	DISTANCE MEAN	DISTANCE STD. DEV.	REMARKS
POWELL	7	6.599×10^{-3}	1.71×10^{-3}	0.478	0.165	
	15	5.251×10^{-3}	0.43×10^{-3}	0.336	0.111	
C.F. DESIGN	7	6.307×10^{-3}	--	0.366	--	1 trial
	15	57.487×10^{-3}	--	0.766	--	1 trial
C.F. RANDOM	7	5.491×10^{-3}	--	0.337	--	9 indef quad
	15	8.018×10^{-3}	2.70×10^{-3}	0.558	0.268	7 indef quad
75% C.F. DESIGN	15	--	--	--	--	1 trial indef quad
50% C.F. DESIGN	15	5.241×10^{-3}	--	0.389	--	1 trial
75% C.F. RANDOM	15	1056.026	1829.075	0.746	0.563	6 indef quad
50% C.F. RANDOM	15	93.852×10^{-3}	145.24×10^{-3}	0.506	0.142	6 indef quad

NOTES:

1. C.F. design technique achieved the best results at 7 observations and the 50% C.F. design technique proved to be the best at 15 observations.

2. At 7 observations C.F. random technique achieved a better mean optimum response than C.F. design technique. However, 9 out of 10 quadratics generated were indefinite.

3. Eliminating the worst 25% (50%) of the test points using the C.F. random technique decreased the number of indefinite quadratics generated. However, the mean optimum response in both cases became worse.

TABLE IV

3 VARIABLES

EXPONENTIAL 2

TRUE OPTIMUM = 0.0212

TECHNIQUE	#OBS	OPTIMUM RESPONSE MEAN	OPTIMUM RESPONSE STD. DEV.	DISTANCE MEAN	DISTANCE STD. DEV.	REMARKS
POWELL	7	0.0380	0.0154	0.606	0.200	
	15	0.0285	0.0088	0.408	0.249	
C.F. DESIGN	7	0.0306	--	0.374	--	1 trial
	15	0.2548	--	0.776	--	1 trial
C.F. RANDOM	7	0.0368	0.0092	0.694	0.238	6 indef quad
	15	0.0943	0.0652	0.641	0.149	6 indef quad
75% C.F. DESIGN	15	7.0428	--	1.128	--	1 trial
50% C.F. DESIGN	15	0.0380	--	0.365	--	1 trial
75% C.F. RANDOM	15	1.22×10^7	1.22×10^7	1.802	1.026	8 indef quad
50% C.F. RANDOM	15	106.38	148.94	0.729	0.302	7 indef quad

NOTES:

1. C.F. design technique was best at 7 observations. Powell's method achieved the best results at 15 observations.

2. C.F. random techniques generated a large number of indefinite quadratics. Eliminating the worst 25% (50%) of the test points neither improved the mean optimum response or distance, nor did it decrease the number of indefinite quadratics.

TABLE V

3 VARIABLES

EXPONENTIAL 3

TRUE OPTIMUM = 0.2157

TECHNIQUE	#OBS	OPTIMUM RESPONSE MEAN	OPTIMUM RESPONSE STD. DEV.	DISTANCE MEAN	DISTANCE STD. DEV.	REMARKS
POWELL	7	0.5255	0.313	0.534	0.205	
	15	0.2969	0.056	0.336	0.110	
C.F. DESIGN	7	0.2232	--	0.155	--	1 trial
	15	0.3423	--	0.464	--	1 trial
C.F. RANDOM	7	$>10^{76}$	$>10^{76}$	2.032	2.252	7 indef quad
	15	0.2873	0.045	0.357	0.126	6 indef quad
75% C.F. DESIGN	15	0.4668	--	0.662	--	1 trial
50% C.F. DESIGN	15	0.2205	--	0.164	--	1 trial
75% C.F. RANDOM	15	23.7196	46.698	0.625	0.475	5 indef quad
50% C.F. RANDOM	15	0.3059	0.073	0.330	0.054	7 indef quad

NOTES:

1. C.F. design technique achieved the best results at 7 observations. At 15 observations, C.F. design eliminating the worst 50% of the test points achieved the best results.

2. C.F. random techniques were extremely erratic generating many indefinite quadratics. In the case of 7 observations, one bad fit caused the mean optimum response to become so large as to be meaningless.

TABLE VI

8 VARIABLES

QUADRATIC 1

TRUE OPTIMUM = -44.61

TECHNIQUE	#OBS	OPTIMUM RESPONSE MEAN	OPTIMUM RESPONSE STD.DEV.	DISTANCE MEAN	DISTANCE STD. DEV.	REMARKS
POWELL	17	-41.97	1.207	1.028	0.047	
	34	-43.50	0.591	0.962	0.047	
C.F. DESIGN	17	-44.18	--	0.644	--	1 trial
	34	-27.26	21.14	1.871	1.992	2 indef quad
C.F. RANDOM	17	--	--	--	--	10 indef quad
	34	-34.30	7.76	1.584	1.844	8 indef quad
75% C.F. DESIGN	34	-43.45	1.569	0.55	0.066	7 indef quad
50% C.F. DESIGN	34	-43.14	--	0.827	--	9 indef quad
75% C.F. RANDOM	34	--	--	--	--	10 indef quad
50% C.F. RANDOM	34	--	--	--	--	10 indef quad

NOTES:

1. C.F. design achieved the best results at 17 observations. At 34 observations, Powell's method proved to be the best technique. However, it is significant to note that the results achieved using C.F. design techniques with 17 observations is better than that achieved by Powell with 34 observations.

2. C.F. random techniques were a complete disaster generating all indefinite quadratic surfaces in 3 out of 4 investigations used.

TABLE VII

8 VARIABLES

QUADRATIC 2

TRUE OPTIMUM = -91.89

TECHNIQUE	#OBS	OPTIMUM RESPONSE MEAN	OPTIMUM RESPONSE STD. DEV.	DISTANCE MEAN	DISTANCE STD.DEV.	REMARKS
POWELL	17	-79.56	9.53	0.997	0.212	
	34	-85.04	4.02	0.942	0.233	
C.F. DESIGN	17	15.31	--	0.763	--	1 trial
	34	-16.92	49.88	1.099	1.314	6 indef quad
C.F. RANDOM	17	--	--	--	--	10 indef quad
	34	676.47	--	3.751	--	9 indef quad
75% C.F. DESIGN	34	1358.07	1308.57	3.153	1.885	8 indef quad
50% C.F. DESIGN	34	--	--	--	--	10 indef quad
75% C.F. RANDOM	34	--	--	--	--	10 indef quad
50% C.F. RANDOM	34	--	--	--	--	10 indef quad

NOTES:

1. Powell's method achieved the best results at both 17 and 34 observations.

2. All curve-fitting techniques were inefficient. A probable cause of this problem was the fact that the quadratic function investigated here had sharply rising contours in the vicinity of the optimum.

TABLE VIII

8 VARIABLES

QUADRATIC 3

TRUE OPTIMUM = -68.25

TECHNIQUE	#OBS	OPTIMUM RESPONSE MEAN	OPTIMUM RESPONSE STD. DEV.	DISTANCE MEAN	DISTANCE STD. DEV.	REMARKS
POWELL	17	-43.29	30.42	1.201	0.454	
	34	-47.67	26.62	1.135	0.448	
C.F. DESIGN	17	154.19	--	1.011	--	1 trial
	34	1681.1	3398.9	14.68	31.56	3 indef quad
C .F. RANDOM	17	--	--	--	--	10 indef quad
	34	--	--	--	--	10 indef quad
75% C.F. DESIGN	34	1617.68	1282.98	2.893	1.495	7 indef quad
50% C.F. DESIGN	34	--	--	--	--	10 indef quad
75% C.F. RANDOM	34	--	--	--	--	10 indef quad
50% C.F. RANDOM	34	--	--	--	--	10 indef quad

NOTES:

1. Powell's method achieved the best results at both 17 and 34 observations.

2. All curve-fitting techniques were inefficient. A probable cause of this problem was the fact that the quadratic function investigated here had sharply rising contours in the vicinity of the optimum.

TABLE IX

8 VARIABLES

EXPONENTIAL 1

TRUE OPTIMUM = 1.3×10^{-9}

TECHNIQUE	#OBS	OPTIMUM RESPONSE MEAN	OPTIMUM RESPONSE STD. DEV	DISTANCE MEAN	DISTANCE STD. DEV.	REMARKS
POWELL	17	1245.08×10^{-9}	3605.4×10^{-9}	1.072	0.214	
	34	8.46×10^{-9}	18.38×10^{-9}	0.925	0.188	
C.F. DESIGN	17	2.3×10^{-9}	--	0.526	--	1 trial
	34	10.6×10^{-9}	--	0.674	--	9 indef quad
C.F. RANDOM	17	--	--	--	--	10 indef quad
	34	--	--	--	--	10 indef quad
75% C.F. DESIGN	34	81.93×10^{-9}	106.26×10^{-9}	0.813	0.038	7 indef quad
50% C.F. DESIGN	34	3.31×10^{-9}	--	0.542	--	9 indef quad
75% C.F. RANDOM	34	--	--	--	--	10 indef quad
50% C.F. RANDOM	34	--	--	--	--	10 indef quad

NOTES:

1. C.F. design technique achieved the best results at 17 observations while Powell's method proved best at 34 observations. However, the results at 17 observations using C.F. design were better than those using Powell's method with 34 observations.

2. At 34 observations, C.F. design eliminating the worst 50% of the test points achieved results superior to Powell's method. However, 9 out of the 10 quadratics generated were indefinite.

3. C.F. random techniques were completely inefficient generating indefinite quadratic surfaces in every situation they were used.

TABLE X

8 VARIABLES

EXPONENTIAL 2

TRUE OPTIMUM = 1.02×10^{-25}

TECHNIQUE	#OBS	OPTIMUM RESPONSE MEAN	OPTIMUM RESPONSE STD. DEV.	DISTANCE MEAN	DISTANCE STD. DEV.	REMARKS
POWELL	17	61.27×10^{-25}	108.13×10^{-25}	1.096	0.190	
	34	14.40×10^{-25}	23.61×10^{-25}	0.949	0.204	
C.F. DESIGN	17	33.4×10^{-25}	--	0.591	--	1 trial
	34	--	--	--	--	10 indef quad
C.F. RANDOM	17	--	--	--	--	10 indef quad
	34	--	--	--	--	10 indef quad
75% C.F. DESIGN	34	18599×10^{-25}	26296×10^{-25}	1.262	0.495	7 indef quad
50% C.F. DESIGN	34	94.3×10^{-25}	--	0.696	--	9 indef quad
75% C.F. RANDOM	34	--	--	--	--	10 indef quad
50% C.F. RANDOM	34	--	--	--	--	10 indef quad

NOTES:

1. At 17 observations C.F. design achieved the best results while Powell's method was superior at 34 observations.

2. C.F. random techniques were completely inefficient generating indefinite quadratic surfaces in every situation they were used.

TABLE XI

8 VARIABLES

EXPONENTIAL 3

TRUE OPTIMUM = 9.91×10^{-6}

TECHNIQUE	#OBS	OPTIMUM RESPONSE MEAN	OPTIMUM RESPONSE STD. DEV.	DISTANCE MEAN	DISTANCE STD. DEV.	REMARKS
POWELL	17	67.95×10^{-6}	50.18×10^{-6}	1.078	0.269	
	34	17.0×10^{-6}	5.11×10^{-6}	0.960	0.277	
C.F. DESIGN	17	12.92×10^{-6}	--	0.516	--	1 trial
	34	$>10^{26}$	$>10^{26}$	1.912	1.903	4 indef quad
C.F. RANDOM	17	--	--	--	--	10 indef quad
	34	1172.35×10^{-6}	1147.92×10^{-6}	1.003	0.349	8 indef quad
75% C.F. DESIGN	34	21.74×10^{-6}	12.09×10^{-6}	0.608	0.104	2 indef quad
50% C.F. DESIGN	34	13.63×10^{-6}	0.836×10^{-6}	0.540	0.025	6 indef quad
75% C.F. RANDOM	34	10^{6}	--	2.482	--	9 indef quad
50% C.F. RANDOM	34	--	--	--	--	10 indef quad

NOTES:

1. C.F. design achieved the best results at 17 observations. Although Powell's method proved to be the best at 34 observations the results achieved were not as good as those achieved using C.F. design with 17 observations.

2. At 34 observations, C.F. design eliminating the worst 25% (50%) of the test points improved the mean optimum response and distance. However, when eliminating the worst 25% a decrease in the number of indefinite quadratics occurred, while elimination of the worst 50% caused an increase in the number of indefinite quadratics.

3. C.F. random techniques were inefficient due mainly to the large number of indefinite quadratics generated.

IV. CONCLUSIONS

Overall, the direct search technique, Powell's method of conjugate directions, proved to be the best technique of those considered in the investigations. A particularly desirable feature of Powell's method was the fact that results achieved using this technique converged to the true optimum response when increasing the number of observations. Thus, Powell's method exhibited the characteristic that the more information (function evaluations) available, the better the results that could be expected to be achieved. A decrease in the distance between the location of the true optimum response and the location of the optimum found using Powell's method was also experienced as the number of observations was increased.

The curve-fitting techniques, as a whole, were quite erratic. The biggest single problem area was the fact that the curve-fitting techniques generated a large number of indefinite quadratic surfaces. This situation was prevalent in both the case where the test points were located by random choice and the case where experimental design was used to locate the test points. However, it should be noted that the problem of indefinite quadratics arose in the curve-fitting techniques employing experimental design only in those cases where the total number of observations exceeded that number required by the experimental design. In this situation, the remaining number of test points were located by random choice. Thus, the results achieved strongly indicate a high correlation between the generation of indefinite quadratic surfaces and the use of random choice to locate any or all of the test points.

A second problem encountered when using curve-fitting techniques was the trend toward higher standard deviations of the optimum responses. This situation resulted from the fact that the curve-fitting techniques generated many bad fits with respect to the optimum response. This problem coupled with a reduction in the sample size due to the generation of indefinite quadratics led to a higher variance of optimum response, thus a high standard deviation.

To overcome these problems the effect of reducing the number of observations considered by the curve-fitting technique by elimination of the worst 25% and 50% of the test points was investigated. This approach did not yield consistent results. In some cases the number of indefinite quadratics generated was reduced, while in other cases the number remained the same or increased. In addition, in many cases where the number of indefinite quadratics was reduced, an increase in the number of bad fits with respect to the optimum response was experienced.

However, curve-fitting techniques do merit consideration under certain circumstances. Provided that the objective function did not have sharply rising contours in the immediate location of the true optimum, curve-fitting techniques employing experimental design, where the number of test points considered was exactly equal to the number required by the experimental design, provided results better than those achieved by Powell's method. In fact for some of the test functions investigated, the results achieved using the curve-fitting techniques employing experimental design were better than those achieved using Powell's method with twice the number of observations.

Thus, on the whole Powell's method is an efficient method of optimization that consistently achieved better results than any of the

curve-fitting techniques investigated. Under certain circumstances, the curve-fitting technique employing experimental design to locate test points provided better results and was more efficient than Powell's method.

V. SUGGESTED FUTURE INVESTIGATIONS

The problem of generating indefinite quadratic fits when using curve-fitting techniques was never overcome. In future investigations one way of addressing this problem would be to constrain the regression procedures so that the coefficients of the quadratic terms are all positive. Another method that could be used, either in connection with the above technique or by itself, is to force the fitted surface through a certain point. As a suggestion it might prove useful to force the fitted surface through the best of the test points. The best test point being chosen as the one with the smallest function value in the case of a minimization problem.

Another interesting question that could be considered is whether a combination of the search and curve-fitting procedures would achieve better results than either procedure achieved by itself. A suggested approach to this question would be to use Powell's search initially to find the general area of the optimum solution. Once this has been achieved, the curve-fit design procedure could be used in this general area to obtain the optimum solution.

APPENDIX A

CURVE-FITTING PROGRAM

The purpose of the curve-fitting program was threefold:

1) determine the location of the test points and their respective functional values,

2) fit a quadratic surface to these test points,

3) determine the location of the optimum response of the resulting "fitted" quadratic surface and compute the true functional value at that point.

A. LOCATION OF TEST POINTS

The program considered two different methods for locating test points in the input variable space, random choice or placement in accordance with an experimental design. Input data to the program provided the basis for selection of the appropriate method. A uniform random number generator was used in connection with the method of random choice. The two experimental designs used, composite factorial and orthogonal, were written into the program to be constructed by the computer as they were needed.

Upon selection of the appropriate method a matrix was constructed, each row of the matrix containing one observation of the input variables. At this stage the number of columns in the matrix was equal to the number of input variables. The size of the matrix was then expanded in the following manner. For each observation the value of each input variable was squared and stored in the corresponding row of the matrix, thus, doubling the number of columns of the matrix. A final column was added

to the matrix by computing the function value corresponding to each observation of the input variables.

The completed matrix was then input to the curve-fitting procedure unless a reduction in the number of observations was called for. In that case the function values were scanned by the computer to determine the worst 25% or 50% of the values as appropriate. These worst function values along with their corresponding input variables and their squares were eliminated from the matrix reducing the number of rows in the matrix input to the curve-fitting procedure.

B. FITTING THE QUADRATIC SURFACE

Multiple linear regression techniques were used to fit a quadratic surface to the test points. This section of the program consisted of three subroutines and the procedures for computation of the regression coefficients. The three subroutines used were CORRE, ORDER, and MINV. These subroutines are included in the IBM 360 Scientific Subroutine Package and are stored internally on the computer.

The purpose of subroutine CORRE was to compute the means, the standard deviations and a matrix of correlation coefficients of the variables included in the regression. In this case, the original input variables, their squares and the function values. The matrix constructed in the initial section of the program served as input to CORRE to accomplish these tasks.

The resulting correlation coefficient matrix from subroutine CORRE was then input to subroutine ORDER. The function of subroutine ORDER was to compute a matrix containing the correlation coefficients of the independent variables and a vector containing the correlation coefficients of the independent variables with the dependent variables. In the

program, the original input variables and their squares were designated the independent variables, and the function value was designated the dependent variable.

To check the validity of the regression the matrix of correlation coefficients among independent variables calculated by subroutine ORDER was input to subroutine MINV. Subroutine MINV computed the determinant of this matrix. If the determinant was non-zero, the regression was valid and the program continued. However, if the determinant was zero the program terminated due to the presence of multi-collinearity in the regression.

Provided the regression was valid, the next step was to compute the regression coefficients of the independent variables. To accomplish this task a portion of subroutine MULTR, also included in the IBM 360 Scientific Subroutine Package, was used. The standard deviations computed by subroutine CORRE, and the matrix of correlation coefficients among the independent variables and the vector of correlation coefficients between the independent variables and the dependent variable computed by subroutine ORDER served as input to this subroutine. The output of the subroutine was a vector containing the regression coefficients.

C. LOCATION OF OPTIMUM RESPONSE

The vector of regression coefficients contained the necessary information to determine the location of the optimum response of the fitted surface. The length of this vector was 2n, where n is the number of input variables. The resulting fitted quadratic was of the form:

$$f = \sum_{i=1}^{n} a_i x_i + \sum_{i=1}^{n} b_i x_i^2 + \text{constant}$$

The value of the a_i's was contained in the first n elements of the vector of regression coefficients and the value of the b_i's was contained in the last n elements.

The location of the optimum (minimum) response was determined as follows. The values:

$$x_i = -\frac{1}{2}\frac{a_i}{b_i}\ ;\ i = 1, \ldots , n,$$

were computed. These values uniquely determined the location of the minimum response of the fitted quadratic provided that the quadratic was positive definite. A vector containing the location of the optimum response of the fitted surface was then substituted into the true objective function to determine the optimum objective response for the curve-fitting procedures.

APPENDIX B

TEST FUNCTIONS

I. THREE VARIABLE FUNCTIONS

A. QUADRATIC 1

$$F = X^T \begin{bmatrix} 2 & 1 & 2 \\ & 4 & 3 \\ & & 6 \end{bmatrix} X + X^T \begin{bmatrix} -6 \\ -8 \\ -13 \end{bmatrix}$$

True Optimum = -7.9875 At $X_1 = 0.625$

$X_2 = 0.4$

$X_3 = 0.725$

B. QUADRATIC 2

$$F = X^T \begin{bmatrix} 29 & 26 & 30 \\ & 29 & 34 \\ & & 45 \end{bmatrix} X + X^T \begin{bmatrix} -80.2 \\ -81.0 \\ -98.4 \end{bmatrix}$$

True Optimum = -59.90 At $X_1 = 0.7$

$X_2 = 0.3$

$X_3 = 0.4$

C. EXPONENTIAL 1

$$F = \exp\left(X^T \begin{bmatrix} 8 & 2 & 3 \\ & 2 & 2 \\ & & 4 \end{bmatrix} X + X^T \begin{bmatrix} -10 \\ -6 \\ -8 \end{bmatrix} \right)$$

True Optimum = 0.004476 At $X_1 = 0.2727$

$X_2 = 0.8636$

$X_3 = 0.3636$

D. EXPONENTIAL 2

$$F = \exp\left(X^T \begin{bmatrix} 8 & 2 & -3 \\ & 2 & -2 \\ & & 4 \end{bmatrix} X + X^T \begin{bmatrix} -9 \\ -3 \\ 1 \end{bmatrix} \right)$$

True Optimum = 0.0212 At $X_1 = 0.6364$

$X_2 = 0.9318$

$X_3 = 0.8182$

E. EXPONENTIAL 3

$$F = \exp\left(X^T \begin{bmatrix} 8 & 2 & -3 \\ & 2 & -2 \\ & & 4 \end{bmatrix} X + X^T \begin{bmatrix} -5 \\ -3 \\ 2 \end{bmatrix} \right)$$

True Optimum = 0.2157 At $X_1 = 0.2273$

$X_2 = 0.8864$

$X_3 = 0.3636$

II. EIGHT VARIABLE FUNCTIONS

A. QUADRATIC 1

$$F = X^T \begin{bmatrix} 4 & 2 & 4 & 2 & 2 & 4 & 2 & 0 \\ & 2 & 3 & 3 & 0 & 1 & 2 & 0 \\ & & 9 & 6 & 5 & 5 & 4 & 2 \\ & & & 7 & 0 & 2 & 1 & 1 \\ & & & & 11 & 6 & 6 & 0 \\ & & & & & 9 & 0 & 0 \\ & & & & & & 12 & 0 \\ & & & & & & & 5 \end{bmatrix} X + X^T \begin{bmatrix} -18.8 \\ -13.6 \\ -38.0 \\ -24.4 \\ -25.8 \\ -23.8 \\ -24.8 \\ -10.4 \end{bmatrix}$$

True Optimum = -44.61 At $X_1 = 0.6$ $X_5 = 0.7$

$X_2 = 0.9$ $X_6 = 0.2$

$X_3 = 0.2$ $X_7 = 0.3$

$X_4 = 0.8$ $X_8 = 0.8$

B. QUADRATIC 2

$$F = X^T \begin{bmatrix} 93 & -53 & -54 & 32 & 13 & 31 & -67 & 43 \\ & 103 & 54 & -40 & -28 & 1 & 46 & -1 \\ & & 50 & -28 & -1 & -20 & 54 & -30 \\ & & & 22 & 8 & 7 & -30 & 14 \\ & & & & 32 & -11 & 3 & -15 \\ & & & & & 25 & -30 & 30 \\ & & & & & & 69 & -45 \\ & & & & & & & 57 \end{bmatrix} X + X^T \begin{bmatrix} -96.2 \\ -75.0 \\ 6.4 \\ 3.2 \\ 4.0 \\ -65.8 \\ 48.0 \\ -89.4 \end{bmatrix}$$

True Optimum = -91.89 At $X_1 = 0.7$ $X_5 = 0.6$

$X_2 = 0.7$ $X_6 = 0.8$

$X_3 = 0.4$ $X_7 = 0.2$

$X_4 = 0.3$ $X_8 = 0.3$

C. QUADRATIC 3

$$F = X^T \begin{bmatrix} 43 & -23 & 16 & -29 & 25 & -22 & 22 & -21 \\ & 50 & -22 & 53 & -45 & 39 & -39 & 37 \\ & & 19 & -28 & 24 & -21 & 21 & -20 \\ & & & 106 & -69 & 60 & -60 & 57 \\ & & & & 76 & -45 & 45 & -43 \\ & & & & & 41 & -40 & 37 \\ & & & & & & 45 & -35 \\ & & & & & & & 38 \end{bmatrix} X + X^T \begin{bmatrix} -\ \ 1.0 \\ -\ 78.0 \\ 32.8 \\ -146.2 \\ 92.0 \\ -\ 85.2 \\ 79.6 \\ -\ 83.0 \end{bmatrix}$$

True Optimum = -68.25 At $X_1 = 0.7$ $X_5 = 0.1$
$X_2 = 0.1$ $X_6 = 0.6$
$X_3 = 0.2$ $X_7 = 0.2$
$X_4 = 0.4$ $X_8 = 0.6$

D. EXPONENTIAL 1

$$F = \exp\left(X^T \begin{bmatrix} 5 & 1 & -3 & 0 & 2 & 0 & 0 & 0 \\ & 2 & -2 & 0 & 3 & 0 & 0 & 0 \\ & & 4 & 0 & -2 & 0 & 0 & 0 \\ & & & 10 & 0 & 4 & -4 & 1 \\ & & & & 6 & 0 & 0 & 0 \\ & & & & & 3 & -2 & 1 \\ & & & & & & 8 & -3 \\ & & & & & & & 4 \end{bmatrix} X + X^T \begin{bmatrix} -11.4 \\ -\ 7.6 \\ 8.0 \\ -16.4 \\ -15.0 \\ -\ 7.4 \\ 1.2 \\ -\ 2.0 \end{bmatrix} \right)$$

True Optimum = 1.3×10^{-9} At $X_1 = 0.9$ $X_5 = 0.7$
$X_2 = 0.7$ $X_6 = 0.3$
$X_3 = 0.3$ $X_7 = 0.6$
$X_4 = 0.9$ $X_8 = 0.4$

E. EXPONENTIAL 2

$$F = \exp\left(1/144\ X^T \begin{bmatrix} 806 & 374 & 473 & 390 & 456 & 388 & 570 & 620 \\ & 395 & 453 & 417 & 585 & 275 & 405 & 452 \\ & & 587 & 529 & 696 & 417 & 583 & 535 \\ & & & 810 & 574 & 297 & 403 & 589 \\ & & & & 1581 & 411 & 644 & 624 \\ & & & & & 385 & 591 & 372 \\ & & & & & & 1574 & 693 \\ & & & & & & & 930 \end{bmatrix} X + X^T \begin{bmatrix} -25.84 \\ -20.95 \\ -27.40 \\ -25.00 \\ -34.75 \\ -21.61 \\ -42.48 \\ -32.72 \end{bmatrix} \right)$$

True Optimum $= 1.02 \times 10^{-25}$ At

$$X_1 = 0.1032 \quad X_5 = 0.3980$$
$$X_2 = 0.1502 \quad X_6 = 0.7504$$
$$X_3 = 0.1759 \quad X_7 = 0.9039$$
$$X_4 = 0.3889 \quad X_8 = 0.8039$$

F. EXPONENTIAL 3

$$F = \exp\left(1/144 \; X^T \begin{bmatrix} 295 & 171 & -170 & 123 & 125 & 128 & 5 & -24 \\ & 687 & -90 & 136 & -138 & 89 & 138 & -191 \\ & & 316 & -140 & -150 & 28 & 196 & -44 \\ & & & 302 & -42 & 149 & -192 & 14 \\ & & & & 329 & 199 & -108 & 114 \\ & & & & & 687 & -125 & 122 \\ & & & & & & 707 & 42 \\ & & & & & & & 269 \end{bmatrix} X + X^T \begin{bmatrix} -4.48 \\ -6.76 \\ -1.63 \\ -2.27 \\ -1.28 \\ -10.40 \\ -7.53 \\ -1.84 \end{bmatrix} \right)$$

True Optimum $= 9.91 \times 10^{-6}$ At

$$X_1 = 0.4004 \quad X_5 = 0.4314$$
$$X_2 = 0.4158 \quad X_6 = 0.7791$$
$$X_3 = 0.6355 \quad X_7 = 0.8928$$
$$X_4 = 0.7187 \quad X_8 = 0.2144$$

BIBLIOGRAPHY

1. Box, M.J., Davies, D. and Swann, W.H., Imperial Chemical Industries Monograph No. 5, Non-Linear Optimization Techniques, Oliver and Boyd, 1969.

2. Powell, M.J.D., "An Efficient Method for Finding the Minimum of a Function of Several Variables without Calculating Derivatives," The Computer Journal, v. 7, No. 2, p. 155-162, July 1964.

3. Brownlee, K.A., Statistical Theory and Methodology in Science and Engineering, 2nd ed., p. 433-439, Wiley, 1967.

4. Huang, D.S., Regression and Econometric Methods, p. 56-61, Wiley 1970.

5. Davies, O.L. and Others, The Design and Analysis of Industrial Experiments, 2nd ed., p. 532-536, Oliver and Boyd, 1956.

INITIAL DISTRIBUTION LIST

		No. Copies
1.	Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2.	Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3.	Chief of Naval Personnel Pers - 11b Department of the Navy Washington, D.C. 20370	1
4.	Naval Postgraduate School Department of Operations Research and Administrative Sciences Monterey, California 93940	1
5.	Asst Professor J.K. Hartman, Code 55Hh Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
6.	Major John A. Murray, USMC Division Operations Analysis Section 2nd Marine Division, FMF Camp Lejeune, North Carolina 28540	1

Security Classification

DOCUMENT CONTROL DATA - R & D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1. ORIGINATING ACTIVITY (Corporate author) Naval Postgraduate School Monterey, California 93940	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE
A Comparison of Two Alternative Unconstrained Non-Linear Optimization Techniques

4. DESCRIPTIVE NOTES (Type of report and, inclusive dates)
Master's Thesis; (March 1972)

5. AUTHOR(S) (First name, middle initial, last name)
John Anthony Murray

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
March 1972	42	5

8a. CONTRACT OR GRANT NO.	9a. ORIGINATOR'S REPORT NUMBER(S)
b. PROJECT NO.	
c.	9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report)
d.	

10. DISTRIBUTION STATEMENT

Approved for public release; distribution unlimited.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY
	Naval Postgraduate School Monterey, California 93940

13. ABSTRACT

Two alternative methods for optimizing an unconstrained non-linear function are investigated and compared. The investigations are made subject to a restriction as to the number of function evaluations available to conduct the optimization procedures. Powell's method of conjugate directions is employed as the direct search method and is considered the reference method. The alternate method is based on fitting a quadratic surface to the available function evaluations and optimizing over the resulting fitted surface. The test functions considered in the investigation were limited to unimodal functions.

Security Classification

KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
?well's Method of Conjugate Directions						
?rect Search						
?rve-Fitting						
?gression						
?timization						

FORM NOV 65 1473 (BACK)
1-807-6821

Security Classification A-31409

Thesis 134274
M9832 Murray
c.1 A comparison of two alternative unconstrained non-linear optimization techniques.

Thesis 134274
M9832 Murray
c.1 A comparison of two alternative unconstrained non-linear optimization techniques.

thesM9832
A comparison of two alternative unconstr
3 2768 000 99359 6
DUDLEY KNOX LIBRARY