A TRANSFORMATION METHOD FOR CONSTRAINED-FUNCTION MINIMIZATION

Stephen K. Park

Langley Research Center
Hampton, Va. 23665

1. Report No. NASA TN D-7983	2. Government Accession No.	3. Recipient's Catalog No.
4. Title and Subtitle A TRANSFORMATION METHOD FOR CONSTRAINED- FUNCTION MINIMIZATION		5. Report Date November 1975
		6. Performing Organization Code
7. Author(s) Stephen K. Park		8. Performing Organization Report No. L-10178
		10. Work Unit No. 506-25-99-02
9. Performing Organization Name and Address NASA Langley Research Center Hampton, Va. 23665		11. Contract or Grant No.
		13. Type of Report and Period Covered Technical Note
12. Sponsoring Agency Name and Address National Aeronautics and Space Administration Washington, D.C. 20546		14. Sponsoring Agency Code

15. Supplementary Notes

16. Abstract

A direct method for constrained-function minimization is discussed. The method involves the construction of an appropriate function mapping all of one finite dimensional space onto the region defined by the constraints. Functions which produce such a transformation are constructed for a variety of constraint regions including, for example, those arising from linear and quadratic inequalities and equalities. In addition, the computational performance of this method is studied in the situation where the Davidon-Fletcher-Powell algorithm is used to solve the resulting unconstrained problem. Good performance is demonstrated for 19 test problems by achieving rapid convergence to a solution from several widely separated starting points.

17. Key Words (Suggested by Author(s)) Computer mathematics Optimization Numerical methods Transformations	18. Distribution Statement Unclassified — Unlimited Subject Category 59
19. Security Classif. (of this report) Unclassified	20. Security Classif. (of this page) Unclassified	21. No. of Pages 36	22. Price* $3.75

A TRANSFORMATION METHOD FOR CONSTRAINED-FUNCTION MINIMIZATION

Stephen K. Park
Langley Research Center

SUMMARY

A direct method for constrained-function minimization is discussed. The method involves the construction of an appropriate function mapping all of one finite dimensional space onto the region defined by the constraints. Functions which produce such a transformation are constructed for a variety of constraint regions including, for example, those arising from linear and quadratic inequalities and equalities. In addition, the computational performance of this method is studied in the situation where the Davidon-Fletcher-Powell algorithm is used to solve the resulting unconstrained problem. Good performance is demonstrated for 19 test problems by achieving rapid convergence to a solution from several widely separated starting points.

INTRODUCTION

One popular method for solving a constrained-function minimization problem is to reformulate the problem — by various devices — as an equivalent unconstrained problem. This, for example, is the philosophy underlying the method of penalty functions. In this report an alternate philosophy is discussed — namely, that of providing an equivalent unconstrained problem by way of a transformation in the space of independent variables.

As an example, suppose F is a function of n real parameters $x_1, x_2, \ldots, x_n$ and it is desired to minimize F subject to a collection of equality and inequality constraints. Let X in R^n denote the set of points $x = (x_1, x_2, \ldots, x_n)$ which (simultaneously) satisfy all the constraints. A transformation in the space of independent parameters involves the construction of a function θ which maps R^p (for an appropriate integer p) onto X.

If such a function θ can be constructed, it provides a representation of each of the n constrained parameters $x = (x_1, x_2, \ldots, x_n)$ in X in terms of p unconstrained parameters, say $z = (z_1, z_2, \ldots, z_p)$ in R^p. Consequently, by way of the parameter transformation $x = \theta(z)$, the original constrained problem becomes the equivalent problem of minimizing $f(z) = F(\theta(z))$ subject to no constraints.

The procedure just outlined is referred to as the (parameter) transformation method. It is clear that this is not a general-purpose method applicable to all problems but rather a fairly specialized one applicable only when X has a reasonably simple structure. However, the method remains attractive since many important types of constraints give rise to rather simple constraint regions and since it allows for considerable ingenuity on the part of the user.

In the late 1950's and early 1960's simple transformations were occasionally used to reformulate optimal control problems with bounded control variables as equivalent calculus of variations problems. (See, e.g., refs. 1 to 3.) This procedure is now known to yield the same information as a direct application of the Pontryagin Maximum Principle. Consequently, the procedure is rarely used now to analyze control problems mathematically. However, it is used in the numerical solution of optimal control problems — at least in those situations where the control variables are merely bounded above or below (or both).

Surprisingly, transformation methods received no attention in the mathematical programing literature (except, in passing, in ref. 4) until Box's 1966 paper (ref. 5). Since that time the use of simple transformations has become rather widespread. (See, e.g., ref. 6.) However, with the exception of reference 7, there apparently has been little effort made toward systematically studying and extending the earlier results. Moreover, with the exception of reference 5, there has been little reported computational experience with parameter transformations when used in conjunction with some of the best currently available unconstrained minimization algorithms.

In this report the transformation method is formulated and discussed as the problem of constructing a mapping whose domain is all of some Euclidean space and whose range is X. The equivalence of the original problem and the resulting transformed problem is demonstrated. Moreover, a variety of constraint regions X are considered corresponding, for example, to linear and quadratic inequality and equality constraints. In each case an appropriate transformation is constructed. In addition, the performance of the transformation method as a computational technique is studied by combining the method with the Davidon-Fletcher-Powell algorithm. The resulting algorithm is used to solve 19 constrained minimization test problems. For each problem good performance is demonstrated by achieving rapid convergence to a solution from several widely separated starting points.

SYMBOLS

A^T	transpose of matrix A
B^n, B^p	set of all vectors in R^n, R^p with $\|x\| \leq 1$
C^∞	set of all functions with continuous partials of all orders

2

$F, F(x)$ real valued function to be minimized

$f, f(z)$ real valued function defined by $f(z) = F(\theta(z))$

n, p positive integers

R^n, R^p set of all real vectors of dimension n and p, respectively

X subset of R^n, the constraint region

x n-dimensional vector with components x_i

x^0 optimal (i.e., minimal) value of x

z p-dimensional vector with components z_i

z^0 optimal (i.e., minimal) value of z

$\rho_i, \lambda_i, \phi_i$ spherical coordinates of data points (see appendix B)

θ function relating z to x (i.e., $x = \theta(z)$) with n components θ_i

$[-1, 1]^n$ set of all vectors in R^n with each component confined to the interval $[-1, 1]$

$< , >$ inner (i.e., dot) product of two vectors

$\| \ \|$ Euclidean norm of a vector

o   denotes composition of two functions; for example, $f = F \circ \theta$ is the function defined by $f(z) = F(\theta(z))$

Abbreviations:

DFP   Davidon-Fletcher-Powell algorithm

DFP/T   modification of Davidon-Fletcher-Powell algorithm

FORMULATION

As in the "Introduction," suppose that F is a real valued function defined for all values of x in R^n.   Assume that a collection of equality and inequality constraints are

present and that X is the (nonempty) set of all those values of x which satisfy all constraints (i.e., X is the feasible region). The problem of finding an (optimal) x^O in X such that $F(x^O) \leq F(x)$ for all values of x in X will be referred to as the original (or constrained) problem.

If θ is a function mapping R^p onto X, then the composition $f = F \circ \theta$ (i.e., $f(z) = F(\theta(z))$) is a real valued function defined for all values of z in R^p. As in the previous paragraph, the second problem to consider is finding an optimal z^O in R^p such that $f(z^O) \leq f(z)$ for all values of z in R^p. This second problem is called the transformed (or unconstrained) problem.

As a simple example, consider the problem of minimizing $F(x_1, x_2)$ subject to the constraint $x_1^2 + x_2^2 = 1$. In this case n = 2 and X is the unit circle. To construct an equivalent unconstrained problem one can use the transformation $x_1 = \sin z$ and $x_2 = \cos z$. In this case p = 1 and the transformed problem is as follows: minimize $f(z) = F(\sin z, \cos z)$ where z is now unconstrained.

These two problems and, in general, any original problem and its associated transformed problem are equivalent in the following sense:

If z in R^p and x in X satisfy $x = \theta(z)$, then z is optimal for the transformed problem if and only if x is optimal for the original problem.

The proof of this result is immediate but interesting since it demonstrates that equivalence is a consequence of the global nature of the optimization problems and that this is true independent of any continuity or differentiability requirements on F and θ. By construction, $f(z) = F(\theta(z)) = F(x)$ whenever z in R^p and x in X satisfy $\theta(z) = x$. Consequently, suppose z^O is optimal for the transformed problem and let $x^O = \theta(z^O)$. Since θ is onto, for any x in X, there is a z in R^p with $\theta(z) = x$. Moreover, since z^O is optimal, $f(z) \geq f(z^O)$. Therefore, because $f(z^O) = F(x^O)$, it follows that $F(x) \geq F(x^O)$; that is, x^O is optimal for the original problem.

Conversely, if x^O is optimal for the original problem and if z^O satisfies $\theta(z^O) = x^O$, then z^O is optimal for the transformed problem. If z is an arbitrary point in R^p, then $x = \theta(z)$ in X satisfies $F(x) \geq F(x^O)$ and, consequently, $f(z) \geq f(z^O)$.

Thus, a parameter transformation neither introduces new (extraneous) global minima (for the original problem) nor does it eliminate old ones. In particular, if a global minima, say z^O, has been obtained for the unconstrained problem, then the image $\theta(z^O)$ is a global minimum for the original constrained problem. Moreover, if the constrained problem has a solution — for example, this must be true if X is closed and bounded (compact) — then such a z^O must exist.

Consistent with the standard assumptions of nonlinear programing it is assumed that F and at least its first partials are continuous on an open set containing X. It also

4

is assumed that each component of θ is continuous with at least continuous first partials for all values of z in R^p. With these assumptions, a parameter transformation is a generalized change of variables from R^p to X. The generalization arises from the fact that — as the previous example illustrates — θ need not be (and, in general, is not) one-to-one. It is this lack of one-to-oneness which permits a transformation between sets $\left(R^p \text{ and } X\right)$ which are, in general, not topologically equivalent. Specifically, in the previous example X was the unit circle which is compact, whereas R^1 is not compact.

TRANSFORMATION CONSTRUCTIONS

For many nonlinear programing problems the construction of a parameter transformation is impractical if not, in fact, impossible. However, in these situations it is frequently possible to effect a decomposition of the problem and thereby eliminate some, but not all, of the constraints. More will be said about this later. For the present, attention will be restricted to those situations where all the constraints are of the same form.

It turns out that all the transformations constructed in this report are C^∞. That is, each component θ_i of θ is continuous with continuous partials of all orders. Thus, there is no loss of generality involved in formulating the question of the existence and construction of a parameter transformation as follows:

Given a subset X of R^n, construct a C^∞ function θ mapping some Euclidean space R^p onto X. By necessity X must be connected and, in general, although not always, it is also compact. In most cases p is the dimension of the set X, although in at least one important case this is not so.

Upper and Lower Bound Constraints

In R^1, if X is compact and connected, then X must be an interval which without loss of generality may be taken as $[-1,1]$. For example, since any one of the functions

$$\theta(z) = \sin \frac{\pi}{2} z \tag{1a}$$

or

$$\theta(z) = \frac{2z}{1 + z^2} \tag{1b}$$

or

$$\theta(z) = z \exp\left(\frac{1 - z^2}{2}\right) \tag{1c}$$

is a C^∞ map from R^1 onto $[-1,1]$, for $n = 1$ the question has an affirmative answer. For convenience, these functions have been normalized so that $\theta(-1) = -1$ and $\theta(1) = 1$.

The functions immediately generalize, by way of a simple affine transformation, to the case where $X = [a,b]$. In fact, the functions generalize to the case where X is the set of all $x = (x_1, x_2, \ldots, x_n)$ which satisfy

$$a_i \leq x_i \leq b_i \qquad\qquad (i = 1, 2, \ldots, n) \qquad (2)$$

where the upper and lower limits b_i and a_i, respectively, are fixed. For example, using equation (1a) yields $\theta = (\theta_1, \theta_2, \ldots, \theta_n)$ where

$$\theta(z) = \frac{b_i - a_i}{2} \sin \frac{\pi z_i}{2} + \frac{b_i + a_i}{2} \qquad\qquad (i = 1, 2, \ldots, n) \qquad (3)$$

is a C^∞ map from R^n onto the right parallelepiped defined by expression (2). In a natural way, equation (3) generalizes further to the case where X is the set of all values of x in R^n with

$$x = w^0 + \sum_{i=1}^{n} \mu_i w^i \qquad\qquad (4)$$

where each μ_i can vary from 0 to 1. The points $w^1, w^2, \ldots, w^n$ in R^n are linearly independent, and w^0 in R^n is arbitrary. If X is given by equation (4), the function defined by

$$\phi(z) = w^0 + \frac{1}{2} \sum_{j=1}^{n} (1 + z_j) w^j$$

is a C^∞ map (in fact, an affine transformation) from R^n to R^n which maps $[-1,1]^n$ onto X. Here, $[-1,1]^n$ denotes the right parallelepiped defined by expression (2) in the case $a_i = -1$ and $b_i = 1$. In this case the function defined by equation (3) maps R^n onto $[-1,1]^n$, and taking the composition of this function with ϕ yields a C^∞ map from R^n onto X. That is,

$$\theta(z) = w^0 + \frac{1}{2} \sum_{j=1}^{n} \left(1 + \sin \frac{\pi}{2} z_j\right) w^j$$

is a C^∞ map from R^n onto the parallelepiped defined by equation (4). In component form,

$$\theta_i(z) = w_i^0 + \frac{1}{2} \sum_{j=1}^{n} \left(1 + \sin \frac{\pi}{2} z_j\right) w_i^j \tag{5}$$

where w_i^j is the ith component of the vector w^j.

Quadratic Constraints

Another generalization of equations (1), specifically of (1b), is given in the following discussion. Let Q be a real, symmetric, positive definite $n \times n$ matrix and let X be the set of x in R^n with

$$\langle x, Qx \rangle \leqq 1 \tag{6}$$

Here, $\langle \, , \, \rangle$ denotes the inner product. Since Q is positive definite, X is convex and compact. In fact, in terms of the norm and inner product induced on R^n by Q, X becomes the unit ball. This observation suggests

$$\theta(z) = \frac{2z}{1 + \langle z, Qz \rangle} \tag{7}$$

as a possible candidate for a C^∞ map from R^n onto the set X given by expression (6).

Notice that, for any value of z in R^n,

$$0 \leqq (1 - \langle z, Qz \rangle)^2 = (1 + \langle z, Qz \rangle)^2 - 4\langle z, Qz \rangle$$

Thus,

$$\langle \theta(z), Q\theta(z) \rangle = \frac{4\langle z, Qz \rangle}{(1 + \langle z, Qz \rangle)^2} \leqq 1$$

and so θ maps R^n into X. To verify that the map is, in fact, onto, consider an arbitrary point $\bar{x}$ in X and define $\bar{z}$ in R^n as

$$\bar{z} = \frac{\bar{x}}{1 + \sqrt{1 - \langle \bar{x}, Q\bar{x} \rangle}} \tag{8}$$

Inserting $\bar{z}$ into equation (7) and performing the indicated operations yields, after some algebraic manipulation, $\theta(\bar{z}) = \bar{x}$ and, thus, θ maps R^n onto X. It is easily verified that θ is C^∞. In component form θ is given by

$$\theta_i(z) = \frac{2z_i}{1 + \langle z, Qz \rangle} \tag{9}$$

In the particular case where Q is the identity matrix, $\langle z, Qz \rangle$ reduces to $\|z\|^2 = z_1^2 + z_2^2 + \ldots + z_n^2$, and so

$$\theta(z) = \frac{2z}{1 + \|z\|^2} \tag{10}$$

is a C^∞ map from R^n onto the unit ball B^n, that is, the points x in R^n with $\|x\| \leqq 1$.

The previous results may be used to study a slightly more general quadratic constraint; that is, X is all points x in R^n with

$$\langle x, Qx \rangle + 2\langle a, x \rangle \leqq 1 \tag{11}$$

where a is a fixed vector in R^p. Since Q is positive definite, a positive definite inverse Q^{-1} exists. Define a second positive definite matrix

$$Q^* = \frac{1}{1 + \langle a, Q^{-1}a \rangle} Q$$

and let X^* denote the set of all values of x in R^n with $\langle x, Q^*x \rangle \leqq 1$. Then,

$$\theta^*(z) = \frac{2z}{1 + \langle z, Q^*z \rangle}$$

is a C^∞ map from R^n onto X^*. Moreover, the affine function

$$\gamma(z) = z - Q^{-1}a$$

maps X^* onto X and, consequently, X is the image of R^n under the C^∞ function $\gamma \circ \theta^*$. That is,

$$\theta(z) = \frac{2\left(1 + \langle a, Q^{-1}a \rangle\right)z}{1 + \langle a, Q^{-1}a \rangle + \langle z, Qz \rangle} - Q^{-1}a \tag{12}$$

is a C^∞ map from R^n onto the set defined by expression (11).

Linear Constraints

All the parameter transformations previously constructed preserve the dimension of the constraint set X. That is, in each case the dimension of X was n and θ was a

map from R^n onto X. However, if X is a polytope with $p+1$ vertices, θ becomes a map from R^p onto X. Typically, although not necessarily, in this case p is larger than n.

If X is a polytope, then it is the convex hull of $p+1$ vectors (vertices) $v^1, v^2, \ldots, v^{p+1}$ in R^n. Such a set is extremely important since it typically arises as the feasible region for a collection of linear inequalities. The set X consists of all vectors x of the form

$$x = \sum_{j=1}^{p+1} \mu_j v^j$$

where $\mu_j \geqq 0$ and $\sum_{j=1}^{p+1} \mu_j = 1$. An equivalent representation is

$$x = v^{p+1} + \sum_{j=1}^{p} \mu_j \left(v^j - v^{p+1} \right)$$

where $\mu_j \geqq 0$ and now $\sum_{j=1}^{p} \mu_j \leqq 1$. Thus, the function

$$\gamma(z) = v^{p+1} + \sum_{j=1}^{p} z_j^2 \left(v^j - v^{p+1} \right)$$

is a C^∞ map from the unit ball B^p onto X. This map, in composition with the map given by equation (10), yields a C^∞ map from R^p onto X. That is,

$$\theta(z) = v^{p+1} + \left(\frac{2}{1 + \|z\|^2} \right)^2 \sum_{j=1}^{p} z_j^2 \left(v^j - v^{p+1} \right) \tag{13}$$

is a C^∞ map from R^p onto the convex hull of the points $v^1, v^2, \ldots, v^{p+1}$ in R^n. In component form this map is

$$\theta_i(z) = v_i^{p+1} + \left(\frac{2}{1 + \|z\|^2} \right)^2 \sum_{j=1}^{p} z_j^2 \left(v_i^j - v_i^{p+1} \right) \tag{14}$$

where v_i^j is the ith component of the vector v^j.

Equation (14) assumes a particularly simple form in the case where X is the region generated by the constraints

$$x_i \geqq 0 \qquad\qquad\qquad (i = 1, 2, \ldots, n)$$

$$\sum_{i=1}^{n} a_i^2 x_i \leqq 1 \qquad\qquad\qquad (a_i \neq 0)$$

In this case $p = n$, the vertex v^{n+1} is the origin, and for $i = 1, 2, \ldots, n$ the vertex v^i is the vector with $1/a_i^2$ in the ith position and zeroes elsewhere. Consequently, equation (14) becomes

$$\theta_i(z) = \left[\frac{2\left(\dfrac{z_i}{a_i}\right)}{1 + \|z\|^2} \right]^2$$

Additional Constructions

Suppose X is the set of all vectors x in R^n which satisfy

$$0 \leqq a_1 x_1 \leqq a_2 x_2 \leqq \ldots \leqq a_n x_n \qquad\qquad (15)$$

where each (fixed) a_i is positive. In this case X is closed and connected but not bounded. The dimension of X is n. The function θ with components θ_i defined by

$$\theta_i(z) = \frac{1}{a_i}\left(z_1^2 + z_2^2 + \ldots + z_i^2\right) \qquad\qquad (16)$$

is a C^∞ map from R^n onto X. This is clearly true since

$$a_1 \theta_1(z) = z_1^2 \geqq 0$$

and, for $i = 2, 3, \ldots, n,$

$$a_i \theta_i(z) - a_{i-1}\theta_{i-1}(z) = z_i^2 \geqq 0$$

Linear Equality Constraints

Consider now the situation where X is defined by the system of linear equality constraints

10

$$Ax = b \qquad (17)$$

Here, x is a (column) vector in R^n, b is a fixed (column) vector in R^m, and A is an $m \times n$ matrix with $m < n$. That is, X is the hyperplane of all vectors x in R^n which satisfy equation (17). Let p (where $p \geq n - m$) be the dimension of the null space of A and assume that X is not empty; that is, b is in the range of A. In this case there are p linearly independent vectors $y^1, y^2, \ldots, y^p$ in R^n which form a basis for the null space of A, and there is a vector c in R^n with $Ac = b$. Since any point in X has the representation

$$x = c + \sum_{j=1}^{p} \mu_j y^j$$

$\left(\text{where } \mu_j \text{ is a scalar} \right)$, it follows that

$$\theta(z) = c + \sum_{j=1}^{p} z_j y^j \qquad (18)$$

is a C^∞ (in fact, an affine) map from R^p onto X. In component form this map is given as

$$\theta_i(z) = c_i + \sum_{j=1}^{p} z_j y_i^j \qquad (19)$$

where again the subscript i denotes the ith component of the vectors c and y^j.

Quadratic Equality Constraints

Finally, in R^n let X be the surface of an n dimensional ellipse, that is, the set of all values of x in R^n such that

$$\left(\frac{x_1}{a_1} \right)^2 + \left(\frac{x_2}{a_2} \right)^2 + \ldots + \left(\frac{x_n}{a_n} \right)^2 = 1$$

where each value of a_i is positive. Using generalized spherical coordinates provides a C^∞ map from R^{n-1} (i.e., $p = n - 1$) onto X. Namely, the function is $\theta = \left(\theta_1, \theta_2, \ldots, \theta_n \right)$ where

$$\theta_1(z) = a_1 \sin z_1 \qquad\qquad (m = 1)$$

$$\vdots$$

$$\theta_m(z) = a_m \cos z_1 \cos z_2 \ldots \cos z_{m-1} \sin z_m \qquad (2 \leqq m \leqq n - 1) \qquad (20)$$

$$\vdots$$

$$\theta_n(z) = a_n \cos z_1 \cos z_2 \ldots \cos z_{n-2} \cos z_{n-1} \qquad (m = n)$$

THE GENERAL NONLINEAR PROGRAMING PROBLEM

In the situation where the constraint region is too complicated — in its entirety — to be represented by a parameter transformation, a decomposition may be possible. Specifically, consider the general nonlinear programing problem of minimizing $F(x)$ where x is in R^n subject to the constraints

$$G_i(x) = 0 \qquad\qquad (i = 1, 2, \ldots, R)$$

$$H_j(x) \leqq 0 \qquad\qquad (j = 1, 2, \ldots, S)$$

This is the original problem.

Suppose G_i and H_j are indexed in such a manner that the constraint region arising from the last $R-r$ equality constraints and $S-s$ inequality constraints can be represented by a parameter transformation. That is, let X be the set of points x in R^n for which

$$G_i(x) = 0 \qquad\qquad (i = r + 1, \ldots, R)$$

and

$$H_j(x) \leqq 0 \qquad\qquad (j = s + 1, \ldots, S)$$

and suppose that θ is a C^∞ map from R^p onto X.

As before, the transformed problem is constructed by defining the functions

$$f(z) = F(\theta(z))$$

$$g_i(z) = G_i(\theta(z)) \qquad\qquad (i = 1, 2, \ldots, r)$$

$$h_j(z) = H_j(\theta(z)) \qquad\qquad (j = 1, 2, \ldots, s)$$

12

where z is in R^p. The transformed problem reads: minimize $f(z)$ where z is in R^p subject to the constraints

$$g_i(z) = 0 \qquad\qquad (i = 1, 2, \ldots, r)$$

$$h_i(z) \leqq 0 \qquad\qquad (j = 1, 2, \ldots, s)$$

Again, it is easy to prove that these two problems are equivalent in the sense that if $x = \theta(z)$ then x is optimal for the original problem if and only if z is optimal for the transformed problem. The proof is omitted.

The rest of this report is devoted to computational considerations. However, there is an interesting aspect of parameter transformations which should be mentioned — namely, their use as a conceptual device to relate the methods of constrained and unconstrained optimization. As an example of this, a derivation of the Lagrange multiplier rule using a parameter transformation is presented in appendix A.

TEST PROBLEMS AND RESULTS

In order to explore the efficacy of parameter transformations as a computational technique, a total of 19 (constrained) nonlinear programing problems were solved. Six of these problems are constrained versions of the standard unconstrained test functions due to Rosenbrock and Wood. For three others the objective function is quadratic. An additional four are constrained test problems due to Box (ref. 5). Of the remaining six, one is a nonlinear least-squares problem, two are discrete formulations of a continuous-time optimal-control problem, two others are geometric problems, and the last is a problem which Dixon (ref. 8) indicates many algorithms have failed to solve.

The algorithm used to solve each of these 19 problems is a FORTRAN adaption of the original Davidon-Fletcher-Powell (DFP) algorithm (ref. 9) and will be denoted by DFP/T. To implement DFP/T the user must supply two subroutines: one which evaluates the n vector θ and the $n{\times}p$ matrix $\partial\theta/\partial z$ at z, and the other which evaluates F and its gradient $\partial F/\partial x$ at $x = \theta(z)$. In each case algebraic expressions for all these partials are available. The gradient $\partial f/\partial z$ is obtained by use of the chain rule.

In order to use DFP/T, one must also supply an initial estimate of the solution in either z-space or in x-space. When x is supplied and not z, it also is necessary to supply a subroutine which solves the equation $x = \theta(z)$ for z (once).

For each of the 19 test problems, convergence is achieved for several starting points. In each case, convergence means that both F and each component of the solution vector agree with their optimum values to at least five significant digits. The results

presented are the number of iterations and the number of function evaluations required to achieve convergence. The number of function evaluations is the number of calls to the subroutine which evaluates F and its gradient $\partial F/\partial x$. All calculations were performed on CDC 6000 series computers.

Problems With Upper and Lower Bound Constraints

A total of six test problems were solved with upper and lower bound constraints

$$a_i \leqq x_i \leqq b_i \qquad\qquad (i = 1, 2, \ldots, n)$$

The results are presented in table I. In each case the parameter transformation given by equation (3) was used. In order of increasing value of n, these six problems were as follows:

<u>Rosenbrock (I)</u>.- In the Rosenbrock (I) problem,

$$F = 100\left(x_1^2 - x_2\right)^2 + \left(1 - x_1\right)^2$$

with constraints

$$-2 \leqq x_1 \leqq 0.5$$

$$-1 \leqq x_2 \leqq 2$$

The solution is $F = 0.25$ at the boundary point $(0.5, 0.25)$.

<u>Quadratic (I)</u>.- In the Quadratic (I) problem,

$$F = 2x_1^2 + 2x_2^2 + x_3^2 + 2x_1 x_2 + 2x_1 x_3 - 8x_1 - 6x_2 - 4x_3 + 9$$

with constraints

$$0 \leqq x_1 \leqq 3$$

$$0 \leqq x_2 \leqq 3$$

$$0 \leqq x_3 \leqq 1.5$$

The solution is $F = 0$ at the interior point $(1,1,1)$.

NLS (Nonlinear Least Square).- In the NLS problem,

$$F(x) = \frac{1}{2} \sum_{i=1}^{108} \left[r_i(x) - \rho_i \right]^2$$

with constraints

$$0 \leq x_1 \leq 3500$$

$$0 \leq x_2 \leq 3500$$

$$0 \leq x_3 \leq 3500$$

The function $r_i(x)$ is given by

$$r_i\left(x_1, x_2, x_3\right) = \left(\frac{\cos^2\phi_i \cos^2\lambda_i}{x_1^2} + \frac{\cos^2\phi_i \sin^2\lambda_i}{x_2^2} + \frac{\sin^2\phi_i}{x_3^2} \right)^{-1/2}$$

where the 108 data points $\left(\rho_i, \lambda_i, \phi_i\right)$ are listed in table V, appendix B. This is a nonlinear parameter estimation problem (n = 3) to determine the three ellipsoidal parameters of a nearly spherical body. The solution is $F = 175.09$ at the interior point (1740.9, 1738.9, 1736.1).

Wood (I).- In the Wood (I) problem,

$$F = 100\left(x_1^2 - x_2\right)^2 + \left(1 - x_1\right)^2 + 90\left(x_3^2 - x_4\right)^2 + \left(1 - x_3\right)^2$$

$$+ 10.1\left[\left(1 - x_2\right)^2 + \left(1 - x_4\right)^2\right] + 19.8\left(1 - x_2\right)\left(1 - x_4\right)$$

with constraints

$$-4 \leq x_1 \leq 0.5$$

$$-2 \leq x_2 \leq 2$$

$$-4 \leq x_3 \leq 0.5$$

$$-2 \leq x_4 \leq 2$$

This problem has, in effect, two widely separated global minima: namely, $F = 5.7419$ at the boundary point $(-1.2748, 1.6341, 0.50000, 0.26296)$ and $F = 5.7425$ at the boundary point $(0.50000, 0.26167, -1.2749, 1.6354)$. The algorithm can converge to either point depending on the starting conditions. In those cases where convergence is to the latter point, the results are indicated by an asterisk.

Rational.- In the Rational problem,

$$F = c_0 - x_1 \left(\frac{c_1 + c_2 x_2 + c_3 x_3 + c_4 x_4 + c_5 x_5}{d_1 + d_2 x_2 + d_3 x_3 + d_4 x_4 + d_5 x_5} \right)$$

with constraints

$$0 \leqq x_i \leqq 1 \qquad\qquad (i = 1, 2, \ldots, 5)$$

where

$$c_0 = 10$$

$$c_1 = 1 \qquad\qquad d_1 = 5$$

$$c_2 = 10 \qquad\qquad d_2 = 3$$

$$c_3 = -5 \qquad\qquad d_3 = -1$$

$$c_4 = 9 \qquad\qquad d_4 = 1$$

$$c_5 = -1 \qquad\qquad d_5 = -2$$

The form of this problem represents a modification of Box's problem A (ref. 5). However, the coefficients are not those used by Box. The solution is $F = 7.2857$ at the corner point $(1,1,0,1,1)$.

Minimum Time (I).- In the Minimum Time (I) problem,

$$F = \frac{1}{2}\left(\frac{V_6 - V_f}{K} \right)^2 + \sum_{i=1}^{6} \tau_i \qquad\qquad (K = 5)$$

where $V_0 = 250$, V_i is defined by

$$V_i = \left(V_{i-1}^2 + 2 x_i \Delta L_i \right)^{1/2} \qquad\qquad (i = 1, 2, \ldots, 6)$$

16

and τ_i is given by

$$\tau_i = \frac{2\Delta L_i}{V_i + V_{i-1}} \qquad (i = 1, 2, \ldots, 6)$$

The constants ΔL_i are

$$\Delta L_1 = \Delta L_2 = 32\,500$$

$$\Delta L_3 = \Delta L_4 = 65\,000$$

$$\Delta L_5 = \Delta L_6 = 80\,000$$

and $V_f = 800$. This problem corresponds to a six-segment parameterization of a Minimum Time trajectory and is discussed in appendix C. The contraints are

$$0 \leq x_i \leq 2 \qquad (i = 1, 2, \ldots, 6)$$

and the solution is $F = 538.64$ at the boundary point $(2,2,2,0.53446,0,0)$.

Problems With Quadratic Constraints

A total of five test problems were solved with a quadratic inequality constraint of the form

$$\sum_{i=1}^{n} \left(\frac{x_i - a_i}{b_i} \right)^2 \leq 1$$

where a_i and b_i are constants. The parameter transformation used is based on equation (10) and is given by

$$\theta_i(z) = \frac{2b_i z_i}{1 + \|z\|^2} + a_i$$

This transformation may be obtained by defining the intermediate variable

$$y_i = \frac{x_i - a_i}{b_i}$$

in which case the inequality constraint becomes $\|y\|^2 \leq 1$. Using equation (10) yields

$$y_i = \frac{2z_i}{1 + \|z\|^2}$$

from which it follows that

$$x_i = \frac{2b_i z_i}{1 + \|z\|^2} + a_i$$

The computational results for these five test problems are presented in table II. In order of increasing value of n, these problems were as follows:

Rosenbrock (II).- In the Rosenbrock (II) problem the same objective function as that in Rosenbrock (I) is used, but now the constraint is

$$\left(\frac{x_1 + 1}{2}\right)^2 + \left(\frac{x_2}{\sqrt{2}}\right)^2 \leq 1$$

The solution is $F = 0.043117$ at the boundary point $(0.79247, 0.62732)$.

Quadratic (II).- In the Quadratic (II) problem the same objective function as that in Quadratic (I) is used, but now the constraint is

$$\left(\frac{x_1}{3}\right)^2 + \left(\frac{x_2}{3}\right)^2 + \left(\frac{x_3}{1.5}\right)^2 \leq 1$$

The solution is $F = 0$ at the interior point $(1,1,1)$.

Wood (II).- In the Wood (II) problem the same objective function as that in Wood (I) is used, but now the constraint is

$$\left(\frac{x_1 + 3}{5}\right)^2 + \left(\frac{x_2 + 1}{5}\right)^2 + \left(\frac{x_3 + 3}{5}\right)^2 + \left(\frac{x_4 + 1}{5}\right)^2 \leq 1$$

As with Wood (I) the problem has, effectively, two widely separated global minima: namely, $F = 4.9812$ at the boundary point $(-1.1311, 1.2848, 0.72903, 0.54314)$ and $F = 4.9821$ at the boundary point $(0.72899, 0.54193, -1.1313, 1.2858)$. Again, which of these two points the algorithm converges to is determined by the initial conditions. In those cases where convergence is to the latter point, the results are indicated by an asterisk.

Minimum Distance.- In the Minimum Distance problem the objective function is

$$F = \sum_{m=1}^{5} \left(x_m - 2\right)^2$$

with the constraint

$$\sum_{m=1}^{5} \left(\frac{x_m}{2^m}\right)^2 \leqq 1$$

This problem has the geometric interpretation of finding that point in the ellipsoid which is closest to (2,2,2,2,2). The solution is $F = 0.12211$ at the boundary point (1.6649, 1.9042,1.9752,1.9937,1.9984).

Minimum Time (II).- The problem in Minimum Time (II) is identical to that in Minimum Time (I) except that the constraints $0 \leqq x_i \leqq 2$ are replaced by the constraint

$$\sum_{m=1}^{6} \left(\frac{x_m - 1}{\sqrt{6}}\right)^2 \leqq 1$$

The solution is $F = 524.16$ at the boundary point (2.5484,1.7752,1.6364,0.94378,0.25772, -0.42912). This problem has a neighboring minimum of $F = 524.53$ at the boundary point (2.5420,1.7720,1.6332,0.94315,0.25963,-0.42388), and for some initial conditions the algorithm converges to this latter point.

Problems With Linear Constraints

A total of seven problems were solved with linear inequality constraints; that is, the feasible region was the convex hull of a finite set of vertices. The results are presented in table III. In each case the parameter transformation used is given by equation (14) and the vertices used to generate the feasible region are listed (as row vectors). In order of increasing value of n, these problems were as follows:

Rosenbrock (III).- The objective function in the Rosenbrock (III) problem is the same as that in Rosenbrock (I), but the constraints are

$$x_2 - 2x_1 \leqq 0$$

$$x_2 + 2x_1 - 4 \leqq 0$$

$$x_2 \geqq 0$$

The feasible region is the triangle generated by the vertices (2,0), (1,2), and (0,0) and, thus, $p = n = 2$. The solution is $F = 0$ at the interior point (1,1).

Box B.- In the Box B problem (ref. 5),

$$F = \frac{x_1(x_1 - 6)x_2^3}{27\sqrt{3}}$$

with constraints

$$x_1 \geqq 0$$

$$x_1 \geqq \sqrt{3}\, x_2 \geqq 0$$

$$6 \geqq \sqrt{3}\, x_2 + x_1 \geqq 0$$

The feasible region is the triangle generated by the vertices (6,0), (3,$\sqrt{3}$), and (0,0) and, thus, $p = 2$. The solution is $F = -1$ at the vertex (3,$\sqrt{3}$).

Quadratic (III).- The objective function in the Quadratic (III) problem is the same as that in Quadratic (I), but the constants are

$$x_i \geqq 0 \qquad\qquad\qquad (i = 1, 2, 3)$$

$$x_1 + x_2 + 2x_3 \leqq 3$$

The feasible region is the three-dimensional simplex generated by the vertices (3,0,0), (0,3,0), (0,0,1.5), and (0,0,0) and, so, $p = 3$. The solution is $F = 1/9$ at the boundary point (4/3,7/9,4/9).

POP (Post Office Parcel).- In the POP problem (ref. 5),

$$F = -x_1 x_2 x_3$$

with constraints

$$0 \leqq x_1 \leqq 42$$

$$0 \leqq x_2 \leqq 42$$

$$0 \leqq x_3 \leqq 42$$

$$0 \leqq x_1 + 2x_2 + 2x_3 \leqq 72$$

20

The feasible region is depicted in figure 1 and is generated by the six indicated vertices (i.e., $p = 5$). The solution is $F = -3456$ at the boundary point $(24,12,12)$.
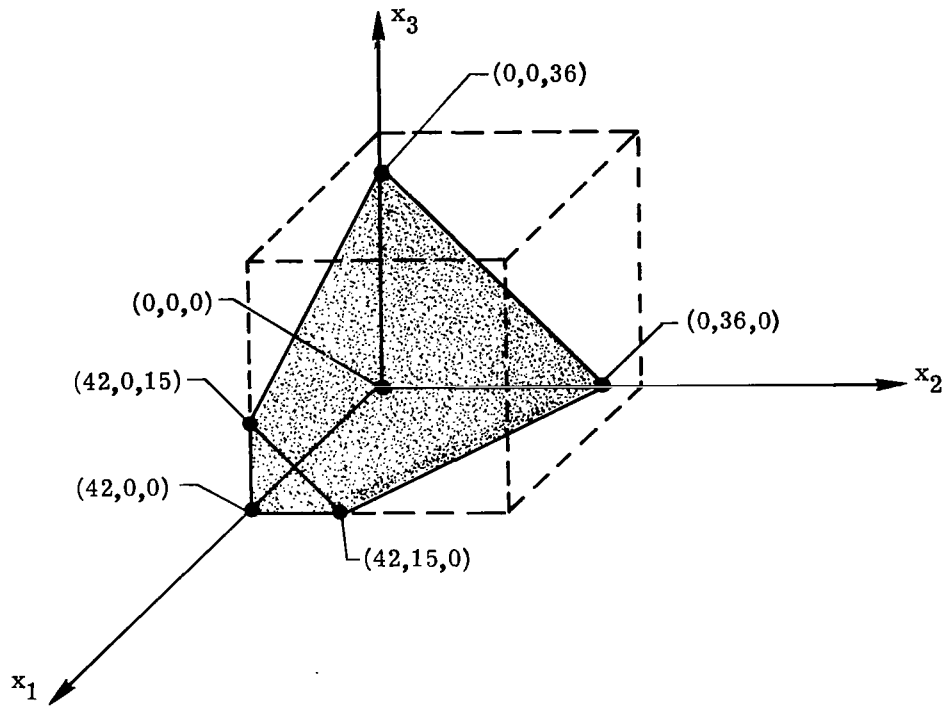

Figure 1.- Constraint region for POP problem.

Modified POP.- In the Modified POP problem (ref. 5),

$$F = -x_1 x_2 x_3$$

with constraints

$$0 \leq x_1 \leq 20$$

$$0 \leq x_2 \leq 11$$

$$0 \leq x_3 \leq 42$$

$$0 \leq x_1 + 2x_2 + 2x_3 \leq 72$$

The feasible region is depicted in figure 2 and is generated by the eight indicated vertices (i.e., p = 7). The solution is F = -3300 at the vertex (20,11,15).
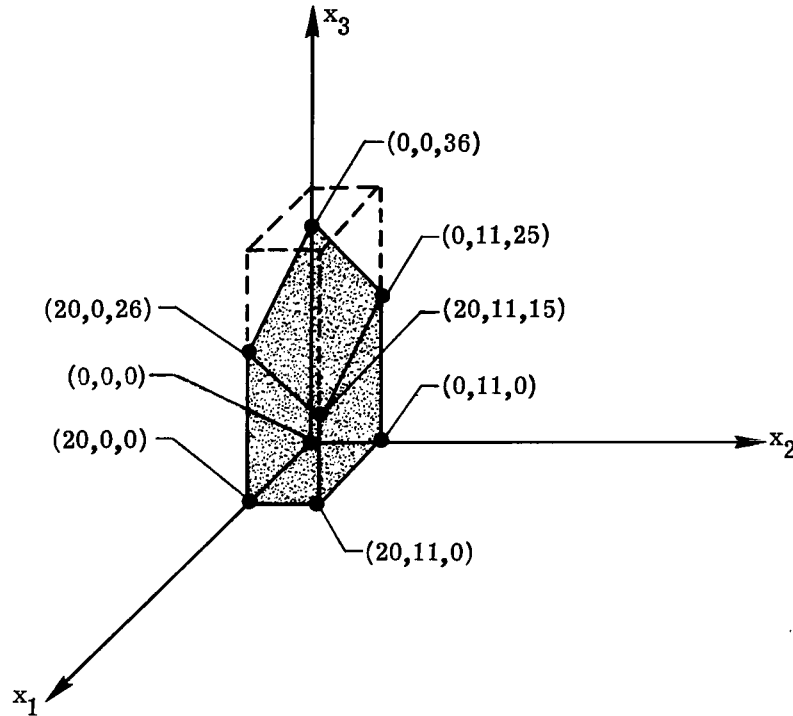
Figure 2.- Constraint region for Modified POP problem.

Wood (III).- The objective function in the Wood (III) problem is the same as that in Wood (I), but the constraints are

$$0 \leqq 2x_4 \leqq x_3 \leqq x_2 \leqq x_1 \leqq 2$$

The feasible region is the simplex generated by the five vertices (2,0,0,0), (2,2,0,0), (2,2,2,0), (2,2,2,1), and (0,0,0,0) and, thus, p = 4. The solution is F = 5.0406 at the boundary point (1.0542,1.0542,0.59805,0.29902).

Maximum Product.- In the Maximum Product problem,

$$F = -x_1 x_2 x_3 x_4 x_5 x_6$$

with the constraints

$$x_i \geqq 0 \qquad\qquad (i = 1, 2, \ldots, 6)$$

and

$$x_1 + \frac{1}{2} x_2 + \frac{1}{3} x_3 + \frac{1}{4} x_4 + \frac{1}{5} x_5 + \frac{1}{6} x_6 \leqq 6$$

The feasible region is the simplex generated by the origin and the seven vectors $v^1 = (1,0,0,0,0,0)$, $v^2 = (0,2,0,0,0,0), \ldots,$ $v^6 = (0,0,0,0,0,6)$, and $v^7 = (0,0,0,0,0,0)$. The solution is $F = -720$ at the boundary point $(1,2,3,4,5,6)$.

In tables I and II the x-space starting points are listed. In contrast, table III lists the z-space starting points. For convenience, table IV provides the x-space starting points corresponding to those z-space starting points listed in table III.

An Open Constraint Region

The nineteenth problem solved involved a constraint region X, unlike any of those considered to this point: namely, it is open. This problem is referred to as RECIP in reference 8 and is as follows:

$$F = \left(x_1 - 5\right)^2 + x_2{}^2 + \frac{x_3{}^2}{x_2 - x_1{}^2}$$

with the constraint

$$x_2 - x_1{}^2 > 0$$

This constraint generates a constraint region which is the interior of a parabolic valley lying with its bottom along the x_3-axis. The parameter transformation used to solve this problem is given by

$$\theta_1(z) = z_1$$

$$\theta_2(z) = z_1{}^2 + \exp\left(z_2\right)$$

$$\theta_3(z) = z_3$$

That θ is, in fact, a map from R^3 onto X follows from the observation that, if $x = \theta(z)$, then $x_2 - x_1{}^2 = \exp\left(z_2\right) > 0$. The solution is $F = 16.50154$ at $(1.2348, 1.5247,0)$ and is obtained in the limit as $z \rightarrow (1.2348,-\infty,0)$.

This problem illustrates just how beneficial a transformation can be. Dixon (ref. 8) tested 16 versions of popular unconstrained function-minimization routines (including DFP)

on RECIP and none succeeded in finding the correct answer. When applied to this problem, DFP/T quickly reduced the function value to $F = 16.590$ (in 7 iterations with 17 function evaluations) and converged to the solution listed in the previous paragraph in 26 iterations with 69 function evaluations. The starting point used was the one given in reference 8; namely, $x = (2,5,1)$.

COMPUTATIONAL EXPERIENCE

In order to summarize the computational results, 19 test problems were solved using a total of 65 different starting points. In every case convergence was achieved. A discussion of such items as robustness and rate of convergence is superseded by the observation that DFP/T is, of course, not really a new algorithm but rather an established algorithm, namely, DFP, in a new coordinate system. Consequently, general comments pertinent to DFP also apply to DFP/T, provided the transformation $x = \theta(z)$ does not render the transformed problem badly scaled.

The particular adaptation of DFP used herein involved the use of Davidon's cubic interpolation (ref. 10) in the one-dimensional search in conjunction with a logic which guarantees a function decrease at each iteration. Except for this guaranteed function decrease no additional attempt is made to perform an exact one-dimensional search.

Since there are points at which the matrix $\partial\theta/\partial z$ has less than maximum rank, one might anticipate that DFP/T will occasionally converge to false minima. In practice, one finds that this almost never happens provided, of course, that the initial estimate is not on the boundary of the constraint region. As evidence of this fact, several starting points are reported in tables I, II, and III for each of the first 18 test problems.

Since DFP/T is an interior method (that is, the sequence of iterates remains interior to the constraint region), it is possible to construct starting points and constraints which will trap the algorithm into local but not global minima. For example, in Rosenbrock (I) if the starting point is (-1.2,1) and the second constraint is replaced by $0.5 \leq x_2 \leq 2$, the algorithm will track the valley $x_2 = x_1^2$ and converge to the boundary point (-0.25, 0.5). Such a phenomenon is not particularly bad, however, since it can be revealed by trying several starting points.

In general, DFP/T performed quite well. For example, when applied to the common unconstrained test functions the number of iterations required for convergence compared favorably. That is, the presence of the constraints does not greatly increase the number of iterations required for convergence, and, in some cases, this number actually is decreased. For those three problems with a quadratic cost function the convergence was always rapid — as one would hope. Finally, for those problems (Box B, POP, and Modified POP) for which some comparable data are available (ref. 5), DFP/T was clearly superior.

Consequently, computational experience indicates that, for those problems where a parameter transformation is possible, its use should be strongly considered.

CONCLUDING REMARKS

The transformation method is not a general-purpose method applicable to all problems. Instead, it is a fairly specialized one applicable only when the constraint region has a reasonably simple structure. However, the method is intuitively appealing and remains attractive since many important types of constraints give rise to constraint regions with simple structure.

The transformation method allows for considerable ingenuity on the part of the user, and this is, in fact, the spirit in which the method is frequently discussed. However, the method is naturally formulated as the mathematical problem of constructing a mapping whose range is known. This formulation not only permits a systematic study of the method but it provides the mechanism for constructing complicated transformations as the composition of a sequence of simple transformations.

Langley Research Center
National Aeronautics and Space Administration
Hampton, Va. 23665
July 29, 1975

APPENDIX A

DERIVATION OF THE LAGRANGE MULTIPLIER RULE

In deriving the Lagrange multiplier rule, consider the problem of minimizing a differentiable function $F(x)$ subject to the constraints

$$Ax = b$$

where the column vectors x, b are in R^n and R^m, respectively, the matrix A is $m \times n$ with $m < n$, and it is assumed that b is in the range of A. For simplicity, it is also assumed that the rank of A is m (i.e., the m rows of A are linearly independent). Equation (19) provides a transformation from the set of n constrained values of x to a set of $n-m$ unconstrained values of z. That is,

$$x_i = \theta_i(z) = c_i + \sum_{j=1}^{n-m} z_j y_i^j$$

If $F(x)$ has a constrained minimum at, say, x^0, then $f(z) = F(\theta(z))$ has an unconstrained minimum at, say, z^0, where $\theta(z^0) = x^0$. At z^0 it is necessary that

$$\frac{\partial f}{\partial z_k} = 0 \qquad\qquad (k = 1, 2, \ldots, n-m)$$

However,

$$\frac{\partial f}{\partial z_k} = \sum_{i=1}^{n} \frac{\partial F}{\partial x_i} \frac{\partial \theta_i}{\partial z_k}$$

$$= \sum_{i=1}^{n} \frac{\partial F}{\partial x_i} y_i^k$$

and, consequently, at x^0, the gradient $\partial F / \partial x$ (a column vector) is orthogonal to the null space of A; that is, at x^0

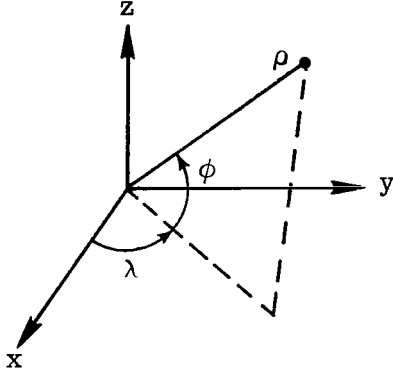$$\left(\frac{\partial F}{\partial x}\right)^T y^k = 0$$

for $k = 1, 2, \ldots, n-m$. Therefore, $\partial F/\partial x$ is in the row space of A $\Big($i.e., the range of $A^T\Big)$ and so there is a (column) vector λ in R^m such that at x^0

$$\frac{\partial F}{\partial x} = A^T \lambda$$

The vector λ is, of course, the Lagrange multiplier for this problem.

APPENDIX B

DESCRIPTION OF THE TEST PROBLEM NLS

The purpose of this appendix is to describe the test problem NLS. By using a spherical coordinate system (as indicated in the accompanying sketch), measurements $(\rho_i, \lambda_i, \phi_i)$ are made at various points on the surface of a nearly spherical body (e.g., the moon). If the body is assumed to satisfy an elliptic model, the points on the surface satisfy the equation

$$\left(\frac{x}{a}\right)^2 + \left(\frac{y}{b}\right)^2 + \left(\frac{z}{c}\right)^2 = 1 \tag{B1}$$

where the (nearly equal) ellipsoidal parameters a,b,c are to be estimated. Using the relations

$$x = \rho \cos \phi \cos \lambda$$

$$y = \rho \cos \phi \sin \lambda$$

$$z = \rho \sin \phi$$

in equation (B1) yields the equation

$$\rho = \left(\frac{\cos^2\phi \, \cos^2\lambda}{a^2} + \frac{\cos^2\phi \, \sin^2\lambda}{b^2} + \frac{\sin^2\phi}{c^2}\right)^{-1/2} \tag{B2}$$

For each measurement $(\rho_i, \lambda_i, \phi_i)$, equation (B2) yields a radius $r_i(a,b,c)$ which is to be compared with ρ_i. Consequently, the parameters a,b,c may be estimated by minimizing the squared sum of residuals

$$\frac{1}{2} \sum_{i=1}^{108} \left(r_i - \rho_i\right)^2$$

28

where

$$r_i(a,b,c) = \left(\frac{\cos^2\phi_i \cos^2\lambda_i}{a^2} + \frac{\cos^2\phi_i \sin^2\lambda_i}{b^2} + \frac{\sin^2\phi_i}{c^2} \right)^{-1/2}$$

The data $\left(\rho_i, \lambda_i, \phi_i\right)$ were statistically generated for testing purposes by using ellipsoidal parameters felt to be characteristic of the moon. The data are listed in table V with ρ_i given in kilometers and λ_i and ϕ_i given in degrees.

APPENDIX C

DESCRIPTION OF THE TEST PROBLEMS MINIMUM TIME (I)

AND MINIMUM TIME (II)

The purpose of this appendix is to describe the test problems Minimum Time (I) and Minimum Time (II). These problems model a point mass moving along a path consisting of n straight-line segments of length ΔL_i. The initial velocity V_o is specified and the point is permitted to accelerate or decelerate at a constant (constrained) rate $\dot{V}_i$ while traveling the length ΔL_i in time τ_i.

Given the lengths ΔL_i with $i = 1, 2, \ldots, n$ and the initial velocity V_o, the velocities V_i and times τ_i are defined recursively — as a function of $\dot{V} = \left(\dot{V}_1, \dot{V}_2, \ldots, \dot{V}_n \right)$ — by the respective equations

$$\left.\begin{aligned} V_i &= \left(V_{i-1}^2 + 2\dot{V}_i \Delta L_i \right)^{1/2} \\[2mm] \tau_i &= \frac{2\Delta L_i}{V_i + V_{i-1}} \end{aligned}\right\} \qquad (i = 1, 2, \ldots, n)$$

The form of these equations is chosen so that $\dot{V}_i = 0$ does not have to be treated as a special case.

The optimization problem is to find that (constrained) collection of accelerations or decelerations $\dot{V}_1, \dot{V}_2, \ldots, \dot{V}_n$ which minimize the transit time plus a weighted squared deviation from a desired final velocity V_f. That is, the objective function is ($n = 6$)

$$F(\dot{V}) = \frac{1}{2}\left(\frac{V_n - V_f}{K} \right)^2 + \sum_{i=1}^{n} \tau_i$$

where $V_f = 800$ and $K = 5$. Calculation of the first partials of this function involves a straightforward but rather tedious process which yields

$$\frac{\partial F}{\partial \dot{V}_j} = \frac{V_n - V_f}{K^2} \frac{\Delta L_j}{V_n} + \frac{\tau_j^2}{2V_{j-1}} - \frac{\Delta L_j}{2} \sum_{i=j}^{n} \left(\frac{1}{V_i} + \frac{1}{V_{i-1}} \right) \frac{\tau_i^2}{\Delta L_i}$$

REFERENCES

1. Miele, Angelo: Flight Mechanics and Variational Problems of a Linear Type. J. Aero/Space Sci., vol. 25, no. 9, Sept. 1958, pp. 581-590.

2. Letov, A. M.: Analytic Controller Design. II. Automat. & Remote Contr., vol. 21, no. 5, 1960, pp. 389-393.

3. Desoer, C. A.: The Bang Bang Servo Problem Treated by Variational Techniques. Inform. & Contr., vol. 2, 1959, pp. 333-348.

4. Dickinson, Allan W.: Nonlinear Optimization: Some Procedures and Examples. Proceedings of 19th National Conference, A.C.M. Publ. P-64, Assoc. Comput. Mach., 1964.

5. Box, M. J.: A Comparison of Several Current Optimization Methods, and the Use of Transformations in Constrained Problems. Computer J., vol. 9, no. 1, May 1966, pp. 67-77.

6. Murray, W., ed.: Numerical Methods for Unconstrained Optimization. Academic Press, Inc., 1972.

7. Rayk, E.; and Petersen, I.: A Method of Mapping for Solving Extremum Problems With Constraints. NASA TT F-12,805, 1970.

8. Dixon, L. C. W.: The Choice of Step Length, a Crucial Factor in the Performance of Variable Metric Algorithms. Numerical Methods for Non-Linear Optimization, F. A. Lootsma, ed., Academic Press, Inc., 1972, pp. 149-170.

9. Fletcher, R.; and Powell, M. J. D.: A Rapidly Convergent Descent Method for Minimization. Computer J., vol. 6, no. 2, pp. 163-168, July 1963.

10. Davidon, William C.: Variable Metric Method for Minimization. AEC Research and Development Report ANL-5990 Rev. (Contract W-31-109-eng-38), Argonne Nat. Lab., Nov. 1959.

TABLE I.- SUMMARY OF DFP/T PERFORMANCE ON SIX TEST PROBLEMS
WITH UPPER AND LOWER BOUND CONSTRAINTS

Problem	Starting points (x-space)	Iterations	Function evaluations
Rosenbrock (I)	(-1.2,1)	21	80
	(-0.4,1.6)	7	22
	(-1.5,-0.5)	11	57
	(0,0)	8	21
Quadratic (I)	(0.1,0.1,0.1)	8	24
	(1.5,1.5,1.5)	6	19
	(4/3,7/9,4/9)	6	13
	(2,1,0.5)	6	13
NLS	(1800,1700,1600)	5	11
	(1750,1700,1650)	3	7
	(1600,1700,1800)	4	9
Wood (I)	(-3,-1,-3,-1)	*26	*97
	(0,0,0,0)	*20	*58
	(-2,1,-2,-1)	*24	*90
	(0,1,0,-1)	*21	*73
Rational	(0.5,0.5,0.5,0.5,0.5)	7	16
	(0.9,0.7,0.5,0.3,0.1)	8	20
	(0.1,0.3,0.5,0.7,0.9)	7	20
	(0.8,0.8,0.2,0.8,0.8)	7	16
Minimum Time (I)	(0.8,0.8,0.8,0.8,0.8,0.8)	17	53
	(1,1,1,1,1,1)	23	70
	(1.8,1.5,1.2,0.9,0.6,0.3)	18	54
	(1.5,1.5,1.5,0.6,0.6,0.6)	20	59

*See discussion of Wood (I) problem in text.

TABLE II.- SUMMARY OF DFP/T PERFORMANCE ON FIVE TEST PROBLEMS WITH QUADRATIC INEQUALITY CONSTRAINTS

Problem	Starting points (x-space)	Iterations	Function evaluations
Rosenbrock (II)	(-1.2,1)	23	104
	(-1,-1)	18	94
	(0,0)	16	64
	(0.5,0)	11	38
Quadratic (II)	(0.1,0.1,0.1)	10	34
	(0,0,0)	9	34
	(4/3,7/9,4/9)	10	38
	(2,1,0)	11	39
Wood (II)	(-3,-1,-3,-1)	*45	*374
	(0,0,0,0)	25	98
	(-2,1,-2,-1)	31	209
	(-1.75,0,-1.75,0)	*21	*70
Minimum Distance	(1,1,1,1,1)	13	39
	(0.0625,0.125,0.25,0.5,1)	13	38
	(0,0,0,0,0)	13	38
	(1,0.5,0.25,0.125,0.0625)	15	43
Minimum Time (II)	(0.8,0.8,0.8,0.8,0.8,0.8)	16	41
	(1,1,1,1,1,1)	16	40
	(1.8,1.5,1.2,0.9,0.6,0.3)	16	47
	(1.5,1.5,1.5,0.6,0.6,0.6)	16	44

*See discussion of Wood (II) problem in text.

TABLE III.- SUMMARY OF DFP/T PERFORMANCE ON SEVEN TEST PROBLEMS
WITH LINEAR INEQUALITY CONSTRAINTS

Problem	Starting points (z-space)	Iterations	Function evaluations
Rosenbrock (III)	(0.7,0.5)	11	71
	(0.5,0.7)	11	62
	(0.2,0.8)	9	52
	(0.8,0.2)	7	25
Box B	(0.7,0.7)	5	13
	(0.5,0.7)	4	9
	(0.2,0.8)	5	11
	(0.7,0.5)	14	31
Quadratic (III)	(0.5,0.5,0.5)	7	15
	(0.2,0.4,0.6)	9	27
	(0.7,0.5,0.3)	7	15
POP	(0.4,0.4,0.4,0.4,0.4)	7	15
	(0.2,0.2,0.4,0.4,0.6)	9	19
	(0.6,0.6,0.2,0.2,0.2)	10	21
Modified POP	(0.35,0.35,0.35,0.35,0.35,0.35,0.35)	13	32
	(0.2,0.2,0.2,0.4,0.4,0.5,0.5)	13	32
	(0.5,0.4,0.3,0.2,0.2,0.2,0.2)	13	33
Wood (III)	(0.6,0.5,0.3,0.2)	16	44
	(0.2,0.2,0.2,0.2)	14	35
	(0.4,0.5,0.5,0.4)	17	62
	(0.2,0.3,0.5,0.7)	16	48
Maximum Product	(0.2,0.2,0.3,0.3,0.5,0.5)	5	12
	(0.5,0.5,0.3,0.3,0.2,0.2)	5	12

TABLE IV.- THE x-SPACE AND z-SPACE STARTING POINTS FOR
SEVEN TEST PROBLEMS WITH LINEAR CONSTRAINTS

Problem	Starting points (z-space)	Starting points (x-space)
Rosenbrock (III)	(0.7,0.5)	(1.63,0.66)
	(0.5,0.7)	(1.31,1.29)
	(0.2,0.8)	(1.02,1.81)
	(0.8,0.2)	(1.87,0.11)
Box B	(0.7,0.7)	(4.50,0.87)
	(0.5,0.7)	(3.93,1.12)
	(0.2,0.8)	(3.06,1.57)
	(0.7,0.5)	(4.88,0.57)
Quadratic (III)	(0.5,0.5,0.5)	(0.98,0.98,0.49)
	(0.2,0.4,0.6)	(0.120,0.79,189)
	(0.7,0.5,0.3)	(1.76,0.90,0.16)
POP	(0.4,0.4,0.4,0.4,0.4)	(24.89,10.07,10.07)
	(0.2,0.2,0.4,0.4,0.6)	(13.02,8.21,19.84)
	(0.6,0.6,0.2,0.2,0.2)	(37.71,8.08,2.41)
Modified POP	(0.35,0.35,0.35,0.35,0.35,0.35,0.35)	(11.36,6.25,14.49)
	(0.2,0.2,0.2,0.4,0.4,0.5,0.5)	(5.95,5.73,21.27)
	(0.5,0.4,0.3,0.2,0.2,0.2,0.2)	(15.68,5.27,7.01)
Wood (III)	(0.6,0.5,0.3,0.2)	(1.96,1.00,0.34,0.05)
	(0.2,0.2,0.2,0.2)	(0.95,0.71,0.48,0.12)
	(0.4,0.5,0.5,0.4)	(1.98,1.59,0.99,0.19)
	(0.2,0.3,0.5,0.7)	(1.99,1.90,1.69,0.56)
Maximum Product	(0.2,0.2,0.3,0.3,0.5,0.5)	(0.31,0.62,2.09,2.79,9.68,11.62)
	(0.5,0.5,0.3,0.3,0.2,0.2)	(1.94,3.87,2.09,2.79,1.55,1.86)

TABLE V.- INPUTS $\left(\rho_i, \lambda_i, \phi_i\right)$ FOR THE TEST PROBLEM NLS

ρ_i, km	λ_i, deg	ϕ_i, deg	ρ_i, km	λ_i, deg	ϕ_i, deg	ρ_i, km	λ_i, deg	ϕ_i, deg
1737.77	149.541	-86.453	1738.63	173.850	-83.543	1737.81	-172.550	-80.418
1738.41	-162.013	-77.227	1738.59	-152.740	-73.997	1737.62	-144.111	-70.739
1738.39	-135.863	-67.452	1737.56	-127.863	-64.136	1737.92	-120.040	-60.791
1737.14	-112.350	-57.414	1738.49	-104.767	-54.006	1736.63	-97.272	-50.565
1737.29	-89.854	-47.091	1737.39	-82.503	-43.584	1737.66	-75.212	-40.046
1737.71	-67.974	-36.476	1737.77	-60.784	-32.877	1738.27	-53.637	-29.251
1739.09	-46.529	-25.600	1738.92	-39.454	-21.927	1739.77	-32.408	-18.233
1739.54	-25.387	-14.523	1739.86	-18.386	-10.800	1740.05	-11.399	-7.067
1741.25	-4.423	-3.327	1741.26	2.547	.416	1741.13	9.517	4.159
1741.02	16.490	7.897	1740.88	23.472	11.628	1740.11	30.468	15.349
1740.83	37.482	19.055	1739.88	44.518	22.745	1738.70	51.582	26.416
1739.68	58.678	30.064	1738.08	65.811	33.688	1739.10	72.986	37.286
1737.77	80.207	40.856	1737.72	87.480	44.397	1736.45	94.811	47.908
1737.47	102.209	51.389	1736.86	109.683	54.839	1736.18	117.246	58.259
1735.73	124.918	61.650	1735.00	132.727	65.012	1734.26	140.722	68.347
1735.00	148.983	71.655	1735.15	157.663	74.936	1734.51	167.086	78.188
1734.94	178.053	81.400	1734.87	-166.825	84.529	1734.91	-133.977	87.314
1734.22	-142.465	86.400	1734.15	-152.161	82.956	1734.30	-150.821	79.318
1734.31	-146.664	75.639	1733.70	-141.395	71.951	1734.13	-135.571	68.267
1734.55	-129.426	64.595	1734.55	-123.072	60.938	1733.78	-116.572	57.301
1734.63	-109.962	53.686	1734.83	-103.265	50.097	1733.79	-96.495	46.534
1735.31	-89.662	42.999	1734.81	-82.773	39.494	1735.79	-75.835	36.018
1735.28	-68.852	32.572	1736.37	-61.827	29.156	1737.29	-54.766	25.768
1736.55	-47.672	22.406	1737.40	-40.548	19.070	1738.23	-33.399	15.757
1738.90	-26.228	12.463	1739.56	-19.041	9.186	1739.92	-11.840	5.921
1740.47	-4.631	2.667	1740.97	2.582	-.583	1742.12	9.794	-3.831
1741.65	17.001	-7.081	1741.71	24.198	-10.339	1742.08	31.381	-13.607
1740.70	38.545	-16.889	1742.60	45.686	-20.189	1741.42	52.800	-23.511
1741.58	59.883	-26.856	1740.64	66.930	-30.227	1741.01	73.939	-33.627
1741.50	80.904	-37.056	1739.78	87.821	-40.515	1740.29	94.685	-44.005
1740.66	101.489	-47.525	1738.78	108.226	-51.075	1739.82	114.881	-54.653
1738.96	121.440	-58.259	1739.47	127.873	-61.888	1738.84	134.137	-65.538
1739.44	140.156	-69.206	1739.38	145.786	-72.885	1738.34	150.720	-76.567
1737.56	154.195	-80.234	1738.71	153.672	-83.836	1738.08	135.133	-87.080
1738.39	91.990	-87.938	1738.41	-176.064	-.101	1736.86	-85.962	-2.068
1742.05	4.063	-.031	1740.50	93.963	1.932	1734.60	-85.923	88.062

"*The aeronautical and space activities of the United States shall be
conducted so as to contribute . . . to the expansion of human knowl-
edge of phenomena in the atmosphere and space. The Administration
shall provide for the widest practicable and appropriate dissemination
of information concerning its activities and the results thereof.*"
—NATIONAL AERONAUTICS AND SPACE ACT OF 1958

NASA SCIENTIFIC AND TECHNICAL PUBLICATIONS

TECHNICAL REPORTS: Scientific and
technical information considered important,
complete, and a lasting contribution to existing
knowledge.

TECHNICAL NOTES: Information less broad
in scope but nevertheless of importance as a
contribution to existing knowledge.

TECHNICAL MEMORANDUMS:
Information receiving limited distribution
because of preliminary data, security classifica-
tion, or other reasons. Also includes conference
proceedings with either limited or unlimited
distribution.

CONTRACTOR REPORTS: Scientific and
technical information generated under a NASA
contract or grant and considered an important
contribution to existing knowledge.

TECHNICAL TRANSLATIONS: Information
published in a foreign language considered
to merit NASA distribution in English.

SPECIAL PUBLICATIONS: Information
derived from or of value to NASA activities.
Publications include final reports of major
projects, monographs, data compilations,
handbooks, sourcebooks, and special
bibliographies.

TECHNOLOGY UTILIZATION
PUBLICATIONS: Information on technology
used by NASA that may be of particular
interest in commercial and other non-aerospace
applications. Publications include Tech Briefs,
Technology Utilization Reports and
Technology Surveys.

Details on the availability of these publications may be obtained from:

SCIENTIFIC AND TECHNICAL INFORMATION OFFICE

NATIONAL AERONAUTICS AND SPACE ADMINISTRATION
Washington, D.C. 20546