A Compact Algorithm for Computing the Stationary Point of a Quadratic Function Subject to Linear Constraints

J. T. BETTS

Configuration Group
The Aerospace Corporation
El Segundo, Calif. 90245

8 July 1978

APPROVED FOR PUBLIC RELEASE;
DISTRIBUTION UNLIMITED

Prepared for
SPACE AND MISSILE SYSTEMS ORGANIZATION
AIR FORCE SYSTEMS COMMAND
Los Angeles Air Force Station
P.O. Box 92960, Worldway Postal Center
Los Angeles, Calif. 90009

78 08 31 023
This final report was submitted by The Aerospace Corporation, El Segundo, CA 90245, under Contract F04701-77-C-0078 with the Space and Missile Systems Organization, Deputy for Advanced Space Programs, P.O. Box 92960, Worldway Postal Center, Los Angeles, CA 90009. It was reviewed and approved for The Aerospace Corporation by J. R. Allder, Engineering Group. First Lieutenant A. G. Fernandez, YCPT, was the Deputy for Advanced Space Programs project engineer.

This report has been reviewed by the Information Office (OI) and is releasable to the National Technical Information Service (NTIS). At NTIS, it will be available to the general public, including foreign nations.

This technical report has been reviewed and approved for publication. Publication of this report does not constitute Air Force approval of the report's findings or conclusions. It is published only for the exchange and stimulation of ideas.

A. G. Fernandez, 1st Lt, USAF
Project Engineer

Robert W. Lindemuth, Lt Col, USAF
Chief, Technology Plans Division

FOR THE COMMANDER

LEONARD E. BALTZELL, Col, USAF, Asst, Deputy for Advanced Space Programs
A Compact Algorithm for Computing the Stationary Point of a Quadratic Function Subject to Linear Constraints

An algorithm is presented for computing the stationary point of a quadratic function of \( n \) variables subject to a set of \( m(m \leq n) \) linear equality constraints. The procedure is compact in the sense that it requires no two-dimensional arrays of computer storage beyond that needed to store the problem data. The use of a Householder orthogonal decomposition by the method should not degrade the numerical conditioning of the original problem. The method is applicable to problems with singular Hessian matrices, and can be adapted for use in a general quadratic programming algorithm.
CONTENTS

1. INTRODUCTION ........................................... 3
2. PRELIMINARIES. ............................................ 5
3. ORTHOGONAL DECOMPOSITION ALGORITHM .................... 9
4. COMPUTATIONAL ALGORITHM ............................... 13
5. SUMMARY ................................................... 17

APPENDIX. THE STATIONARY POINT OF A QUADRATIC
FUNCTION SUBJECT TO LINEAR CONSTRAINTS. . . . . 19
1. INTRODUCTION

This report presents an algorithm for computing the stationary point of a quadratic function of \( n \) variables subject to a set of \( m (m \leq n) \) linear equality constraints. The procedure is compact in the sense that it requires no two-dimensional arrays of computer storage beyond that needed to store the problem data. The use of a Householder orthogonal decomposition by this method should not degrade the numerical conditioning of the original problem. This method is applicable to problems with singular Hessian matrices, and can be adapted for use in a general quadratic programming algorithm.

In the subsequent sections of this report, the identifying numbers of equations in the text are enclosed with parentheses, and the identifying numbers of references are enclosed with brackets.
2. PRELIMINARIES

Define the quadratic function

\[ f(x) = \frac{1}{2} x^T A x + b^T x \]  \hspace{1cm} (1)

and the linear constraints

\[ C x = d \]  \hspace{1cm} (2)

where \( A \) is an \( n \times n \) symmetric matrix, \( b \) is an \( n \)-vector, \( C \) is an \( m \times n \) matrix of rank \( m \), \( d \) is an \( m \)-vector, and \( x \) is an \( n \)-vector for \( m \leq n \). Define the solution \( x^* \) to be the vector which: (a) satisfies the constraints, (b) minimizes the norm of the gradient of \( f \) restricted to the constraint surface and, (c) minimizes the length of the orthogonal projection of \( x \) on the constraint surface.

When \( A \) is positive (negative) definite, the solution defines the unique stationary point which corresponds to the minimum (maximum) of \( f \) restricted to the constraint surface. While the stated problem may be of interest by itself, typically it may appear as a subproblem in a more general application. For example, many quadratic programming algorithms solve a series of problems of this type with different constraint sets. Furthermore, an algorithm designed to optimize a non-quadratic function subject to nonlinear constraints may pose a series of quadratic-linear problems to approximate the behavior of the actual functions. Consequently, it is desirable to develop a computational algorithm which will compute a solution to the problem without restricting the rank of \( A \). The computed solution should be the unique solution when it exists, and should be uniquely defined by the algorithm when a unique solution does not exist.

When the minimum norm of the projected gradient is zero, the solution to the stated problem is a stationary point of the Lagrangian function
\[ L(x, \lambda) = \frac{1}{2} x^T A x + b^T x + \lambda^T (C x - d) \]  

where \( \lambda \) is the \( m \)-vector of Lagrange multipliers. Setting the derivative of this function with respect to \( x \) and \( \lambda \) equal to zero yields the set of linear equations

\[
\begin{bmatrix}
A & C^T \\
C & 0
\end{bmatrix}
\begin{bmatrix}
x^* \\
\lambda^*
\end{bmatrix}
= 
\begin{bmatrix}
-b \\
d
\end{bmatrix}.
\]  

(4)

The optimal solution \( x^* \) and the corresponding multiplier values \( \lambda^* \) can be obtained by solving the system (4). It is not necessary to assume that \( A \) is of full rank. When the problem has a unique solution, the system may be solved using a suitable algorithm for linear equations. If the possibility of a non-unique solution exists, the system may be solved as a linear least squares problem. This approach has been used \(^1,^2\) utilizing the linear least squares algorithm of Hanson and Lawson \(^3\). A defect in this approach is the need to store the \( (n + m) \times (n + m) \) coefficient array.

An approach which can be implemented using only the storage required for \( A \) and \( C \) can be derived by inverting the coefficient matrix in a partitioned

---


form. It is easily demonstrated that

\[ \begin{bmatrix} A & C^T \\ C & O \end{bmatrix}^{-1} = \begin{bmatrix} B_1 & B_2 \\ B_2^T & B_3 \end{bmatrix} \] (5)

where

\[ B_1 = A^{-1} - A^{-1} C^T M^{-1} C A^{-1} \] (6)

\[ B_2 = A^{-1} C^T M^{-1} \] (7)

\[ B_3 = M^{-1} \] (8)

and

\[ M = C A^{-1} C^T \] (9)

The partitioned form of the inverse plays an important role in a number of quadratic programming algorithms, including those of Goldfarb\(^4\) and Fletcher\(^5\), as well as in the constrained minimization algorithm of Murtagh and Sargent\(^6\). Two significant points deserve comment regarding this approach. First, if it is assumed that \( A^{-1} \) exists, the submatrices in (5) can be computed directly. Goldfarb, Murtagh, and Sargent assume that \( A \) is

---


positive definite, which ensures that \( A \) is of full rank. If the algorithm is part of a more general nonlinear programming algorithm as in [1] and [2], it is restrictive to assume that \( A \) is of full rank. For example, the approach could not be used as part of an algorithm to optimize a linear function subject to nonlinear constraints. In these general applications, it is really only necessary that the Hessian matrix restricted to the constraint surface be definite. Fletcher does not assume \( A \) is definite, noting that the partitions \( B_1, B_2, \) and \( B_3 \) must exist if the solution to the problem is unique. However, to compute the initial submatrices in the computer implementation of his quadratic programming algorithm\(^7\), it is necessary to invert the full \((n+m) \times (n+m)\) matrix.

Even if \( A \) is assumed to be definite, the partitioned approach to the problem suffers from a second defect. This occurs when \( A = I, M = C^T C \), which is referred to as the normal matrix. In this case, the condition number of \( M \) is the square of the condition number of \( C \) and it is generally recognized that the formation of \( M \) is to be avoided.

In summary, direct solution of the system (4) is not compact from a storage standpoint. The various forms of solving the partitioned system (5), although compact, require operations which can degrade the numerical conditioning of the given problem and are arbitrarily restrictive with regard to Hessian matrix \( A \). A new algorithm will be proposed which is compact, does not degrade the numerical conditioning, and makes no restrictions concerning the rank of \( A \).

3. ORTHOGONAL DECOMPOSITION ALGORITHM

In this section, an algorithm is developed for solving the stated constrained stationary point problem using an orthogonal decomposition of the constraint matrix. The algorithm is an extension of the linearly constrained linear least squares algorithm LSE given in [3], and makes use of Theorem (3.19) and Theorem (2.3) stated therein.

Define the orthogonal decomposition of $C$ by

$$ C = RK^T $$

where $K$ is an $n \times n$ orthogonal matrix and $R$

$$ R = \begin{bmatrix} R_{11} & 0 \\ \end{bmatrix} $$

where $R_{11}$ is an $m \times m$ nonsingular triangular matrix. Substituting (10) into (2)

$$ RK^T x = d $$

Define the $n$-vector $y$ by

$$ y = K^T x $$

and the partitions of $K$ and $y$

$$ K = \begin{bmatrix} K_1 & K_2 \\ \frac{m}{n-m} & \end{bmatrix} $$

$$ y = \begin{bmatrix} y_1 \\ y_2 \end{bmatrix} $$
From (11), (13), and (15), one can write (12) as

$$ R K^T x = Ry = [R_{11} \ 0] y_2 = R_{11} y_1 = d $$  \hspace{1cm} (16)

Since $R_{11}$ is non-singular, (16) determines the m-vector $y_1$. Call the solution $\hat{y}_1$. The (n-m)-vector $y_2$ is arbitrary.

Pre-multiply (13) by $K$ to give

$$ Ky = K K^T x = x $$  \hspace{1cm} (17)

where $KK^T = I$, since $K$ is orthogonal. From (14) and (15), (17) becomes

$$ x = Ky = [K_1 \ K_2] [y_1 \ y_2] = K_1 y_1 + K_2 y_2 $$  \hspace{1cm} (18)

Observe that all points satisfying the constraints in Eq. (2) can be represented as functions of the n-m parameters $y_2$ when the solution of (16) $\hat{y}_1$ is substituted in Eq. (18). In fact if there were no other conditions to satisfy, a reasonable choice for the arbitrary parameters $y_2$ would be zero, in which case $x$ is the minimum norm solution to the constraints.

Instead of setting $y_2 = 0$, the choice of $y_2$ shall be determined by a different criterion. Substitute the parametric representation of $x$ from (18) with $y_1 = \hat{y}_1$ into (1) to obtain

$$ f = \hat{y}_1^T (K_1 \hat{y}_1 + K_2 y_2)^T A (K_1 \hat{y}_1 + K_2 y_2) $$

$$ + b^T (K_1 \hat{y}_1 + K_2 y_2) $$  \hspace{1cm} (19)

The gradient with respect to the variables $y_2$ is

$$ \nabla f = K_2^T A (K_1 \hat{y}_1 + K_2 y_2) + K_2^T b. $$ \hspace{1cm} (20)

-10-
Let us define $\hat{y}_2$ to be the value of $y_2$ which minimizes

$$\| \nabla f \| = \| K_2^T A K_2 y_2 + (K_2^T b + K_2^T A K_1 \psi_1) \|$$

and is of minimum norm, i.e., minimizes $\| y_2 \|$. If a stationary point of $f$ restricted to the constraint surface exists, then $\| \nabla f \| = 0$ and $\hat{y}_2$ defines the optimal value. If the matrix $K_2^T A K_2$ is indefinite, the minimum norm criterion uniquely determines $\hat{y}_2$. In fact, when the rank of $K_2^T A K_2$ is zero as is the case for a linear objective function, the solution which minimizes the norm of $\| y_2 \|$ is just $\hat{y}_2 = 0$. Notice also that the formation of the matrix $K_2^T A K_2$ should not degrade the conditioning of the original problem.

Observe also that a solution which minimizes $\| y_2 \|$ minimizes $\| K_2 y_2 \|$ since $\| y_2 \| = \| K_2 y_2 \|$, and $K_2 y_2$ is just the orthogonal component of $x$ in the constraint surface.

In summary, the original constrained optimization problem is replaced by a lower dimensional unconstrained least squares problem in the variables $y_2$, after choosing the variables $y_1$ to satisfy the constraints. The method has the property that the unique minimum length solution of the derived unconstrained problem defines the unique solution of the constrained problem when it exists. When the constrained problem has no unique solution, the algorithm computes a unique point which satisfies the constraints, minimizes the norm of the gradient on the constraint surface, and minimizes the length of the orthogonal component in the constraint surface.
4. COMPUTATIONAL ALGORITHM

In this section a computational procedure based on the approach derived in Section 3 is developed. The procedure is organized so that no additional two-dimensional arrays are needed. Specifically, the original problem data stored in A, C and b is destroyed by the algorithm. Quantities written with a tilde can replace quantities without a tilde in storage, and quantities written with a circumflex can overwrite quantities written with a tilde.

Step 1. Compute the orthogonal matrix \( K \) and postmultiply \( C \) by it to triangularize \( C \), i.e.,

\[
CK = \begin{bmatrix} \tilde{C}_1 & 0 \end{bmatrix}_m \begin{bmatrix} m & n-m \end{bmatrix}
\] (22)

Step 2. Compute

\[
\tilde{A} = K^T_A
\] (23)

Step 3. Form the last \( n-m \) rows of the matrix \( \tilde{A} \) where

\[
\hat{A} = \tilde{A}K
\] (24)

Observe that from (23) and (24)

\[
\hat{A} = K^T_AK = \begin{bmatrix} K_1^TAK_1 & K_1^TAK_2 \end{bmatrix}_m \begin{bmatrix} m & n-m \end{bmatrix}
\]

\[
= \begin{bmatrix} \hat{A}_{11} & \hat{A}_{12} \\ \hat{A}_{21} & \hat{A}_{22} \end{bmatrix}
\] (25)
Step 4. Compute
$$\tilde{b} = K^T b$$ (26)

Step 5. Solve the lower triangular system
$$\tilde{c}_1 y_1 = d$$ (27)
for the $m$-vector $\hat{y}_1$.

Step 6. Compute
$$\tilde{b}_2 = -\tilde{b}_1 - A_{21} \hat{y}_1$$ (28)
where
$$\tilde{b} = \begin{bmatrix} b_1 \mid m \\ \hat{b}_2 \mid n-m \end{bmatrix}, \quad \tilde{b}_1 = \begin{bmatrix} \hat{b}_1 \mid m \\ \hat{b}_2 \mid n-m \end{bmatrix}$$ (29)

Step 7. Determine $\hat{y}_2$ as the minimum length solution of the linear least square problem
$$\min \| A_{22} y_2 - \hat{b}_2 \|.$$ (30)
Observe that this process is equivalent to solving (21)

Step 8. Construct the solution vector
$$x = K \hat{y}$$ (31)
using $\hat{y}_1$ as computed in Step 5, and $\hat{y}_2$ from Step 7.
The algorithm described has been implemented in the subroutine HSQP. The FORTRAN listing of this subroutine is found in [8]. The subroutine makes extensive use of the subroutines HFTI and H12 which implement the algorithms referred to as HFTI, H1, and H2 in [3]. The subroutine HFTI computes the minimum length solution to a linear least squares problem. HFTI requires storage for the problem data and three one-dimensional work arrays. The subroutine H12 implements algorithm H1 and H2 for the construction and application of a Householder transformation. Using H12 it is not necessary to explicitly form the orthogonal matrix $K$ of Eq. (10). Instead, the elements necessary to construct the matrix can be stored in the upper triangular portion of the original matrix $C$ and some one-dimensional work arrays. Successive applications of the matrix $K$ to other vectors implicitly reconstruct the original transformations.

The total storage required for subroutine HSQP, including that required to specify the problem data, is $N_1 = n(m + n) + 5n - m + 4$. In contrast, any algorithm which solves (4) directly will require at least $N_2 = (n + m)(n + m + 1)$ storage locations. Consequently, for some problems $N_2$ can be nearly twice as large as $N_1$. The algorithm is used repeatedly as part of the general nonlinear programming algorithm described in [1]. In particular, all of the extrapolation steps used in the constraint phase of this algorithm employ HSQP. Computational experience with the algorithm includes its use to solve the set of 17 equality constrained and 34 inequality constrained problems in [1], as well as a number of larger engineering applications. One typical application is described in [9]. The largest engineering application of the algorithm to date occurred in an optimum solid rocket motor design problem which involved 48 variables and 83 constraints.


5. SUMMARY

An algorithm for computing the stationary point of a quadratic function of \( n \) variables subject to \( m \) linear equality constraints is developed. The algorithm has been implemented in FORTRAN. The implementation is compact since it requires no two-dimensional arrays beyond that needed to define the problem. The algorithm avoids mathematical operations which would degrade the conditioning of the original problem by utilizing an orthogonal decomposition of the constraint matrix. The solution generated by the algorithm is characterized by three properties: (a) the constraints are satisfied, (b) the norm of the gradient of the objective function restricted to the constraint surface is minimized and, (c) among all solutions satisfying the first two properties, the minimum length solution is chosen. When the stated problem has a unique solution, satisfaction of the first two properties defines the point. Nevertheless, the algorithm is not restricted to problems with definite Hessian matrices. The algorithm has been successfully tested as part of a general nonlinear programming algorithm.
APPENDIX

THE STATIONARY POINT OF A QUADRATIC FUNCTION
SUBJECT TO LINEAR CONSTRAINTS

This algorithm implements the method developed in the preceding sections of this report.

SUBROUTINE HSQP(A,B,C,D,M,N,TAU,G,H,U,IP,MAXRA,MAXRC,DJNORM,X,
$\quad \text{K RANK}$)

DIMENSION B(1),G(1),D(1),H(1),U(1),IP(1),DJNORM(1),X(1)
DIMENSION A(MAXRA,1),C(MAXRC,1)


PURPOSE: GIVEN AN M X N MATRIX C (OF RANK M), AN M VECTOR D,
AN N X N SYMMETRIC MATRIX A, AND AN N VECTOR B, FIND THE
STATIONARY POINT X OF THE QUADRATIC
$J = 0.5*(X^T)*A*X + (B^T)*X$

SUBJECT TO THE CONSTRAINTS
$C*X = D.$

IF A STATIONARY POINT DOES NOT EXIST THE ALGORITHM WILL FIND
A POINT WHICH SATISFIES THE CONSTRAINTS AND MINIMIZES THE
NORM OF THE GRADIENT OF J PROJECTED ON THE CONSTRAINT SURFACE.

ALGORITHM: ORTHOGONAL DECOMPOSITION OF C MATRIX USING
HOUSEHOLDER TRANSFORMATIONS, FOLLOWED BY APPLICATION OF THE
OPTIMALITY CONDITIONS IN THE PREDUCED VARIABLES.

INPUT:

A  N X N SYMMETRIC HESSIAN MATRIX
B  N DIMENSIONAL GRADIENT VECTOR
C  M X N JACOBIAN MATRIX (RANK M)
D  M DIMENSIONAL CONSTRAINT VECTOR
N  THE NUMBER OF VARIABLES
TAU PSEUDORANK TEST PARAMETER. FOR A MACHINE WITH K
SIGNIFICANT FIGURES AN APPROPRIATE VALUE IS
TAU = 1.E-(K+2).
G  AUXILIARY STORAGE (LENGTH M)
H  AUXILIARY STORAGE (LENGTH N-M)
U  AUXILIARY STORAGE (LENGTH N-M)
IP AUXILIARY STORAGE (LENGTH N-M)
MAXRA MAXIMUM ROW DIMENSION OF A (MAXRA = N)
MAXRC MAXIMUM ROW DIMENSION OF C (MAXRC = M)

OUTPUT:

DJNORM PROJECTED GRADIENT NORM (ZERO IF X IS A STATIONARY
POINT, NEGATIVE IF THERE IS AN INPUT ERROR)
X COMPUTED STATIONARY POINT
KRANK PSEUDORANK OF PROJECTED HESSIAN MATRIX (K2**T)*A*K2.
WHEN KRANK .LT. N-M THE PROJECTION OF X ON THE
CONSTRAINT SURFACE HAS MINIMUM NORM.

NOTE: THE INPUT VALUES OF A, B, C, AND D ARE DESTROYED.

INITIALIZATION

KRANK = 0
MP1 = M + 1
NMM = N - M
DJNORM(1) = -1.

CHECK FOR INPUT ERRORS

IF(N.EQ.0. OR.N.GT.MAXRA. OR.M.GT.MAXRC. OR.M.GT.N) RETURN
IF THE PROBLEM IS UNCONstrained GO TO STEP 7

IF(M.EQ.0) GO TO 100

-----------------------------------------------------------------------------------------------------------------------------

STEP 1. COMPUTE ORTHOGONAL MATRIX K. TRIANGULARIZE C.

DO 10 I = 1, M
CALL H12(I, I+1, N, C(I, 1), MAXRC, G(I), C(I+1, 1), MAXRC, 1, M-I)
10 CONTINUE

IF(QI.EQ.N) GO TO 50

-----------------------------------------------------------------------------------------------------------------------------

STEP 2. COMPUTE ATILDA = (K**T).*A

DO 20 I = 1, M
CALL H12(2, I, I+1, N, C(I, 1), MAXRC, G(I), A, 1, MAXRA, N)
20 CONTINUE

-----------------------------------------------------------------------------------------------------------------------------

STEP 3. FORM THE LAST N-M ROWS OF AHAT = ATILDA*K; I.E.

COMPUTE A21HAT = (K2**T).*A*K1 AND A22HAT = (K2**T).*A*K2

DO 30 I = 1, M
CALL H12(2, I, I+1, N, C(I, 1), MAXRC, G(I), A(MP1, 1), MAXRA, 1, NMN)
30 CONTINUE

-----------------------------------------------------------------------------------------------------------------------------

STEP 4. COMPUTE BTILDA = (K**T).*B

DO 40 I = 1, M
CALL H12(2, I, I+1, N, C(I, 1), MAXRC, G(I), B, 1, 1, 1)
40 CONTINUE

-----------------------------------------------------------------------------------------------------------------------------

STEP 5. COMPUTE Y1HAT BY SOLVING THE LOWER TRIANGULAR SYSTEM C*Y1 = D. STORE IN X.

50 CONTINUE
\[ X(I) = \frac{D(I)}{C(I,1)} \]

**IF** \( M = 1 \) **GO TO** 80

**DO** 70 \( I = 2, N \)

**IM1** = \( I - 1 \)

\[ X(I) = \frac{D(I)}{} \]

**DO** 60 \( J = 1, IM1 \)

\[ X(I) = X(I) - C(I,J) \times X(J) \]

**CONTINUE**

\[ X(I) = \frac{X(I)}{C(I,I)} \]

**CONTINUE**

**CONTINUE**

**WHEN THERE ARE NO DEGREES OF FREEDOM GO TO STEP 8**

**IF** \( M = N \) **GO TO** 140

**STEP 6. COMPUTE** \( B2HAT = -B2TILDA - A21HAT \times Y1HAT \)

**DO** 93 \( I = MP1, N \)

\[ B(I) = -B(I) \]

**DO** 93 \( J = 1, M \)

\[ B(I) = B(I) - A(I,J) \times X(J) \]

**CONTINUE**

**STEP 7. SOLVE** \( A22HAT \times Y2 = B2HAT \) **FOR** \( Y2 \) **USING HFTI**

**CONTINUE**

**COMPUTE PSEUDORANK TEST PARAMETER EPS**

\[ \text{EPS} = \frac{ \text{TAU} }{ \text{MAX1} \left( \text{EPS}, \frac{\text{TAU} \times \text{SQRT(COLNRM)}}{\text{COLNRM}} \right) } \]

**CALL HFTI(A(MP1, MP1), MAXRA, NNN, NNN, B(MP1), 1, 1, EPS, KRANK, DJNORM),**
$H, U, IP$

DO 130 I = MP1, N
X(I) = R(I)
130 CONTINUE

IF THE PROBLEM IS UNCONSTRAINED, RETURN.
IF (M EQ. 0) RETURN

--------------------------------------------------------

STEP 8. COMPUTE $X = K \cdot Y$

CONTINUE
DO 150 K = 1, M
I = MP1 - K
CALL H12 (2, I, I+1, N, C(I, 1), MAXRC, G(I), X, 1, 1, 1)
150 CONTINUE
RETURN
END