MATHEMATICS OF COMPUTATION
VOLUME 41, NUMBER 163
JULY 1983, PAGES 165-170

The Numerical Solution of Equality-Constrained Quadratic Programming Problems*

By Nira Dyn and Warren E. Ferguson, Jr.

Abstract. This paper proves that a large class of iterative schemes can be used to solve a certain constrained minimization problem. The constrained minimization problem considered involves the minimization of a quadratic functional subject to linear equality constraints. Among this class of convergent iterative schemes are generalizations of the relaxed Jacobi, Gauss-Seidel, and symmetric Gauss-Seidel schemes.

1. Introduction. In this paper we will present several iterative schemes which solve the following constrained minimization problem.

Problem 1. Find the real n-vector x_* which minimizes $f(x) \equiv \frac{1}{2}x^TAx - x^Tr$ subject to the constraints $g(x) \equiv E^Tx - s = 0$.

Here A is a real symmetric nonnegative definite $n \times n$ matrix, E is a real $n \times m$ matrix with full column rank, r is a real n-vector, and s is a real m-vector.

As discussed in Section 2, the theory of quadratic programming [7] states that under reasonable conditions on A and E the solution of Problem 1 exists and is unique. Furthermore, under these conditions on A and E the solution x_* of Problem 1 forms part of the solution (x_*, λ_*) of the following problem.

Problem 2: Find the real n-vector x_* and the real m-vector λ_* which solves the linear system

$$\begin{pmatrix} A & E \\ E^T & 0 \end{pmatrix}\begin{pmatrix} x \\ \lambda \end{pmatrix} = \begin{pmatrix} r \\ s \end{pmatrix}.$$

In Section 3 we describe the convergence of a large class of iterative schemes used to solve Problem 2, and hence Problem 1. Although our iterative schemes are generally applicable to these problems, they are typically efficient only when A is a large sparse matrix and there are only a moderate number of constraints. In this situation the usual methods used to solve these problems become inefficient.

Our work was motivated by the work of [12], in which a variant of one of the iterative schemes described in this paper was used to numerically construct a smooth surface from aggregated data. This application is also analyzed in [4], [5]. The numerical solution of Problems 1 and 2 has also been considered by other authors; in particular we mention the work presented in [6], [8], [10]. The numerical solution

Received January 5, 1982.

1980 *Mathematics Subject Classification*. Primary 49D40, 65F10, 65N20.

Key words and phrases. Constrained minimization, quadratic programming, iterative schemes, indefinite systems of equations.

*Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.

0025-5718/83/0000-1294/$02.25

of quadratic programs subject to inequality constraints by iterative methods has also been considered in [1], [2], [3], [9].

The authors would like to thank the Mathematics Research Center at the University of Wisconsin–Madison for providing a stimulating intellectual atmosphere during our stay in 1978 and 1979.

2. Preliminaries. In the previous section we stated that under reasonable conditions on A and E the solution x_* of Problem 1 is part of the solution (x_*, λ_*) of Problem 2. This statement is contained in the following theorem.

THEOREM 2.1. *Assume that*
(a) *A is a real symmetric nonnegative definite matrix,*
(b) *E is a real matrix with full column rank, and*
(c) *A and E^T have no nontrivial null vectors in common.*
Then the solutions of Problems 1 *and* 2 *exist and are unique. Furthermore if x_* is the solution of Problem* 1, *then $(x_*, (E^TE)^{-1}E^T(r - Ax_*))$ is the solution of Problem* 2; *if (x_*, λ_*) is the solution of Problem* 2, *then x_* is the solution of Problem* 1.

Proof. See the treatment of quadratic programming given in [7]. □

COROLLARY 2.2. *Under the assumptions of Theorem* 2.1 *the matrix*

$$\begin{pmatrix} A & E \\ E^T & 0 \end{pmatrix}$$

is nonsingular.

Proof. Observe that this matrix is the coefficient matrix of the linear system in Problem 2. Under the assumptions of Theorem 2.1 this linear system has only unique solutions. Therefore, as shown in [11], the coefficient matrix is nonsingular. □

The iterative schemes we use to solve Problem 2 are all based upon a splitting

$$A = B - C$$

of the matrix A, and they have the following form. Given an initial iterate (x_0, λ_0), define (x_k, λ_k) for $k = 1, 2, 3, \ldots$ to be the solution of the linear system

$$(1) \qquad \begin{pmatrix} B & E \\ E^T & 0 \end{pmatrix}\begin{pmatrix} x_k \\ \lambda_k \end{pmatrix} = \begin{pmatrix} c & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} x_{k-1} \\ \lambda_{k-1} \end{pmatrix} + \begin{pmatrix} r \\ s \end{pmatrix}.$$

Of course, for this linear stationary iterative method to be well defined it is necessary and sufficient that the matrix

$$\begin{pmatrix} B & E \\ E^T & 0 \end{pmatrix}$$

be nonsingular. This problem is addressed by the following theorem.

THEOREM 2.3. *In addition to the assumptions of Theorem* 2.1 *let*
(d) $A = B - C$,
(e) *B be a real nonsingular matrix, and*
(f) *$2A + C + C^T$ be a positive definite matrix.*
Then the iterative scheme (1) *is well defined.*

Proof. Since B is nonsingular, then

$$\begin{pmatrix} B & E \\ E^T & 0 \end{pmatrix} = \begin{pmatrix} B & 0 \\ E^T & -E^TB^{-1}E \end{pmatrix} \begin{pmatrix} I & B^{-1}E \\ 0 & I \end{pmatrix}.$$

Therefore it follows that the matrix on the left-hand side of the equality is nonsingular if and only if $E^TB^{-1}E$ is nonsingular. To prove that $E^TB^{-1}E$ is nonsingular, let us prove that $E^TB^{-1}E$ has no nontrivial null vector. If $E^TB^{-1}E\lambda = 0$, then

$$\begin{aligned} 0 &= \lambda^TE^TB^{-1}E\lambda = (B^{-1}E\lambda)^TB^T(B^{-1}E\lambda) \\ &= \tfrac{1}{2}(B^{-1}E\lambda)^T(B + B^T)(B^{-1}E\lambda) = \tfrac{1}{2}(B^{-1}E\lambda)^T(2A + C + C^T)(B^{-1}E\lambda). \end{aligned}$$

This implies $B^{-1}E\lambda = 0$, since $2A + C + C^T$ is a positive definite matrix, and so $\lambda = 0$, because E is a matrix with full column rank. □

Let us now describe one procedure for solving the linear system (1) for (x_k, λ_k).

Step 1. Solve
$$B\hat{x}_k = Cx_{k-1} + r$$
for $\hat{x}_k$.

Step 2. Solve
$$(E^TB^{-1}E)\lambda_k = E^T\hat{x}_k - s$$
for λ_k.

Step 3. Solve
$$B(x_k - \hat{x}_k) = -E\lambda_k$$
for $x_k - \hat{x}_k$.

In the next section we will see that if assumption (f) of Theorem 2.3 is slightly strengthened, then the iterative scheme (1) is not only well defined but also convergent.

3. Convergence of the Iterative Schemes. One set of conditions which guarantees the convergence of the iterative scheme (1) is described in the following theorem.

THEOREM 3.1. *Assume that*
(a) *A is a real symmetric nonnegative definite matrix,*
(b) *E is a real matrix with full column rank,*
(c) *A and E^T have no nontrivial null vectors in common,*
(d) $A = B - C$,
(e) *B is a real nonsingular matrix, and*
(f′) *$A + C + C^T$ is a positive definite matrix.*
Then the iterative scheme (1) *is well defined and convergent.*

Proof. From Theorem 2.1 we deduce that a solution of Problem 2 exists and is unique. Since A is a nonnegative definite matrix and $A + C + C^T$ is a positive definite matrix, then $2A + C + C^T$ is a positive definite matrix. From Theorem 2.3 we therefore deduce that the iterative scheme (1) is well defined.

As shown in [13], the iterative scheme (1) is convergent if and only if each eigenvalue of the matrix

$$\begin{pmatrix} B & E \\ E^T & 0 \end{pmatrix}^{-1} \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix} \tag{2}$$

has magnitude less than one. Let us therefore show that if μ is a nonzero eigenvalue of the matrix (2), then the magnitude of μ is less than one.

Since μ is an eigenvalue of the matrix (2), then there are complex vectors u, v not both zero for which

$$\mu\begin{pmatrix} B & E \\ E^T & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix} = \begin{pmatrix} C & 0 \\ 0 & 0 \end{pmatrix}\begin{pmatrix} u \\ v \end{pmatrix}. \tag{3}$$

Let us now argue that $u \neq 0$ and $u^H A u > 0$. Since A is a real symmetric nonnegative definite matrix, then clearly $u^H A u \geqslant 0$, and $u^H A u = 0$ only if $Au = 0$. However, $Au = 0$ only if $u = 0$, for (3) states that $E^T u = 0$, and by hypothesis A and E^T have no nontrivial null vectors in common. But $u = 0$ only if $v = 0$, for by hypothesis E has full column rank, and (3) implies that $Ev = 0$ when $u = 0$. Since u, v are not both zero, then we conclude $u \neq 0$ and $u^H A u > 0$.

Let us now establish the fact that

$$\left\{1 - |\mu|^2\right\} u^H A u = |1 - \mu|^2 u^H (A + C + C^T) u. \tag{4}$$

We begin with the identity

$$u^H A u - (\mu u)^H A(\mu u) = (u - \mu u)^H A(u - \mu u) + 2\mathrm{Re}\left\{(u - \mu u)^H A(\mu u)\right\}. \tag{5}$$

Using (3), and the fact that $A = B - C$, we find that

$$\begin{aligned}(u - \mu u)^H A(\mu u) &= (u - \mu u)^H (B - C)(\mu u) = (u - \mu u)^H (\mu B u - \mu C u) \\ &= (u - \mu u)^H (Cu - \mu E v - \mu C u) = (u - \mu u)^H C(u - \mu u),\end{aligned}$$

which reduces (5) to the result stated in (4).

By hypothesis, $A + C + C^T$ is a positive definite matrix. Since $u \neq 0$, we know that $u^H(A + C + C^T)u > 0$, and so (4) implies that either $|\mu| < 1$ or $\mu = 1$. However, it is impossible that $\mu = 1$, for if $\mu = 1$, then (3) would imply that u, v are both zero, because by Corollary 2.2 the matrix

$$\begin{pmatrix} A & E \\ E^T & 0 \end{pmatrix}$$

is nonsingular. □

Let us now describe several iterative schemes whose convergence is assured by Theorem 3.1.

COROLLARY 3.2. *Assume that*

(a) *A is a real symmetric nonnegative definite matrix,*

(b) *E is a real matrix with full column rank,*

(c) *A and E^T have no nontrivial null vectors in common,*

(g) *$A = D - L - L^T$, where D is a nonsingular diagonal matrix and L is a strictly lower triangular matrix.*

Then the iterative scheme (1) *is convergent for the following choices of B and C.*

$$B = \frac{1}{\omega}D, \qquad C = \frac{1-\omega}{\omega}D + L + L^T \tag{1}$$

with $\omega > 0$ chosen so small that $2D/\omega - A$ is a positive definite matrix.

$$B = \frac{1}{\omega}D - L, \qquad C = \frac{1-\omega}{\omega}D + L^T \tag{2}$$

with $0 < \omega < 2$.

$$B = \left(\frac{2-\omega}{\omega}\right)^{-1}\left(\frac{1}{\omega}D - L\right)D^{-1}\left(\frac{1}{\omega}D - L\right)^T, \tag{3}$$

$$C = \left(\frac{2-\omega}{\omega}\right)^{-1}\left(\frac{1-\omega}{\omega}D + L\right)D^{-1}\left(\frac{1-\omega}{\omega}D + L\right)^T$$

with $0 < \omega < 2$.

Proof. It is clear that assumptions (a)—(e) of Theorem 3.1 are valid for each of the above choices for B and C. Therefore the iterative scheme (1) will be convergent if assumption (f′) of Theorem 3.1 is valid for each of the above choices of B and C. For the first choice of B and C we find that

$$A + C + C^T = 2D/\omega - A,$$

and so assumption (f′) of Theorem 3.1 is valid if $\omega > 0$ is chosen so small that $2D/\omega - A$ is positive definite. For the second choice of B and C we find that

$$A + C + C^T = (2 - \omega)D/\omega,$$

and so assumption (f′) of Theorem 3.1 is valid if $0 < \omega < 2$. For the third choice of B and C we find that

$$A + C + C^T = B + C,$$

where, for $0 < \omega < 2$, B is a symmetric positive definite matrix and C is a symmetric nonnegative definite matrix, and so assumption (f′) of Theorem 3.1 is valid. □

We note that the first, second, and third choices of B and C described in Corollary 3.2 correspond, respectively, to the usual JOR, SOR, and SSOR splittings of A described in [14]. Under further assumptions Corollary 3.2 can be extended to the line and block versions of JOR, SOR, and SSOR. Furthermore, there is an obvious generalization of Theorem 3.1 to complex matrices.

School of Mathematical Sciences
Tel-Aviv University
Tel-Aviv, Israel

Department of Mathematics
Southern Methodist University
Dallas, Texas 75275

1. R. W. COTTLE, G. H. GOLUB & R. S. SACHER, "On the solution of large, structured linear complementarity problems: The block tridiagonal case," *Appl. Math. Optim.*, v. 4, 1978, pp. 347–363.

2. R. W. COTTLE & J. S. PANG, "On the convergence of a block successive overrelaxation method for a class of linear complementarity problems," *Math. Prog. Studies.* (To appear.)

3. R. W. COTTLE, *Application of a Block Successive Over-Relaxation Method to a Class of Constrained Matrix Problems*, Tech. Report SOL 81-20, Stanford University, November 1981.

4. N. DYN & W. FERGUSON, *Numerical Construction of Smooth Surfaces from Aggregated Data*, University of Wisconsin—Madison, Math. Res. Center, TSR #2129, October 1980.

5. N. DYN & G. WAHBA, *On the Estimation of Functions of Several Variables from Aggregated Data*, University of Wisconsin—Madison, Math. Res. Center, TSR #1974, July 1979; *SIAM J. Math. Anal.* (To appear.)

6. P. E. GILL & W. MURRAY, *Numerical Methods for Constrained Optimization*, Academic Press, New York, 1974.

7. G. HADLEY, *Nonlinear and Dynamic Programming*, Addison-Wesley, Reading, Mass., 1964.

8. D. G. LUENBERGER, "Hyperbolic pairs in the method of conjugate gradients," *SIAM J. Appl. Math.*, v. 17, 1969, pp. 1263–1267; "The conjugate residual method for constrained minimization problems," *SIAM J. Numer. Anal.*, v. 7, 1970, pp. 390–398.

9. O. L. Mangasarian. "Solution of symmetric linear complementarity problems by iterative methods," *J. Optim. Theory Appl.*, v. 22, 1977, pp. 465–485; *Sparsity-Preserving SOR Algorithms for Separable Quadratic and Linear Programming*, University of Wisconsin-Madison, Math. Res. Center, TSR #2260, August 1981.

10. C. C. Paige & M. A. Saunders, "Solution of sparse indefinite systems of equations," *SIAM J. Numer. Anal.*, v. 12, 1975, pp. 617–629.

11. G. W. Stewart, *Introduction to Matrix Computations*, Academic Press, New York, 1973.

W. Tobler, "Smooth pycnophylactic interpolation for geographical regions," *J. Amer. Statist. Assoc.*, v. 74, 1979, pp. 519–530.

13. B. Wendroff, *Theoretical Numerical Analysis*, Academic Press, New York, 1966.

14. D. M. Young, *Iterative Solution of Large Linear Systems*, Academic Press, New York, 1971.