Computer Sciences Department
The University of Wisconsin
1210 West Dayton Street
Madison, Wisconsin 53706

THE SOLUTION OF A QUADRATIC PROGRAMMING PROBLEM USING SYSTEMATIC OVERRELAXATION

by

Colin W. Cryer*

Technical Report #73

October 1969

*Sponsored by the Office of Naval Research under Contract No.: N00014-67-A-0128-0004.

THE SOLUTION OF A NONLINEAR PROGRAMMING PROBLEM USING SYSTEMATIC OVERRELAXATION

by

Colin W. Cryer

1. INTRODUCTION

Let $\underline{A} = (a_{ij})$ be a real symmetric positive definite $n \times n$ matrix and $\underline{b} = (b_i)$ a real column n-vector. We shall be concerned with the following problem:

Problem 1

Find real column n-vectors $\underline{x} = (x_i)$ and $\underline{y} = (y_i)$ such that

$$\underline{A}\,\underline{x} - \underline{y} = \underline{b}, \tag{1.1}$$

$$\underline{x}^T\underline{y} = 0, \tag{1.2}$$

$$\underline{x} \geq 0, \quad \underline{y} \geq 0 . \tag{1.3}$$

It is known that Problem 1 is equivalent to a quadratic programming problem, Problem 2 (see section 2). Both Problems 1 and 2 have been extensively studied from the viewpoint of linear and quadratic programming (Cottle and Dantzig [1], Hadley [5, p. 212], and Lemke [6]) and there are many methods available for solving these problems.

Our interest in Problem 1 arose because problems of this type occur when the method of Christopherson is used to solve free boundary problems for journal bearings (Cryer [3]). In such cases Problem 1 has certain features which are unusual in nonlinear programming problems:

(i) $\underline{A}$ is a large matrix, perhaps a $10,000 \times 10,000$ matrix.

(ii) $\underline{A}$ is a "finite-difference" matrix. Typically, each row of $\underline{A}$ will have no more than five non-zero elements. However, $\underline{A}^{-1}$ is a full matrix.

(iii) Because of the physical significance of the solution vector $\underline{x}$, most of the components x_i may be expected to be positive.

When these features are present, the conventional methods for solving Problems 1 and 2 have substantial disadvantages.

In Section 3 we introduce a method for solving Problem 1 which is particularly suitable when $\underline{A}$ is a "finite-difference" matrix, since the method is a modified version of S.O.R. (systematic overrelaxation). In Section 3 we prove that this method converges, and in Section 4 we study how the rate of convergence depends upon the relaxation parameter.

ACKNOWLEDGEMENT

It is a pleasure to thank Professor J. B. Rosen for his help.

2. EXISTENCE AND UNIQUENESS OF SOLUTION

It is convenient to introduce the following quadratic programming problem:

Problem 2

Find a column n-vector $\underline{x}$ which maximizes

$$f(\underline{x}) = \underline{b}^T\underline{x} - (\underline{x}^T\underline{A}\,\underline{x})/2, \tag{2.1}$$

subject to the constraints

$$\underline{x} \geq 0. \tag{2.2}$$

Theorem 2.1

Problems 1 and 2 are equivalent: if $\{\underline{x}, \underline{y}\}$ is a solution of Problem 1 then $\underline{x}$ is a solution of Problem 2; if $\underline{x}$ is a solution of Problem 2 then $\{\underline{x}, \underline{A}\,\underline{x} - \underline{b}\}$ is a solution of Problem 1.

There exists a unique solution to Problem 2 (and hence to Problem 1).

Proof: Since $\underline{A}$ is positive definite, the equivalence of Problems 1 and 2 follows from the Kuhn-Tucker theory (Hadley [5, pages 212-214]).

Since $\underline{A}$ is positive definite, $f(\underline{x})$ is strictly concave (Hadley [5, p. 213]). Hence, since $\underline{x} = 0$ is a "feasible" solution of Problem 2, there exists a unique solution to Problem 2. The proof of the theorem is therefore complete.

It is natural to ask whether Theorem 2.1 remains true if $\underline{A}$ is not positive definite. As the following two examples show, if no conditions are imposed upon $\underline{A}$ and $\underline{b}$ then Problem 1 may either have no solution or more than one solution.

<u>Example 1:</u> $n = 1$, $\underline{A} = 0$. Then Problem 1 has no solution if $b_1 > 0$.

<u>Example 2:</u> $n = 2$, $\underline{A} = \begin{pmatrix} -1, & 2 \\ 2, & -1 \end{pmatrix}$, $\underline{b} = \begin{pmatrix} -1 \\ -1 \end{pmatrix}$. Then Problem 1 has at least two solutions, namely,

$$\underline{x} = \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \quad \underline{y} = \begin{pmatrix} 0 \\ 3 \end{pmatrix},$$

and

$$\underline{x} = \begin{pmatrix} 0 \\ 0 \end{pmatrix}, \quad \underline{y} = \begin{pmatrix} 1 \\ 1 \end{pmatrix}.$$

3. APPLICATION OF S.O.R.

We study the following algorithm for solving Problem 1:

Algorithm 1

Choose a column n-vector $\underline{x}^{(0)} = (x_i^{(0)})$ where $\underline{x}^{(0)} \geq 0$. Choose a relaxation parameter ω, where $0 < \omega < 2$.

Generate a sequence of column n-vectors $\underline{x}^{(k)} = (x_i^{(k)})$, $\underline{r}^{(k)} = (r_i^{(k)})$, $\underline{y}^{(k)} = (y_i^{(k)})$, $k = 1, 2, \ldots,$ using the equations,

$$r_i^{(k+1)} = b_i - \sum_{j=1}^{i-1} a_{ij} x_j^{(k+1)} - \sum_{j=i}^{n} a_{ij} x_j^{(k)}, \qquad (3.1)$$

$$x_i^{(k+1)} = \max\{0, x_i^{(k)} + \omega r_i^{(k+1)}/a_{ii}\}, \qquad (3.2)$$

$$y_i^{(k+1)} = -r_i^{(k+1)} + a_{ii}(x_i^{(k+1)} - x_i^{(k)}). \qquad (3.3)$$

(We remind the reader that we have assumed that $\underline{A}$ is positive definite so that $a_{ii} > 0$ for $1 \leq i \leq n$.)

Algorithm 1 is a generalization of methods used by Christopherson [2] and Gnanadoss and Osborne [4]; a brief account of the history of the algorithm is given in Cryer [3].

Algorithm 1 can be interpreted in two ways. On the one hand, Algorithm 1 consists of applying S.O.R. to the equations $\underline{A}\,\underline{x} = \underline{b}$ with the proviso that the vectors $\underline{x}^{(k)}$ should be nonnegative. On the other hand, as will be seen in the proof of Theorem 3.1, $f(\underline{x}^{(k+1)}) \geq f(\underline{x}^{(k)})$, so that Algorithm 1 is a method for maximizing $f(\underline{x})$ subject to the restraint that $\underline{x} \geq 0$. Of course, it is not surprising that two interpretations of Algorithm 1 exist, since it has been known for a long time (Temple [10]) that there is a connection between relaxation methods and the minimization of quadratic forms.

Theorem 3.1

Let $\underline{x}^{(k)}$ and $\underline{y}^{(k)}$ be generated by Algorithm 1.

Then $\underline{x}^{(k)} \to \underline{x}$ and $\underline{y}^{(k)} \to \underline{y}$, where $\{\underline{x}, \underline{y}\}$ is the solution of Problem 1.

Proof: The method of proof is similar to that used by Schechter [8,9].

For any column n-vector $\underline{u}$ let

$$G(\underline{u}) = -2f(\underline{u}) = \underline{u}^T \underline{A}\,\underline{u} - 2\underline{u}^T \underline{b}\ . \tag{3.4}$$

Then, direct computation shows that if $\underline{u}$ and $\underline{v}$ are column n-vectors,

$$G(\underline{u}) - G(\underline{v}) = (\underline{u} - \underline{v})^T A(u - v) + 2(u - v)^T (\underline{A}\,\underline{v} - \underline{b})\ . \tag{3.5}$$

It is convenient to introduce the vectors $\underline{x}^{(k,\ell)} = (x_i^{(k,\ell)})$ where

$$x_i^{(k+1,\ell)} = \begin{cases} x_i^{(k+1)}, & \text{if } 1 \leq i \leq \ell\ , \\ x_i^{(k)}, & \text{if } \ell < i \leq n\ , \end{cases} \tag{3.6}$$

for $k \geq 0$ and $0 \leq \ell \leq n$. Then,

$$\underline{x}^{(k+1,0)} = \underline{x}^{(k)}, \quad \underline{x}^{(k+1,n)} = \underline{x}^{(k+1)}, \tag{3.7}$$

and, from (3.1),

$$r_i^{(k+1)} = [\underline{b} - \underline{A}\,\underline{x}^{(k+1,i-1)}]_i . \tag{3.8}$$

Let

$$\omega_{k+1,i} = \begin{cases} [x_i^{(k+1)} - x_i^{(k)}]\, a_{ii}/r_i^{(k+1)}, & \text{if } r_i^{(k+1)} \neq 0, \\ \omega, & \text{if } r_i^{(k+1)} = 0 . \end{cases} \tag{3.9}$$

Then, noting (3.2),

$$x_i^{(k+1)} = x_i^{(k)} + \omega_{k+1,i}\, r_i^{(k+1)}/a_{ii}, \tag{3.10}$$

and

$$0 \leq \omega_{k+1,i} \leq \omega . \tag{3.11}$$

Using (3.5), (3.6), (3.8), and (3.10), we find that

$$\begin{aligned} &G(\underline{x}^{(k+1,i)}) - G(\underline{x}^{(k+1,i-1)}) \\ &= [\underline{x}^{(k+1,i)} - \underline{x}^{(k+1,i-1)}]^T \underline{A}\, [\underline{x}^{(k+1,i)} - \underline{x}^{(k+1,i-1)}] \\ &\qquad + 2[\underline{x}^{(k+1,i)} - x^{(k+1,i-1)}]^T [A\,\underline{x}^{(k+1,i)} - \underline{b}], \end{aligned}$$

and $$G(\underline{x}^{(k+1,\, i)}) - G(\underline{x}^{(k+1,\, i-1)})$$

$$= a_{ii}[x_i^{(k+1)} - x_i^{(k)}]^2 - 2[x_i^{(k+1)} - x_i^{(k)}]\, r_i^{(k+1)} \;,$$

$$= -\omega_{k+1,\, i}\,(2 - \omega_{k+1,\, i})\,[r_i^{(k+1)}]^2/a_{ii} \;. \tag{3.12}$$

Remembering that $0 < \omega < 2$, it follows from (3.11) and (3.12) that $G(\underline{x}^{(k+1,\, i)}) \le G(\underline{x}^{(k+1,\, i-1)})$. Therefore, the sequence $\{G(\underline{x}^{(k,\, i)})\}$ is monotone decreasing. But $\underline{A}$ is positive definite so that $G(\underline{u})$ is strictly convex and hence bounded below. Consequently there is a constant, G_∞ say, such that

$$G(\underline{x}^{(k,\ell)}) \downarrow G_\infty \;. \tag{3.13}$$

Next, we prove that

$$|x_i^{(k+1)} - x_i^{(k)}| \le [a(-1 + 2/\omega)]^{\frac{1}{2}}\,[G(\underline{x}^{(k+1,\, i-1)}) - G(\underline{x}^{(k+1,\, i)})]^{\frac{1}{2}} \;, \tag{3.14}$$

where

$$a = \min_i \; a_{ii} \;. \tag{3.15}$$

If $x_i^{(k+1)} = x_i^{(k)}$, then (3.14) is trivially true, so that we need only consider the case when $x_i^{(k+1)} \ne x_i^{(k)}$. But then, from (3.10), $\omega_{k+1,\, i} \ne 0$, so that, from (3.10), (3.11), (3.12), and (3.15),

$$G(\underline{x}^{(k+1,\, i-1)}) - G(\underline{x}^{(k+1,\, i)})$$

$$= (-1 + 2/\omega_{k+1,\, i})\, a_{ii}\,[x_i^{(k+1)} - x_i^{(k)}]^2 \;,$$

$$\ge (-1 + 2/\omega)\; a\; [x_i^{(k+1)} - x_i^{(k)}]^2 \;. \tag{3.16}$$

Inequality (3.14) follows immediately from (3.16).

Noting (3.13), it follows from (3.14) that

$$x_i^{(k+1)} - x_i^{(k)} \rightarrow 0 \quad \text{as} \quad k \rightarrow \infty, \quad 1 \le i \le n. \tag{3.17}$$

Now let $\underline{x}$ be any limit point of the sequence $\{\underline{x}^{(k)}\}$. Then there is an increasing sequence of integers $\{k_p\}$, $p = 1, 2, \ldots,$ such that

$$\underline{x}^{(k_p)} \rightarrow \underline{x} \quad \text{as} \quad p \rightarrow \infty . \tag{3.18}$$

From (3.1), (3.3), and (3.18), we have that, as $p \rightarrow \infty$,

$$\left.\begin{aligned} \underline{r}^{(k_p)} &\rightarrow \underline{r} = \underline{b} - \underline{A}\,\underline{x}, \\ \underline{y}^{(k_p)} &\rightarrow \underline{y} = -\underline{r} . \end{aligned}\right\} \tag{3.19}$$

We assert that

$$\underline{x} \ge 0, \quad \underline{r} \le 0. \tag{3.20}$$

That $\underline{x} \ge 0$ follows immediately from the fact that $\underline{x}^{(k)} \ge 0$ for all k. To prove that $\underline{r} \le 0$ suppose that this is not the case. Then there is an $\epsilon > 0$ and integers i_0 and k'_0 such that $r_{i_0}^{(k_p)} \ge \epsilon$ for $k_p \ge k'_0$. Hence, from (3.2),

$$x_{i_0}^{(k_p)} - x_{i_0}^{(k_p-1)} \ge \epsilon\omega/a_{i_0 i_0},$$

for $k_p \ge k'_0$. But this contradicts (3.17).

Next we show that

$$\underline{r}^T \underline{x} = 0 . \tag{3.21}$$

Suppose that this is not the case. Then, noting (3.20), we see that there is an $\epsilon > 0$ and integers i_0 and k' such that $r_{i_0}^{(k_p)} \le -\epsilon$ and $x_{i_0}^{(k_p)} \ge \epsilon$ for $k_p \ge k'$. It follows from (3.2) that if $k_p \ge k'$ then $x_{i_0}^{(k_p - 1)} \ge x_{i_0}^{(k_p)}$ and

$$|x_{i_0}^{(k_p)} - x_{i_0}^{(k_p - 1)}| \ge \omega\, \epsilon / a_{i_0 i_0} .$$

But this contradicts (3.17).

From (3.19), (3.20), and (3.21), it follows that $\{\underline{x}, \underline{y}\}$ satisfies (1.1) through (1.3) and is the (unique) solution of Problem 1.

To complete the proof of the theorem we must show that the sequence $\{\underline{x}^{(k)}\}$ has at least one limit point. But this is a consequence of the fact that (see (3.13)) $\underline{x}^{(k)} \in R$ for all k, where R is the compact set

$$R = \{\underline{x};\ G(\underline{x}) \le G(\underline{x}^{(0)})\} .$$

4. DETERMINATION OF THE OPTIMUM RELAXATION PARAMETER

It is natural to ask how the convergence of Algorithm 1 depends upon ω, and whether there is a value of ω for which the rate of convergence is maximized. In this section we partially answer these questions.

Since we make use of the theory of S.O.R., we first summarize this theory.

Let

$$\underset{\sim}{\tilde{A}}\,\underset{\sim}{\tilde{x}} = \underset{\sim}{\tilde{b}}, \tag{4.1}$$

and

$$\underset{\sim}{\tilde{A}} = \underset{\sim}{\tilde{D}} - \underset{\sim}{\tilde{E}} - \underset{\sim}{\tilde{F}}, \tag{4.2}$$

where $\underset{\sim}{\tilde{x}}$ is a column m-vector, $\underset{\sim}{\tilde{D}}$ is a diagonal $m \times m$ matrix, and $\underset{\sim}{\tilde{E}}$ and $\underset{\sim}{\tilde{F}}$ are respectively strictly upper and lower triangular $m \times m$ matrices. Let

$$\mathcal{L}_{\omega}(\underset{\sim}{\tilde{A}}) = (\underset{\sim}{\tilde{D}} - \omega\underset{\sim}{\tilde{E}})^{-1}\ \{(1-\omega)\underset{\sim}{\tilde{D}} + \omega\underset{\sim}{\tilde{F}}\}\ . \tag{4.3}$$

For a given relaxation parameter ω and initial guess $\underset{\sim}{\tilde{x}}^{(0)}$, let $\underset{\sim}{\tilde{x}}^{(k)}$, $k = 1, 2, \ldots,$ denote the iterates generated by S.O.R, applied to (4.1). Let

$$\underset{\sim}{\tilde{e}}^{(k)} = \underset{\sim}{\tilde{x}}^{(k)} - \underset{\sim}{\tilde{x}}\ . \tag{4.4}$$

Then, (Varga [11, p. 59]),

$$\underset{\sim}{\tilde{e}}^{(k+1)} = \mathcal{L}_{\omega}(\underset{\sim}{\tilde{A}})\,\underset{\sim}{\tilde{e}}^{(k)}\ . \tag{4.5}$$

From (4.5) it can be seen that $\underline{\tilde{e}}^{(k)}$ depends upon $\underline{\tilde{A}}$, $\underline{\tilde{e}}^{(0)}$, and ω. The asymptotic rate of convergence corresponding to $\underline{\tilde{A}}$ and ω is (Varga [11, p. 67]),

$$R_\infty[\mathcal{L}_\omega(\underline{\tilde{A}})] = -\log[\rho[\mathcal{L}_\omega(\underline{\tilde{A}})]], \tag{4.6}$$

where $\rho[\mathcal{L}_\omega(\underline{\tilde{A}})]$ denotes the spectral radius of $\mathcal{L}_\omega(\underline{\tilde{A}})$. Equivalently,

$$R_\infty[\mathcal{L}_\omega(\underline{\tilde{A}})] = -\log[\sup_{\underline{\tilde{e}}^{(0)}} \limsup_{k\to\infty} \|\underline{\tilde{e}}^{(k)}\|^{1/k}]. \tag{4.7}$$

The optimum relaxation parameter, $\omega_b = \omega_b(\underline{\tilde{A}})$, is defined by means of the relation (Varga [11, p. 109]),

$$R_\infty[\mathcal{L}_{\omega_b}(\underline{\tilde{A}})] = \max_{0<\omega<2} R_\infty[\mathcal{L}_\omega(\underline{\tilde{A}})]. \tag{4.8}$$

For certain classes of matrices $\underline{\tilde{A}}$, notably 2-cyclic consistently ordered matrices, $\omega_b(\underline{\tilde{A}})$ is known in terms of the eigenvalues of the Jacobi matrix corresponding to $\underline{A}$ (Varga [11, p. 110]).

Next we introduce some notation. We set

$$Z = \{1, 2, \ldots, n\}. \tag{4.9}$$

Let $T \subset Z$, $\underline{B} = (B_{ij})$ be an $n \times n$ matrix, and $\underline{z} = (z_i)$ be an n-vector. Then $|T|$ denotes the number of elements of T; $\underline{B}(T)$ is the $|T| \times |T|$ submatrix of $\underline{B}$ obtained by deleting those elements B_{ij} for which $i \notin T$ or $j \notin T$; and $\underline{z}(T)$ is the $|T| \times 1$ subvector of $\underline{z}$ obtained by deleting those elements z_i for which $i \notin T$. Finally,

$$Z(\underline{z}) = \{i \in Z;\ z_i \neq 0\}. \tag{4.10}$$

We are now ready to consider Algorithm 1. Let $\{\underline{x}, \underline{y}\}$ be the solution of Problem 1, and let $\{\underline{x}^{(k)}, \underline{y}^{(k)}\}$ be generated using Algorithm 1. We set

$$\underline{e}^{(k)} = \underline{x}^{(k)} - \underline{x} \; . \tag{4.11}$$

From (3.1) through (3.3) and (4.11) we see that $\underline{e}^{(k)}$ depends upon $\underline{\tilde{A}}$, $\underline{b}$, $\underline{e}^{(0)}$, and ω. By Theorem 3.1, $\underline{e}^{(k)} \rightarrow 0$ as $k \rightarrow \infty$ for any $\underline{e}^{(0)}$.

Corresponding to (4.7) and following Ortega and Rockoff [7] we define the asymptotic rate of convergence corresponding to $\underline{A}$, $\underline{b}$, and ω, to be

$$R(\underline{A}, \underline{b}, \omega) = -\log \{\sup_{\underline{e}^{(0)}} \limsup_{k \rightarrow \infty} \|\underline{e}^{(k)}\|^{1/k}\} \; . \tag{4.12}$$

Lemma 4.1

Given $\underline{x}^{(0)}$ there is an integer k_0 such that for $k \geq k_0$,

$$\left.\begin{array}{lll} x_i^{(k)} > 0, & & \text{if} \quad i \in X, \\ x_i^{(k)} = 0 \quad \text{and} \quad y_i^{(k)} > 0, & & \text{if} \quad i \in Y, \end{array}\right\} \tag{4.13}$$

where $X = Z(\underline{x})$ and $Y = Z(\underline{y})$.

Proof: Let $\underline{x}^{(0)}$ be given. Let $\underline{r}^{(k)}$ be as in Algorithm 1. Since $\underline{x}^{(k)} \rightarrow \underline{x}$ and $\underline{r}^{(k)} \rightarrow -\underline{y}$, it follows that there is an $\epsilon > 0$ and an integer k_1 such that if $k \geq k_1$ then

$$x_i^{(k)} > 0 \,, \quad \text{if} \quad i \in X,$$

$$y_i^{(k)} > 0 \quad \text{and} \quad r_i^{(k)} < -\epsilon \,, \quad \text{if} \quad i \in Y \,.$$

Noting (3.2), we see that (4.13) holds if $k_0 \geq k_1 + 1 + [\max_i \; a_{ii} \, x_i^{(k_1)}]/(\epsilon \, \omega)$.

Theorem 4.2

Let $\underline{A}$ and $\underline{b}$ be such that

$$|x_i| + |y_i| > 0, \qquad 1 \leq i \leq n, \tag{4.14}$$

where $\{\underline{x}, \underline{y}\}$ is the solution of Problem 1. Let $X = Z(\underline{x})$.

Then,

$$R(\underline{A}, \underline{b}, \omega) = \begin{cases} \infty, & \text{if} \quad X \quad \text{is empty,} \\ R_\infty(\mathcal{L}_\omega[\underline{A}(X)]), & \text{otherwise.} \end{cases} \tag{4.15}$$

Proof: Let $\underline{x}^{(0)}$ be given. Then it follows from Lemma 4.1, (3.1), (3.2), and (4.14), that, for $k \geq k_0$,

$$\underline{e}^{(k+1)}(X) = \mathcal{L}_\omega(\underline{A}(X)) \; \underline{e}^{(k)}(X) \,,$$
$$\underline{e}^{(k+1)}(Y) = 0 \,.$$

The theorem follows from (4.7) and (4.12).

Condition (4.14) is satisfied by "almost all" $\underline{A}$ and $\underline{b}$, and the following theorem covers an important subclass of the remaining problems.

Theorem 4.3

Let

$$0 < \omega \le 1 \tag{4.16}$$

and

$$a_{ij} \le 0, \quad \text{for} \quad i \ne j . \tag{4.17}$$

Then, $R[\underline{A}, \underline{b}, \omega] \le R_\infty[\mathcal{L}_\omega(A[T])]$, where $T = Z - Z(\underline{y})$.

Proof: Let $\underline{x}^{(0)}$ be given. Using an idea due to Gnanadoss and Osborne [4] we see from Lemma 4.1, (3.1), and (3.2), that for $k \ge k_0$,

$$\left.\begin{aligned} \underline{e}^{(k+1)}(T) &= \underline{C}^{(k+1)}\, \underline{e}^{(k)}(T) , \\ \underline{e}^{(k+1)}(Y) &= 0 , \end{aligned}\right\} \tag{4.18}$$

Here, $\underline{C}^{(k+1)}$ is a $|T| \times |T|$ matrix such that

$$\underline{C}^{(k+1)} = \prod_{\ell=1}^{|T|} (\underline{H}^{(k+1,\ell)}\, \underline{L}^{(\ell)}) ; \tag{4.19}$$

$\underline{L}^{(\ell)}$ are $|T| \times |T|$ matrices such that

$$\mathcal{L}_\omega[\underline{A}(T)] = \prod_{\ell=1}^{|T|} \underline{L}^{(\ell)} ; \tag{4.20}$$

and $\underline{H}^{(k+1,\ell)}$ is a $|T| \times |T|$ diagonal matrix with diagonal elements equal to either 0 or 1 . In particular, when $T = Z$ then $\underline{L}^{(\ell)} = (L_{ij}^{(\ell)})$ and $\underline{H}^{(k+1,\ell)} = \text{diag}\,(H_{ii}^{(k+1,\ell)})$ where

$$L_{ij}^{(\ell)} = \begin{cases} 1, & \text{if } i = j \text{ and } i \neq \ell, \\ 1-\omega, & \text{if } i = j = \ell, \\ -\omega a_{\ell j}/a_{\ell \ell}, & \text{if } i = \ell \text{ and } j \neq \ell, \\ 0, & \text{otherwise,} \end{cases}$$

$$H_{ii}^{(k+1,\ell)} = \begin{cases} 1, & \text{if } i \neq \ell, \\ 1, & \text{if } i = \ell \text{ and } x_{\ell}^{(k+1)} > 0, \\ 0, & \text{otherwise.} \end{cases}$$

From (4.16) and (4.17) it follows that $\underline{L}^{(\ell)} \geq 0$; that is, the elements of $\underline{L}^{(\ell)}$ are non-negative. Hence, we see from (4.19) and (4.20) that $0 \leq \underline{C}^{(k+1)} \leq \mathcal{L}_{\omega}[\underline{A}(T)]$. The theorem follows immediately.

On the basis of Theorems 4.2 and 4.3 we make the following conjecture:

Conjecture 4.4

$R[\underline{A}, \underline{b}, \omega] \leq R_{\infty}[\mathcal{L}_{\omega}(\underline{A}[T])]$, where $T = Z - Z(\underline{y})$, and $\{\underline{x}, \underline{y}\}$ is the solution of Problem 1 .

Theorems 4.2 and 4.3 provide some help in choosing ω so as to maximize the rate of convergence of Algorithm 1 .

If (4.14) holds, then we see from Theorem 4.2 that we should set $\omega = \omega_{opt}$ where

$$\omega_{opt} = \omega_b [\underline{A}(X)]. \tag{4.21}$$

Of course, (4.21) does not give ω_{opt} explicitly since, in general, neither X nor $\omega_b[\underline{A}(X)]$ is known explicitly.

If Problem 1 is derived from Christopherson's method then (4.21) is very useful. For, in this case, $\underline{A}$ is 2-cyclic and consistently ordered (Cryer [3]). Furthermore, for any $S \subset Z$, $\omega_b(\underline{A}(S)) \leq \omega_b(\underline{A})$. Remembering that it is in general better to overestimate ω_{opt} rather than underestimate ω_{opt} (Varga [11, p. 114]) we see that in this case it is a good strategy to set $\omega = \omega_b(\underline{A})$.

If (4.14) does not hold then we can say much less about the choice of ω. However, if Problem 1 is derived from Christopherson's method then (4.17) is satisfied and, for any $T \subset Z$, $R_\infty[\mathcal{L}_\omega(\underline{A}(T))]$ is a monotone decreasing function of ω for $0 < \omega \leq 1$. Hence, the results of Theorem 4.3 suggest that we should choose $\omega \geq 1$.

REFERENCES

[1] Cottle, R. W., and G. B. Dantzig: Complementary pivot theory of mathematical programming. In Mathematics of the Decision Sciences, Part I. Dantzig, G. B., and A. F. Veinott, Jr., editors. Providence, R. I.: American Mathematical Society, 1968.

[2] Christopherson, D. G.: A new mathematical method for the solution of film lubrication problems. Proc. Institution Mechanical Engineers 146, 126-135 (1941).

[3] Cryer, C. W.: The method of Christopherson for solving free boundary problems for infinite journal bearings by means of finite differences. Technical Report No. 72, Computer Sciences Dept., University of Wisconsin, Madison, Wisconsin, October 1969.

[4] Gnanadoss, A. A., and M. R. Osborne: The numerical solution of Reynolds' equation for a journal bearing. Quart. J. Mech. Appl. Math. 17, 241-246 (1964).

[5] Hadley, G.: Nonlinear and Dynamic Programming. Reading, Mass.: Addison-Wesley, 1964.

[6] Lemke, C. E.: On complementary pivot theory. In Mathematics of the Decision Sciences, Part I, Dantzig, G. B. and A. F. Veinott, Jr., editors. Providence, R. I.: American Mathematical Society, 1968.

[7] Ortega, J. M., and M. L. Rockoff: Nonlinear difference equations and Gauss-Seidel type iterative methods. SIAM J. Numer. Anal. 3, 497-513 (1966).

[8] Schechter, S.: Iteration methods for nonlinear problems. Trans. Amer. Math. Soc. 104, 179-189 (1962).

[9] ____________: Relaxation methods for convex problems. SIAM J. Numer. Anal. 5, 601-612 (1968).

[10] Temple, G.: The general theory of relaxation methods applied to linear systems. Proc. Roy. Soc. London A169, 476-500 (1939).

[11] Varga, R. S.: Matrix Iterative Analysis. Englewood Cliffs, New Jersey: Prentice-Hall, 1962.

Unclassified
Security Classification

DOCUMENT CONTROL DATA - R & D

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

1. ORIGINATING ACTIVITY *(Corporate author)* Computer Sciences Department University of Wisconsin	2a. REPORT SECURITY CLASSIFICATION Unclassified 2b. GROUP

3. REPORT TITLE

THE SOLUTION OF A QUADRATIC PROGRAMMING PROBLEM USING SYSTEMATIC OVERRELAXATION.

4. DESCRIPTIVE NOTES *(Type of report and inclusive dates)*
Computer Sciences Technical Report, October 1969

5. AUTHOR(S) *(First name, middle initial, last name)*

Colin W. Cryer

6. REPORT DATE	7a. TOTAL NO. OF PAGES	7b. NO. OF REFS
October 1969	18	11

8a. CONTRACT OR GRANT NO. N00014-67-A-0128-0004 b. PROJECT NO c. d.	9a. ORIGINATOR'S REPORT NUMBER(S) Computer Sciences Technical Report #73 9b. OTHER REPORT NO(S) *(Any other numbers that may be assigned this report)*

10. DISTRIBUTION STATEMENT

Releasable without limitations on dissemination.

11. SUPPLEMENTARY NOTES	12. SPONSORING MILITARY ACTIVITY Mathematics Branch Office of Naval Research Washington, D. C. 20360

13. ABSTRACT

Let $\underline{A}$ be a real symmetric positive definite $n \times n$ matrix and $\underline{b}$ a real column n-vector. We consider the following problem: Find real column n-vectors $\underline{x}$ and $\underline{y}$ such that

$$\underline{A}\,\underline{x} = \underline{b},$$

$$\underline{x}^T\underline{y} = 0, \quad \underline{x} \geq 0, \quad \underline{y} \geq 0.$$

Problems of this type occur when the method of Christopherson is used to solve free boundary problems for journal bearings. In such cases, $\underline{A}$ is a "finite-difference" matrix.

We present a method for solving the above problem which is a modification of systematic overrelaxation. This method is particularly suitable when $\underline{A}$ is a finite-difference matrix.

DD FORM 1473 1 NOV 65

Unclassified
Security Classification

Unclassified

Security Classification

14 KEY WORDS	LINK A		LINK B		LINK C	
	ROLE	WT	ROLE	WT	ROLE	WT
Systematic overrelaxation						
Quadratic programming						
Nonlinear programming						
Fundamental problem of quadratic programming						
Iterative methods.						

Unclassified

Security Classification