KERNFORSCHUNGSANLAGE JÜLICH
GESELLSCHAFT MIT BESCHRÄNKTER HAFTUNG
Zentralinstitut für Angewandte Mathematik

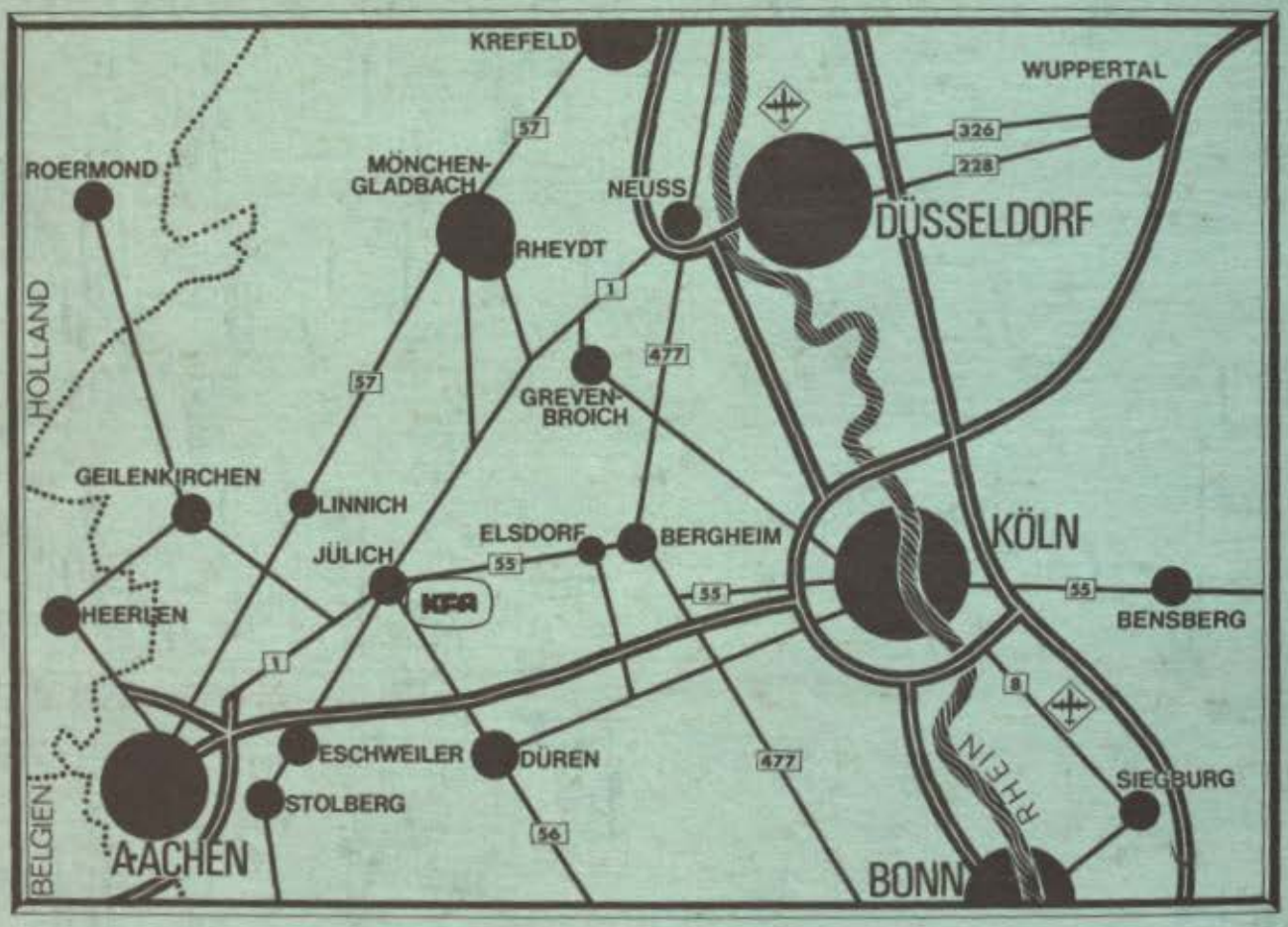
Quadratic Programming by Successive Overrelaxation

by

U. Eckhardt

Jül - 1064 - MA
April 1974

Als Manuskript gedruckt



Berichte der Kernforschungsanlage Jülich - Nr. 1064

Zentralinstitut für Angewandte Mathematik Jül - 1064 - MA

Dok.: Numerical Mathematics
Spline Function
Finite Element Method
Inequality
Quadratic Programming
Iterative Method

Im Tausch zu beziehen durch: ZENTRALBIBLIOTHEK der Kernforschungsanlage Jülich GmbH, Jülich, Bundesrepublik Deutschland

Quadratic Programming by Successive Overrelaxation

by

U. Eckhardt

Abstract

The idea of solving the definite linear complementarity problem by successive overrelaxation was originally proposed by Cryer. Hildreth and d'Esopo presented a Gauss-Seidel-like iterative method to solve a quadratic programming problem.

In this paper a detailed discussion of Cryer's method applied to quadratic programming problems is given. The convergence behaviour is treated without assumptions on solvability of the problem.

Numerical examples indicate the efficiency of the method.

Contents

1. Introduction (1)

2. Reduction to a linear complementarity problem (3)

3. Solution of the linear complementarity problem: Some known methods (5)

4. The method: Preliminaries (7)

5. Convergence of the method (9)

6. Instability (13)

7. Numerical considerations (15)

8. Example (17)

References (21)

1. Introduction

The aim of this paper is to present a method for solving the following problem

Find the minimum of

(1.1) $\quad x^T Cx - c^T x,$

subject to the constraints

(1.2) $\quad Ax \leq a$

(1.3) $\quad Bx = b.$

Here C is assumed to be a positive definite symmetric (d,d)-matrix, c and x are d-vectors, A is a (n,d)-matrix, a is a n-vector and the inequality sign is to be understood componentwise. The matrix B is (k,d) and b is a k-vector. Under these conditions the problem has a unique solution if and only if there is an $x \in R^d$ satisfying (1.2-3).

Such problems arise quite naturally in numerical analysis. If one wants for example to find the spline function or more generally the piecewise polynomial (i.e. finite element function) which is nearest to a given set of data in the least squares sense and fulfils some side conditions (positivity, monotonicity or convexity) given by inequality constraints then one has to solve a problem like (1.1-3). (1.1) is the condition of minimum distance to the given data, (1.2) are the (discretized) inequality conditions, and (1.3) are the continuity requirements generally imposed on splines or finite element functions [1, 12, 22].

The same type of problem arises if a variational problem is discretized by splines or finite elements and if in addition some inequalities are imposed to hold for the solution (contact problems, variational inequalities) [11, 18].

There exists a vast amount of similar problems in the literature [2, 4, 5, 7, 8, 14, 16, 20, 23, 24]. The main common features of all these problems are:

1. The matrices of the problem are large,

2. The matrices are sparse or composed by sparse matrices,

3. In (1.2), equality holds only for relatively few inequalities,

4. Usually there are good approximate solutions to the problem known in advance.

2. Reduction to a Linear Complementarity Problem

First we assume (without loss of generality) that B has maximal rank. That means there are matrices B_1, B_2 such that (after eventually rearranging the components of x)

$$B = (B_1, B_2) \tag{2.1}$$

and B_1 is nonsingular and quadratic. Partitioning in the same manner

$$A = (A_1, A_2), \tag{2.2}$$

$$C = \begin{bmatrix} C_{11}, & C_{12} \\ C_{12}^T, & C_{22} \end{bmatrix},$$

$$x = (x_1, x_2)^T,$$

$$c = (c_1, c_2)^T,$$

we can eliminate x_1 by

$$x_1 = b - B_1^{-1} B_2 x_2 \tag{2.3}$$

and get the reduced problem

$$x_2^T \tilde{C}\, x_2 - \tilde{c}\, x_2 \longrightarrow \text{minimum} \tag{2.4}$$

$$\text{subject to} \quad \tilde{A}\, x_2 \leq a,$$

where

$$\tilde{A} = A_2 - A_1\, B_1^{-1}\, B_2 \tag{2.5}$$

$$\tilde{a} = a - A_1\, B_1^{-1}\, b$$

$$\tilde{C} = B_2^T\, B_1^{-1^T}\, C_{11}\, B_1^{-1}\, B_2 - 2\, C_{12}^T\, B_1^{-1}\, B_2 + C_{22}$$

$$\tilde{c} = 2\, B_2^T\, B_1^{-1^T}\, C_{11}\, B_1^{-1}\, b - 2\, C_{12}\, B_1^{-1}\, b + B_2^T\, B_1^{-1^T}\, c_1 - c_2 .$$

This reduction can be easily performed by simple elimination. Matrix $\tilde{C}$ is positive definite and symmetric. In some cases (see § 8) C is only positive semidefinite and $\tilde{C}$ is positive definite.

The solution of (2.4) is characterized by the Kuhn-Tucker conditions ([3], Chapter III):

x_2 is a solution of (2.4) if and only if there are $y \geq \Theta_n$ (zero element of R^n), $u \geq \Theta_n$ such that

$$\begin{aligned} \tilde{A}\, x_2 \qquad\quad + y &= \tilde{a} \\ -2\,\tilde{C}\, x_2 - \tilde{A}^T u \quad &= \tilde{c} \\ y \geq \Theta_n, \quad u \geq \Theta_n, \quad & y^T u = 0. \end{aligned} \tag{2.6}$$

Now we eliminate x_2 in (2.6) by

$$x_2 = -\frac{1}{2}\,\tilde{C}^{-1}\tilde{A}^T u - \frac{1}{2}\,\tilde{C}^{-1}\,\tilde{c}\,. \tag{2.7}$$

This yields the following linear complementarity problem

$$\begin{aligned} M\,u + y &= p \\ u \geq \Theta_n, \quad y \geq \Theta_n, \quad & u^T y = 0, \end{aligned} \tag{2.8}$$

where

$$\begin{aligned} M &= -\frac{1}{2}\,\tilde{A}\,\tilde{C}^{-1}\,\tilde{A}^T \\ p &= \tilde{a} + \frac{1}{2}\,\tilde{A}\,\tilde{C}^{-1}\,\tilde{c}. \end{aligned} \tag{2.9}$$

The (n,n) matrix M is only negative semidefinite. It should be kept in mind that y (and x) are uniquely determined if the problem has a solution. u however is not necessarily unique.

Having computed u and y, x is given by (2.3) and (2.7). The matrices $\tilde{A}$ and $\tilde{C}^{-1}$ are usually sparse matrices which is generally not the case with M. Moreover, M is a (n,n)-matrix while $\tilde{A}$ is (n,d-k) and $\tilde{C}$ is (d-k, d-k). Therefore it seems adviceable to store M in the form (2.9) in order to take advantage of the structure of $\tilde{A}$ und $\tilde{C}^{-1}$ and of the sparsity of u.

3. Solution of the Linear Complementarity Problem: Some Known Methods

Among the classical methods for solution of a quadratic programming problem like (1.1-3) the Wolfe method [25] is the most prominent one (see [3], § 14.3). This method is an elimination method like the simplex method for linear programming. The main drawbacks of such methods in our case are:

1. Nothing can be said about the number of iteration steps necessary to solve the problem,
2. The structure of M is destroyed after a few iteration steps. Even symmetry cannot maintained during the iteration process,
3. There is no possibility to take advantage of known "good" solutions,
4. As generally with elimination methods, rounding errors tend to accumulate from one iteration to another.

There are many other methods for solving (1.1-3) by an elimination process. Recently Lemke ([17], see also [9]) proposed a very elegant method to solve the complementarity problem (2.8). This method, although very attractive from a theoretical standpoint, suffers from the drawbacks mentioned above.

There has also been a large amount of iterative methods in the literature. Hildreth [15] and d'Esopo [6] proposed to solve (2.8) by a Gauss-Seidel-like method. Unfortunately they needed very restrictive restrictions on the generality of the problem. The most serious assumption is that there should be a solution of (1.2-3) (moreover, even a "strong" solution). This assumption cannot guaranteed in practice. As it is often the case with practical problems, even a small data error in most cases causes (1.2-3) to have no solution.

A very interesting method for a very interesting problem was given by Fridman and Chernina [11]. M, however, is assumed to be negative definite. The problem attacked by these authors is the finite dimensional contact problem which gives a very nice physical interpretation to the linear complementarity problem (2.8).

Cryer [4] used in a similar context a variant of the SOR method (Successive Overrelaxation method). He also assumed M to be definite. This as-

sumption is too restrictive for most cases arising in practice.

In this context we also mention the iterative methods of Habetler and Price [13], Seĭsov [20] and Torsti and Aurela [23].

The method given here is essentially Cryer's method for negative semidefinite M. In contrast to other authors it is not assumed that the problem has a solution at all. If this is not the case, the method will give a clear indication for it.

4. The Method: Preliminaries

There is a close relationship between (2.8) and the following problem:

Find the maximum of

$$\phi(u) = \frac{1}{2}\, u^T M u - p^T u \tag{4.1}$$

subject to $u \geq \Theta_n$.

Obviously, (2.8) gives exactly the Kuhn-Tucker conditions to (4.1). Without loss of generality we consider only those $u \in R^n$ with $\phi(u) \geq 0$. Then

$$p^T u \leq \frac{1}{2} \cdot u^T M u \leq 0. \tag{4.2}$$

If all the rows of A are different from Θ_{d-k} then $m_{ii} < 0$ for all diagonal elements of M. Without loss of generality we may assume

$$m_{ii} = -1 \quad \text{for all } i. \tag{4.3}$$

As usual we decompose

$$M = A_L - I + A_R \tag{4.4}$$

where I is the identity matrix, A_L is a strictly lower triangular matrix and $A_R = A_L^T$.

Now we formulate the method (Cryer, [4]):

0. Given $u^0 \geq \Theta_n$, $0 < \omega < 2$ fixed.

1. For each r let

$$z^{r+1} = -u^r - \omega \cdot (A_L\, u^{r+1} - u^r + A_R\, u^r - p), \tag{4.5}$$

where u^{r+1} is uniquely determined by

$$z^{r+1} = -u^{r+1} + y^{r+1}, \tag{4.6}$$

$$u^{r+1} \geq \Theta_n,\ y^{r+1} \geq \Theta_n,\ u^{r+1^T} y^{r+1} = 0.$$

2. Apply an appropriate termination criterion (see § 7). If it fails, let r := r + 1 and go to 1.

We note that this method is monotone in the following sense:

Lemma 4.1: Let $0 < \omega < 2$ and define for $k = 1, \ldots, n + 1$

$$\tilde{u}_j^{r;k} = \begin{cases} u_j^{r+1} & \text{for } j < k \\ u_j^{r} & \text{for } j \geq k. \end{cases} \tag{4.7}$$

Similarly we define $\tilde{y}_j^{r;k}$. Then $\tilde{u}^{r;1} = u^r$, $\tilde{u}^{r;n+1} = u^{r+1}$ and

$$-\tilde{u}^{r;k+1} + \tilde{y}^{r;k+1} = -\tilde{u}^{r;k} - \omega \cdot e_k^T (M\,\tilde{u}^{r;k} - p) \cdot e_k \tag{4.8}$$

(e_k = k-th unit vector).

Then

$$\phi(\tilde{u}^{r;k}) \leq \phi(\tilde{u}^{r;k+1}),$$

consequently

$$\phi(u^r) \leq \phi(u^{r+1}).$$

Proof: Observing

$$\frac{\partial \phi(u)}{\partial u_\alpha} = e_\alpha^T \cdot (M\,u - p)$$

we put for fixed r and k

$$\phi_k = e_k^T \cdot (M\,\tilde{u}^{r;k} - p).$$

In computing $\tilde{u}^{r;k+1}$ only one component of $\tilde{u}^{r;k}$ is changed, viz.

$$\tilde{u}^{r;k+1} = \tilde{u}^{r;k} + \mu \cdot \phi_k \cdot e_k, \quad 0 \leq \mu \leq \omega,$$

thus

$$\phi(\tilde{u}^{r;k+1}) - \phi(\tilde{u}^{r;k}) = -\frac{1}{2} \cdot \mu \cdot \phi_k^2 + \mu \cdot \phi_k \cdot e_k^T\,(M\,\tilde{u}^{r;k} - p) = \tag{4.9}$$

$$= \frac{1}{2} \cdot \mu \cdot (2 - \mu) \cdot \phi_k^2 > 0$$

for $0 < \mu < 2$. Thus the Lemma is proved.

5. Convergence of the Method

For the first convergence theorem we need no explicit assumptions on the solvability of the problem (2.8):

Theorem 5.1: If there ia a bounded subsequence $\{u^r\nu\}$ of $\{u^r\}$ then

$$\lim \frac{1}{\omega} \cdot y^r\nu = y^*$$

and y^* together with any accumulation point u^* of $\{u^r\nu\}$ solves (2.8).

Proof: Let $||u^r\nu|| \leq C_1$ then $\phi(u^r\nu) \leq C_2$, consequently, by monotonicity (Lemma 4.1) $\phi(\tilde{u}^{r;k}) \leq C_2$ for all r and k.

Moreover

$$\phi_k^r = e_k^T \cdot (M \tilde{u}^{r;k} - p)$$

is also bounded: $\phi_k^r \leq C_3$.

Let $\phi^* = \lim \phi(u^r)$ and, for any given $\varepsilon > 0$

$$\phi^* - \phi(u^r) < \varepsilon$$

$$\Longrightarrow \phi^* - \phi(\tilde{u}^{r;k}) < \varepsilon.$$

(4.9) implies

$$0 \leq \frac{1}{2} \cdot \mu \cdot (2 - \mu) \cdot (\phi_k^r)^2 < \varepsilon.$$

We consider three cases

1. $\mu = 0$: $\tilde{u}_k^{r;k+1} = \tilde{u}_k^{r;k} = 0$, $\phi_k^r < 0$.

2. $0 < \mu < \omega$: $\tilde{u}_k^{r;k+1} = 0$, $\phi_k^r < 0$, $\mu = -\tilde{u}_k^{r;k}/\phi_k^r$

$$\Longrightarrow 0 \leq -\frac{1}{2} \cdot (2 - \omega)\, \tilde{u}_k^{r;k} \cdot \phi_k^r < \varepsilon.$$

3. $\mu = \omega$: $\frac{1}{2} \cdot \omega \cdot (2 - \omega) \cdot (\phi_k^r)^2 < \varepsilon$,

(5.1) $\Longrightarrow$ $\phi_k^r < \sqrt{2\varepsilon/(\omega \cdot (2 - \omega))}$.

Since $\{u^{r_\nu}\}$ is bounded we have again in this case

$$|\phi_k^{r_\nu} \cdot u_k^{r_\nu}| < C_1 \cdot \sqrt{2\varepsilon/(\omega \cdot (2 - \omega))}.$$

Summarizing we have

$$|\phi_k^{r_\nu} \cdot u_k^{r_\nu}| < \max(C_1 \cdot \sqrt{2\varepsilon/(\omega \cdot (2 - \omega))},\ 2\varepsilon/(2 - \omega)) \tag{5.2}$$

for all three cases.

Now let $\{u^{r_{\nu_j}}\}$ be a subsequence of $\{u^{r_\nu}\}$ converging to u^*. Firstly we note

$$u^* \geq \Theta_n.$$

Since $\{\phi_k^{r_\nu}\}$ is bounded for all k there is a subsequence converging to a limit - y^*.

By (5.1)

$$y^* \geq \Theta_n,$$

and by (5.2)

$$u^{*T}\, y^* = 0.$$

According to the definition of ϕ_k^r we have

$$y^* = -M u^* + p,$$

thus u^*, y^* is a solution to (2.8). By (4.8)

$$\tilde{y}_k^{r;k+1} = -\tilde{u}_k^{r;k} - \omega \cdot e_k^T (M \tilde{u}^{r;k} - p)$$

if $\tilde{u}_k^{r;k+1} = 0$. Since y^* is unique and by construction we conclude that y^* is an accumulation point of every convergent subsequence of $\frac{1}{\omega} y^{r_\nu}$.

If the assertion of Theorem 5.1 is not true, then

$$\lim_{r \to \infty} ||u^r|| = \infty.$$

Now we go into a detailed discussion of the case that there is an unbounded subsequence of $\{u^r\}$. For this purpose we define

Definition 5.1: Let $U = \{u \in R^n \mid M u = \Theta_d\}$ and V the orthogonal complement of U.

Lemma 5.1: If $v \in V$ then

$$v^T M v \leq \lambda_o \cdot ||v||^2 \leq 0,$$

where $\lambda_o < 0$ is the largest eigenvalue of M different from zero.

Proof: Recall that M is symmetric and negative semidefinite and

$$v^T M v = 0 \Longleftrightarrow \tilde{A}^T v = \Theta_n \Longleftrightarrow M v = \Theta_n.$$

■

It is possible to decompose every $u \in R^n$ according to

(5.3) $$u = u_1 + u_2, \quad u_1 \in U, \quad u_2 \in V.$$

If we assume additionally $\phi(u) \geq 0$ then by (4.2) and Lemma 5.1:

$$p^T u \leq \frac{1}{2} \cdot u^T M u = \frac{1}{2} \cdot u_2^T M u_2 \leq \frac{1}{2} \cdot \lambda_o \cdot ||u_2||^2 \leq 0$$

or

$$-\frac{1}{2} \cdot \lambda_o \cdot ||u_2||^2 \leq -p^T u \leq ||p|| \cdot ||u||,$$

thus

(5.4) $$\frac{||u_2||^2}{||u||} \leq \frac{2 \cdot ||p||}{-\lambda_o}.$$

Theorem 5.2: Let $\{u^{r_\nu}\}$ be a subsequence of $\{u^r\}$ with $\lim_{\nu \to \infty} u^{r_\nu} = \infty$. If $v^\nu = u^{r_\nu} / ||u^{r_\nu}||$ then

$$\lim_{\nu \to \infty} M v^\nu = \Theta_n$$

and for each accumulation point v^* of $\{v^\nu\}$ is

$$v^* \geq \Theta_n, \quad v^* \neq \Theta_n \quad \text{and} \quad p^T v^* \leq 0.$$

Proof: For $||u^{r_\nu}|| \longrightarrow \infty$ we decompose as above

$$u^{r_\nu} = u_1^{r_\nu} + u_2^{r_\nu}, \quad u_1^{r_\nu} \in U, \quad u_2^{r_\nu} \in V.$$

By (5.4)

$$\frac{||u_2^{r_\nu}||^2}{||u^{r_\nu}||^2} \leq \frac{2 \cdot ||p||}{-\lambda_o} \quad \frac{1}{||u^{r_\nu}||^2} \longrightarrow 0.$$

Thus

$$\lim_{\nu \to \infty} M v^\nu = \lim M u_2^{r_\nu} / ||u^{r_\nu}|| = \Theta_n.$$

Let v^* be an accumulation point of $\{v^\nu\}$. Then obviously

$$v^* \geq \Theta_n, \quad v^* \neq \Theta_n \quad \text{and} \quad p^T v^* < 0 \quad \text{by (4.2)}.$$

We note that the conditions of the Theorem can be verified numerically in a very convenient manner. It suffices to pick out a divergent subsequence $\{u^{r_\nu}\}$ (if there is any). $M v^\nu$ then converges to Θ_n. If $||M v^\nu||$ is sufficiently small, all the other assertions of the Theorem are fulfilled automatically by v^ν.

If there is a $v \geq \Theta_n$ with $M v = \Theta_n$ and $p^T v < 0$ then, by a well-known theorem (Gale's theorem, [19], p. 33), (2.8) has no solution.

Corollary 5.1: If for any accumulation point v^* of $u^r / ||u^r||$

$$v^* \geq \Theta_n, \; M v^* = \Theta_n \text{ and } p^T v^* < 0,$$

then (2.8) has no solution.

In this case is $\lim_{r \to \infty} ||u^r|| = \infty$.

The last assertion of the Corollary is a direkt consequence of Theorem 5.1.

In the next paragraph we study the remaining case, namely that there is an accumulation point v^* as in Theorem 5.2 with $p^T v^* = 0$.

6. Instability

Definition 6.1: If there is a $v \in R^n$ with

(6.1) $\quad v \neq \Theta_n,\ v \geq \Theta_n,\ M v = \Theta_n$ and $p^T v = 0$

then problem (2.8) is termed instable.

In this case, even small changes of p can cause $p^T v$ to become negative and then the problem certainly has no solution by Gale's theorem. This sort of instability can cause serious numerical difficulties with any existing method (see [10] for details). It is therefore adviceable to treat this case with great care.

The problem becomes relevant with the iteration method described if there is at least one unbounded subsequence $\{u^{r_\nu}\}$ of $\{u^r\}$ and if for any accumulation point v^* of $u^{r_\nu}/||u^{r_\nu}||$ (6.1) holds. It is possible, however, to reduce the size and complexity of problem (2.8) considerably once v^* is known. For let u^*, y^* be any solution to (2.8) then

$$v^{*T} M u^* + v^{*T} y^* = p^T v^*$$

implies

$$v^{*T} y^* = 0.$$

Since $v^* \neq \Theta_n$, $v^* \geq \Theta_n$ and $y^* \geq \Theta_n$, we conclude

$$y_j^* = 0 \quad \text{whenever } v_j^* > 0.$$

If we denote

$$J_o = \{j \in \{1, \ldots, n\} \mid v_j^* > 0\}$$

and

$$w^o = v^*$$

then it is sufficient to solve the following problem

(6.2)
$$M u + y = p$$
$$u_j \geq 0 \text{ for } j \notin J_o,$$
$$y_j \geq 0 \text{ for all } j,$$
$$y_j = 0 \text{ for } j \in J_o$$

and no restrictions on u_j for $j \in J_o$.

Problem (6.2) can be reduced to a smaller complementarity problem by eliminating all u_j, $j \in J_o$. This procedure offers two advantages:

1. The conditions (6.1) for v^* obtained numerically by the method can be tested carefully by elimination,
2. The problem is reduced to a simpler and smaller problem having normally no further instabilities.

With the reduced problem, the iteration process is set forth until it terminates in one of the following cases:

1. Convergence in the sense of Theorem 5.1.
2. The assertion that there is no solution as in Corollary 5.1.
3. A new instability is detected by finding a v_1^* satisfying (6.1).

In the third case we define J_1 as above and

$$J_1 = J_o \cup J_1, \quad w^1 = w^o + v_1^*.$$

Again we eliminate to reduce the problem as described. The process then eventually stops with case 1 and any index set J_k together with a solution u^*, y^* with $u_j^* \geq 0$ for $j \notin J_k$, $y_j^* \geq 0$ for all j, $y_j^* = 0$ for $j \in J_k$. Furthermore we have a vector w^k with $w_j^k > 0$ for $j \in J_k$, $w_j^k = 0$ for $j \notin J_k$ and $Mw^k = \Theta_n$, $p^T w^k = 0$. Then obviously

$$u = u^* + \mu \cdot w^k \geq \Theta_n$$

if $\mu \geq 0$ is sufficiently large, and u, y^* solves problem (2.8).

The most prominent case of stability is characterized by the fact that there is a u and y componentwise positive with $Mu + y = p$ (Slater condition, see [19], 5.4.3). In [10] stability is discussed in a more general context.

7. Numerical Considerations

The method described in the preceding paragraphs can be programmed very easily and there are no difficulties in using special structures as sparsity of $\tilde{A}$, $\tilde{C}^{-1}$ and u. Unfortunately, there is no indication for the best choice of the overrelaxation factor ω. Numerical evidence and theoretical considerations by Cryer [4] suggest to choose $\omega = 1.8$.

It is necessary with iterative methods to formulate an appropriate stopping criterion. The method described offers such a criterion in a quite natural way because at each step ist is necessary to compute

$$\eta_k^{r;k} = e_k^T \cdot (M\, \tilde{u}^{r;k} - p).$$

Let

$$\tilde{y}_k^{r;k} = \begin{cases} -\eta_k^{r;k} \text{ if } \eta_k^{r;k} < 0 \text{ and } \tilde{u}_k^{r;k} = 0 \\ 0 \text{ elsewhere.} \end{cases}$$

Then $\eta^{r;k} + \tilde{y}^{r;k} =: \varepsilon^{r;k}$ is the defect of the approximate solution $\tilde{u}^{r;k}$. The iteration may be stopped if $||\varepsilon^{r;k}||$ is smaller than any prescribed accuracy.

If the problem has no solution, the vectors u^r are not bounded. Consequently, if $||u^r||$ is greater than a prescribed positive constant C, we can apply the test for unsolvability of Corollary 5.1 which means no additional computational efford because $e_k^T M\, \tilde{u}^{r;k}$ has to be computed anyway at each step. Three possibilities for the outcome of this test are given:

1. $||M\, v^r||$ is not smaller than a given $\varepsilon > 0$. Then we continue the iteration with a greater constant C.

2. $||M\, v^r||$ is small and $p^T v^r$ is decisively negative. Then we can suspect that the problem has no solution.

3. $||M\, v^r||$ is small and $|p^T v^r|$ too. Then we reduce the problem as described in § 6.

It seems to be somewhat inconsistent to start an iteration procedure by an elimination process as described in § 2. However, these preparatory elimination steps are very adviceable because they can be performed with great care (multiple precision arithmetic, pivoting) in order to avoid numerical instabilities. In the quadratic programming code using Lemke's method written in the Zentralinstitut für Angewandte Mathematik of the Nuclear Research Center Jülich exactly the same preparatory algorithm is carried out. This elimination process reduces the number of variables and thus the complexity of the problem by a number of pivot steps known in advance. So it is possible to detect e.g. inconsistencies or linear dependent rows in (1.3), to make a simple check wether $\tilde{C}$ is indeed nonsingular and finally to store equations (2.3) and (2.7) on disk or tape in order to use them after solutions u^*, y^* of (2.8) are obtained for computing the corresponding x^*.

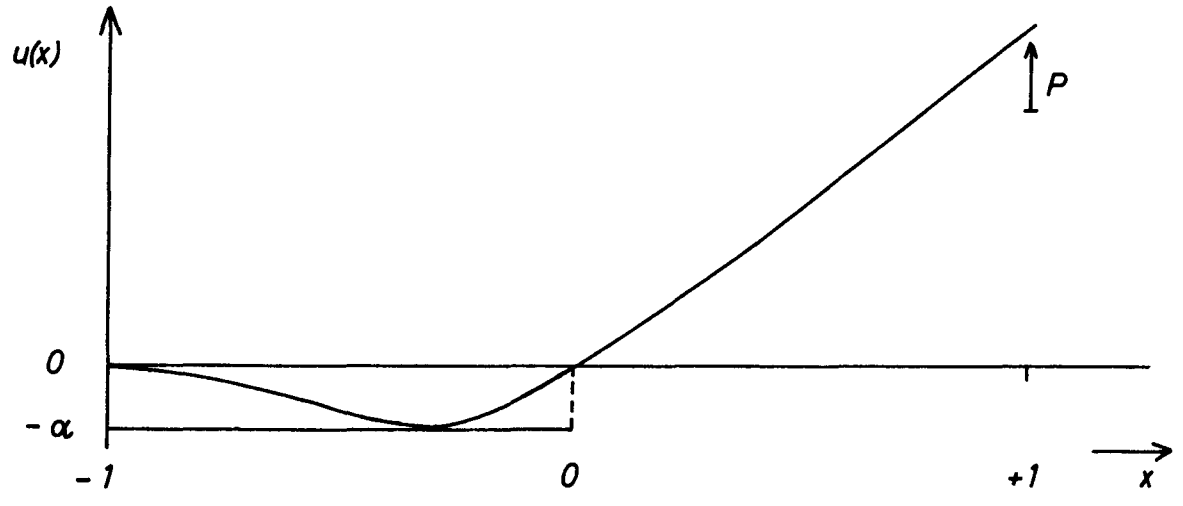
In a similar manner as described it is possible to treat the following modification of Cryer's method

$$(7.1) \qquad z^{r+1} = z^r - \omega\cdot(A_L\, u^{r+1} - z^r + A_R\, u^r - p).$$

Numerical tests indicated in some cases a slight superiority of this variant over Cryer's original method.

8. Example

We consider the following simple example:



A thin homogenous elastic bar is clamped at one end ($x = -1$) and the other end ($x = +1$) is free. The free end is subjected to a certain force P. As an additional condition, the bar is restricted to the "tube" given by $0 \geq u(x) \geq -\alpha$ for $-1 \leq x \leq 0$. The corresponding variational problem is

$$\int_{-1}^{+1} \left[\frac{1}{2} (u''(x))^2 - u(x) \right] dx - u(1) \longrightarrow \text{Minimum},$$

subject to

$$0 \geq u(x) \geq -\alpha \text{ for } -1 \leq x \leq 0,$$
$$u(-1) = u'(-1) = 0.$$

The Euler equations of the variational problem involve derivatives of u of the order four whereas u is only *two* times continuously differentiable. Consequently, the differential equation formulation of this problem needs some caution in order to define the solution of it in a correct manner.

The variational problem is discretized by replacing the solution by some spline function u(x), e.g. a third degree spline with continous first derivative. The parameters of u(x) play the role of the variables x in problem (1.1-3). The side conditionsare replaced by a finite number of inequalities

$$0 \geq u(x_i) \geq -\alpha$$

for $-1 \leq x_i \leq 0$. So we have two discretization parameters, N_s for the

number of spline intervals and N_I for the number of inequalities in the discretization of the side conditions. Only N_I has influence on the size of M. The matrix C constructed in this way is only semidefinite because u'' lacks the constant and linear terms of the spline polynomials. After the first elimination step, the matrix $\tilde{C}$ is definite, as required. The matrices $\tilde{A}$ and $\tilde{C}^{-1}$ are sparse matrices, but M is full.

For the numerical results I am very much indebted to Mr. Herbert Winter who carried them out very carefully on the IBM/370-168 of the ZAM in the KFA Jülich. The results of a systematic comparison of different methods (Fridman and Chernina [11], Cryer with $\omega = 1.0$ and $\omega = 1.8$ and the modification (7.1) with $\omega = 1.8$) are shown in Table I. These results clearly indicate the superiority of Cryer's method. Similar results were found with other examples although these were not treated in a systematic manner. If there were problems having no solution, mainly caused by data errors, this fact was detected very quickly by the method.

Any comparision with elimination methods is necessarily not fair because this means to give weights to the advantages and drawbacks of the different classes of methods. In any case, however, there was no significant argument from the viewpoint of computing time favouring elimination methods. This statement, of course, is only true for problems with sparse matrices.

Table I. Comparison of Different Methods

Method	- Log of desired accuracy	N_I	N_s	Iteration count	Computing time min	Computing time sec
CRYER	5	20	4	1 814		2
$\omega = 1.0$	6	100	8	16 783	6	0
	6	100	10	31 917	11	26
	5	100	8	14 167	5	16
	5	100	10	13 982	5	0
CRYER	5	20	4	520		1
$\omega = 1.8$	3	100	8	387		12
	3	100	10	416		13
	5	100	8	4 006	1	26
	5	100	10	5 568	1	59
Modification (7.1)	5	20	8	573		1
$\omega = 1.8$	5	100	10	4 393	1	34
FRIDMAN	5	20	4	1 814		2
CHERNINA	3	100	10	528		15
	3	100	8	530		15
	5	100	8	15 853	5	38
	5	100	10	15 043	5	31

Comments to Table I.

N_I Number of inequalities (1.2)

N_S Number of spline intervals used

In the notation of § 1 is

$d = 4\ N_s$

$k = 2\ N_s$

$n = N_I$.

The iteration count means in the case of the method of Fridman and Chernina the number of cycles.

The computing time is the running time of the program (without compilation time). The significance of the computing times given is limited because of system modifications during the tests.

In the tests, no efford was made to use special structures of the problems. The computation time can be reduced in the case $N_I = 100$ by more than a factor 5 - 6 by using sparsity of the matrices involved.

References:

1. AMOS, D. E., SLATER, M. L.
Polynomial and spline approximation by quadratic programming.
Comm. ACM 12, 379 - 381 (1969).

2. CANNON, J. R., CECCHI, Maria M.
The numerical solution of some biharmonic problems by mathematical programming techniques.
SIAM J. Numer. Anal. 3, 451 - 466 (1966).
Zbl. 255.65038

3. COLLATZ, L., WETTERLING, W.
Optimierungsaufgaben.
Heidelberger Taschenbücher Band 15.
Berlin, Heidelberg, New York: Springer-Verlag 1966.

4. CRYER, C. W.
The solution of a quadratic programming problem using systematic overrelaxation.
SIAM J. Control 9, 385 - 392 (1971).
MR 45.7971

5. de DONATO, O., FRANCHI, A.
A modified gradient elastoplastic analysis by quadratic programming.
Comput. Methods Appl. Mech. & Eng. 2, 107 - 131 (1973).

6. D'ESOPO, D. A.
A convex programming procedure.
Naval Res. Logist. Quart. 6, 33 - 42 (1959).
MR 21.3273

7. DU VAL, Patrick
The unloading problem for plane curves.
Amer. J. Math. 62, 307 - 311 (1940).

8. DUVAUT, G., LIONS, J. L.
Les inéquations en méchanique et en physique.
Paris: Dunod 1972.

9. EAVES, B. C.
On quadratic programming.
Management Sci. 17, 698 - 711 (1971).
Zbl. 242.90040

10. ECKHARDT, U.
Theorems on the dimension of convex sets.
To appear

11. FRIDMAN, V. M., CHERNINA, V. S.
An iteration process for the solution of the finite-dimensional contact problem.
USSR Comput. Math. math. Phys. 7 (1), 210 - 214 (1967).

12. GOLUB, G. H., SAUNDERS, M. A.
Linear least squares and quadratic programming.
In: J. ABADIE, ed.: Integer and Nonlinear Programming.
Amsterdam, London: North-Holland Publishing Company 1970, pp. 229 - 256.

13. HABETLER, G. J., PRICE, A.L.
An iterative method for generalized complementarity problem.
J. Optimization Theory Appl. 11, 36 - 48 (1973).
Zbl. 251.90040

14. HERAKOVICH, C. T., HODGE, P. G.
Elastic-plastic torsion of hollow bars by quadratic programming.
Int. J. Mech. Sci. 11, 53 - 63 (1969).

15. HILDRETH, C. G.
A quadratic programming procedure.
Naval Res. Logist. Quart. 4, 79 - 85 (1957).
MR 19.619

16. INGLETON, A. W.
A problem in linear inequalities.
Proc. London Math. Soc. (3) 16, 519 - 536 (1966).

17. LEMKE, C. E.

On complementary pivot theory.

In: G. B. DANTZIG and A. F. VEINOTT, eds.:
Mathematics of the Decision Sciences, Part I.
Lectures in Applied Mathematics, Vol. 11.

Providence: American Mathematical Society 1968, pp. 95 - 114.

MR 38.958

18. MANCINO, O. G., STAMPACCHIA, G.

Convex programming and variational inequalities.

J. Optimization Theory Appl. 9, 3 - 23 (1972).

MR 46.4944

19. MANGASARIAN, O. L.

Nonlinear Programming.

New York, St. Louis, San Francisco, London, Sydney, Toronto, Mexico, Panama: McGraw-Hill Book Company 1969.

20. RUST, B. W., BURRUS, W. R.

Mathematical programming and the numerical solution of linear equations.

Modern Analytic and Computational Methods in Science and Mathematics, Number 38.

New York, London, Amsterdam: American Elsevier Publishing Comp., Inc. 1972.

21. SEĬSOV, Ju. B.

An iteration method for solution of the problems of quadratic programming.

Izv. Akad. Nauk Turkmen. SSR, Ser. Fiz.-Tehn. Him. Geol. Nauk 1971, no. 2, 91 - 94.

MR 44.5108

22. SHRAGER, Richard I.

Quadratic programming for nonlinear regression.

Comm. ACM 15, 41 - 45 (1972).

23. SZABO, A. A., CHUNG-TA, Tsai

The quadratic programming approach to the finite element method.

International J. for Numerical Methods in Engineering, 5, 375 - 381 (1973).

24. TORSTI, J. J., AURELA, A. M.

A fast quadratic programming method for solving illconditioned systems of equations.

J. Math. Anal. and Appl. 38, 193 - 204 (1972).

25. WOLFE, Ph.

The simplex method for quadratic programming.

Econometrica 27, 382 - 398 (1959).