NORMAL SOLUTIONS OF LINEAR PROGRAMS

by

O. L. Mangasarian

Computer Sciences Technical Report #498

March 1983

UNIVERSITY OF WISCONSIN-MADISON
COMPUTER SCIENCES DEPARTMENT

NORMAL SOLUTIONS OF LINEAR PROGRAMS

O. L. Mangasarian

Technical Report #498
March 1983

ABSTRACT

The solvability of a linear program is characterized in terms of the existence of a fixed projection on the feasible region, of all sufficiently large positive multiples of the gradient of the objective function. This projection turns out to be the normal solution obtained by projecting the origin on the optimal solution set. By seeking the solution with least 2-norm which minimizes the 1-norm infeasibility measure of a system of linear inequalities or of the optimality conditions of a linear program, one is led to a simple minimization problem of a convex quadratic function on the nonnegative orthant which is guaranteed to be solvable by a successive overrelaxation (SOR) method. This normal solution is an exact solution if the original system is solvable, otherwise it is an error-minimizing solution. New computational results are given to indicate that SOR methods can solve very large sparse linear programs that cannot be handled by an ordinary linear programming package.

AMS (MOS) Subject Classifications: 90C05, 90C06, 90C20, 15A39

Key Words: Linear programming, linear inequalities, large scale systems, quadratic programming

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This material is based on work sponsored by National Science Foundation Grant MCS-8200632.

NORMAL SOLUTIONS OF LINEAR PROGRAMS

O. L. Mangasarian

1. Introduction

A *normal solution* to a linear program is an exact solution with some least norm property if the linear program is solvable, otherwise it is an *approximate* solution with some least norm property also. By an *approximate* solution we mean a point which minimizes a measure of satisfaction of the optimality conditions of the linear program. By considering normal solutions we are led to:

(i) Iterative successive overrelaxation (SOR) methods capable of solving very large linear programs.

(ii) Approximate solutions to poorly posed or unsolvable linear programs.

(iii) A stable solution or approximate solution, to a linear program, endowed with a least norm property.

For solvable linear programs our normal solution is essentially equivalent to that of Tikhonov and Arsenin [16] which they obtain by solving an asymptotic problem [16, Theorem 1, p. 226], whereas our solution is obtained by solving a simpler exact problem, problem (2.2) for any $\varepsilon \in (0,\bar{\varepsilon}]$ for some $\bar{\varepsilon} > 0$ (Theorem 2.1). Tikhonov and Arsenin's weaker asymptotic result comes about because they square the objective function of the linear program in their regularization problem [16, p. 226] and thereby lose an essential exact feature of our problem (2.2). Tikhonov and Arsenin also do not consider the

---

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041. This material is based on work sponsored by National Science Foundation Grant MCS-8200632.

important case of possibly unsolvable linear programs (Section 4), nor do they give explicit computational methods for solving their asymptotic problem.

We outline now our principal results and their relation to other work. In Section 2 we consider normal solutions of solvable linear programs. In Theorem 2.1 we give a complete characterization of the solvability of a linear program in terms of a 2-norm projection on the feasible region of a sufficiently large but finite positive multiple of the gradient of the objective function. This projection turns out to be fixed and equal to the unique 2-norm projection of the origin on the optimal solution set. Part of Theorem 2.1, a(i), follows readily from [13, Theorem 1], while its converse, part a(ii), which is essential for a comprehensive justification of the linear programming SOR Algorithm 2.3, has not been available before. Theorem 2.2, which follows from Theorem 2.1 and quadratic programming duality, characterizes the solvability of a linear program in terms of the solvability of a convex quadratic function minimization on the nonnegative orthant (2.7) without any a priori assumptions regarding the solvability of the linear program (2.1) as was the case in [11,12]. In addition, Theorem 2.2 gives the complete basis for the linear programming SOR Algorithm 2.3 and its convergence (Theorem 2.4) thereby sharpening earlier convergence results [11,12].

In Section 3 we turn our attention to a system of possibly inconsistent linear inequalities (3.1) and reduce it to the problem of finding the unique least 2-norm solution of the problem of minimizing the 1-norm infeasibility measure of (3.1). The principal advantage of this approach is that it leads to the SOR Algorithm 3.1 which, unlike most other iterative procedures [1,15,3] which require an a priori consistency assumption, will converge no matter whether the original system (3.1) is consistent or not. In either case Algorithm 3.1 will give an exact or approximate solution with least 2-norm (Theorem 3.2). Among the potential useful applications of this approach is in the image reconstruction techniques of tomography [7,8] which require the solution of enormous sparse systems of linear equations with nonnegative variables. Most current iterative techniques for the tomography problem [7,8,2] need an a priori assumption regarding the consistency of the original system. In contrast our Algorithm 3.1 needs no such assumption.

In Section 4 we consider possibly unsolvable linear programs and reduce their solution to finding the least 2-norm primal-dual solution which minimizes the 1-norm of the optimality conditions of the given linear program. This approach leads to an SOR algorithm that is guaranteed to work whether the original linear program is solvable or not. In either case it will give an exact or approximate solution with least 2-norm.

Finally in Section 5 we give some numerical comparisons for one version of our linear programming SOR algorithm with the XMP version [14] of the revised simplex method for medium and large size sparse linear programs. These comparisons indicate that SOR methods can solve very large sparse linear programs that <u>cannot be solved</u> by an ordinary linear programming package.

We briefly describe the notation used. All matrices and vectors are real. For the $m \times n$ matrix A we denote row i by A_i, column j by $A_{\cdot j}$ and the element in row i and column j by A_{ij}. For x in the real n-dimensional Euclidean space R^n, x_i denotes element i for i=1,....,n, and x_+ denotes the vector with components $(x_+)_i = \max\{x_i, 0\}$, i=1,....,n. Vectors are either row or column vectors depending on the context. For x and y in R^n, xy denotes the scalar product $\sum_{i=1}^{n} x_i y_i$, while $\|x\|_t$ for $1 \leqq t \leqq \infty$ denotes the t-norm $(\sum_{i=1}^{n} |x_i|^t)^{\frac{1}{t}}$, $\|x\|$ denotes an arbitrary but fixed norm on R^n and $\|A\|$ denotes the subordinate matrix norm $\max_{\|x\|=1} \|Ax\|$ for an $m \times n$ matrix. The vector e will denote a vector of ones in any real Euclidean space. R^n_+ will denote the nonnegative orthant $\{x | x \in R^n,\ x \geqq 0\}$. For a point c in R^n, a closed set X in R^n and a number $t \in [1,\infty]$ the t-norm projection $p_t(c,X)$ of the point c on X is defined by

$$\|c - p_t(c,X)\|_t = \min_{x \in X} \|c - x\|_t$$

For a function $f: R^n \to R$ which is twice differentiable on R^n, $\nabla f(x)$ denotes the n-dimensional gradient vector at x with components $\nabla_{x_i} f(x)$, i=1,....,n, and $\nabla^2 f(x)$ denotes the $n \times n$ Hessian matrix at x with elements $(\nabla^2 f(x))_{ij}$, i, j=1,....,n.

2. Normal solutions of solvable linear programs

We consider here the linear program

(2.1) maximize cx subject to $x \in X := \{x \mid x \in R^n, Ax \leqq b, x \geqq 0\}$

where b and c are given vectors in R^m and R^n respectively and A is a given $m \times n$ real matrix. Let $\bar{X}$ denote the (possibly empty) optimal solution set of (2.1). We shall assume throughout this section that this linear program is feasible, that is X is nonempty. We begin with the following fundamental and geometrically plausible result.

2.1 Theorem Let the linear program (2.1) be feasible. Then

a. (i) $\max_{x \in X} cx$ has a solution $\Rightarrow \exists \bar{\varepsilon} > 0: p_2(\frac{c}{\varepsilon}, X) = p_2(0, \bar{X})$ for all $\varepsilon \in (0, \bar{\varepsilon}]$

(ii) $\max_{x \in X} cx$ has a solution and $\bar{x} = p_2(0, \bar{X})$ $\Leftarrow \exists \bar{\varepsilon} > 0, \bar{x}: p_2(\frac{c}{\varepsilon}, X) = \bar{x}$ for all $\varepsilon \in (0, \bar{\varepsilon}]$

where $p_2(x, X)$ denotes the 2-norm projection of x on X.

b. $$\sup_{x \in X} cx = \infty \iff \|p_2(\tfrac{c}{\varepsilon}, X)\| \to \infty \text{ as } \varepsilon \to 0+$$

Proof

a(i): By noting that $p_2(\frac{c}{\varepsilon}, X)$ is a solution of either of the equivalent problems

(2.2) $$\min_{x \in X} \|x - \tfrac{c}{\varepsilon}\|_2^2 \iff \min_{x \in X} -cx + \tfrac{\varepsilon}{2}\|x\|_2^2$$

the implication of a(i) follows from Theorem 1 of [13].

a(ii): Since $\bar{x} = p_2(\frac{c}{\varepsilon}, X)$ for all $\varepsilon \in (0, \bar{\varepsilon}]$, then there exists $(u(\varepsilon), v(\varepsilon)) \in R^{m+n}$ such that $(\bar{x}, u(\varepsilon), v(\varepsilon))$ satisfies the Karush-Kuhn-Tucker conditions [9] for (2.2), that is

(2.3) $\quad \varepsilon\bar{x} - c + A^T u(\varepsilon) - v(\varepsilon) = 0,\ v(\varepsilon)\bar{x} = 0,\ A\bar{x} \leqq b,\ u(\varepsilon)(A\bar{x}-b) = 0,$

$$(\bar{x},\ u(\varepsilon),\ v(\varepsilon)) \geqq 0,\ \forall\varepsilon \in (0,\bar{\varepsilon}]$$

By the fundamental theorem for the existence of basic feasible solutions for linear equations with nonnegative variables [6, Theorem 2.11], and the complementarity conditions $u(\varepsilon)(A\bar{x}-b) = 0$, $v(\varepsilon)\bar{x} = 0$, it follows that there exist $(u(\varepsilon), v(\varepsilon))$ satisfying (2.3) such that all elements of $(u(\varepsilon), v(\varepsilon))$ not corresponding to some subset of k linearly independent columns of $[A^T\ -I]$ are zero. Since the rank of $[A^T\ -I]$ is n, it follows that we can take $k = n$ and denote by $B(\varepsilon)$ this "basis" matrix of n linearly independent columns of $[A^T\ -I]$. Hence it follows for such a "basic" solution $(u(\varepsilon), v(\varepsilon))$ satisfying (2.3) that

$$\|u(\varepsilon)\ v(\varepsilon)\| \leqq (\|c\| + \bar{\varepsilon}\|\bar{x}\|)\|B(\varepsilon)^{-1}\| \quad \forall\varepsilon \in (0,\bar{\varepsilon}]$$

Since $[A^T\ -I]$ contains a finite number of basis matrices it follows that for some basis matrix B, $\|B(\varepsilon)^{-1}\| \leqq \|B^{-1}\|$ for all $\varepsilon \in (0,\bar{\varepsilon}]$ and consequently

(2.4) $$\|u(\varepsilon)\ v(\varepsilon)\| \leqq (\|c\| + \bar{\varepsilon}\|\bar{x}\|)\|B^{-1}\| \quad \forall\varepsilon \in (0,\bar{\varepsilon}]$$

Now let $\{\varepsilon^i\}$ be a sequence of positive numbers in $(0,\bar{\varepsilon}]$ converging to 0. Then there exists a sequence $\{(u(\varepsilon^i), v(\varepsilon^i))\}$ satisfying (2.3) and (2.4) and hence it is bounded and has an accumulation point $(\bar{u},\bar{v})$ satisfying

(2.5) $\quad -c + A^T\bar{u} - \bar{v} = 0,\ \bar{v}\bar{x} = 0,\ A\bar{x} \leqq b,\ \bar{u}(A\bar{x}-b) = 0,\ (\bar{x},\bar{u},\bar{v}) \geqq 0$

These are the Karush-Kuhn-Tucker conditions for the linear program (2.1) and hence $\bar{x}$ solves (2.1). Since $(\bar{x}, u(\varepsilon), v(\varepsilon))$ also satisfies (2.3) which are also the Karush-Kuhn-Tucker conditions for

$$\min \frac{1}{2}\|x\|_2^2 \quad \text{subject to} \quad Ax \leqq b,\ x \geqq 0,\ cx \geqq c\bar{x}$$

with optimal $x = \bar{x}$ and optimal multiplier vector of $(\frac{u(\varepsilon)}{\varepsilon}, \frac{v(\varepsilon)}{\varepsilon}, \frac{1}{\varepsilon})$, it follows that $\bar{x} = p_2(0,\bar{X})$.

(b) ($\Leftarrow$): If not then the linear program (2.1) has a solution and by part a(i) of this theorem $\exists\bar{\varepsilon} > 0$: $p_2(\frac{c}{\varepsilon},X) = p_2(0,\bar{X})$ for all $\varepsilon \in (0,\bar{\varepsilon}]$. This however contradicts the hypothesis that $\|p_2(\frac{c}{\varepsilon},X)\| \to \infty$ as $\varepsilon \to 0+$.

($\Rightarrow$): If not then, for a sequence of positive numbers $\{\varepsilon^i\}$ converging to zero, the sequence $\{\|p_2(\frac{c}{\varepsilon^i},X)\|\}$ is bounded. By defining $x(\varepsilon^i) := p_2(\frac{c}{\varepsilon^i},X)$ we get that $x(\varepsilon^i)$ and some $(u(\varepsilon^i), v(\varepsilon^i)) \in R^{m+n}$ satisfy the Karush-Kuhn-Tucker conditions for $\min\limits_{x\in X} \|x - \frac{c}{\varepsilon^i}\|_2^2$ for $i=1,2,\ldots\ldots,$ that is

$$\text{(2.6)}\quad \varepsilon^i x(\varepsilon^i) - c + A^T u(\varepsilon^i) - v(\varepsilon^i) = 0,\ v(\varepsilon^i)\ x(\varepsilon^i) = 0,\ Ax(\varepsilon^i) \leqq b,\ u(\varepsilon^i)(Ax(\varepsilon^i) - b) = 0,$$
$$(x(\varepsilon^i),\ u(\varepsilon^i),\ v(\varepsilon^i)) \geqq 0$$

By the same argument as in the proof of part a(ii) of this theorem we can show that the sequence $\{(u(\varepsilon^i), v(\varepsilon^i))\}$ satisfying (2.6) can be taken as bounded since $\{x(\varepsilon^i)\}$ is also bounded. Thus the sequence $\{(x(\varepsilon^i), u(\varepsilon^i), v(\varepsilon^i))\}$ is bounded and has an accumulation point $(\bar{x},\bar{u},\bar{v})$ satisfying (2.5). Hence $\bar{x}$ solves (2.1) which contradicts the hypothesis that $\sup\limits_{x\in X} cx = \infty$. □

By noting that the quadratic programming dual [9] to (2.2) is

$$\text{(2.7)} \qquad \underset{(u,v)\in R_+^{m+n}}{\text{minimize}}\ \frac{1}{2}\|A^T u - v - c\|_2^2 + \varepsilon bu$$

where the primal and dual variables x and (u,v) are related by

(2.8)
$$x = \frac{1}{\varepsilon}(-A^T u + v + c),$$

the following theorem is a direct consequence of Theorem 2.1.

2.2 <u>Theorem</u> The linear program (2.1) is solvable if and only if there exists an $\bar{\varepsilon} > 0$ such that for each $\varepsilon \in (0,\bar{\varepsilon}]$ the quadratic program (2.7) has a solution $(u(\varepsilon), v(\varepsilon))$ and such that the vector $\bar{x}$ defined by

(2.9)
$$\bar{x} := \frac{1}{\varepsilon}(-A^T u(\varepsilon) + v(\varepsilon) + c) \quad \varepsilon \in (0,\bar{\varepsilon}]$$

is independent of ε, in which case $\bar{x} = p_2(0,\bar{X})$.

If we define the objective function of (2.7) by

(2.10)
$$f(z) := \frac{1}{2}\|A^T u - v - c\|_2^2 + \varepsilon bu, \; z := \begin{pmatrix} u \\ v \end{pmatrix}$$

then we can prescribe an SOR procedure for solving (2.7) which in view of Theorem 2.2 solves the linear program (2.1). The SOR procedure is essentially a gradient projection algorithm of the following type

(2.11)
$$z_j^{i+1} = (z_j^i - \omega(\nabla^2 f(z^i))_{jj}^{-1} \nabla_{z_j} f(z_1^{i+1}, \ldots\ldots, z_{j-1}^{i+1}, z_j^i, \ldots\ldots, z_{m+n}^i))_+ ,$$
$$0 < \omega < 2, \; j=1,\ldots\ldots,m+n.$$

More specifically [12] the following SOR algorithm for solving the linear program (2.1) follows directly from (2.11) and (2.10).

2.3 <u>LPSOR(A, b, c) Algorithm</u> Choose $(u^0, v^0) \in R^{m+n}$, $\omega \in (0,2)$ and $\varepsilon > 0$. Having (u^i, v^i) determine (u^{i+1}, v^{i+1}) as follows:

$$u_j^{i+1} = \left(u_j^i - \frac{\omega}{\|A_j\|_2^2}\left(A_j\left(\sum_{\substack{\ell=1\\ \text{for } j>1}}^{j-1} (A^T)_{\cdot\ell}\, u_\ell^{i+1} + \sum_{\ell=j}^{m} (A^T)_{\cdot\ell}\, u_\ell^i - v^i - c\right) + \varepsilon b_j\right)\right)_+ \quad \text{if } A_j \neq 0$$

$$u_j^{i+1} = 0 \quad \text{if } A_j = 0 \qquad\qquad j=1,\ldots\ldots,m$$

$$v^{i+1} = (v^i - \omega(-A^T u^{i+1} + v^i + c))_+$$

Note that Algorithm 2.3 is sparsity-preserving for it works with the <u>rows</u> of A only and the product AA^T need not be computed.

The following convergence theorem which follows from Theorem 2.2 above and [12, Theorem 2] sharpens previous LPSOR convergence theorems [11, Theorem 3.2] and [12, Theorem 4].

2.4 <u>LPSOR(A, b, c) Convergence Theorem</u>

(a) The linear program (2.1) has a solution if and only if there exists a real positive number $\bar{\varepsilon}$ such that for each $\varepsilon \in (0,\bar{\varepsilon}]$, each accumulation point $(u(\varepsilon), v(\varepsilon))$ of the sequence $\{(u^i,v^i)\}$ generated by the LPSOR(A, b, c) Algorithm 2.3 solves (2.7) and the corresponding $\bar{x}$ determined by (2.9) is independent of ε, in which case $\bar{x} = p_2(0,\bar{X})$.

(b) If the linear program (2.1) has a solution and its constraints satisfy the Slater constraint qualification, that is $Ax < b$ for some $x \geqq 0$, then the sequence $\{(u^i,v^i)\}$ of the LPSOR(A, b, c) Algorithm 2.3 is bounded and has an accumulation for each $\varepsilon \in (0,\bar{\varepsilon}]$ for some $\bar{\varepsilon} > 0$.

Computational results for the Algorithm 2.3 are given in Section 5.

3. Normal solutions of possibly inconsistent linear inequalities

We consider in this section the possibly inconsistent system of linear inequalities

$$(3.1) \qquad Ax \leqq b,\ x \geqq 0$$

where A is a given $m \times n$ matrix and b is a given vector in R^m. If we try to "solve" the above system by an SOR [10] procedure applied to the obvious 2-norm minimization problem

$$(3.2) \qquad \min_{x \geqq 0} \|(Ax-b)_+\|_2^2 = \min_{(x,y) \geqq 0} \|Ax+y-b\|_2^2 =: \min_{(x,y) \geqq 0} \theta(x,y)$$

one needs the condition

$$(3.3) \qquad \nabla\theta(x,y) = \begin{bmatrix} A^T(Ax+y-b) \\ \\ Ax+y-b \end{bmatrix} > 0, \quad \text{for some} \quad (x,y) \in R^{n+m}$$

to guarantee boundedness of the SOR iterates [10, Theorem 2.2], which by the Gordan Theorem [9, Theorem 2.4.5] is equivalent to the condition that

$$(3.4) \qquad Ax \leqq 0,\ 0 \neq x \geqq 0, \quad \text{has no solution}$$

Unfortunately this condition is <u>not</u> satisfied in general, as is the case when the feasible region is nonempty and unbounded. To avoid this difficulty we use the SOR procedure of Section 2 to find the 2-norm projection of the origin in R^{n+m} on the nonempty solution set of the linear program

$$(3.5) \qquad \min_{(x,y) \in R_+^{n+m}} \{ey \mid Ax - y \leqq b\}$$

which is the equivalent of the problem of minimizing the 1-norm feasibility of (3.1)

$$\min_{x\in R^n_+} \|(Ax-b)_+\|_1 \tag{3.6}$$

The key feature of this approach is that the SOR procedure will work no matter whether the system (3.1) is consistent or not. In either case the SOR procedure will obtain the unique solution $(\bar{x},\bar{y})$ of (3.5) with least 2-norm. In terms of the original inequalities (3.1), $\bar{x}$ is the unique solution of (3.6) which minimizes $\|x, (Ax-b)_+\|_2$. Needless to say, if (3.1) is consistent then $\bar{x}$ is the unique 2-norm projection of the origin in R^n on the nonempty feasible region determined by (3.1). To obtain an SOR procedure for solving (3.5) we take the dual of the quadratic perturbation of (3.5)

$$\min_{(x,y)\in R^{n+m}_+} \{ey + \frac{\varepsilon}{2}\|x,\ y\|_2^2 \,\big|\, Ax-y \leqq b\} \tag{3.7}$$

which turns out to be [9]

$$\min_{(u,v,w)\in R^{m+n+m}_+} \frac{1}{2}\|A^Tu-v\|_2^2 + \frac{1}{2}\|u+w-e\|_2^2 + \varepsilon bu =: \min_{(u,v,w)\in R^{m+n+m}_+} \psi(u,v,w) \tag{3.8}$$

with (x,y) related to (u,v,w) by

$$x = \frac{1}{\varepsilon}(-A^Tu+v), \quad y = \frac{1}{\varepsilon}(u+w-e) \tag{3.9}$$

Since

$$\nabla\psi(u,v,w) = \begin{bmatrix} A(A^Tu-v)+u+w-e+\varepsilon b \\ -(A^Tu-v) \\ u+w-e \end{bmatrix}, \quad \nabla^2\psi(u,v,w) = \begin{bmatrix} AA^T+I & -A & I \\ -A^T & I & 0 \\ I & 0 & I \end{bmatrix} \tag{3.10}$$

it follows that $\nabla\psi(0, e, \lambda e) > 0$ for sufficiently large λ and consequently the iterates of the SOR algorithm of [10, Algorithm 3.2, Remark 3.2] applied to 3.8 will have an accumulation for all positive values of ε. In particular we have the following algorithm and convergence theorem.

3.1 LISOR(A, b) Algorithm Choose $(u^0, v^0, w^0) \in R_+^{m+n+m}$, $\omega \in (0,2)$ and $\varepsilon > 0$. Having (u^i, v^i, w^i) determine $(u^{i+1}, v^{i+1}, w^{i+1})$ as follows:

$$u_j^{i+1} = \Big(u_j^i - \frac{\omega}{1+\|A_j\|_2^2}\Big(A_j\Big(\sum_{\substack{\ell=1\\ \text{for } j>1}}^{j-1} (A^T)_{\cdot\ell}\, u_\ell^{i+1} + \sum_{\ell=j}^{m} (A^T)_{\cdot\ell}\, u_\ell^i - v^i\Big) + u_j^i + w_j^i - 1 + \varepsilon b_j\Big)\Big)_+$$

$$j=1,\ldots\ldots,m$$

$$v^{i+1} = (v^i + \omega(A^T u^{i+1} - v^i))_+$$

$$w^{i+1} = (w^i - \omega(u^{i+1} + w^i - e))_+$$

3.2 LISOR(A, b) Convergence Theorem For each $\varepsilon > 0$ the iterates (u^i, v^i, w^i) of the LISOR(A, b) algorithm are bounded and have an accumulation point $(u(\varepsilon), v(\varepsilon), w(\varepsilon))$. For all $\varepsilon \in (0,\bar{\varepsilon}]$ for some $\bar{\varepsilon} > 0$, the point $(\bar{x},\bar{y})$

$$\text{(3.11)} \qquad \bar{x} := \frac{1}{\varepsilon}(-A^T u(\varepsilon) + v(\varepsilon)), \quad \bar{y} := \frac{1}{\varepsilon}(u(\varepsilon) + w(\varepsilon) - e)$$

is independent of ε and is the unique solution of (3.5) with least 2-norm, and $\bar{x}$ is the unique solution (3.6) with least $\|x, (Ax - b)_+\|_2$.

Proof That the iterates (u^i, v^i, w^i) have an accumulation point which solves (3.8) for all $\varepsilon > 0$ follows from Theorem 2.2 of [10]. That for $\varepsilon \in (0,\bar{\varepsilon}]$ for some $\bar{\varepsilon} > 0$, $(\bar{x},\bar{y})$ defined by (3.11) is the unique solution of (3.5) with least 2-norm follows from the duality equivalence of (3.7) and (3.8) and from Theorem 2.1. Since problems (3.5) and (3.6) are equivalent and

$\bar{y} = (A\bar{x} - b)_+$ for a solution of (3.5), it follows that $\bar{x}$ is the unique solution of (3.6) with least value of $\|x, (Ax - b)_+\|_2$. □

We note here Eremin's algorithm [5] which is one of the few iterative algorithms capable of handling inconsistent inequalities. Eremin gives no computational experience and the presence in his algorithm of a positive stepsize λ_i satisfying $\lambda_i \to 0$ and $\sum_{i=1}^{\infty} \lambda_i = \infty$ may cause slow convergence.

An interesting application of the above method is to the problem of image reconstruction techniques [7,8] where the fundamental problem is to solve the system

(3.12) $$Bx = d \quad x \geqq 0$$

where typically the $m \times n$ matrix B may be of order 28000 × 6000 with less than 1% of nonzero elements [8]. Iterative methods are well suited for such large sparse problems. Unfortunately such methods often require assumptions that are rarely verifiable. Typically such methods assume a priori that the system (3.12) has a solution [7,8]. In contrast our proposed LISOR method requires no assumptions whatsoever when applied to the equivalent problem

(3.13) $$Bx \leqq d, \ -Bx \leqq -d, \ x \geqq 0$$

In particular LISOR$\left(\begin{pmatrix} B \\ -B \end{pmatrix}, \begin{pmatrix} d \\ -d \end{pmatrix}\right)$ will lead to the unique solution $\bar{x}$ of $\min_{x \in R^n_+} \|Bx - d\|_1$ with least $\|x, Bx - d\|_2$.

4. Normal solutions of possibly unsolvable linear programs

We consider here again the linear program (2.1) but make no assumptions whatsoever regarding its feasibility or solvability. The idea here is to apply the LISOR Algorithm 3.1 to the equivalent linear complementarity problem for (2.1) [4]

$$(4.1) \qquad Mz + q \geqq 0, \ qz \leqq 0, \ z \geqq 0$$

where

$$(4.2) \qquad z = \begin{pmatrix} x \\ u \end{pmatrix} \in R^{n+m}, \ M = \begin{pmatrix} 0 & A^T \\ -A & 0 \end{pmatrix}, \ q = \begin{pmatrix} -c \\ b \end{pmatrix}$$

and u is the dual variable. Direct application of the convergence Theorem 3.2 to the $\mathrm{LISOR}\left(\begin{pmatrix} -M \\ q \end{pmatrix}, \begin{pmatrix} q \\ 0 \end{pmatrix}\right)$ algorithm gives that the iterates of $\mathrm{LISOR}\left(\begin{pmatrix} -M \\ q \end{pmatrix}, \begin{pmatrix} q \\ 0 \end{pmatrix}\right)$ are bounded for all $\varepsilon > 0$ and that for all $\varepsilon \in (0, \bar{\varepsilon}]$ for some $\bar{\varepsilon} > 0$ they lead to a $\bar{z}$ which is independent of ε and such that $\bar{z}$ is the unique solution of

$$(4.3) \qquad \min_{z \in R_+^{n+m}} \|(-Mz - q, \ qz)_+\|_1$$

with least $\|z, (-Mz - q, \ qz)_+\|_2$. Hence $\bar{z} = \begin{pmatrix} \bar{x} \\ \bar{u} \end{pmatrix}$, is an exact solution of the primal-dual pair with least 2-norm if the linear program (2.1) is solvable. Else, it is the unique solution of (4.3) with least 2-norm for the vector in $R^{2(n+m)+1}$ composed of the primal-dual variable z, the primal-dual infeasibility $(-Mz - q)_+$, and the primal-dual objective function inequality $(qz)_+$.

5. Computational results

Computational experiments have been carried so far on the LPSOR Algorithm 2.3 only. Results on medium-sized problems were given in [11]. We give below new computational results for randomly generated large sparse problems carried on the VAX 11/780 with double accuracy floating point addition time of 4.6 μs and multiplication time of 6.0 μs. Comparisons were made with Marston's XMP revised simplex linear programming code [14]. The results shown in Table 1 are all for a matrix A with essentially a "tridiagonal" structure and fully dense last column and row. The XMP accuracy was to within 12-figure accuracy of the current objective function when it managed to obtain a solution. The accuracy of the LPSOR was measured by the ∞-norm of the primal infeasibility of the numerical solution and the relative deviation of the computed maximum value from the true maximum. The table indicates that for the accuracy obtained, the LPSOR method becomes competitive with the simplex method as the problem size gets larger and that for very large problems, SOR methods may be the only viable methods of solution.

Table 1

Comparison of the Revised Simplex Code XMP and LPSOR for Solving 2.1
m = no. of inequality constraints, n = no. of nonnegative variables

		XMP		LPSOR		
m	n	Iteration No.	Hr:Min:Sec	Iteration No.	Hr:Min:Sec	Relative Accur.
100	200	123	0:00:11	180	0:00:17	10^{-6}
500	1,000	746	0:03:12	520	0:05:11	10^{-9}
1,000	1,000	2,309	0:42:02	1,640	0:26:12	10^{-4}
2,500	10,000	Could not solve [a)]		480	0:37:25	10^{-4}
5,000	20,000	Could not solve [b)]		660	1:17:53	10^{-4}

[a)] Program was killed after more than 3 hours of CPU time.

[b)] Program used virtual memory space much larger than physical memory, so it ran inefficiently and had to be killed within 10 minutes of CPU time which corresponded to over 8 hours of real elapsed time.

Acknowledgement

I wish to acknowlege the original and tireless programming effort of my former student David P. Anderson in obtaining all the computational results of this paper.

REFERENCES

1. S. Agmon: "The relaxation method for linear inequalities", Canadian Journal of Mathematics 6, 1954, 382-392.

2. Y. Censor: "Row-action methods for huge and sparse systems and their applications" SIAM Review 23, 1981, 444-466.

3. Y. Censor & T. Elfving: "New methods for linear inequalities", Linear Algebra and Its Applications 42, 1982, 199-211.

4. R. W. Cottle & G. B. Dantzig: "Complementary pivot theory of mathematical programming", Linear Algebra and Its Applications 1, 1968, 103-125.

5. I. I. Eremin: "Iteration method for Čebyšev approximations for sets of incompatible linear inequalities", Soviet Mathematics Doklady 3, 1962, 570-572.

6. D. Gale: "The theory of linear economic models", McGraw-Hill, New York 1960.

7. G. T. Herman: "Image reconstruction from projections: The fundamentals of computerized tomography", Academic Press, New York 1980.

8. G. T. Herman, A. Lent & P. H. Lutz: "Relaxation methods for image reconstruction", Communications of the ACM 21, 1978, 152-158.

9. O. L. Mangasarian: "Nonlinear programming", McGraw Hill, New York, 1969.

10. O. L. Mangasarian: "Solution of symmetric linear complementarity problems by iterative methods", Journal of Optimization Theory and Applications 22, 1977, 465-485.

11. O. L. Mangasarian: "Iterative solution of linear programs", SIAM Journal on Numerical Analysis 18, 1981, 606-614.

12. O. L. Mangasarian: "Sparsity-preserving SOR algorithms for separable quadratic and linear programming problems", University of Wisconsin Computer Sciences Technical Report 438, 1981, to appear in "Mathematical programming with data perturbations", (A. V. Fiacco, ed.), special issue of Computers and Operations Research, Pergamon, New York 1983.

13. O. L. Mangasarian & R. R. Meyer: "Nonlinear perturbation of linear programs", SIAM Journal on Control and Optimization 17, 1979, 745-757.

14. R. E. Marsten: "The design of the XMP linear programming library", MIS Technical Report No. 80-2, Management Information Systems, The University of Arizona, Tuscon, Arizona 85721.

15. T. S. Motzkin & I. J. Schoenberg: "The relaxation method for linear inequalities", Canadian Journal of Mathematics 6, 1954, 393-404.

16. A. N. Tikhonov and V. Y. Arsenin: "Solutions of ill-posed problems", Halsted Press, Wiley, New York 1977.