AD-A096 658 WISCONSIN UNIV-MADISON MATHEMATICS RESEARCH CENTER F/G 12/1
LEAST-SQUARES SOLUTION OF LINEAR INEQUALITIES.(U)
NOV 80 S - HAN DAAG29-80-C-0041
UNCLASSIFIED MRC-TSR-2141 NL

1 OF 1

AD A
096658

END
DATE
FILMED
4-81
DTIC

LEVEL II (2)

AD A096658

MRC Technical Summary Report # 2141

LEAST-SQUARES SOLUTION OF LINEAR INEQUALITIES

S.-P. Han

See 1413

Mathematics Research Center
University of Wisconsin–Madison
610 Walnut Street
Madison, Wisconsin 53706

November 1980

(Received September 9, 1980)

Approved for public release
Distribution unlimited

Sponsored by

U. S. Army Research Office
P. O. Box 12211
Research Triangle Park
North Carolina 27709

National Science Foundation
Washington, D.C. 20550

UNIVERSITY OF WISCONSIN-MADISON
MATHEMATICS RESEARCH CENTER

LEAST-SQUARES SOLUTION OF LINEAR INEQUALITIES

S.-P. Han

Technical Summary Report #2141

November 1980

ABSTRACT

The paper deals with a system of linear inequalities in a finite dimensional space. When the system is inconsistent, we are interested in vectors that satisfy the system in a least-squares sense. We characterize such least-squares solutions and propose a method to find one of these solutions. It is shown that for any given system of linear inequalities and for any starting point, the method can produce a solution in a finite number of iterations. Computational results are very satisfactory in terms of accuracy and number of iterations.

AMS (MOS) Subject Classification: 65F10, 65F20

Key Words: Linear inequalities, Least-squares solution

Work Unit: 5 (Operations Research)

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.
This material is based upon work supported by the National Science Foundation under Grant No. ENG 79-03881.

SIGNIFICANCE AND EXPLANATION

The linear least-square problem for equations has been thoroughly studied and efficient computational methods for solving it exist. In this paper we consider its extension to the case of linear inequalities. Such a problem formulation can be used to find a least-squares solution to a linear programming problem, even when the linear programming problem has no solution.

We characterize such least-squares solutions and propose a method to compute one such solution. It is shown that for any given system of linear inequalities the method can produce a least-squares solution in a finite number of steps. Numerical results for the method are extremely satisfactory.

The responsibility for the wording and views expressed in this descriptive summary lies with MRC, and not with the author of this report.

LEAST-SQUARES SOLUTION OF LINEAR INEQUALITIES

S.-P. Han

1. INTRODUCTION

We are concerned with a system of linear inequalities $Ax \leqq b$, where A is an $m \times n$ matrix and b an m-vector. When the system $Ax \leqq b$ is inconsistent, as in the case of equalities, it is often desirable to have a vector that satisfies the system in the least-squares sense. More specifically, we are interested in a vector x that minimizes the quantity $\|(Ax - b)_+\|_2$, where $(Ax - b)_+$ is the m-vector whose i-th component is $\max \{(Ax - b)_i, 0\}$.

The problem to find a least-squares solution to the system $Ax \leqq b$ is fundamental. It is a natural extension of the equality linear least-squares problem and abounds with applications. For instance, a least-squares solution to the system derived from the Karush-Kuhn-Tucker conditions of a linear program can often supply useful information even when the given linear program is infeasible. A technique for solving such a least-squares problem can also be used to check the feasibility or to find a feasible point for a linearly constrained problem.

There are basically two direct approaches for tackling this problem. The first one is to use an unconstrained minimization method to minimize the following function

$$f(x) := \tfrac{1}{2}(Ax - b)_+^T(Ax - b)_+ \tag{1.1}$$

Unfortunately, the function f is not twice differentiable and the powerful Newton-like methods are not applicable.

The second approach is to transform the problem into its equivalent quadratic programming problem

$$\begin{aligned} &\min_{(x,z)} && \tfrac{1}{2}z^Tz \\ &\text{s.t.} && Ax - b \leqq z. \end{aligned} \tag{1.2}$$

Sponsored by the United States Army under Contract No. DAAG29-80-C-0041.
This material is based upon work supported by the National Science Foundation under Grant No. ENG 79-03881.

But, most general quadratic programming techniques are not very suitable for the problem (1.2) because they fail to exploit its special structure. Furthermore, since the solutions to (1.2) are usually not unique, some numerical difficulties may also arise.

For the least-squares problem we propose in this paper a method that can avoid the difficulties mentioned above. The method is iterative in nature but has a finite convergence property. Its numerical results are extremely good.

We note here that the theorems and the method considered in this paper can be straightforwardly extended to the mixed-type system $Ax \leq b$, $Bx = c$. The exclusion of equalities is merely to facilitate our presentation.

In Section 2 we will study some properties of the least-squares solutions. In Section 3 we will present the method and discuss its relationship with other approaches. Finite convergence theorems will be established for the method in Section 4. Numerical results will be given in Section 5.

2. THE PROBLEM

Let A be an $m \times n$ matrix and b be an m-vector, we consider the system of linear inequalities $Ax \leqq b$. For any n-vector x there is a corresponding m-vector $Z(x)$ defined by

$$Z(x) := (Ax - b)_+ .$$

The vector $Z(x)$ may be called the residual of the vector x and is a measurement of violation of the inequalities $Ax \leqq b$ at x. A least-squares solution x^* of $Ax \leqq b$ is any vector that minimizes the quantity $\|Z(x)\|_2$, or equivalently the function $f(x)$ defined in (1.1). We note that the function f is differentiable and convex and its gradient $\nabla f(x)$ is $A^T(Ax - b)_+$. Therefore we have the following necessary and sufficient condition for a least-squares solution.

THEOREM 2.1 A n-vector x^* is a least-squares solution of the system $Ax \leqq b$ if and only if

$$A^T(Ax - b)_+ = 0. \tag{2.1}$$

Condition (2.1) may be called the normal equation of the system $Ax \leqq b$ because it is analogous to the normal equation in the equality case.

The relationship between a least-squares solution and the quadratic program (1.2) is fundamental and useful. We state it in the following theorem. The proof is trivial and hence omitted.

THEOREM 2.2 A n-vector x^* is a least-squares solution of $Ax \leqq b$ and and $z^* = Z(x^*)$ if and only if (x^*,z^*) is a solution of the quadratic program (1.2).

We next show the existance of a least-squares solution.

THEOREM 2.3 A least-squares solution of $Ax \leqq b$ exists.

PROOF: Because of Theorem 2.2 we only need to show that a solution of the quadratic program (1.2) exists. We observe that the feasible region $X = \{(x,z) \mid Ax - z \leqq b\}$ is a nonempty polyhedral set and the objective function

$\frac{1}{2}z^Tz$ is quadratic convex and bounded below on X. Therefore, it follows from Corollary 27.3.1 in [6] that a solution exists. □

It is shown by Mangasarian [5] that McCormick's second order sufficient optimality condition is also necessary for a solution of a quadratic program to be locally unique. If we apply this result to (1.2) then we have the following theorem on the uniqueness of least-squares solution to the system $Ax \leqq b$. We first note that, for a given index set $I \subset \{1,\ldots,m\}$, we use A_I to denote the submatrix of A that consists of rows in I. The subvector z_I is defined similarly.

<u>THEOREM 2.4</u> Let x^* be a least-squares solution of $Ax \leqq b$ and let $I = \{i | a_i^T x^* = b_i\}$. Then x^* is the unique least-squares solution to the system $Ax \leqq b$ if and only if there is no nonzero x satisfying $A_I x \leqq 0$.

<u>PROOF</u>: Recall that McCormick's second order sufficient optimality condition for Problem (1.2) is that $z^Tz > 0$ for any nonzero pair (x,z) satisfying

$$A_I x - z_I \leqq 0$$

$$z^{*T} z \leqq 0,$$

where $z^* = (Ax^* - b)_+$. Clearly, this condition is equivalent to the system $A_I x \leqq 0$ having no nonzero solution. Therefore, our result follows directly from Theorem 2.1 of [5]. □

We note here that the uniqueness condition in the above theorem is also equivalent to the following: the matrix A_I has full column rank and there exists a positive vector u satisfying $A_I^T u = 0$. These conditions are too restrictive and least-squares solutions of most problems in practice are usually not unique. However, for a given system $Ax \leqq b$, the optimal residual vector z^* is always unique. We give this result below.

<u>THEOREM 2.5</u> There exists a unique m-vector z^* such that x^* is a least-squares solution of $Ax \leqq b$ if and only if $(Ax^* - b)_+ = z^*$. Furthermore, $z^* \geqq 0$ and $A^T z^* = 0$.

PROOF: By Theorem 2.3 there exists a solution, (x^*, z^*) say, of the quadratic program (1.2). We first show that for any solution $(\bar{x}, \bar{z})$ of (1.2) we have that $\bar{z} = z^*$. Suppose $\bar{z} \neq z^*$. Let ρ be the optimal value of (1.2). Hence

$$\rho = \frac{1}{2} z^{*T} z^* = \frac{1}{2} \bar{z}^T \bar{z}.$$

Let $\hat{x} = \lambda x^* + (1-\lambda)\bar{x}$ and $z = z^* + (1-\lambda)\bar{z}$ for some $\lambda \in (0,1)$. Clearly, $(\hat{x}, \hat{z})$ is feasible to (1.2). Because the function $\frac{1}{2} z^T z$ is strictly convex in z, we have that

$$\frac{1}{2} \hat{z}^T \hat{z} < \frac{\lambda}{2} z^{*T} z^* + \frac{(1-\lambda)}{2} \bar{z}^T \bar{z} = \rho.$$

This contradicts that ρ is the optimal value. Hence, we have $\bar{z} = z^*$.

It immediately follows from Theorem 2.2 and the uniqueness of z^* that if x is a least-squares solution then the residual $Z(x) := (Ax - b)_+ = z^*$.

Conversely, if for some x we have that $(Ax - b)_+ = z^*$, then (x, z^*) is a solution of (1.2). It follows again from Theorem 2.2 that x is a least-squares solution. The last statement follows form the Karush-Kuhn-Tucker condition of (1.2). □

Let z^* be the vector defined in the previous theorem and let $P = \{i \mid z_i^* > 0\}$ and $L = \{i \mid z_i^* = 0\}$. Then the equation $(Ax - b)_+ = z^*$ is equivalent to

$$\left\{ \begin{array}{l} A_P x = b_P + z^*_P \\ A_L x \leqq b_L \end{array} \right\}. \qquad (2.2)$$

Hence, we have the following corollary.

COROLLARY 2.6 There exists an m-vector $z^* \geqq 0$ such that the system (2.2) is consistent and x^* is a least-squares solution of $Ax \leqq b$ if and only if x^* satisfies (2.2).

Using a similar argument as in the proof of Theorem 2.4, we can show that the residual vector of the equality linear least-squares problem is unique. Since this result will be used later, we state it below as a corollary of Theorem 2.4.

COROLLARY 2.7 Let B be an $m \times n$ matrix and c be an m-vector. There exists a unique m-vector r^* such that a vector x^* is a least-squares solution of the linear equation $Bx = c$ if and only if $Bx^* = c + r^*$. Furthermore, $B^T r^* = 0$.

3. THE METHOD

The proposed method is an iterative process. Having an estimate x of a solution, we produce a new estimate $\bar{x}$ by first finding the index set I of active and violated inequalities at x. That is $I = \{i \mid a_i^T x \geqq b_i\}$ where a_i^T is the i-th row of A and b_i is the i-th componant of b. We then generate a search direction $\bar{d}$ by

$$\bar{d} = -A_I^+(A_I x - b_I), \tag{3.1}$$

where A_I^+ is the pseudo-inverse of A_I. Recall that A_I^+ is the unique matrix satisfying the following Moore-Penrose conditions:

$$\begin{array}{llll} \text{(a)} & A_I A_I^+ A_I = A_I & \text{(b)} & (A_I A_I^+)^T = A_I A_I^+ \\ \text{(c)} & A_I^+ A_I A_I^+ = A_I^+ & \text{(d)} & (A_I^+ A_I)^T = A_I^+ A_I \end{array} \tag{3.2}$$

Once the search direction $\bar{d}$ is found, we determine a stepsize $\bar{\lambda}$ by minimizing the function $\theta(\lambda) := f(x + \lambda\bar{d})$. Here, $\bar{\lambda}$ is chosen to be the smallest minimizer of $\theta(\lambda)$. Then the new estimate $\bar{x}$ is set to be $\bar{x} = x + \bar{\lambda}\bar{d}$. We repeat this process until a point x^* satisfying $A^T(Ax^* - b)_+ = 0$ is found.

We note here that to compute $\bar{d}$ it is not necessary to construct the pseudo-inverse A_I^+ explicitly. The direction can be generated by applying a singular value decomposition to A_I. We also note that the function $\theta(\lambda)$ is of the form $\theta(\lambda) = \sum_{i=1}^{m} (\alpha_i\lambda - \gamma_i)_+^2$. Therefore, if $\alpha_i \leqq 0$ for all i then $\theta(\lambda)$ becomes a constant when λ is larger than a certain number; and if $\alpha_i > 0$ for some i then $\theta(\lambda) \to \infty$ when $\lambda \to \infty$. Hence, the stepsize $\bar{\lambda}$ can not be infinite. Because θ is convex and piesewise quadratic, the stepsize $\bar{\lambda}$ can be very accurately and effeciently computed.

Recall that the vector $\bar{d}$ defined in (3.1) is the minimum-norm least-squares solution of the linear equation $A_I d = b_I - A_I x$. To gain more insight into the method, we give the following theorem.

THEOREM 3.2 Let $\hat{x} = x + \bar{d}$, then $\hat{x}$ is the closest least-squares solution of the system $A_I y = b_I$ to the point x in 2-norm.

PROOF. We first show that $\hat{x}$ is a least-squares solution of $A_I y = b_I$. It follows from (3.2) that

$$A_I^T A_I A_I^+ = A_I^T (A_I A_I^+)^T = (A_I A_I^+ A_I)^T = A_I^T .$$

This implies that

$$\begin{aligned} A_I^T A_I \hat{x} &= A_I^T A_I (x + \bar{d}) \\ &= A_I^T A_I x - A_I^T A_I A_I^+ (A_I x - b_I) \\ &= A_I^T A_I x - A_I^T (A_I x - b_I) \\ &= A_I^T b_I . \end{aligned}$$

Hence, $\hat{x}$ is a least-squares solution of $A_I y = b_I$.

Let $\bar{y}$ be any least-squares solution of $A_I y = b_I$ and let $d = \bar{y} - x$. We want to show that $\|\bar{d}\|_2 \leqq \|d\|_2$. Because $\bar{d}$ is the minimum norm least-squares solution of $A_I y = b - A_I x$, we only need to show that d is a least-squares solution of $A_I y = b - A_I x$. This follows from

$$\begin{aligned} A_I^T A_I d &= A_I^T A_I \bar{y} - A_I^T A_I x \\ &= A_I^T b_I - A_I^T A_I x \\ &= A_I^T (b_I - A_I x) . \end{aligned}$$

□

The above theorem shows that $\hat{x}$ is the projection of x onto the set of all least-squares solution of $A_I y = b_I$. From this observation the vector $\bar{d}$ is a very natural choice for a search direction.

The method is also related to Newton's method for minimizing the function f. If the set $\{i \mid a_i^T x = b_i\}$ is empty then the function f is twice differentiable at x and its Hessian is $\nabla^2 f(x) = A_I^T A_I$. If we further assume that A_I has full column rank, then $\nabla^2 f(x)$ is nonsingular and the direction generated by Newton's method is given by

$$d = -(A_I^T A_I)^{-1}\nabla f(x) = -(A_I^T A_I)^{-1} A^T (Ax - b)_+$$

$$= -(A_I^T A_I)^{-1} A_I^T (A_I x - b_I).$$

Because, in this case we have $A_I^+ = (A_I^T A_I)^{-1} A_I^T$, our method and Newton's are identical.

4. FINITE CONVERGENCE

In this section we will show that for any given system of linear inequalities $Ax \leqq b$, our method will produce a least-squares solution in a finite number of iterations.

We first show that the direction $\bar{d}$ generated in (3.1) is a descent direction for the function f at x.

LEMMA 4.1 $\nabla f(x) = -A_I^T A_I \bar{d}$.

PROOF It follows from Theorem 3.1 that the vector $\hat{x} = x + \bar{d}$ satisfies the normal equation $A_I^T A_I \hat{x} = A_I^T b_I$. Thus

$$\begin{aligned}
\nabla f(x) = A^T(Ax - b)_+ &= A_I^T(A_I x - b_I) \\
&= A_I^T A_I(\hat{x} - \bar{d}) - A_I^T b_I \\
&= -A_I^T A_I \bar{d} + A_I^T A_I \hat{x} - A_I^T b_I \\
&= -A_I^T A_I \bar{d}. \qquad \square
\end{aligned}$$

From Lemma 4.1 we immediately have the following theorem.

THEOREM 4.2 $\nabla f(x)^T \bar{d} = -\|A_I \bar{d}\|_2^2$.

We then show that for any sequence $\{x^k\}$ generated by the method we have $\lim_{k\to\infty} \nabla f(x^k) = 0$.

LEMMA 4.3 For any u, v in R^n

$$\|\nabla f(u) - \nabla f(v)\|_2 \leqq \|A\|_2^2 \|u - v\|_2 .$$

PROOF

$$\begin{aligned}
\|\nabla f(u) - \nabla f(v)\|_2 &= \|A^T(Au - b)_+ - A^T(Av - b)_+\|_2 \\
&\leqq \|A^T\|_2 \|(Au - b)_+ - (Au - b)_+\|_2 \\
&\leqq \|A^T\|_2 \|Au - Av\|_2 \\
&\leqq \|A\|_2^2 \|u - v\|_2 \qquad \square
\end{aligned}$$

THEOREM 4.4 Let $\{x^k\}$ be generated by the method from any starting point x^0 Then either for some $\bar{k} < \infty$, $\nabla f(x^{\bar{k}}) = 0$ or $\lim_{k\to\infty} \nabla f(x^{\bar{k}}) = 0$.

PROOF: We note that if $\bar{d} = 0$ then by Lemma 4.1 we have $\nabla f(x) = 0$ and x is a solution. Hence, we assume that x is not a solution and $\bar{d} \neq 0$. This also implies that $I = \{i | a_i^T x \geqq \beta_i\} \neq \emptyset$.

Let
$$c_1 = \| A \|_2^2 ,$$
$$c_2 = \max_{I \neq \emptyset} \| A_I^+ \|_2$$
$$\hat{\lambda} = \frac{-\nabla f(x)^T \bar{d}}{c_1 \| \bar{d} \|_2^2} .$$

Then we have that

$$\begin{aligned} f(x + \hat{\lambda}\bar{d}) - f(x) &= \hat{\lambda}\int_0^1 \nabla f(x + t\hat{\lambda}\bar{d})^T \bar{d}\, dt \\ &= \hat{\lambda}[\nabla f(x)^T \bar{d} + \int_0^1 ((\nabla f(x + t\hat{\lambda}\bar{d}) - \nabla f(x))^T \bar{d}) dt] \\ &\leqq \hat{\lambda}[\nabla f(x)^T \bar{d} + c_1 \hat{\lambda} \| \bar{d} \|_2^2 \int_0^1 t\, dt] \\ &= \frac{\hat{\lambda}}{2} \nabla f(x)^T \bar{d} \\ &= \frac{-(f(x)^T \bar{d})^2}{2c_1 \| \bar{d} \|_2^2} . \end{aligned}$$

We also note that it follows from (3.2.c) that

$$\begin{aligned} \| \bar{d} \|_2 = \| A_I^+ (A_I x - b_I) \|_2 &= \| A_I^+ A_I \bar{d} \|_2 \\ &\leqq \| A_I^+ \|_2 \| A_I \bar{d} \|_2 \\ &\leqq c_2 \| A_I \bar{d} \|_2 . \end{aligned}$$

On the other hand, it follows from Theorem 4.2 that $\nabla f(x)^T \bar{d} = -\| A_I \bar{d} \|_2^2$.

Therefore, we have that

$$f(x) - f(x + \hat{\lambda}\bar{d}) \geqq \frac{1}{2c_1 c_2^2} \| A_I \bar{d} \|_2^2 .$$

Because of the choice of the stepsize $\bar{\lambda}$ and $\bar{x} = x + \bar{\lambda}\bar{d}$, we have that

$$f(x) - f(\bar{x}) \geqq f(x) - f(x + \hat{\lambda}\bar{d}) \geqq \frac{1}{2c_1c_2^2} \| A_I\bar{d}\|_2^2 .$$

Since the sequence $\{f(x^k)\}$ is monotone decreasing and bounded below, we have

$$\infty > \sum_{k=0}^{\infty} (f(x^k) - f(x^{k+1})) \geqq \frac{1}{2c_1c_2^2} \sum_{k=0}^{\infty} \| A_{I_k}\bar{d}^k\|_2 .$$

This implies that

$$\lim_{k\to\infty} A_{I_k}\bar{d}^k = 0$$

Then it follows from $\nabla f(x^k) = -A_{I_k}^T A_{I_k}\bar{d}^k$ that

$$\lim_{k\to\infty} \nabla f(x^k) = 0 \qquad \square$$

By Theorem 2.4 there exists a unique vector z^* such that x is a least-squares solution of $Ax \leqq b$ is and only if $(Ax - b)_+ = z^*$. If the sequence $\{x^k\}$ is generated by the method and $z^k = (Ax^k - b)_+$, we want to show that $\lim_{k\to\infty} z^k = z^*$. To do this, we need the following lemma. The proof of the lemma can be found in [2].

<u>LEMMA 4.5</u> Let B and E be any $m \times n$ and $\ell \times n$ matrices and let $\{g^k\}$ and $\{h^k\}$ be sequences of vectors in R^m and in R^ℓ respectively. If $g^k \to \bar{g}$ and $h^k \to \bar{h}$ and for each k the system $Bx = g^k$, $Ex \leqq h^k$ is consistent then the system $Bx = \bar{g}$, $Ex \leqq \bar{h}$ is also consistent.

<u>THEOREM 4.6</u> Let $\{x^k\}$ be generated by the method and $z^k = (Ax^k - b)_+$, then

$$\lim_{k\to\infty} z^k = z^* ,$$

where z^* is the optimal residual vector defined in Theorem 2.4.

<u>PROOF</u>: We have that $f(x^k) = \frac{1}{2}(z^k)^T z^k$ and $\{f(x^k)\}$ is a monotone decreasing sequence and bounded below. Hence, there exists a $\rho > 0$ such that for each k

$$z^k \in U = \{z \mid \|z\|_2 < \rho\}$$

Let $\{z^{k_j}\}$ be any convergent subsequence of $\{z^k\}$ and let $z^{k_i} \to \hat{z}$. We want to show $z = z^*$. Let $P = \{i \mid \hat{z}_i > 0\}$ and $L = \{i \mid \hat{z}_i = 0\}$. Because $z^{k_j} = (Ax^{k_j} - b)_+$, we have that for all sufficiently large j, the system

$$\begin{cases} A_P x = b_P + z_P^{k_j} \\ A_L x \leqq b_L + z_L^{k_j} \end{cases}$$

has x^{k_j} as a solution and hence is consistent. Then it follows from Lemma 4.5 that the system

$$\begin{cases} A_P x = b_P + \hat{z}_P \\ A_L x \leqq b_L \end{cases}$$

is also consistent and has a solution, $\hat{x}$ say. We now show that $\hat{x}$ is a least-squares solution of $Ax \leqq b$. Notice that $\hat{z} = (A\hat{x} - b)_+$. It follows from Theorem 4.4 that $A^T z^k = \nabla f(x^k) \to 0$, which, in conjunction with $z^{k_j} \to \hat{z}$, implies that

$$\nabla f(\hat{x}) = A^T(A\hat{x} - b)_+ = A^T\hat{z} = 0.$$

Hence, $\hat{x}$ is a least-squares solution. By Theorem 2.4 the optimal residual vector z^* is unique, hence $\hat{z} = z^*$.

We now show that $z^k \to z^*$. If it is not the case, then there exists an $\varepsilon > 0$ and a subsequence $\{z^{k_j}\}$ such that $\| z^{k_j} - z^* \| \geqq \varepsilon$ for each j. Because $\{ z^{k_j}\}$ is in the compact set U, it has a convergent subsequence. By the first part of the proof this subsequence converges to z^*, which contradicts $\| z^{k_j} - z^* \| \geqq \varepsilon$ for each j. Hence, we have $z^k \to z^*$. □

In the sequel we use $I(x)$ to denote the index set $\{i \mid a_i^T x \geqq b_i\}$ at x. We will show that if the sequence $\{x^k\}$ generated by the method does not terminate after a finite number of steps, then for all sufficiently large k we have that $I(x^k) \subsetneqq I(x^{k+1})$. This is a contradiction because we have only finite index sets. To establish these results we need some lemmas.

LEMMA 4.7 There exists an $\varepsilon > 0$ such that if I is an index set and $I = I(x)$ for some x satisfying $\|(Ax - b)_+ - z^*\| \leqq \varepsilon$ then the system $A_I y = b_I + z_I^*$ is consistent.

PROOF. If it is not true, then there exists a sequence $\{x^k\}$ satisfying $(Ax^k - b)_+ \to z^*$ and for each k, the system $A_{I_k} y = b_{I_k} + z^*_{I_k}$ is inconsistent, where $I_k = I(x^k)$. Because there are only finite index sets, we may assume that I_k is fixed and for some I, $I_k = I$ for each k.

Let $z^k = (Ax^k - b)_+$, then the system $A_I y = b_I + z_I^k$ has x^k as a solution and hence is consistent. It follows from Lemma 4.5 and $z^k \to z^*$ that $A_I y = b_I + z_I^*$ is consistent, which is a contradiction. Hence the lemma is proved. $\square$

LEMMA 4.8 There exists and $\varepsilon > 0$ such that if x is a point satisfying $\|(Ax - b)_+ - z^*\|_2 \leqq \varepsilon$, $\bar{d}$ is the direction generated in (3.1) and $\hat{x} = x + \bar{d}$ then $I(\hat{x}) \supset I(x)$.

PROOF. Choose a positive number ε sufficiently small so that Lemma 4.7 holds and $\|(Ax - b)_+ - z^*\|_2 \leqq \varepsilon$ implies that $I(x) \supset P := \{i \mid z_i^* > 0\}$. Let x be any point satisfying $\|(Ax - b)_+ - z^*\|_2 \leqq \varepsilon$ and let $I = I(x)$. Then by Lemma 4.7 we have that the system $A_I y = b_I + z_I^*$ has a solution, $\bar{y}$ say. It follows from $I \supset P$ and $A^T z^* = A_I^T z_I^* = 0$ that $A_I^T A_I \bar{y} = A_I^T b_I + A_I^T z^*_I = A_I^T b_I$. Hence, $\bar{y}$ is a least-squares solution of $A_I y = b_I$ and z_I^* is its residual vector. Notice that $\hat{x} = x + \bar{d}$ is also a least-squares solution of $A_I y = b_I$. It follows from Corollary 2.6 that the residual vector is unique. Hence, z_I^* is also the residual vector of $\hat{x}$ and we have $A_I \hat{x} = b_I + z_I^*$. Because $z_I^* \geqq 0$, we have $I(\hat{x}) \supset I = I(x)$. This completes our proof. $\square$

LEMMA 4.9 Let x not be a least-squares solution of $Ay \leqq b$ and let $\hat{x} = x + \bar{d}$ and $\bar{x} = x + \bar{\lambda}\bar{d}$. Then

(a) $I(\hat{x}) = I(x) \Rightarrow \bar{x}$ is a least-squares solution of $Ay \leqq b$

(b) $I(\hat{x}) \supsetneq I(x) \Rightarrow I(\bar{x}) \supsetneq I(x)$.

<u>PROOF</u>. (a) If $I(\hat{x}) = I(x) = I$, then

$$A^T(A\hat{x} - b)_+ = A_I^T(A_I\hat{x} - b_I) = 0.$$

Hence, $\hat{x}$ is a least-squares solution of $Ay \leqq b$. Because $f(\bar{x}) \leqq f(\hat{x})$, the point $\bar{x}$ must also be a least-squares solution of $Ay \leqq b$.

(b) We first show that the stepsize $\bar{\lambda} \in (0,1]$. It follows from Theorem 4.2 that $\theta'(0) < 0$. Therefore, because θ is convex, we only need to show $\theta'(1) \geqq 0$.

Let $I(\hat{x}) = I \cup J$, where $J \cap I = \emptyset$ and $J \neq \emptyset$. Then, we have that

$$\begin{aligned}\theta'(1) = \bar{d}^T\nabla f(\hat{x}) &= \bar{d}^T A^T(A\hat{x} - b)_+ \\ &= \bar{d}^T A_I^T(A_I\hat{x} - b_I) + \bar{d}^T A_J^T(A_J\hat{x} - b_J). \\ &= \bar{d}^T A_J^T(A_J\hat{x} - b_J).\end{aligned}$$

Since $A_J\hat{x} \geqq b_J$ and $A_Jx < b_J$, we have

$$A_J\bar{d} = A_J\hat{x} - A_Jx > b_J - b_J = 0.$$

Therefore, it follows from $A_J\hat{x} - b_J \geqq 0$ and $A_J\bar{d} > 0$ that $\theta'(1) \geqq 0$.

Hence, we have $\bar{\lambda} \in (0,1]$.

We next show that $I(\hat{x}) \supsetneq I(x)$ implies $I(\bar{x}) \supset I(x)$. We have that

$$\begin{aligned}A_I\bar{x} &= (1 - \bar{\lambda})A_Ix + \bar{\lambda}A_I\hat{x} \\ &\geqq (1 - \bar{\lambda})b_I + \bar{\lambda}b_I = b_I.\end{aligned}$$

Hence, $I(\bar{x}) \supset I(x)$.

We then show that $I(\bar{x}) \neq I(x)$. Suppose $I(\bar{x}) = I(x) = I$. Then, we have

$$\begin{aligned}\theta'(\bar{\lambda}) &= \bar{d}^T A^T (A\bar{x} - b)_+ \\ &= \bar{d}^T A_I^T (A_I(x + \bar{\lambda}\bar{d}) - b_I) \\ &= \bar{\lambda}\bar{d}^T A_I^T A_I \bar{d} + \bar{d}^T A_I^T (A_I x - b_I) \\ &= (\bar{\lambda} - 1)\bar{d}^T A_I^T A_I \bar{d}.\end{aligned}$$

Because $\theta'(\bar{\lambda}) = 0$ and $A_I\bar{d} \neq 0$, we have that $\bar{\lambda} = 1$, which implies $\bar{x} = \hat{x}$. However, it is impossible because, by our assumption, $I(\hat{x}) \neq I$ but $I(\bar{x}) = I$. This completes our proof. □

Our convergence theorem is given below.

<u>THEOREM 4.10</u> For any $m \times n$ matrix A and any m-vector b, the method produces from any starting point x^0 a least-squares solution of the system $Ax \leqq b$ in a finite number of iterations.

<u>PROOF</u>: Let $\{x^k\}$ be the sequence of points generated by the method from a given starting point x^0. If the sequence $\{x^k\}$ does not terminate at a least-squares solution after a finite steps, then by Theorem 4.6 we have that $\lim_{k\to\infty} z^k = z^*$. Let $\varepsilon > 0$ be defined as in Lemma 4.8 and let $\bar{k}$ be chosen so that for any $k \geqq \bar{k}$, we have $\|z^k - z^*\| \leqq \varepsilon$. Then it follows from Lemma 4.8 and Lemma 4.9 that $I(x^k) \subsetneq I(x^{k+1})$ for all $k \geqq \bar{k}$. This is certainly impossible, hence $\{x^k\}$ must terminate in a finite number of iterations. □

5. REMARKS AND NUMERICAL RESULTS

The linear least-squares problem of equalities has been widely studied and efficient computational methods exist [see, for example, 1,3 and 7]. However, with very wide range of applications, the linear least-squares problem of inequalities has not been throughly investigated as it should. This paper is an attempt to provide for this problem a useful computational method.

A very interesting application of the linear least-squares problem is to solve linear programming problems. Consider the following linear program

$$\begin{aligned} &\min_y \ c^T y \\ &\text{s.t. } By \geqq d \\ &\qquad\ y \geqq 0 . \end{aligned}$$

Its dual program is

$$\begin{aligned} &\max_u \ d^T u \\ &\text{s.t. } B^T u \leqq c \\ &\qquad\quad u \leqq 0. \end{aligned}$$

By a duality theorem, the linear program and its dual can be solved by finding a pair $(\bar{y},\bar{u})$ that satisfies the following inequalities:

$$Ax \leqq b, \tag{5.1}$$

where

$$A = \begin{bmatrix} c^T & d^T \\ -B & 0 \\ 0 & B^T \\ -I & 0 \\ 0 & -I \end{bmatrix} \qquad b = \begin{bmatrix} 0 \\ -d \\ c \\ 0 \\ 0 \end{bmatrix} \qquad \text{and} \qquad x = \begin{bmatrix} y \\ u \end{bmatrix}.$$

Then our method can apply to the system (5.1). When the linear program has no solution, a least-squares solution to (5.1) can be found. It is also interesting to note that, by doing so, we can transform a linear programming problem into unconstrained minimalization of the convex and differentiable function f defined in (1.1). A similar function was studied by Mangasarian [4]. His function is designed to find a least norm solution for a linear program. In his case the least-squares measurement is taken in the domain space rather than the range space.

The method has been tested for many randomly generated problems. The search direction d is computed by using a singular value decomposition subroutine in LINPACK. The method is very satisfactory in accuracy and number of iterations. For each test problem a point with $\|A^T(Ax - b)_+\|_2 \leq 10^{-25}$ is found. The number of iterations is usually much less than the size of the problem.

In the following table, m is the number of inequalities and n is the number of variables. The numbers in the table indicate the numbers of iterations. The computation is done in Cyber 175 System at the University of Illinois at Urbana.

Table

m \ n	10	10	20	20	30	30	40	40	50	50
10	5	2	2	2	3	2	2	2	2	3
20	4	9	3	3	4	3	3	4	3	2
30	4	4	8	11	2	4	5	6	4	3
40	5	5	10	10	9	8	4	6	3	5

ACKNOWLEDGEMENT:

The author should like to thank Professor O.L. Mangasarian for his valuable comments on a draft of this paper.

REFERENCES

[1] Golub, G.H., *Numerical methods for solving linear least-squares problems*, Numer. Math. 7 (1965) pp. 206-216.

[2] Han, S.P., Solving quadratic programs by an exact penalty function, MRC Technical Report, to appear.

[3] Lawson, C.L. and Hanson, R.J., Solving Least Squares Problems, Prentice-Hall, New Jersey, 1974.

[4] Mangasarian, O.L., Least norm linear programming solution as an unconstrained minimization problem, MRC Technical Report, to appear.

[5] Mangasarian, O.L., *Locally unique solution of quadratic programs, linear and nonlinear complementarity problems*, Mathematical Programming 19 (1980) pp. 200-212.

[6] Rockafellar, R.T., Convex Analysis, Princeton University Press, 1970.

[7] Stewart, G.W., Introduction to Matrix Computations, Academic Press, New York, 1973.

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered) (14) MRC-TSR-2141

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER TSR# 2141	2. GOVT ACCESSION NO. AD-A096658	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Least-Squares Solution of Linear Inequalities.		5. TYPE OF REPORT & PERIOD COVERED Summary Report - no specific reporting period
		6. PERFORMING ORG. REPORT NUMBER
7. AUTHOR(s) (10) S.-P. Han		8. CONTRACT OR GRANT NUMBER(s) (15) NSF Grant ENG 79-03881 DAAG29-80-C-0041 NSF-ENG79-03881
9. PERFORMING ORGANIZATION NAME AND ADDRESS Mathematics Research Center, University of Wisconsin 610 Walnut Street Madison, Wisconsin 53706		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS 5 (Operations Research)
11. CONTROLLING OFFICE NAME AND ADDRESS See Item 18 below (12) 23		12. REPORT DATE (11) November 1980
		13. NUMBER OF PAGES 19
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) (9) Technical summary rept.		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES U.S. Army Research Office and National Science Foundation P.O. Box 12211 Washington, D.C. 20550 Research Triangle Park North Carolina 27709		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Linear inequalities, Least-squares solution		

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The paper deals with a system of linear inequalities in a finite dimensional space. When the system is inconsistent, we are interested in vectors that satisfy the system in a least-squares sense. We characterize such least-squares solutions and propose a method to find one of these solutions. It is shown that for any given system of linear inequalities and for any starting point, the method can produce a solution in a finite number of iterations. Computational results are very satisfactory in terms of accuracy and number of iterations.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED 221200

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ATE

LMED

-8