JOURNAL OF APPROXIMATION THEORY 24, 1–17 (1978)

Algorithms for Solving Overdetermined Systems of Linear Equations in the l_p-Metric, $0 < p < 1$

R. W. OWENS*

Department of Mathematics, Lewis and Clark College, Portland, Oregon 97219

AND

V. P. SREEDHARAN

Department of Mathematics, Michigan State University, East Lansing, Michigan 48823

Communicated by Oved Shisha

Received January 10, 1977

We investigate overdetermined systems of m linear equations in d unknowns. We equip $\mathbb{R}^m$ with the p-homogeneous metric $\| x \|_p = \sum_{j=1}^{m} | x_j |^p$, $0 < p < 1$, and seek approximate solutions of the linear system which minimize the error vector in this metric. After showing that the number of points at which a solution of this problem can occur is finite, we present several algorithms for solving the given approximation problem globally and locally. The algorithms apply to the interesting l_1-case as well.

1. INTRODUCTION

In this paper, we consider the system of linear equations

$$Ax = b,$$

where A is an $m \times d$ real matrix, $m > d$, $x \in \mathbb{R}^d$, and $b \in \mathbb{R}^m$. For $y \in \mathbb{R}^m$ and $0 < p < 1$, let

$$\| y \|_p = \sum_{j=1}^{m} | y_j |^p. \tag{1.1}$$

Given A, b and p, the problem that we study here, referred to as problem (P), is

(P) Find $x \in \mathbb{R}^d$ minimizing $\{\| b - Ax \|_p \mid x \in \mathbb{R}^d\}$.

* Portions of this paper are taken from this author's Ph.D. thesis at Michigan State University.

0021-9045/78/0241-0001$02.00/0

To solve the problem, a dual problem, to be denoted (P*), is introduced, a relation between (P) and (P*) proved, and a characterization of the solutions of problem (P*) established. Although, in general, problem (P*) cannot be solved in a computationally feasible manner, it can always be solved in a finite number of steps. Moreover, efficient algorithms for solving particular cases and exchange algorithms for finding local solutions of problem (P*) are outlined.

2. Basic Notions

It should be noted that $\|\cdot\|_p$, $0 < p < 1$, is a p-homogeneous metric but not a norm on $\mathbb{R}^m$, although the triangle inequality

$$\|y + z\|_p \leqslant \|y\|_p + \|z\|_p, \qquad y, z \in \mathbb{R}^m, \tag{2.1}$$

holds. $\|\cdot\|_p$ is p-homogeneous in the sense that

$$\|\alpha y\|_p = |\alpha|^p \|y\|_p \qquad \alpha \in \mathbb{R}, \quad y \in \mathbb{R}^m. \tag{2.2}$$

Also,

$$\|y\|_p = 0 \qquad \text{iff} \quad y = 0. \tag{2.3}$$

The usual l_p norm, i.e., the pth root of (1.1), fails to satisfy the triangle inequality if one chooses $0 < p < 1$. However, we still refer to $\|\cdot\|_p$ as the l_p norm and problem (P) as the l_p problem.

Let $(\cdot \mid \cdot)$ denote the usual inner product on $\mathbb{R}^m$, i.e.,

$$(x \mid y) = \sum_{j=1}^{m} x_j y_j, \qquad x, y \in \mathbb{R}^m.$$

A^T is the transpose of the matrix A. Set

$$K = \text{Image}(A) = \{Ax \mid x \in \mathbb{R}^d\},$$

and

$$K^\perp = \text{Ker } A^T = \{x \in \mathbb{R}^m \mid (x \mid k) = 0, \forall k \in K\}.$$

Throughout this paper, we assume that the dimension of K is n and that $b_1, \dots, b_n$ form a basis for K. Let E: $\mathbb{R}^m \to K^\perp$ be the orthogonal projection of $\mathbb{R}^m$ onto $K^\perp$, where orthogonality is with respect to the inner product $(\cdot \mid \cdot)$, and set $s = Eb$. Let

$$\rho = d(b, K) = \inf\{\|b - k\|_p \mid k \in K\}.$$

We assume $b \notin K$, or equivalently $\rho > 0$, since the problem is trivial otherwise. Let (s) denote the linear span of the vector s and let

$$D = (b_1, \ldots, b_n, s) = (r_1{}^T, \ldots, r_m{}^T)^T,$$

where r_i is the ith row of the $m \times (n+1)$ matrix D. Finally, let $B = \{w \in K \oplus (s) \mid \| w \|_p \leqslant 1\}$.

Observe that $s - b \in K$, and consequently $d(s, K) = d(b, K)$. The existence of a solution of problem (P) follows from the continuity of the l_p norm and the assumption that $\dim K = n < \infty$.

3. Problem (P*)

Given problem (P), we associate a dual problem

(P*) Find $z \in K \oplus (s)$, $\| z \|_p \leqslant 1$, maximizing $(s \mid w)$ over all $w \in K \oplus (s)$, $\| w \|_p \leqslant 1$.

Problem (P*), when $\mathbb{R}^m$ is equipped with a norm, was considered by Sreedharan in [4]. The relation between problems (P) and (P*) is given in the following theorem which extends Theorem 2.4 of [4].

THEOREM 3.1. *Let z solve problem* (P*). *Then*

(i) $\rho^{1/p}(s \mid z) = (s \mid s)$, (3.1.1)

and

(ii) $b - \rho^{1/p} z \in K$. (3.1.2)

Proof.

$$\begin{aligned}
(s \mid z) &= \max\{(s \mid w) \mid w \in K \oplus (s), \| w \|_p \leqslant 1\}, \\
&= \max\{(s \mid k + \beta s) \mid k \in K, \beta \in \mathbb{R}, \| k + \beta s \|_p \leqslant 1\}, \\
&= \max\{(s \mid s)\, \beta \mid \beta \in \mathbb{R}, k \in K, \| k + \beta s \|_p \leqslant 1\}, \\
&= (s \mid s) \max\{\beta \in \mathbb{R}\backslash\{0\} \mid k \in K, \| k + s \|_p \leqslant 1/| \beta |^p\}, \\
&= (s \mid s) \max \left\{ \frac{1}{\| k + s \|_p^{1/p}} \,\middle|\, k \in K \right\}, \\
&= (s \mid s) \frac{1}{\min\{\| k + s \|_p^{1/p} \mid k \in K\}}, \\
&= (s \mid s)/\rho^{1/p}.
\end{aligned}$$

Thus, $\rho^{1/p}(s \mid z) = (s \mid s)$ which is (3.1.1).

Let $t \in K^{\perp}$. We show that $(t \mid b - \rho^{1/p}z) = 0$. Write $t = u + \alpha s$, where $(u \mid s) = 0$ and $\alpha \in \mathbb{R}$. Since $z \in K \oplus (s)$ and $u \in K^{\perp} \cap (s)^{\perp}$, $(u \mid z) = 0$. Also, $(u \mid b) = (Eu \mid b) = (u \mid Eb) = (u \mid s) = 0$. Thus, $(u \mid b - \rho^{1/p}z) = 0$. Next, since E is the orthogonal projection of $\mathbb{R}^m$ onto $K^{\perp}$, and $s = Eb$, we have

$$(s \mid s) = (s \mid b), \tag{3.1.3}$$

$$\begin{aligned} (s \mid b - \rho^{1/p}z) &= (s \mid b) - \rho^{1/p}(s \mid z), \\ &= (s \mid s) - (s \mid s), \end{aligned}$$

by (3.1.3) and (3.1.1). Thus, $(t \mid b - \rho^{1/p}z) = 0$. Since $t \in K^{\perp}$ was arbitrary, $b - \rho^{1/p}z \in K$. This concludes the proof of Theorem 3.1.

Remark 3.2. It is interesting to note that the above theorem and its proof did not use the explicit expression (1.1) for $\| \cdot \|_p$. We used only the facts that $\| \cdot \|_p$ satisfies (2.1), (2.3) and that (2.2) holds with $p > 0$.

4. Definitions and Lemmas

We now develop several ideas that will be needed in the next section to show that problem (P*) can always be solved in a finite number of steps.

Let X be a real linear metric space, i.e., a real vector space on which a translation invariant metric is defined so that the metric space structure is compatible with the linear space structure. Denote by X^* the algebraic dual of X.

Definition 4.1. Let X be a real linear metric space, $A \subset X$, $a \in A$, and H a nontrivial hyperplane in X, i.e., $H = \{x \in X \mid f(x) = 0\}$, where $f \in X^* \backslash \{0\}$. We say that $H + a = \{h + a \mid h \in H\}$ supports A at a if either

$$f(x) \geqslant f(a), \ \forall x \in A \qquad \text{or} \qquad f(x) \leqslant f(a), \ \forall x \in A. \tag{4.1.1}$$

Lemma 4.2. *Let X be a real linear metric space with* $\dim X \geqslant 1$, $f \in X^* \backslash \{0\}$, $A = f^{-1}(0)$, $z \in X \backslash A$, *Z a subspace of X with $z \in Z$, and $A_1 = A \cap Z$. Then A_1 is a hyperplane in Z.*

Proof. Since $f(Z) \neq 0$, $A_1 = A \cap Z = \{x \in Z \mid f(x) = 0\}$, the kernel of a nonzero linear functional, is a hyperplane in Z.

Lemma 4.3. *Let Y be a subspace of $\mathbb{R}^m$, f a linear functional on Y with $f \not\equiv 0$ on Y, $H = \{x \in Y \mid f(x) = 0\}$, and $B_Y = \{x \in Y \mid \| x \|_p \leqslant 1\}$. Let $z \in Y$ satisfy* (i) $\| z \|_p = 1$, (ii) $z_i \neq 0$, $i = 1, \ldots, m$, (iii) *$H + z$ supports B_Y at z. Then* $\dim Y = 1$.

Proof. $\dim Y \geqslant 1$, since $z \in Y\backslash\{0\}$. Suppose $\dim Y > 1$. Since $\dim H = \dim Y - 1 > 0$, there exist x, $-x \in B_Y$ with $f(-x) < 0 < f(x)$, and since $H + z$ supports B_Y at z by hypothesis, $f(z) \neq 0$. Choose $x \in H\backslash\{0\}$ and define

$$\begin{aligned} \delta_i &= 1 \qquad \text{if} \quad x_i = 0, \\ &= | z_i/x_i | \qquad \text{if} \quad x_i \neq 0, \end{aligned} \qquad i = 1,\ldots, m, \tag{4.3.1}$$

$$\delta = \tfrac{1}{2} \min\{\delta_i \mid i = 1,\ldots, m\}. \tag{4.3.2}$$

Then $| z_i + \epsilon x_i | > 0$ for $i = 1,\ldots, m$ and $| \epsilon | < \delta$. Let

$$g(\epsilon) = \| z + \epsilon x \|_p , \qquad -\delta < \epsilon < \delta. \tag{4.3.3}$$

Now

$$\frac{d^2 g}{d\epsilon^2} = p(p-1) \sum_{j=1}^{m} x_j^2 \, | z_j + \epsilon x_j |^{p-2} < 0, \tag{4.3.4}$$

since $x \neq 0$ and $0 < p < 1$. Thus g does not have a local minimum for $\epsilon = 0$. Hence there exists γ, $0 < | \gamma | < \delta$, such that

$$\| z + \gamma x \|_p < 1. \tag{4.3.5}$$

Now there exists $\epsilon > 0$ such that

$$\| \alpha z + \gamma x \| \leqslant 1 \qquad \text{for all} \quad \alpha \in (1 - \epsilon, 1 + \epsilon). \tag{4.3.6}$$

Let $u = (1 - \epsilon/2) z + \gamma x$ and $v = (1 + \epsilon/2) z + \gamma x$. By (4.3.6), $u, v \in B_Y$, and

$$f(u) = (1 - \epsilon/2) f(z) \qquad \text{and} \qquad f(v) = (1 + \epsilon/2) f(z). \tag{4.3.7}$$

Thus, either $f(u) < f(z) < f(v)$ or $f(v) < f(z) < f(u)$. In either case, the hypothesis that $H + z$ supports B_Y at z is contradicted. Hence $\dim Y = 1$.
Q.E.D.

Before proceeding, we define three symbols which will be used extensively in the remainder of this paper.

DEFINITION 4.4. Let $z \in B$ with $\| z \|_p = 1$. Define

$$J(z) = \{j \mid z_j = 0\}, \tag{4.4.1}$$

and let

$$C(z) = (r_j), \; j \in J(z), \tag{4.4.2}$$

be a matrix with rows r_j, where r_j was defined in Section 2. When $J(z)$ is empty, $C(z)$ is taken to be the zero row vector of $n + 1$ entries. To make

$C(z)$ unique, specify that if $i, j \in J(z)$ and $i < j$, then r_i appears above r_j in $C(z)$. Finally, we denote by $N(z)$ the kernel of $C(z)$, i.e.,

$$N(z) = \{x \in \mathbb{R}^{n-1} \mid C(z)\, x = 0\}. \tag{4.4.3}$$

THEOREM 4.5. *Let $z \in B$ with $\| z \|_p = 1$, f a nontrivial linear functional on $K \oplus (s)$, $H = f^{-1}(0)$, and let $H + z$ support B at z. Then* $\dim N(z) = 1$.

Proof. There exists $\beta \in \mathbb{R}^{n-1}$ such that $D\beta = z$, where D is defined in Section 2. Since $z \neq 0$, $\beta \neq 0$ and by the definition of $C(z)$, $\beta \in N(z)$. This shows that $\dim(N(z)) \geqslant 1$.

Without loss of generality, let $I = \{1, ..., \mu\}$ and $J(z) = \{\mu + 1, ..., m\}$. Set $K^* = \{Dx \mid x \in N(z)\}$, denote by f^* the restriction of f to K^*, and let $H^* = H \cap K^*$ and $B^* = B \cap K^*$. Clearly, $z \in B^*$. Since the rank of D is $n + 1$, $Dx = 0$ implies that $x = 0$ by the Rank–Nullity Theorem of linear algebra, and hence $\dim K^* = \dim N(z)$.

Since $\dim K^* \geqslant 1$, by Lemma 4.2, H^* is a hyperplane in K^*. Also, $H^* + z$ supports B^* at z since $f^* = f \mid K^*$. Finally, each $x \in K^*$ satisfies $x_{\mu+1} = \cdots = x_m = 0$. This suggests dropping the $m - \mu$ trailing zeroes and considering the problem in $\mathbb{R}^\mu$. We make this more precise by setting

$$\tilde{K} = \{(x_1, ..., x_\mu) \in \mathbb{R}^\mu \mid (x_1, ..., x_\mu, 0, ..., 0) \in K^*\},$$

$$\tilde{f}\colon \tilde{K} \to \mathbb{R} \qquad \text{by} \quad \tilde{f}(x) = \tilde{f}(x_1, ..., x_\mu) = f^*(x_1, ..., x_\mu, 0, ..., 0),$$

$$\tilde{H} = \{x \in \tilde{K} \mid \tilde{f}(x) = 0\},$$

$$\tilde{B} = \{x \in \tilde{K} \mid \| x \|_p \leqslant 1\}, \qquad \text{where now} \quad \| x \|_p = \sum_{j=1}^{\mu} | x_j |^p,$$

$$\tilde{z} = (z_1, ..., z_\mu).$$

Notice that $\tilde{z}$ has no coordinates equal to zero and $\sum_{j=1}^{\mu} | \tilde{z}_j |^p = 1$. Also, $\tilde{H}$ is a hyperplane in $\tilde{K}$, $\tilde{H} + \tilde{z}$ supports $\tilde{B}$ at $\tilde{z}$, and $\dim \tilde{K} = \dim K^*$ which equals $\dim N(z)$. Thus by Lemma 4.3, $\dim \tilde{K} = 1$, and hence $\dim N(z) = 1$.
Q.E.D.

Geometrically, the points on the unit ball at which a hyperplane can support the unit ball correspond very closely to the corners of a convex polyhedron. Since these points will be of interest in the solution of problem (P*), we make the following definition.

DEFINITION 4.6. $z \in B$ with $\| z \|_p = 1$ is called a corner point of B or simply a corner if $\dim N(z) = 1$.

5. Main Theorems

We are now prepared to prove that the solution of problem (P*) is a corner point of B and that there are only a finite number of corners.

THEOREM 5.1. *If z solves problem* (P*), *then z is a corner point of B.*

Proof. Problem (P*) requires us to find $z \in K \oplus (s)$, $\| z \|_p = 1$, such that $(s \mid z) = \max\{(s \mid w) \mid w \in K \oplus (s), \| w \|_p \leqslant 1\}$, i.e., find $z \in B$ such that $K + z$ supports B at z. Thus if z solves problem (P*), by Theorem 4.5 z must be a corner point of B.

COROLLARY 5.2. *If z solves problem* (P*), *then z has at least n coordinates equal to zero.*

Proof. $\dim N(z) = 1$ implies that $C(z)$, defined in (4.4.2), has at least n rows, and hence z has at least n coordinates equal to zero.

COROLLARY 5.3. *If D satisfies the Haar condition*, i.e., *each $n + 1$ rows of D are linearly independent, then a solution of problem* (P*) *has exactly n coordinates equal to zero.*

Proof. Using Definition 4.6, $\dim N(z) = 1$. Now the Haar condition forces $C(z)$ to be an $n \times (n + 1)$ matrix so that z has exactly n coordinates equal to zero.

COROLLARY 5.4. *Suppose that $n = m - 1$ and* $| s_j | = \max\{| s_i | \mid i = 1,..., m\}$. *Then a solution of problem* (P*) *is* $z = e_j \operatorname{sgn} s_j$, where *$e_j$ is the usual unit basis vector in $\mathbb{R}^m$.*

Proof. By Corollary 5.2, z must have at least $n = m - 1$ coordinates equal to zero, and since $\| z \|_p = 1$, z must be one of the vectors $\pm e_k$, $1 \leqslant k \leqslant m$. $(\pm e_k \mid s) = \pm s_k$ is clearly maximized by $e_j \operatorname{sgn} s_j$, where $| s_j | = \max\{| s_i | \mid i = 1,..., m\}$. Hence $e_j \operatorname{sgn} s_j$ solves problem (P*).

LEMMA 5.5. *If x, y are corner points of B with $J(y) \subset J(x)$, then $x = \pm y$, and hence $J(x) = J(y)$.*

Proof. $J(y) \subset J(x)$ implies that $N(x) \subset N(y)$, and since both $N(x)$ and $N(y)$ are one-dimensional subspaces of $\mathbb{R}^m$, $N(x) = N(y)$. From this it follows that x is a scalar multiple of y. But since $\| x \|_p = 1 = \| y \|_p$, $x = \pm y$.

Q.E.D.

For convenience, we call two corner points x, y of B different if $x \neq \pm y$, i.e., if they are neither equal nor antipodal, i.e., $J(x) \neq J(y)$.

THEOREM 5.6. *There are at most $\binom{m}{n}$ different corner points of B. Moreover, if D satisfies the Haar condition, then there are exactly $\binom{m}{n}$ different corner points of B.*

Proof. Let the set of $\binom{m}{n}$ distinct n element subsets of $\{1,..., m\}$ be denoted by E_1, and let $u(1),..., u(q)$ be a complete enumeration of all the different corner points of B. The existence of such a $q \in \mathbb{N}$ will be established in the course of the proof of this theorem. In fact, we shall show that to each corner point $u(i)$ of B we can assign a distinct $I \in E_1$ showing that the number of different corner points of B is at most $\binom{m}{n}$.

We inductively define subsets F_i and E_i of E_1 by setting $F_i = \{I \in E_i \mid I \subset J(u(i))\}$ and $E_{i+1} = E_i \backslash F_i$, $i \geqslant 1$. We assert that $F_i \neq \varnothing$ whenever $u(i)$ is a corner point of B. If not, there exists a first index k for which $F_k = \varnothing$. Since $u(k)$ is a corner point of B, $C(u(k))$ must contain n linearly independent rows, say r_j, $j \in I \subset J(u(k))$, such that $N(u(k)) = \{x \in \mathbb{R}^{n+1} \mid (r_j \mid x) = 0, j \in I\}$ is one-dimensional. Notice also that $u_j(k) = 0$ for all $j \in I$ since $I \subset J(u(k))$. By assumption,

$$\varnothing = F_k = E_1 \backslash (F_1 \cup \cdots \cup F_{k-1}).$$

Hence $I \subset F_l$ for some l, $1 \leqslant l \leqslant k-1$. This shows that $I \subset J(u(l))$ and hence $u_i(l) = 0$ for all $i \in I$. Since $N(u(l))$ is also one-dimensional and $I \subset J(u(l))$, we conclude that

$$N(u(l)) = \{x \in \mathbb{R}^{n+1} \mid (r_j \mid x) = 0, j \in I\}.$$

Moreover, since $\| u(k)\|_p = \| u(l)\|_p$, $u(k) = \pm u(l)$ contradicting the assumption that $u(k)$ and $u(l)$ are different corner points of B. Hence $F_i \neq \varnothing$, $i = 1,..., q$.

By construction, the F_i are mutually exclusive subsets of E_1. Hence to each corner point $u(i)$ of B one can assign a distinct $I \in F_i$. Thus there are at most $\binom{m}{n}$ different corner points of B.

If D satisfies the Haar condition, then by Corollary 5.3 each corner point of B has exactly n coordinates equal to zero. Each of the $\binom{m}{n}$ choices of n coordinates from the m yields an $n \times (n+1)$ matrix M for which dim (Ker M) $= 1$, and hence each of the $\binom{m}{n}$ possible choices produces a different corner point of B. Q.E.D.

Corollary 5.2 was first proved by Motzkin and Walsh [2, Theorem 6] in the following form:

Let E consist of the real points $x_1,..., x_m$ $(m > n+1)$, let $F(x)$ be defined on E, let p $(0 < p < 1)$ be given, and let the functions $\psi_1(x),..., \psi_{n+1}(x)$ satisfy Condition

A. Then every function $P(x) \equiv \sum_{j=1}^{n+1} \alpha_j \psi_j(x)$ of best approximation measured by the deviation

$$\sum_{k=1}^{m} \mu_k \mid F(x_k) - P(x_k) \mid^p \qquad (\mu_k > 0)$$

coincides with $F(x)$ in at least $n + 1$ points of E.

Condition A says that the $m \times (n + 1)$ matrix $(\psi_j(x_i))$ has rank $n + 1$.

In the same paper, Motzkin and Walsh observe that

> Theorem 6 implies that every extremal polynomial $P(x)$ is found by interpolation to $F(x)$ in $n + 1$ points of E; there exists but a finite number of polynomials interpolating to $F(x)$ in $n + 1$ points of E, so every extremal polynomial can be found merely by comparing their measures of approximation.

In our terminology, Motzkin and Walsh assert that there are but a finite number of points $z \in B$ which have at least n coordinates equal to zero, so by checking these points one can solve problem (P*) and hence problem (P).

Without making a further assumption about the class of interpolating functions, i.e., about our matrix D, there need not be only a finite number of polynomials interpolating $F(x)$ at $n + 1$ points of E. As Corollary 5.3 indicates, assuming the Haar condition is sufficient to establish Motzkin and Walsh's assertion. When the Haar condition is violated, however, one can easily construct counterexamples to the assertion. Suppose $m = 3$, $F(x_1) = F(x_2) = 0, F(x_3) = \psi_1(x_1) = \psi_1(x_2) = \psi_2(x_1) = \psi_2(x_2) = \psi_2(x_3) = 1$, and $\psi_1(x_3) = -2$. The matrix

$$\psi = (\psi_j(x_i)) = \begin{pmatrix} 1 & 1 \\ 1 & 1 \\ -2 & 1 \end{pmatrix}$$

clearly has rank $n + 1 = 2$, so Condition A is satisfied. For any $\alpha \in \mathbb{R}$

$$P(x) = -\alpha\psi_1(x) + \alpha\psi_2(x)$$

satisfies $P(x_1) = P(x_2) = 0$, so that $P(x)$ interpolates $F(x)$ in $n + 1 = 2$ points of E. Clearly, there are an infinite number of these interpolating polynomials showing that the observation of Motzkin and Walsh is incorrect. Theorems 5.1 and 5.6 guarantee that the given approximation problem can always be solved in a finite number of steps.

6. Solution of Problem (P*)

The formal similarity of problem (P*) with linear programming problems suggests an exchange algorithm moving from one corner to an adjacent corner always increasing the value of the objective function $(s \mid w)$ until a solution of problem (P*) is found. Following the example of the simplex method, such an algorithm is easily described. The corner point of B obtained by such an algorithm, however, need not solve problem (P*) because of the nonconvexity of the l_p-unit ball, $0 < p < 1$. Figure 1 shows an example in which a local maximum need not be a global maximum. However, guided by the above analogy, we investigated several exchange algorithms. We describe them in the next section.

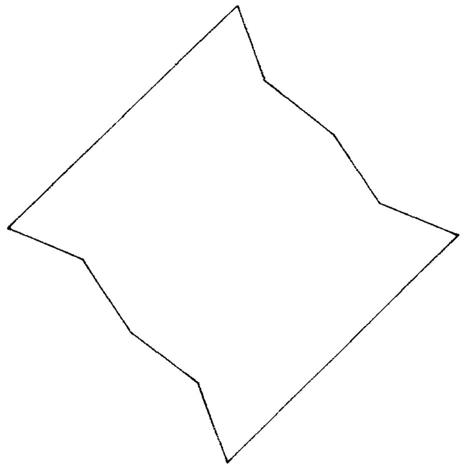

FIG. 1. Intersection of the five-dimensional $l_{0.2}$-unit ball with the plane spanned by the vectors

$$\begin{pmatrix} 50 \\ 1 \\ 5 \\ 5 \\ 0 \end{pmatrix} \quad \begin{pmatrix} 0 \\ -5 \\ -5 \\ -1 \\ 50 \end{pmatrix}.$$

Indeed, one sure method of solving problem (P*), and hence (P) also, is to find all of the corner points of B and compare their inner product with s.

Remark 6.1. Given $z \in K \oplus (s)$, there exists a unique $\beta \in \mathbb{R}^{n+1}$ such that $z = D\beta$. Also, since $s \in K^{\perp}$, $(z \mid s) = (D\beta \mid s) = \beta_{n+1}(s \mid s)$. Recall that problem (P*) requires us to

$$\text{maximize } (s \mid w) \text{ over all } w \in K \oplus (s), \| w \|_p = 1.$$

In view of the above observation, this is the same as

$$\text{maximize } \beta_{n+1} \text{ over all } \beta \in \mathbb{R}^{n+1}, \| D\beta \|_p = 1, \tag{6.1.1}$$

where $\beta = (\beta_1, ..., \beta_{n+1})$.

Using (6.1.1) as our formulation of problem (P*), to solve the problem we need only find all those $\beta \in \mathbb{R}^{n+1}$ for which $D\beta$ is a corner point of B, select one whose $(n+1)$-coordinate is the largest, say β^*, and then $z = D\beta^*$ solves problem (P*). Using Theorem 3.1, one can then solve problem (P).

Before outlining in more detail such an algorithm, we define two functions that will be useful in ensuring that each corner point of B is found exactly once.

DEFINITION 6.2. Let

$$U = \{u \in \mathbb{N}^n \mid u = (u_1, \dots, u_n).\ 1 \leqslant u_1 < u_2 < \cdots < u_n \leqslant m\}$$

$$T = \left\{1, \dots, \binom{m}{n}\right\}.$$

Define $\psi\colon T \to U$ by the following rules. Given $t \in T$,

(1) Set $t_0 = t$, $u_0 = 0$, and $i = 1$.

(2) Find $u_i \in \mathbb{N}$ such that $u_{i-1} < u_i \leqslant m - n + 1$ and

$$1 \leqslant t_{i-1} - \sum_{j=1+u_{i-1}}^{u_i - 1} \binom{m-j}{n-i} \leqslant \binom{m-u_i}{n-i}.$$

(3) Set

$$t_i = t_{i-1} - \sum_{j=1+u_{i-1}}^{u_i-1} \binom{m-j}{n-i}$$

and increment i by 1.

(4) Repeat steps 2 and 3 until u_n has been found.

Then $\psi(t) = u \in U$, where the components $u_1, \dots, u_n$ of u were found above. By convention, $\sum_{j=\mu}^{\nu} (\cdot) = 0$ if $\mu > \nu$.

One can show [3] that ψ is a one-to-one function, and hence ψ has an inverse. By rearranging step 3 of Definition 6.2, we find

$$\psi^{-1}(u) = 1 + \sum_{i=1}^{n} \left[\sum_{j=1+u_{i-1}}^{u_i-1} \binom{m-j}{n-i} \right],$$

where $u_0 = 0$ and $\sum_{j=\mu}^{\nu} (\cdot) = 0$ if $\mu > \nu$.

DEFINITION 6.3. Let T and U be as in Definition 6.2. Define $\Phi\colon U \to T$ by $\Phi = \psi^{-1}$.

ALGORITHM 6.4. (1) Set $q = 1$, $p_i = 1$, $\beta_i = 0$ for $i = 1, \dots, \binom{m}{n}$.

(2) Compute $\psi(q) = (k_1, \dots, k_n)^T$.

(3) Construct

$$A = \begin{pmatrix} r_{k_1} \\ \vdots \\ r_{k_n} \end{pmatrix}$$

and set $I = \{k_1, ..., k_n\}$.

(4) Select $i \in \{1, ..., m\} \backslash I$, and form the matrix $C = \binom{A}{r_i}$.

(5) If C contains $n + 1$ linearly independent rows, then go to step 7, otherwise go to step 6.

(6) Set A equal to C, $I = I \cup \{i\}$, and return to step 4.

(7) Solve $Ax = 0$, $(r_i \mid x) = 1$.

(8) Compute Dx and $\gamma = \| Dx \|_p^{1/p}$.

(9) Set $\beta_q = | x_{n+1} | / \gamma$ and $z(q) = (\operatorname{sgn} x_{n+1}/\gamma)\, Dx$.

(10) Form all possible sets containing exactly n elements of $J(Dx)$.

(11) For $j = \{j_1, ..., j_n\}$ found in step 10, where $1 \leqslant j_1 < j_2 < \cdots < j_n \leqslant m$, compute $\Phi(j) = t$ and set $p_t = 0$.

(12) If $p_i = 0$, $i = 1, ..., \binom{m}{n}$, then go to step 13. Otherwise, let q be the smallest integer k, $1 \leqslant k \leqslant \binom{m}{n}$, such that $p_k = 1$. Return to step 2.

(13) Select k, $1 \leqslant k \leqslant \binom{m}{n}$, with

$$\beta_k = \max \left\{ \beta_i \;\middle|\; 1 \leqslant i \leqslant \binom{m}{n} \right\}.$$

Then $z(k)$ solves problem (P*) and

$$\max\{(s \mid w) \mid w \in k \oplus (s), \| w \|_p = 1\} = \beta_k(s \mid s).$$

In step 5, one must eventually answer the question in the affirmative since the rank of D is $n + 1$ by hypothesis. The question itself can be answered in a number of ways. For example, one might orthogonalize the rows of C and check whether any zero rows occur. This method will also help when step 7 is reached since one then knows which rows of C yield a nonsingular matrix G with which to solve $Gx = e_n$. Steps 10 and 11 are present to exploit Lemma 5.5, which says that some of the original $\binom{m}{n}$ possible corner points may in fact be redundant. In step 9, one need not save all of the β_i and $z(i)$, but only the current largest β_i and the corresponding $z(i)$ which would make step 13 unnecessary.

Algorithm 6.4 solves problem (P*) for any choice of positive integers m, n with $m > n$. The price paid for this flexibility is a considerable amount of index manipulation. In the cases where dim K is either very small or nearly equal to m, we can avoid much of this work by developing special algorithms designed to solve only problems with a particular fixed choice of dim K.

ALGORITHM 6.5. If $\dim K = m - 1$, then Corollary 5.4 gives the solution of problem (P*).

Corollary 5.2 and some algebraic manipulations lead to the following special algorithms for solving problem (P*).

ALGORITHM 6.6. Let $\dim K = m - 2$, $a \in (K \oplus (s))^{\perp} \setminus \{0\}$, $s = (s_1, \ldots, s_m)^T$, $a = (a_1, \ldots, a_m)^T$, and $z = (z_1, \ldots, z_m)^T$. With

$$\begin{aligned} f_{ij} &= \frac{|a_j s_i| + |a_i s_j|}{(|a_i|^P + |a_j|^P)^{1/P}}, & a_i^2 + a_j^2 \neq 0, \\ &= 0, & a_i^2 + a_j^2 = 0, \end{aligned} \tag{6.6.1}$$

find μ, ν such that $f_{\mu\nu} = \max\{f_{ij} \mid 1 \leqslant i, j \leqslant m, i \neq j\}$. Set

$$z_\mu = \frac{|a_\nu| \operatorname{sgn} s_\mu}{(|a_\mu|^p + |a_\nu|^p)^{1/p}},$$

$$z_\nu = \frac{|a_\mu| \operatorname{sgn} s_\nu}{(|a_\mu|^p + |a_\nu|^p)^{1/p}}.$$

Then a solution of problem (P*) is $z = z_\mu e_\mu + z_\nu e_\nu$.

ALGORITHM 6.7. Let $\dim K = m - 3$ and $s = (s_1, \ldots, s_m)$. Find linearly independent vectors $(a_1, \ldots, a_m)$, $(b_1, \ldots, b_m) \in (K \oplus (s))^{\perp}$. Given distinct i, j, k between 1 and m, let

$$A_i = \begin{vmatrix} a_i & a_k \\ b_i & b_k \end{vmatrix}, \quad A_j = \begin{vmatrix} a_j & a_k \\ b_j & b_k \end{vmatrix}, \quad B_i = \begin{vmatrix} A_j & A_i \\ b_j & b_i \end{vmatrix}, \quad \text{and} \quad B_k = b_k A_j.$$

Find λ, μ, ν such that $g_{\lambda\mu\nu} = \max\{ g_{ijk} \mid 1 \leqslant i, j, k \leqslant m;$ all distinct$\}$, where

$$\begin{aligned} g_{ijk} &= (|A_j B_k s_i| + |A_i B_k s_j| + |A_j B_i s_k|)/D & \text{if } D \neq 0 \\ &= 0 & \text{if } D = 0, \end{aligned}$$

$$D = (|A_j B_k|^P + |A_i B_k|^P + |A_j B_i|^P)^{1/P}.$$

Set

$$z_\lambda = |A_\mu B_\nu| \operatorname{sgn} s_\lambda / D,$$

$$z_\mu = |A_\lambda B_\nu| \operatorname{sgn} s_\mu / D,$$

$$z_\nu = |A_\mu B_\lambda| \operatorname{sgn} s_\nu / D.$$

Then $z = Z_\lambda e_\lambda + Z_\mu e_\mu + Z_\nu e_\nu$ solves problem (P*).

ALGORITHM 6.8. Suppose that $K = (a)$, where $a = (a_1, ..., a_m)$ and $s = (s_1, ..., s_m)$. Set

$$\alpha_i = \frac{-s_i \operatorname{sgn} a_i}{\| as_i - a_i s \|_p^{1/p}} \quad \text{and} \quad \beta_i = \frac{| a_i |}{\| as_i - a_i s \|_p^{1/p}}.$$

Let μ satisfy $\beta_\mu = \max\{\beta_i \mid 1 \leqslant i \leqslant m\}$. Then a solution of problem (P*) is $z = \alpha_\mu a + \beta_\mu s$.

7. LOCAL SOLUTIONS

Algorithms 6.4 through 6.8 have two distinguishing features—one good and the other bad. On the one hand, they always work, i.e., they give the correct solution of problem (P*). On the other hand, Algorithm 6.4 in particular can involve a tremendous amount of work since every corner point of B must be computed. Consequently, unless m and n are fairly small numbers or the Haar condition is so flagrantly violated that the actual number of corner points of B is reasonably small, Algorithm 6.4 does not represent a computationally feasible method for finding the solution of problem (P*).

DEFINITION 7.1. Let x, y be corner points of B. We say that x and y are adjacent if $\{r_j \mid j \in J(x) \cap J(y)\}$ contains $n - 1$ linearly independent vectors.

The idea behind this definition is most readily seen if we assume that the Haar condition holds. In that situation, the fact that x and y are adjacent corner points of B implies that both $J(x)$ and $J(y)$ have exactly n elements and $\{r_j \mid j \in J(x) \cap J(y)\}$ contains $n - 1$ linearly independent row vectors; i.e., $J(x) \cap J(y)$ contains exactly $n - 1$ elements. Thus there is an $i \in J(x)$ and a $j \in J(y)$ such that $J(x) = J(y) \cup \{j\} \backslash \{i\}$ and $J(y) = J(x) \cup \{i\} \backslash \{j\}$. In terms of coordinates, all but one of the zero coordinates of either x or y is also a zero coordinate of the other.

Remark 7.2. Adjacent corner points of B can be much farther apart than one might expect a term like *adjacent* to allow. For example, if K is one-dimensional, then each two corner points of B are adjacent since a corner point need only have one coordinate equal to zero.

DEFINITION 7.3. A corner point z of B is called a local solution of (P*) if $(z \mid s) \geqslant (x \mid s)$ for all corner points x adjacent to z.

It follows from the definition of adjacent corner points of B that there can exist corner points of B which are not adjacent. Consequently, a local solution of problem (P*) need not be a solution of problem (P*).

We now present an exchange algorithm similar to the simplex method. The solution found in this manner may, however, only be a local solution of problem (P*).

ALGORITHM 7.4. (1) Find a corner point $z = D\beta$ of B and set $I = J(z)$.

(2) Select n linearly independent rows $r_{i_1},\ldots, r_{i_n}$ of D with $i_1,\ldots, i_n \in I$.

(3) Pick $\mu \in \{i_1,\ldots, i_n\}$.

(4) Relabel $r_{i_1},\ldots, r_{i_n}$ as $\bar{\rho}_1,\ldots, \bar{\rho}_n$ with $\bar{\rho}_n = r_\mu$.

(5) Orthogonalize the $\bar{\rho}_j$ by (i) $\rho_1 = \bar{\rho}_1$, and (ii) for $i = 2,\ldots, n$, $\rho_i = \bar{\rho}_i - \sum_{j=1}^{i-1} ((\rho_j \mid \bar{\rho}_i)/(\rho_j \mid \rho_j))\, \rho_j$.

(6) Pick $k \in \{1,\ldots, m\}\backslash I$.

(7) Set $\tilde{\beta} = \gamma_k \| D\gamma_k \|^{-1/p} \operatorname{sgn}(\gamma_k)_{n+1}$, where $\gamma_k = \rho_n - ((\rho_n \mid r_k)/(\beta \mid r_k))\beta$. We assert that $\tilde{z} = D\tilde{\beta}$ is a corner point of B. By construction, $\| \tilde{z} \|_p = 1$ and $\{r_{i_1},\ldots, r_{i_n}, r_k\}$ are $n + 1$ linearly independent vectors. $\{i_1,\ldots, i_n, k\}\backslash\{\mu\} \subset J(\tilde{z})$, so $\dim N(\tilde{z}) \leqslant 1$. But $\dim N(\tilde{z}) \geqslant 1$ since $\tilde{\beta} \in N(\tilde{z})\backslash\{0\}$. Thus $\dim N(\tilde{z}) = 1$ showing that $\tilde{z}$ is a corner point of B.

(8) (i) If $\tilde{\beta}_{n+1} > \beta_{n+1}$, replace β by $\tilde{\beta}$ and z by $\tilde{z}$. Set $I = J(z)$ and then return to step 2.

(ii) If $\tilde{\beta}_{n+1} \leqslant \beta_{n+1}$, return to step 6 and try another $k \in \{1,\ldots, m\}\backslash I$ until all have been tried.

(iii) When all $k \in \{1,\ldots, m\}\backslash I$ have been tried in (ii), return to step 3 to choose another $\mu \in \{i_1,\ldots, i_n\}$.

(iv) When all $\mu \in \{i_1,\ldots, i_n\}$ have been tried in (iii), z is a local solution of problem (P*) with value β_{n+1}.

Step 1 can be accomplished in the same manner in which corner points of B were found in Algorithm 6.4. Experience with a few examples indicates that a good starting corner to find in step 1 is that z which has zero coordinates where the coordinates of the vector s are the smallest in absolute value. In many cases, this corner point of B actually solves problem (P*).

Step 8(i) ensures that Algorithm 7.4 eventually terminates since the value of the objective function $\beta_{n+1} = (s \mid z)$ is nondecreasing and there are only a finite number of corner points. Lemma 5.5 guarantees that in step 8(iii) we need only check the n coordinates listed rather than all of the zero coordinates. The assertion in step 8(iv) that the point z found by Algorithm 7.4 is a local solution of problem (P*) follows directly from Definition 7.3.

If $\dim(K \oplus (s))$ is close to $\dim \mathbb{R}^m$, then the computations involved in choosing $n + 1$ linearly independent row vectors (step 2) and orthogonalizing them (step 5) can become tedious. However, when n is nearly equal to m, then $\dim\{(K \oplus (s))^\perp\} = m - n - 1$ is quite small. An exchange algorithm designed to exploit this fact can be found in [3].

8. The l_1-Case

We observe that many of the results obtained for $0 < p < 1$ also hold for the case where $p = 1$, thus providing a method of solving the l_1 problem. With $p = 1$, Theorem 3.1 is a special case of Theorem 2.4 in [4]. If we alter Lemma 4.3 so that (iii) reads $H + z$ supports B_Y at z, i.e., either $f(x) < f(z)$ for all $x \in B_Y \backslash \{z\}$ or $f(x) > f(z)$ for all $x \in B_Y \backslash \{z\}$, then Lemma 4.3 is true for $p = 1$. In the proof, choose x, δ, γ, and ϵ as before, but obtain a contradiction either to the hypothesis that $H + z$ supports B_Y at z or to the strictness of the support. Theorem 4.5 follows immediately for $p = 1$ if we again assume that $H + z$ properly supports B at z. Theorem 5.1 and all three of its corollaries, Lemma 5.5, and Theorem 5.6 all hold as previously stated with $p = 1$.

Algorithm 7.4 solves problem (P*) when $p = 1$ since the l_1-unit ball being convex eliminates the possibility of finding a local solution that is not a global solution of problem (P*).

9. Numerical Results

Algorithm 6.4 was programmed in FORTRAN IV for a CDC 6500, and two examples were studied. The first example appears in [1, p. 44]. The overdetermined system of linear equations is

$$\begin{aligned} x + y &= 3, \\ x - y &= 1, \\ x + 2y &= 7, \\ 2x + 4y &= 11.1, \\ 2x + y &= 6.9, \\ 3x + y &= 7.2. \end{aligned}$$

This system poses special difficulties because the solution of the l_1 problem is not unique. All points on the segment joining

$$P_1 = (1.77, 1.89) \qquad \text{and} \qquad P_2 = (2.51667, 1.51667)$$

solve the l_1 problem with a minimal l_1 error vector of length 4.7.

For $p = 1$, the algorithm found both corner point solutions. For $p = n/10$, $n = 1, 2, \dots, 9$, the l_p problem has the unique solution P_2. For each case, the algorithm took less than one second to compute the solutions.

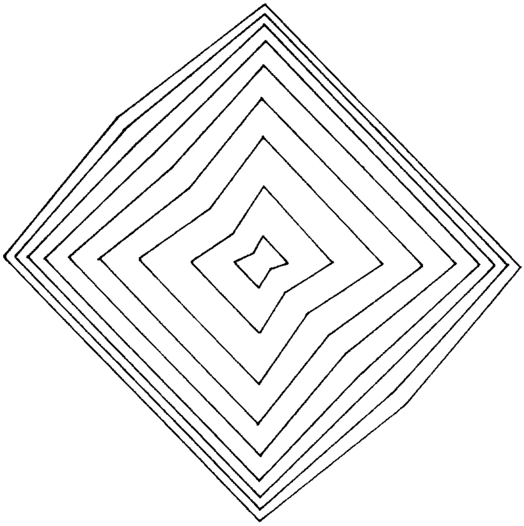

FIG. 2. Intersection of the three-dimensional l_p-unit ball, $p = 1, 0.9, 0.8, \ldots, 0.2$, with the plane spanned by the vectors

$$\begin{pmatrix} 5 \\ 1 \\ -5 \end{pmatrix}, \quad \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

For a second example, we chose K to be the subspace spanned by the single vector

$$\begin{pmatrix} 5 \\ 1 \\ -5 \end{pmatrix}, \text{ and took } b \text{ to be } \begin{pmatrix} 1 \\ 0 \\ 1 \end{pmatrix}.$$

The intersections of $K \oplus (s)$ and the three-dimensional l_p-unit balls for $p = n/10$, $n = 2, \ldots, 10$, are shown in Fig. 2. s runs diagonally from the upper left to the lower right, passing through the corner points shown. The solution of problem (P*) "jumps" near $p = 0.777$, as the figure indicates.

REFERENCES

1. E. W. CHENEY, "Introduction to Approximation Theory," McGraw–Hill, New York, 1966.
2. T. S. MOTZKIN AND J. L. WALSH, Least pth power polynomials on a real finite point set, *Trans. Amer. Math. Soc.* **78** (1955), 67–81.
3. R. W. OWENS, "Algorithms for Solving Overdetermined Systems of Linear Equations in the l_p Sense," Ph.D. Thesis, Michigan State University, 1975.
4. V. P. SREEDHARAN, Least squares algorithms for finding solutions of overdetermined systems of linear equations which minimize error in a smooth strictly convex norm, *J. Approximation Theory* **8** (1973), 46–61.