JOURNAL OF APPROXIMATION THEORY 73, 180–198 (1993)

Least Squares Methods to Minimize Errors in a Smooth, Strictly Convex Norm on $\mathbb{R}^m$

R. W. OWENS

Department of Mathematical Sciences, Lewis and Clark College, Portland, Oregon 97219, U.S.A.

AND

V. P. SREEDHARAN

Department of Mathematics, Michigan State University, East Lansing, Michigan 48824

Communicated by E. W. Cheney

Received June 7, 1990; accepted in revised form January 13, 1992

An algorithm for computing solutions of overdetermined systems of linear equations in n real variables which minimize the residual error in a smooth, strictly convex norm in a finite dimensional space is given. The algorithm proceeds by finding a sequence of least squares solutions of suitably modified problems. Most of the time, each iteration involves one line search for the root of a nonlinear equation, though some iterations do not have any root seeking line search. Convergence of the algorithm is proved, and computational experience on some numerical examples is also reported. © 1993 Academic Press, Inc.

1. INTRODUCTION

In this paper we propose an algorithm for approximating solutions of the overdetermined system of linear equations

$$Ax = b,$$

where A is an $m \times n$ real matrix, m, $n \geqslant 1$, and $b \in \mathbb{R}^m$; no special assumption is made on the rank of A. We seek a minimizer of

$$\|b - Ax\|,$$

the norm of the residual vector $r(x) = b - Ax$, when the norm is smooth and strictly convex. In each iteration of the proposed algorithm, we modify

0021-9045/93 $5.00

the constant vector b twice and solve the resulting linear systems in the least squares sense. We assume throughout that the system is inconsistent.

The algorithm and results of this paper easily extend to the slightly more general problem of minimizing $\|b-z\|$, $z \in K$, where K is a subspace of a smooth, strictly convex, finite dimensional Banach space X and $b \in X \backslash K$. Little, however, is gained by such an extension since, first, the extension is immediate, and second, the most likely use of the algorithm is for solving overdetermind systems of linear equations in the l_p sense, i.e., minimizing $\|b-Ax\|_p$, where $\|\cdot\|_p$ is an l_p norm, $1<p<\infty$.

Many algorithms exist for solving overdetermined systems of linear equations in the l_p sense, $1 \leqslant p \leqslant \infty$. See, for example, Fletcher, Grant, and Hebden [5], Kahng [7], Merle and Späth [9], Owens [10], Späth [11], Sreedharan [13, 14], and Wolfe [18]. Any general purpose minimization algorithm could, at least theoretically, be used to minimize $\|b-Ax\|_p$, $x \in \mathbb{R}^n$, but such an approach would not exploit the special structure of the problem and is unlikely to be particularly efficient. When $p=1$ or ∞, linear programming can be used effectively. See, for example, Barrodale and Young [1], Chvátal [4], Wagner [17], and Zukhovitskiy and Aydeyeva [19]. And when $p=2$, very efficient techniques for solving the l_p problem exists. See, for example, Björck [2], Golub and Van Loan [6], Lawson and Hanson [8], and Sreedharan [16].

For $1<p<\infty$, the function $g(x)=\|b-Ax\|_p^p=\Sigma\,|b_i-(Ax)_i|^p$, $x \in \mathbb{R}^n$, is positive and differentiable, and hence, due to the convexity of the norm, g is continuously differentiable. So minimizing g is equivalent to finding a root of g'. The more effective the root finding method employed in $\mathbb{R}^n$, the better this approach works, but the better the root finding method is, generally, the more derivatives of the function g one needs. For $p \geqslant 2$, Newton-type algorithms are available, but for $1<p<2$, convergence cannot be guaranteed without adding the restriction that $r_i(x) \neq 0$, $1 \leqslant i \leqslant m$, a condition that cannot be know a priori. It is tempting, nevertheless, to employ second order methods to solve l_p problems with $1<p<2$. See, for example, [11], where such methods are applied to specific examples, and the numerical answers obtained are then checked a posteriori to see whether they are the "right" answers. Such tests lend credence to the applicability of second order methods, even when the relevant hypotheses are violated, but such findings do not prove the convergence of the algorithm.

The algorithm presented in this paper is shown to converge for any smooth, strictly convex norm on $\mathbb{R}^m$. In particular, the algorithm is applicable for the troublesome l_p norms when $1<p<2$. Moreover, our algorithm performs well in l_p problems, $1<p<2$, when compared with algorithms whose convergence in general cannot be established, such as those presented in [11].

In Section 2, we specify notation and terminology, and we introduce the duality theory needed later in the paper. The algorithm is given in Section 3. We establish the feasibility and convergence of the algorithm in Section 4. Finally, in Section 5, we present some numerical results using the usual l_p norm.

2. Preliminaries and Duality Theory

Throughout this paper, $m \geqslant 1$. The standard inner product on $\mathbb{R}^m$ is denoted by $\langle \cdot, \cdot \rangle$, i.e.,

$$\langle u, v \rangle = \Sigma u_i v_i, \qquad u, v \in \mathbb{R}^m,$$

$u = (u_1, ..., u_m)$ and $v = (v_1, ..., v_m)$.

Let E be the orthogonal projection of $\mathbb{R}^m$ onto $K = \mathrm{Ker}(A^T)$, where orthogonality is with respect to the standard inner product and A^T denotes the transpose of the matrix A. $K^\perp = \mathrm{Im}(A)$. Let $s = Eb$. Our earlier assumption guarantees that $s \neq 0$. Note that s is the least squares residual.

In terms of this notation, the problem which we address, referred to as problem (P), is

$$\text{Minimize } \|b - z\| \text{ subject to } z \in K^\perp. \tag{P}$$

We assume that the norm $\|\cdot\|$ is smooth and strictly convex. The norm is said to be *smooth* if and only if through each point of unit norm there passes precisely one hyperplane supporting the closed unit ball $B = \{x \in \mathbb{R}^m \mid \|x\| \leqslant 1\}$. The norm is said to be *strictly convex* if and only if the unit sphere $S = \{x \in \mathbb{R}^m \mid \|x\| = 1\}$ has no line segments on it.

In order to introduce a dual problem (P′), we define the dual norm $\|\cdot\|'$ on $\mathbb{R}^m$ by

$$\|y\|' = \max\{\langle x, y \rangle \mid \|x\| = 1, x \in \mathbb{R}^m\}.$$

Given $y \neq 0$, we define y', a $\|\cdot\|$-dual, and y^*, a $\|\cdot\|'$-dual, by

$$\|y'\| = 1, \langle y', y \rangle = \|y\|' \qquad \text{and} \qquad \|y^*\|' = 1, \langle y^*, y \rangle = \|y\|.$$

It is well know that if the norm $\|\cdot\|$ is strictly convex, then $\|\cdot\|$-duals are unique, see, e.g., [12]. The mapping $v \mapsto v'$ is continuous and positively homogeneous of degree zero on $\mathbb{R}^m \backslash \{0\}$; see, e.g., [14]. Similarly, if the norm $\|\cdot\|$ is smooth, then $\|\cdot\|'$-duals are unique and the mapping $v \mapsto v^*$ is continuous and positively homogeneous of degree zero on $\mathbb{R}^m \backslash \{0\}$. For readers who wish to refresh their memory about this or need an explicit proof of the fact the norm $\|\cdot\|$ is smooth if and only if the norm $\|\cdot\|'$ is strictly convex, we refer to Section 2 of [13]. Moreover, for $v \neq 0$, we have the relations

$$v'^* = v/\|v\|' \qquad \text{and} \qquad v^{*\prime} = v/\|v\|,$$

when the norm $\|\cdot\|$ is smooth and strictly convex, respectively. For proof of these, see [14]. We use the continuity of the prime- and star-dual mappings and the special form of their compositions often, and without explicit reference, throughout this paper.

For the usual l_p norm, $1 < p < \infty$, which is given by

$$\|v\|_p = (\Sigma\, |v_i|^p)^{1/p},$$

we have that $\|\cdot\|_p' = \|\cdot\|_q$, where $p + q = pq$. Given $v \neq 0$, then v', its $\|\cdot\|_p$-dual, and v^*, its $\|\cdot\|_p'$-dual, have components v_i' and v_i^* given by

$$v_i' = (|v_i|/\|v\|_q)^{q-1} \operatorname{sgn} v_i, \qquad i = 1, ..., m,$$

$$v_i^* = (|v_i|/\|v\|_p)^{p-1} \operatorname{sgn} v_i, \qquad i = 1, ..., m,$$

respectively. Expressions for the $\|\cdot\|$- and $\|\cdot\|'$-duals with respect to more general weighted l_p norms are given in [14].

With problem (P) we associate the following dual problem (P′):

$$\text{Maximize } \langle b, y \rangle \text{ subject to } y \in K,\ \|y\|' = 1. \qquad \text{(P}'\text{)}$$

Problems (P) and (P′) are easily seen to be equivalent, respectively, to the following problems:

$$\text{Minimize } \|s - z\| \text{ subject to } z \in K^{\perp}, \text{ where } s = Eb \qquad \text{(P)}$$

and

$$\text{Maximize } \langle s, y \rangle \text{ subject to } y \in K,\ \|y\|' = 1. \qquad \text{(P}'\text{)}$$

We use these versions throughout the sequel.

Because of finite dimensionality, problems (P) and (P′) each have a solution. If the norm $\|\cdot\|$ is strictly convex, then the solution of problem (P) is unique, whereas if the norm $\|\cdot\|$ is smooth, then the solution of problem (P′) is unique. Note that the latter result follows from the property we observed earlier, i.e., the norm $\|\cdot\|$ is smooth if and only if the norm $\|\cdot\|'$ is strictly convex.

If we assume that the norm $\|\cdot\|$ is strictly convex, we have the following duality theorem.

2.1. THEOREM [12, Theorem 2.1]. *Let y be a solution of problem* (P′). *Then the system of equations $Ax = b - \langle b, y\rangle y'$ is consistent. Furthermore, any solution of this system is a solution of problem* (P), *and the norm $\|\cdot\|$ of the corresponding residual for problem* (P) *equals $\langle b, y\rangle$.*

Paraphrasing this theorem in terms of K and $K^\perp$ yields the following statement.

2.2. THEOREM. *Let y be a solution of problem* (P′). *Then $u = s - \langle s, y\rangle y'$ is the unique point in $K^\perp$ nearest to s; $v = b - \langle b, y\rangle y'$ is also the unique point in $K^\perp$ nearest to b. Furthermore, the minimal norm $\|\cdot\|$ residual t of problem* (P) *is given by*

$$t = \langle s, y\rangle y'$$

and

$$d(s, K^\perp) = d(b, K^\perp) = \|t\| = \langle s, y\rangle,$$

where we have adopted the notation $d(a, X) = \inf\{\|a - x\| \mid x \in X\}$, for any subset $X \subset \mathbb{R}^m$ and $a \in \mathbb{R}^m$.

Henceforth we assume that the norm $\|\cdot\|$ on $\mathbb{R}^m$ is both smooth and strictly convex. The reader will easily see that stronger theorems can be proven by assuming only smoothness or strict convexity as the case warrants. But this stronger assumption makes many statements simpler and more elegant.

2.3. THEOREM. *Let $s = Eb \neq 0$. Suppose that $t = s + u$, with $u \in K^\perp$. Then t is the minimal norm $\|\cdot\|$ residual of problem* (P) *if and only if $t^* \in K$, in which case t^* is the maximizer for problem* (P′).

Proof. Given the assumption that $t^* \in K$, we show that t^* solves problem (P′) and that t is the minimal norm $\|\cdot\|$ residual of problem (P). If $y \in K$, $\|y\|' = 1$, then

$$\langle s, y\rangle = \langle t, y\rangle \leqslant \|t\|.$$

Furthermore, when $y = t^*$, equality holds, so we have

$$\max\{\langle s, y\rangle \mid y \in K, \|y\|' = 1\} = \langle s, t^*\rangle = \|t\|,$$

completing the proof of the "if" part.

Conversely, let y be a maximizer for problem (P′). Then by Theorem 2.2,

t, which is the minimal norm $\|\cdot\|$ residual of problem (P), satisfies the equality $\|t\| = \langle s, y\rangle$. Let $v = t/\|t\|$, so that $\|v\| = 1$. We now have

$$\begin{aligned}\langle v, y\rangle &= \|t\|^{-1}\langle t, y\rangle\\ &= \|t\|^{-1}\langle s+u, y\rangle\\ &= \|t\|^{-1}\langle s, y\rangle = 1.\end{aligned}$$

This show that $v = y'$, which in turn implies that $t^* = v^* = y'^* = y$. Since $t^* = y$, t^* solves problem (P′) and belongs to K.

2.4. LEMMA. *Let $0 \neq s \in K$ and $y \in K$ with $\|y\|' = 1$ and $\langle s, y\rangle > 0$. Define $t \in \mathbb{R}^m$ by*

$$t = s + \langle s, y\rangle (I - E)\, y'. \tag{2.4.1}$$

Then $Et^* \neq 0$, i.e., $t^* \notin K^\perp$.

Proof. First note that since s is non-zero, t defined by (2.4.1) is non-zero, so that t^* is well defined. Suppose $Et^* = 0$, i.e., $t^* \in K^\perp$. We shall arrive at a contradiction.

$$\begin{aligned}\|t\| &= \langle t, t^*\rangle\\ &= \langle s, t^*\rangle + \langle s, y\rangle\langle y', t^*\rangle - \langle s, y\rangle\langle Ey', t^*\rangle\\ &= \langle s, y\rangle\langle y', t^*\rangle, \qquad \text{since} \quad Et^* = 0\\ &\leqslant \langle s, y\rangle. \end{aligned} \tag{2.4.2}$$

This shows by well known duality theory that y is a maximizer for problem (P′) and that t is the minimal norm $\|\cdot\|$ residual for problem (P). So the inequality in (2.4.2) is in fact an equality. But equality can appear in (2.4.2) if and only if $\langle y', t^*\rangle = 1$, since $\langle s, y\rangle > 0$. This shows that $t^* = y$ and hence $y \in K \cap K^\perp$, which implies that $y = 0$, a contradiction to $\|y\|' = 1$.

The following characterization of an optimal solution of problem (P′) is crucial for this paper since it motivates the algorithm presented in the next section. The algorithm approximates the solution of problem (P′) by constructing, at each iteration, a vector y_k that approximately satisfies the hypotheses of the theorem.

2.5. THEOREM. *Let $s = Eb \neq 0$. Suppose that $t \in \mathbb{R}^m$ is such that*

$$t = s + \langle s, y\rangle (I - E)\, y', \qquad \langle s, y\rangle > 0,$$

where

$$y = Et^*/\|Et^*\|'.$$

Then y is the maximizer of problem (P′), *t is the minimal norm $\|\cdot\|$ residual of problem* (P), *and $y = t^*$.*

Proof. Note that $Et^* \neq 0$ by Lemma 2.4. First observe that for all $z \in K$ with $\|z\|' = 1$ we have

$$\begin{aligned} \langle s, z \rangle &= \langle t, z \rangle, \qquad \text{since} \quad z \in K \quad \text{and} \quad (I-E)\, y' \in K^{\perp} \\ &\leqslant \|t\| \, \|z\|' = \|t\|. \end{aligned} \tag{2.5.1}$$

We now show that $\langle s, y \rangle = \|t\|$. We have

$$\begin{aligned} \|t\| &= \langle t, t^* \rangle \\ &= \langle s, t^* \rangle + \langle s, y \rangle \langle y', t^* \rangle - \langle s, y \rangle \langle Ey', t^* \rangle \\ &= \langle s, t^* \rangle + \langle s, y \rangle \langle y', t^* \rangle - \langle s, y \rangle \langle y', Et^* \rangle \\ &= \langle s, t^* \rangle + \langle s, y \rangle \langle y', t^* \rangle - \|Et^*\|' \, \langle s, y \rangle \langle y', y \rangle \\ &= \langle s, t^* \rangle + \langle s, y \rangle \langle y', t^* \rangle - \langle s, Et^* \rangle \\ &= \langle s, y \rangle \langle y', t^* \rangle \leqslant \langle s, y \rangle. \end{aligned} \tag{2.5.2}$$

In view of (2.5.1) and (2.5.2) we have established

$$\|t\| = \langle s, y \rangle = \max\{ \langle s, z \rangle \mid z \in K, \ \|z\|' = 1 \},$$

i.e., y is the maximizer for problem (P′) and t is the minimal norm $\|\cdot\|$ residual for problem (P). Also, since equality holds in (2.5.2), $y = t^*$.

3. Algorithm

Step 0. Let $\varepsilon > 0$ be a stopping rule parameter. Let $y_1 \in K$, $\|y_1\|' = 1$, $\langle s, y_1 \rangle > 0$. A convenient starting y_1 is $y_1 = s/\|s\|'$. Set $k = 1$.

Step 1. Let $t_k = s + \langle s, y_k \rangle (I - E)\, y_k'$.

Step 2. If $1 - \langle s, y_k \rangle / \|t_k\| \leqslant \varepsilon$, go to Step 9. Otherwise continue.

Step 3. If $\|(I - E)\, t_k^*\|_2 \leqslant \varepsilon$, set $y_k = t_k^*$ and go to Step 9. Otherwise continue.

Step 4. Let $r_k = Et_k^* / \|Et_k^*\|'$.

Step 5. If $\langle s, r_k \rangle \langle r_k', y_k \rangle \geqslant \langle s, y_k \rangle$, set $y_{k+1} = r_k$, $\alpha_k = 1$, and go to Step 8. Otherwise continue.

Step 6. Compute $\alpha_k \in [0, 1]$ such that

$$\begin{aligned} &\langle s, \alpha_k r_k + (1 - \alpha_k)\, y_k \rangle \langle (\alpha_k r_k + (1 - \alpha_k)\, y_k)', r_k - y_k \rangle \\ &\quad = \|\alpha_k r_k + (1 - \alpha_k)\, y_k\|' \, \langle s, r_k - y_k \rangle. \end{aligned}$$

Step 7. Let $y_{k+1} = (\alpha_k r_k + (1-\alpha_k) y_k)/\|\alpha_k r_k + (1-\alpha_k) y_k\|'$.

Step 8. Increment k by 1 and return to Step 1.

Step 9. Accept y_k as the solution of problem (P′) and t_k as the minimal norm $\|\cdot\|$ residual of problem (P). Find a least squares solution of the system of linear equations $Ax = b - t_k$, and accept this solution x as the solution of problem (P).

4. Feasibility and Convergence

4.1. Lemma. *Let $0 \neq s \in K$, and let $y \in K$ with $\|y\|' = 1$ and $\langle s, y\rangle > 0$. Define t by*

$$t = s + \langle s, y\rangle (I - E) y'. \tag{4.1.1}$$

Define r by

$$r = Et^*/\|Et^*\|'. \tag{4.1.2}$$

Let $h = r - y$. Consider φ: $[0, 1] \to \mathbb{R}$ defined by

$$\varphi(\alpha) = \langle s, y + \alpha h\rangle / \|y + \alpha h\|'. \tag{4.1.3}$$

Then

$$\varphi'(1) = \langle s, r\rangle\langle r', y\rangle - \langle s, y\rangle \tag{4.1.4}$$

and

$$\|Et^*\|' \,\varphi'(0) = \|t\| - \langle s, y\rangle\langle y', t^*\rangle \geqslant 0. \tag{4.1.5}$$

Furthermore, the last inequality is strict if $t^* \notin K$.

Proof. By Lemma 2.4, r is well defined. By (5.4.3) of [15],

$$\varphi'(\alpha) = \frac{\langle s, h\rangle}{\|y + \alpha h\|'} - \frac{\langle s, y + \alpha h\rangle\langle (y + \alpha h)', h\rangle}{(\|y + \alpha h\|')^2}. \tag{4.1.6}$$

So

$$\begin{aligned} \varphi'(1) &= \frac{\langle s, h\rangle}{\|y + h\|'} - \frac{\langle s, y + h\rangle\langle (y + h)', h\rangle}{(\|y + h\|')^2} \\ &= \langle s, r - y\rangle - \langle s, r\rangle\langle r', r - y\rangle \\ &= \langle s, r\rangle\langle r', y\rangle - \langle s, y\rangle, \end{aligned} \tag{4.1.7}$$

which is (4.1.4).

By (4.1.6)

$$\begin{aligned}\varphi'(0) &= \langle s, r-y\rangle - \langle s, y\rangle\langle y', r-y\rangle \\ &= \langle s, r\rangle - \langle s, y\rangle\langle y', r\rangle \qquad (4.1.8)\end{aligned}$$

and hence

$$\begin{aligned}\|Et^*\|' \varphi'(0) &= \langle s, Et^*\rangle - \langle s, y\rangle\langle y', Et^*\rangle \\ &= \langle s, t^*\rangle - \langle s, y\rangle\langle Ey', t^*\rangle \\ &= \langle s - \langle s, y\rangle Ey', t^*\rangle \\ &= \langle t - \langle s, y\rangle y', t^*\rangle, \qquad \text{by} \quad (4.1.1) \\ &= \|t\| - \langle s, y\rangle\langle y', t^*\rangle.\end{aligned}$$

Also by (4.1.1), $\langle t, y\rangle = \langle s, y\rangle$ and so

$$\begin{aligned}\langle s, y\rangle\langle y', t^*\rangle &= \langle t, y\rangle\langle y', t^*\rangle \\ &\leqslant \|t\|.\end{aligned}$$

Strict inequality holds unless $y = t^*$, which is excluded since $t^* \notin K$. This proves (4.1.5) and the remark about strict inequality.

4.2. LEMMA. *Let $0 \neq s \in K$, and let $y \in K$ with $\|y\|' = 1$ and $\langle s, y\rangle > 0$. Define t by*

$$t = s + \langle s, y\rangle(I - E)\, y' \qquad (4.2.1)$$

and

$$r = Et^*/\|Et^*\|'. \qquad (4.2.2)$$

We have

(i) *If $t^* \notin K$, then r and y are linearly independent*; *and*

(ii) *r and y are linearly dependent if and only if $r = y$.*

Proof. By Lemma 2.4, $Et^* \neq 0$, so r is well defined. r and y are linearly dependent if and only if $r = \pm y$, since $\|r\|' = 1 = \|y\|'$.

Let us prove (i). If $r = y$, by Theorem 2.5, $t^* \in K$, a contradiction to our hypothesis. So consider the possibility $r = -y$. By Lemma 4.1, $\varphi'(0) > 0$. Inserting $r = -y$ in (4.1.8), we see

$$-\langle s, y\rangle - \langle s, y\rangle\langle y', -y\rangle > 0,$$

i.e., $0 > 0$, a contradiction, completing the proof of (i).

The "if" part of (ii) being clear, we turn to the verification of the "only if" part of (ii). To do this we need to rule out the possibility $r = -y$. Assume for the purpose of arriving at a contradiction that $r = -y$. Since r and y are linearly dependent, by (i), $t^* \in K$. So $Et^* = t^*$ and $r = Et^*/\|Et^*\|' = t^*$, which shows that $t^* = -y$. Now

$$\begin{aligned} \|t\| = \langle t, t^* \rangle = \langle s, t^* \rangle, &\qquad \text{by} \quad (4.2.1) \\ = -\langle s, y \rangle. & \end{aligned}$$

This shows that $\langle s, y \rangle < 0$, a contradiction.

4.3. LEMMA. *The sequence $(\langle s, y_k \rangle)$ generated by the algorithm is strictly increasing.*

Proof. The algorithm will not terminate at the kth iteration if and only if a duality gap is detected in Step 2 and $t_k^* \notin K$, by Step 3. Let

$$\varphi(\alpha) = \langle s, \alpha r_k + (1-\alpha) y_k \rangle / \|\alpha r_k + (1-\alpha) y_k\|',$$

so that

$$\varphi(0) = \langle s, y_k \rangle \qquad \text{and} \qquad \varphi(1) = \langle s, r_k \rangle.$$

Since $t_k^* \notin K$, by Lemma 4.1

$$\|Et_k^*\|' \, \varphi'(0) = \|t_k\| - \langle r_k, y_k \rangle \langle y_k', t_k^* \rangle > 0,$$

which shows that $\varphi'(0) > 0$.

Now if the question in Step 5 is answered affirmatively at the kth iteration, then due to (4.1.4), $\varphi'(1) \geqslant 0$. By Corollary 5.7 of [15], this implies that 1 is a global maximizer of φ on $[0, 1]$. So $\varphi(1) > \varphi(0)$, the strict inequatlity being a consequence of the fact that $\varphi'(0) > 0$. Thus we have shown that $\langle s, r_k \rangle > \langle s, y_k \rangle$, i.e., $\langle s, y_{k+1} \rangle > \langle s, y_k \rangle$.

On the other hand, if the question in Step 5 is answered negatively at the kth iteration, then in view of (4.1.4), this will be due to $\varphi'(1) < 0$. In this case, there exists $\alpha_k \in (0, 1)$ maximizing φ on $[0, 1]$. This α_k is determined in Step 6 of the algorithm via the condition $\varphi'(\alpha_k) = 0$, where (4.1.6) is used. By Step 7 of the algorithm we get

$$\langle s, y_{k+1} \rangle > \langle s, y_k \rangle. \tag{4.3.1}$$

Once more the strict inequality is a consequence of $\varphi'(0) > 0$. Thus we always have (4.3.1), which is the lemma.

4.4. COROLLARY. *The sequence $(\langle s, y_k \rangle)$ converges to a positive limit ρ. Furthermore $\langle s, y \rangle = \rho$ for every cluster point y of the sequence (y_k).*

Proof. The inequality $\langle s, y_k \rangle \leqslant \|s\|$ proves that the sequence $(\langle s, y_k \rangle)$ is bounded and hence $\langle s, y_k \rangle \to \rho$. The corollary is now clear.

4.5. LEMMA. *Let $0 \neq s \in K$ and (y_k) a sequence from K with $\|y_k\|' = 1$, for all k. Suppose there exists $\lambda > 0$ such that $\langle s, y_k \rangle \geqslant \lambda$, for all k. Then the sequence (t_k) given by*

$$t_k = s + \langle s, y_k \rangle (I - E)\, y'_k \tag{4.5.1}$$

is such that the sequence (Et_k^*) is well defined and bounded away from zero.

Proof. As in the proof of Lemma 2.4, one sees that the t_k are non-zero and so the t_k^* are well defined for every k. By Lemma 2.4, $Et_k^* \neq 0$, for all k. If the conclusion of the present lemma were false, we could find a subsequence (k') such that $Et_{k'}^* \to 0$. Since $\|y_{k'}\|' = 1$, for all k', we can pass to another subsequence, again denoted by (k'), such that $y_{k'} \to y$, where $y \in K$ and $\|y\|' = 1$. From (4.5.1) we conclude that $t_{k'} \to t$, where t is given by

$$t = s + \langle s, y \rangle (I - E)\, y', \qquad \langle s, y \rangle > 0. \tag{4.5.2}$$

Equation (4.5.2) implies that $t \neq 0$ and so $t_{k'}^* \to t^*$. Since $Et_{k'}^* \to 0$, we must have $Et^* = 0$. This contradicts Lemma 2.4 in view of (4.5.2).

4.6. THEOREM. *The algorithm generates sequences (y_k) and (t_k), which either terminate at or converge to the maximizer of problem (P′) and the minimal norm $\|\cdot\|$ residual of problem (P), respectively.*

Proof. We emphasize that we are assuming the norm $\|\cdot\|$ to be both smooth and strictly convex. The algorithm terminates at the kth iteration if and only if $t_k^* \in K$. In this case, by Theorem 2.3, t_k is the minimal $\|\cdot\|$ residual of problem (P) and one sees that y_k in Step 9 of the algorithm solves problem (P′).

Consider the situation in which $t_k^* \notin K$, for all k, so we have an infinite sequence (y_k). Let y be any cluster point of (y_k), so that there exists a subsequence $(y_{k'})$ of (y_k) such that $y_{k'} \to y$. We distinguish two cases; they are not mutually exclusive but are jointly exhaustive.

Case 1. Suppose there exists an infinity of indices among the k' for which the question in Step 5 of the algorithm is answered affirmatively. Denote this subsequence once more by (k'). By Lemma 4.5, $\|Et_k^*\|'$ is bounded away from zero, so there exists r, with $\|r\|' = 1$, such that, passing to a further subsequence if necessary, again denoted by (k'), we have $y_{k'+1} = r_{k'} \to r$. This, in particular, implies that r is a cluster point of the sequence (y_k). Writing the inequality in Step 5 of the algorithm for each k' and allowing $k' \to \infty$, we get the inequality

$$\langle s, r\rangle\langle r', y\rangle \geqslant \langle s, y\rangle. \tag{4.6.1}$$

Since r and y are cluster points of (y_k), by Corollary 4.4, $\langle s, r\rangle = \langle s, y\rangle$. This yields the inequality $\langle r', y\rangle \geqslant 1$. But $\|r'\| = 1$ and $\|y\|' = 1$, so $\langle r', y\rangle = 1$, which in turn implies that $y = r$. Writing the expressions in Steps 1 and 4 of the algorithm for each k' and allowing $k' \to \infty$, we find

$$t = s + \langle s, y\rangle(I - E)\, y' \tag{4.6.2}$$

and

$$r = Et^*/\|Et^*\|', \tag{4.6.3}$$

where $t = \lim t_{k'}$. Since we have shown $y = r$, Theorem 2.5, in view of (4.6.2) and (4.6.3), shows that y is a maximizer for problem (P′).

Case 2. Among the chosen subsequence (k'), suppose there exists an infinity of indices, again denoted k', for which the question in Step 5 of the algorithm is answered negatively. By passing to a further subsequence if necessary, denoted once more by k', we can assume that $\alpha_{k'} \to \alpha \in [0, 1]$. Let $r_{k'} \to r$ and $t_{k'} \to t$. This results in (4.6.2) and (4.6.3) holding once more. We want to show that y is a maximizer for problem (P′) and that $y = t^*$. If the vectors r and y are linearly dependent, then by Lemma 4.2(ii) we see that $r = y$, and hence by Theorem 2.5, y is a maximizer for problem (P′) and t a minimal norm $\|\cdot\|$ residual of problem (P).

In view of this, we need to consider only the situation where the vectors r and y are linearly independent. In this case, we claim that $\alpha = 0$ and $y = t^*$ is a maximizer for problem (P′). Let us verify this. Due to the linear independence of r and y

$$\rho = \|\alpha r + (1 - \alpha)\, y\|' > 0. \tag{4.6.4}$$

Writing Steps 6 and 7 of the algorithm for the indices k' and allowing $k' \to \infty$, we get

$$\rho w = \alpha r + (1 - \alpha)\, y \tag{4.6.5}$$

and

$$\langle s, w\rangle\langle w', r - y\rangle = \langle s, r - y\rangle, \tag{4.6.6}$$

where w is the limit of the sequence $(y_{k'+1})$.

If $\alpha > 0$, we derive a contradiction. It follows from (4.6.5) and (4.6.6) that

$$\langle s, w\rangle\langle w', \rho w - y\rangle = \langle s, \rho w - y\rangle,$$
$$\rho\langle s, w\rangle - \langle s, w\rangle\langle w', y\rangle = \rho\langle s, w\rangle - \langle s, y\rangle,$$

i.e.,

$$\langle s, w\rangle\langle w', y\rangle = \langle s, y\rangle. \tag{4.6.7}$$

Since $y_{k'+1} \to w$ and $y_{k'} \to y$, by Corollary 4.4, $\langle s, w\rangle = \langle s, y\rangle$. We have therefore shown that

$$\langle w', y\rangle = 1, \tag{4.6.8}$$

which in turn implies that $y = w$. By (4.6.5) this means that $\rho y = \alpha r + (1-\alpha) y$, which contradicts our current assumption that r and y are linearly independent. We have just seen that assumption $\alpha > 0$ is untenable, so we conclude that $\alpha = 0$.

Since $\alpha = 0$, by (4.6.4) and (4.6.5), $\rho = 1$ and $w = y$. Inserting this in (4.6.6) yields

$$\langle s, y\rangle\langle y', r - y\rangle = \langle s, r - y\rangle, \tag{4.6.9}$$

i.e.,

$$\langle s, y\rangle\langle y', r\rangle = \langle s, r\rangle. \tag{4.6.10}$$

In this case also, t and r are given by (4.6.2) and (4.6.3) above. So by (4.6.2),

$$\begin{aligned} \|t\| &= \langle t, t^*\rangle \\ &= \langle s, t^*\rangle + \langle s, y\rangle\langle y', t^*\rangle - \langle s, y\rangle\langle Ey', t^*\rangle. \end{aligned} \tag{4.6.11}$$

Now

$$\begin{aligned} \langle s, y\rangle\langle Ey', t^*\rangle &= \langle s, y\rangle\langle y', Et^*\rangle \\ &= \|Et^*\|' \langle s, y\rangle\langle y', r\rangle \\ &= \|Et^*\|' \langle s, r\rangle, \qquad \text{by virtue of (4.6.10)} \\ &= \langle s, Et^*\rangle = \langle s, t^*\rangle. \end{aligned}$$

Inserting this in (4.6.11) shows that

$$\|t\| = \langle s, y\rangle\langle y', t^*\rangle. \tag{4.6.12}$$

By a now familiar argument, (4.6.12) tells us that y is a maximizer for problem (P′) and that $y = t^*$.

So we have shown that every cluster point of (y_k) is a maximizer for problem (P′). But maximizers for problem (P′) are unique. Since (y_k) is bounded, in addition, this shows that the entire original sequence (y_k) converges to y, the maximizer of problem (P′). Consequently the entire sequence (t_k) also converges to t, the minimal norm $\|\cdot\|$ residual of problem (P), concluding the proof of the theorem.

5. Numerical Results

Since the chief application of the algorithm is to solving overdetermined systems of linear equations in the l_p sense, $1 < p < \infty$, we discuss briefly our computational experience with four test problems of this type.

The algorithm was programmed in Microsoft BASIC for a Macintosh Plus and was executed using the "decimal" version of the software. This yields up to 14 digits of precision. In all examples, we used

$$1 - \langle s, y_k \rangle / (\min\{\|t_i\| \mid 1 \leqslant i \leqslant k\}) < \varepsilon = 10^{-10},$$

i.e., the relative duality gap being less than 10^{-10}, as our stopping criterion. While the stopping parameters employed yield more correct decimal digits of the l_p norm of the residual than are displayed, it was necessary for the x_k to be determined to the precision given in the tables. Note that $\langle s, y_k \rangle$ is guaranteed by Lemma 4.3 to increase, but $\|t_k\|$ need not be strictly

TABLE I

p	Number of iterations	x_1	x_2	ρ
7	81	1.500985	−0.499662	0.02911148
6	46	1.501486	−0.499711	0.02990692
5	19	1.502293	−0.499813	0.03105973
4	9	1.503757	−0.500057	0.03287394
3	4	1.506860	−0.500710	0.03612070
2	0	1.514762	−0.502571	0.04321596
1.8	3	1.517371	−0.503201	0.04566427
1.6	7	1.519679	−0.503750	0.04879303
1.5	16	1.520005	−0.503800	0.05079019
1.4	11	1.520126	−0.503787	0.05327584
1.3	15	1.520215	−0.503744	0.05643632
1.2	23	1.520187	−0.503621	0.06054325
1.1	65	1.520037	−0.503390	0.06597922

decreasing. Consequently, the relative duality gap, $1 - \langle s, y_k \rangle / \|t_k\|$, may not accurately measure the actual duality gap at the kth step, $\min\{\|t_i\| \mid 1 \leqslant i \leqslant k\} - \max\{\langle s, y_i \rangle \mid 1 \leqslant i \leqslant k\}$. In our computations, we used the actual relative duality gap rather than the one given in Step 2 of the algorithm.

In each problem, the algorithm was started with the l_2 solution as called for in Step 0 of the algorithm. Faster convergence should be expected if the computed solution of one l_p problem were used as the starting point for another l_p problem when the p values were close.

The first example comes from Barrodale and Young [1] and has been used by us [10, 15] before for testing various l_p solving algorithms. The system of equations is

$$x_1 = 1.52$$

$$x_1 + x_2 = 1.025$$

$$x_1 + 2x_2 = 0.475$$

$$x_1 + 3x_2 = 0.01$$

$$x_1 + 4x_2 = -0.475$$

$$x_1 + 5x_2 = -1.005.$$

The result for this example are given in Table I.

The second example, form Cheney [3], has also been used before. This

TABLE II

p	Number of iterations	x_1	x_2	ρ
12	5	2.031705	1.975186	1.1060004
10	4	2.033339	1.970344	1.1341845
8	4	2.035508	1.962903	1.1793407
7	4	2.037010	1.957411	1.2134151
6	4	2.039030	1.949829	1.2609743
5	4	2.041951	1.938681	1.3313999
4	3	2.046597	1.920704	1.4451697
3	3	2.055106	1.887133	1.6565790
2	0	2.074118	1.807843	2.1659214
1.8	2	2.080147	1.780725	2.3631959
1.6	4	2.086123	1.751942	2.6312029
1.5	5	2.088349	1.740083	2.8048045
1.4	7	2.089486	1.732762	3.0168590
1.3	15	2.089512	1.730458	3.2823540
1.25	116	2.089300	1.73037	3.4418323

system poses special difficulties because the solution of the l_1 problem is not unique. The system of equations is

$$x_1 + x_2 = 3$$

$$x_1 - x_2 = 1$$

$$x_1 + 2x_2 = 7$$

$$2x_1 + 4x_2 = 11.1$$

$$2x_1 + x_2 = 6.9$$

$$3x_1 + x_2 = 7.2.$$

The results for this example are given in Table II.

TABLE III

[11, Example 3]			[11, Example 9]			
1	39	144	1	84	46	354
1	47	220	1	73	20	190
1	45	138	1	65	52	405
1	47	145	1	70	30	263
1	65	162	1	76	57	451
1	46	142	1	69	25	302
1	67	170	1	63	28	288
1	42	124	1	72	36	385
1	67	158	1	79	57	402
1	56	154	1	75	44	365
1	64	162	1	27	24	209
1	56	150	1	89	31	290
1	59	140	1	65	52	346
1	34	110	1	57	23	254
1	42	128	1	59	60	395
1	48	130	1	69	48	434
1	45	135	1	60	34	220
1	17	114	1	79	51	374
1	20	116	1	75	50	308
1	19	124	1	82	34	220
1	36	136	1	59	46	311
1	50	142	1	67	23	181
1	39	120	1	85	37	274
1	21	120	1	55	40	303
1	44	160	1	63	30	244
1	53	158				
1	63	144				
1	29	130				
1	25	125				
1	69	175				

In the tables, ρ is the l_p norm of the residual.

The final two exapmles are from Späth [11, Sect. 2.3, Examples 3 and 9]. Example 3 involves 30 equations in 2 unknowns, i.e., $m = 30$ and $n = 2$, and Example 9 involves 25 equations in 3 unknowns, i.e., $m = 25$ and $n = 3$. The data are given in Table III with the first $n - 1$ columns being the columns of the matrix A and the last column being the vector b. Our computed results agree with those in [11], but for completeness we include Table IV which gives those results, including the number of iterations required by Späth's algorithm for the values of p reported in [11]. Since we used different stopping criteria than Späth, in many cases the number of iterations needed simply to replicate Späth's results was less than the number of iterations reported in Table IV.

The data indicate that the algorithm solves the overdetermined system of linear equations in the l_p sense in a small number of iterations for, roughly, $1.3 < p < 4$. In particular, the algorithm converged rapidly for many values of p between 1 and 2.

A number of computational difficulties limited the performance of the algorithm; all seem to be related to the "flatness" of the l_p unit ball for p

TABLE IV

Example 3

p	Späth	New	ρ	x_1	x_2
1.1	21	17	223.1692	97.97677	0.9657306
1.2	7	8	183.0499	98.44222	0.9539866
1.4	9	7	137.2612	98.60850	0.9507508
1.7	4	4	105.8416	98.57653	0.9576799
2.0	1	0	91.61574	98.71472	0.9708704
2.5	6	7	80.92868	100.3369	0.9963743
4	9	21	68.09558	109.0846	1.035749

Example 9

p	Späth	New	ρ	x_1	x_2	x_3
1.1	24	135	622.6518	71.09920	0.6307726	4.931349
1.2	23	25	504.3896	71.69126	0.6075366	4.963526
1.4	12	49	364.0865	72.05922	0.5684263	5.038356
1.7	8	4	260.0769	72.39818	0.5111204	5.155066
2.0	1	0	206.8967	77.98254	0.4173621	5.216591
2.5	6	5	161.2907	86.68675	0.2974348	5.255274
4	10	Did not converge in 300 iterations	114.9529	102.4825	0.1023005	5.283938

large or very near 1. From $1/p + 1/p = 1$, we see that when p is large, $q - 1$ is positive and small, and when p is close to 1, q is large. In either case, the computation of the prime- and the star-dual vectors used in the algorithm calls for raising numbers to exponents that are either very small and positive or very large. Both are computationally unstable.

A second difficulty involved finding α_k in Step 6. Letting $\psi: [0, 1] \to \mathbb{R}$ by

$$\psi(\alpha) = \langle s, \alpha r_k + (1-\alpha) y_k \rangle \langle (\alpha r_k + (1-\alpha) y_k)', r_k - y_k \rangle$$
$$- \|\alpha r_k + (1-\alpha) y_k\|' \langle s, r_k - y_k \rangle,$$

we seek α_k, the root of $\psi(\alpha) = 0$. The root finding method used to find α_k affects how long the algorithm will take to converge, but not significantly the number of iterations of the algorithm are needed for convergence, so we did not expend a great deal of effort optimizing our root finding sub-algorithm. We accepted α as the desired root if either $|\psi(\alpha)| < 10^{-18}$ or the length of the interval on which the root lay was less than 10^{-12}. For p large or close to 1, the accurate determination of α_k became increasingly difficult.

Finally, while the algorithm forces $\langle s, y_k \rangle$ to increase at each interation, we do not know by how much it will increase, nor do we have a guarantee that the norm of the residual, $\|t_k\|_p$, will decrease at each iteration. For p large or near 1, $\langle s, y_{k+1} \rangle - \langle s, y_k \rangle$ was small compared to the duality gap $\|t_k\|_p - \langle s, y_k \rangle$, and $\|t_k\|_p$, while converging to its minimum value, appeared to oscillate between two sequences of slowly decreasing values which differ by roughly one order of magnitude.

ACKNOWLEDGMENTS

The work of the first author was done while on leave at Michigan State University. Thanks are due to a painstaking referee for his helpful comments and remarks.

REFERENCES

1. I. BARRODALE AND A. YOUNG, Algorithms for best L_1 and L_∞ linear approximation on a discrete set, *Numer. Math.* **8** (1966), 295–306.
2. A. BJÖRCK, Solving linear least squares problems by Gram–Schmidt orthogonalization, *BIT* **7** (1967), 1–21.
3. E. W. CHENEY, "Introduction to Approximation Theory," 2nd ed., McGraw–Hill, New York, 1982.
4. V. CHVÁTAL, "Linear Programming," Freeman, NewYork, 1983.
5. R. FLETCHER, J. A. GRANT, AND M. D. HEBDEN, The calculation of linear best L_p approximations, *Comput. J.* **14** (1971), 276–279.
6. G. H. GOLUB AND C. F. VAN LOAN, "Matrix Computations," Johns Hopkins Univ. Press, Baltimore, MD, 1985.

7. S. W. KAHNG, Best L_p approximation, *Math. Comp.* **26** (1972), 505–508.
8. C. L. LAWSON AND R. J. HANSON, "Solving Least Squares Problems," Prentice–Hall, Englewood Cliffs, NJ, 1974.
9. G. MERLE AND H. SPÄTH, Computational experiences with discrete L_p-approximation, *Computing* **12** (1974), 315–321.
10. R. W. OWENS, An algorithm for best approximate solutions of $Ax = b$ with a smooth strictly convex norm, *Numer. Math.* **29** (1977), 83–91.
11. H. SPÄTH, "Mathematical Algorithms for Linear Regression," Academic Press, Orlando, FL, in press.
12. V. P. SREEDHARAN, Solutions of overdetermined linear equations which minimize error in an abstract norm, *Numer. Math.* **13** (1969), 146–151.
13. V. P. SREEDHARAN, Least squares algorithms for finding solutions of overdetermined linear equations which minimize error in an abstract norm, *Numer. Math.* **17** (1971), 387–401.
14. 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. Approx. Theory* **8** (1973), 46–61.
15. V. P. SREEDHARAN, An algorithm for non-negative norm minimal solutions, *Numer. Funct. Anal. Optim.* **9** (1987), 193–232.
16. V. P. SREEDHARAN, A note on the modified Gram–Schmidt process, *Internat. J. Comput. Math.* **24** (1988), 277–290.
17. H. M. WAGNER, Linear programming techniques for regression analysis, *J. Amer. Statist. Assoc.* **54** (1959), 206–212.
18. J. M. WOLFE, On the convergence of an algorithm for discrete L_p approximation, *Numer. Math.* **32** (1979), 439–459.
19. S. I. ZUKHOVITSKIY AND L. I. AYDEYEVA, "Linear and Convex Programming," Saunders, Philadelphia, PA, 1966.