MATHEMATICS OF COMPUTATION, VOLUME 27, NUMBER 123, JULY, 1973

The Calculation of Best Linear One-Sided L_p Approximations

By G. A. Watson

Abstract. The calculation of linear one-sided approximations is considered, using the discrete L_p norm. For $p = 1$ and $p = \infty$, this gives rise to a linear programming problem, and for $1 < p < \infty$, to a convex programming problem. Numerical results are presented, including some applications to the approximate numerical solution of ordinary differential equations, with error bounds.

1. Introduction. Let b be a real m-dimensional column vector, and let A be a real $m \times n$ matrix with $n < m$. Then the problem:

find a vector $\boldsymbol{\alpha} = (\alpha_1, \alpha_2, \cdots, \alpha_n)^T$ to minimize

$$\left(\sum_{i=1}^{m} |r_i|^p\right)^{1/p} \tag{1.1}$$

where

$$\mathbf{r} = \mathbf{b} - A\boldsymbol{\alpha}, \tag{1.2}$$

is a linear discrete L_p approximation problem. The expression (1.1) defines the vector norm $||\mathbf{r}||_p$ for $1 \leqq p < \infty$, with $||\mathbf{r}||_\infty = \max_{1\leqq i\leqq m} |r_i|$.

Such problems arise, for example, as discretizations of the continuous L_p approximation problem: find $\boldsymbol{\alpha}$ to minimize

$$\left[\int_a^b |r(x)|^p \, dx\right]^{1/p} \tag{1.3}$$

where

$$r(x) = b(x) - \sum_{i=1}^{n} \alpha_i \phi_i(x), \qquad a \leqq x \leqq b, \tag{1.4}$$

and $b(x), \phi_i(x) \in C[a, b]$ for $i = 1, 2, \cdots, n$. The expression (1.3) defines $||r(x)||_p$ for $1 \leqq p < \infty$, with $||r(x)||_\infty = \max_{a\leqq x\leqq b} |r(x)|$.

Both the discrete and continuous problems have been considered by a number of authors, and good algorithms are available for obtaining solutions under very general circumstances (see, for example, [1], [2], [8], [14], [23], [24]).

Now suppose that we require that $\boldsymbol{\alpha}$ be constrained to lie in the set R, where

Received June 19, 1972.

AMS (MOS) subject classifications (1970). Primary 65D15, 41A50.

Key words and phrases. One-sided approximation, L_p approximation, linear programming, convex programming.

$$R = \{\alpha : r_i \geqq 0, i = 1, 2, \cdots, m\}. \tag{1.5}$$

In this case, the approximation problem is said to be *one-sided.* The class of problems which we have defined is a subset of the class of all approximation problems in which the difference between the approximation and the approximand is one-signed in the region of interest. Such problems occur frequently in analysis, and examples have been discussed by Bojanic and De Vore [3], De Vore [6] and Marsaglia [12].

For all values of p such that $1 \leqq p \leqq \infty$, the discrete one-sided L_p problem may be conveniently posed and solved (where a solution exists) as a mathematical programming problem. Formally, we may write the problem as:

$$\begin{aligned} &\text{minimize} \quad f(\alpha) = \left(\sum_{i=1}^{m} r_i^p\right)^{1/p} \\ &\text{subject to} \quad \mathbf{r} = \mathbf{b} - A\alpha \geqq 0. \end{aligned} \tag{1.6}$$

The cases $p = 1$ and $p = \infty$ give rise to linear programming problems, and some properties of these problems are considered in Section 2. Section 3 is concerned with the remaining values of p: in this case, problem (1.6) is a convex programming problem. Finally, in Section 4, some numerical results are presented. These include some applications to the approximate numerical solution of ordinary differential equations, with error bounds.

It is clear that a solution to the one-sided L_p problem will exist if and only if R is nonempty, and this will be assumed. Further, the special case of problem (1.6) when the set of equations $\mathbf{b} - A\alpha = 0$ are consistent is of little interest. We will therefore assume in what follows that no solution $\mathbf{r} = 0$ is possible.

2. The Cases $p = 1, \infty$. In this section, it is necessary to make use of some standard linear programming results. These will be quoted without reference, but details may be obtained in, for example, Hadley [10].

It is clear that the one-sided L_1 problem can immediately be posed as the linear programming problem:

$$\begin{aligned} &\text{minimize} \quad \mathbf{e}^T\mathbf{b} - \mathbf{e}^T A\alpha \\ &\text{subject to} \quad A\alpha \leqq \mathbf{b} \end{aligned}$$

where $\mathbf{e}$ is a vector each component of which is 1. This may be solved conveniently by going to the dual problem:

$$\begin{aligned} &\text{maximize} \quad \mathbf{e}^T\mathbf{b} - \mathbf{b}^T\mathbf{w} \\ &\text{subject to} \quad A^T\mathbf{w} = \mathbf{e}^T A, \qquad \mathbf{w} \geqq 0. \end{aligned} \tag{2.1}$$

An algorithm for the continuous one-sided L_1 problem is given by Lewis [11], based on the solution of a sequence of discrete one-sided L_1 problems in this form. Provided that the matrix A has rank n, a solution to (2.1) can readily be obtained by standard techniques, for example, the simplex algorithm. In this case, the solution will be such that at least n values r_i are zero there. This is an immediate consequence of the result that if a variable is in the dual basis, then the corresponding primal constraint holds with equality.

Remark. The simplex procedure can of course be applied to (2.1) when the rank of A is less than n, and will identify the redundant constraints. These may then be removed, leaving a reduced problem which will satisfy the above rank condition for a different value of n. By 'simplex algorithm', we mean here its implementation as an automatic procedure, starting from this point.

As in the standard (two-sided) L_1 problem, the question of uniqueness of the solution vector is not straightforward. Even when the matrix A satisfies the Haar condition, i.e., all $n \times n$ submatrices of A are nonsingular, examples can readily be constructed which have nonunique solutions.

The case $p = \infty$ (the one-sided Chebyshev approximation problem) has been considered by Watson [20]. The notation of that paper, with regard to partitioned vectors and matrices, is adhered to in what follows. It was shown that the one-sided Chebyshev problem may be posed as the linear programming problem:

$$\text{(2.2)} \qquad \begin{aligned} &\text{minimize} \quad \mathbf{e}_{n+1}^T \begin{bmatrix} \alpha \\ h \end{bmatrix} \\ &\text{subject to} \quad \begin{bmatrix} A & \mathbf{e} \\ -A & 0 \end{bmatrix} \begin{bmatrix} \alpha \\ h \end{bmatrix} \geqq \begin{bmatrix} \mathbf{b} \\ -\mathbf{b} \end{bmatrix}, \end{aligned}$$

where $\mathbf{e}_{n+1}$ is a vector each component of which is zero except the $(n + 1)$st which is 1.

Again, it is more efficient to solve the dual problem, which is

$$\text{(2.3)} \qquad \begin{aligned} &\text{maximize} \quad z = [\mathbf{b}^T \; -\mathbf{b}^T]\mathbf{w} \\ &\text{subject to} \quad \begin{bmatrix} A^T & -A^T \\ \mathbf{e}^T & 0 \end{bmatrix} \mathbf{w} = \mathbf{e}_{n+1}, \qquad \mathbf{w} \geqq 0. \end{aligned}$$

We now state some results given in [20] which characterize the linear programming solution of the one-sided Chebyshev problem.

LEMMA 1. *Let the rank of A be n. Then problem* (2.3) *may be solved by the simplex algorithm.*

LEMMA 2. *At such a solution, $n + 1$ equations of the set* (1.2) *hold with $r_i = 0$ or $r_i = \max_i r_i$. The matrix A^σ formed by the corresponding rows of A has rank n.*

LEMMA 3. *We write*

$$\mathbf{b}^\sigma - A^\sigma \alpha = h\mathbf{g},$$

where $g_i = 0$ or 1 and $h = \max_i r_i$. Then there exists a nontrivial vector $\boldsymbol{\lambda}$ such that $\boldsymbol{\lambda}^T A^\sigma = 0$, whose nonzero components satisfy

$$\lambda_i > 0 \quad \text{if } g_i = 1,$$

$$\lambda_i < 0 \quad \text{if } g_i = 0.$$

We use these results to prove

LEMMA 4. *If A satisfies the Haar condition, the solution to the one-sided Chebyshev problem is unique.*

Proof. Let $\alpha = \mathbf{x}$ be a solution found by the simplex method of linear programming, and let $\alpha = \mathbf{y}$ be any other solution. Then there are $n + 1$ equations such that

$$\mathbf{b}^\sigma - A^\sigma \mathbf{x} = h\mathbf{g}$$

where $h = \max_i r_i$, and (re-ordering if necessary)

$$g_i = 1, \qquad i = 1, 2, \cdots, t,$$

$$g_i = 0, \qquad i = t+1, \cdots, n+1.$$

Since $\mathbf{y}$ is also a solution, it follows that

$$\mathbf{b}^\sigma - A^\sigma \mathbf{y} = h\mathbf{v}$$

where $0 \leqq v_i \leqq 1$.

Now, there exists a vector $\boldsymbol{\lambda}$ satisfying the conditions of Lemma 3. Further, by the Haar condition, $\lambda_i \neq 0$, $i = 1, 2, \cdots, n+1$.

Then

$$\boldsymbol{\lambda}^T \mathbf{g} = \boldsymbol{\lambda}^T \mathbf{v}$$

and so

$$\sum_{i=1}^{t} \lambda_i \leqq \sum_{i=1}^{t} \lambda_i v_i \leqq \sum_{i=1}^{t} \lambda_i.$$

Thus $\mathbf{v} = \mathbf{g}$. Since it is a consequence of Lemma 3 that the matrix $[A^\sigma \ \mathbf{g}]$ is nonsingular, it follows that the solution is unique.

3. The Cases $1 < p < \infty$. For $1 < p < \infty$, the problem (1.6) represents the minimization of a strictly convex function subject to linear inequality constraints, and this is a convex programming problem. The theory of such problems is well-developed, and the reader is referred, for example, to Fiacco and McCormick [7]. Before proceeding, we note that the problem (1.6) may be simplified by replacing the objective function by its pth power; clearly, the same values of $\boldsymbol{\alpha}$ solve both problems. Thus, we deal subsequently with the problem:

$$\text{(3.1)} \qquad \text{minimize} \quad \sum_{i=1}^{m} r_i^p$$

$$\text{subject to} \quad \mathbf{r} = \mathbf{b} - A\boldsymbol{\alpha} \geqq 0.$$

An immediate consequence of the strict convexity of the objective function is

LEMMA 5. *The vector $\boldsymbol{\alpha}$ solving problem* (3.1) *is unique.*

The solution may be characterized by making use of the Kuhn-Tucker conditions. A direct application of the theory to this problem gives

LEMMA 6. *Let D_r be a diagonal matrix with (i, i) element r_i^{p-1}. Then the vector $\boldsymbol{\alpha}$ solves problem* (3.1) *if and only if*

(1) $\mathbf{r} \geqq 0$,

(2) *there exists a vector* $\boldsymbol{\lambda} \geqq 0$ *such that*

(i) $\boldsymbol{\lambda}^T \mathbf{r} = 0$,

(ii) $\boldsymbol{\lambda}^T A = p\mathbf{e}^T D_r A$.

There exist a number of methods for minimizing a nonlinear function subject to linear constraints (see, for example, the review paper by Fletcher [9]). Among the best known of these are the reduced gradient method of Wolfe ([21], [22]), and methods based on the gradient projection algorithm of Rosen [17], for example that due to Murtagh and Sargent [13]. In particular, when the objective function is convex, as in this case, it is possible to guarantee convergence.

We mention, finally, some results which are precise analogues of those for two-sided approximation, and which can readily be proved in a similar manner. Let $\boldsymbol{\alpha}^{(p)}$ denote the solution to the one-sided L_p problem for $1 < p < \infty$, and let

$$\mathbf{r}^{(p)} = \mathbf{b} - A\boldsymbol{\alpha}^{(p)}.$$

LEMMA 7. *If* $p < q$, *then* $\|\mathbf{r}^{(p)}\|_p \geqq \|\mathbf{r}^{(q)}\|_q$.

Since the sequence of points $\{\boldsymbol{\alpha}^{(p)}\}$ is bounded, it possesses convergent subsequences, which we may also call $\{\boldsymbol{\alpha}^{(p)}\}$. We have

LEMMA 8. *Let*

$$\lim_{p\to\infty} \boldsymbol{\alpha}^{(p)} = \boldsymbol{\alpha}^*,$$

$$\lim_{p\to 1} \boldsymbol{\alpha}^{(p)} = \bar{\boldsymbol{\alpha}}.$$

Then $\boldsymbol{\alpha}^*$ *and* $\bar{\boldsymbol{\alpha}}$ *respectively solve the one-sided Chebyshev and one-sided* L_1 *problems.*

4. Numerical Results. All the numerical results of this section were obtained on the Elliott 4130 of the University of Dundee, using single length arithmetic. For floating-point computation, this gives about 11 decimal places.

The numerical solution of the convex programming problem of Section 3 was considered using both the reduced gradient method of Wolfe, and Rosen's gradient projection method. Both methods are iterative by nature, and move from point to point within the set R reducing the objective function at each step. They differ in the way in which directions of motion are obtained.

The application of these methods to problem (3.1) as it stands runs into numerical difficulties for large p because of the size of the objective function, and some form of scaling strategy is essential. The procedure adopted for both methods involved the following modifications to the basic algorithms. For full details of these algorithms, see [17] and [22].

1. Choose an $\boldsymbol{\alpha} \in R$, and set k to a prescribed value.
2. Compute

$$\mathbf{r} = \mathbf{b} - A\boldsymbol{\alpha},$$

$$F = \sum_{i=1}^{m} (kr_i)^p,$$

and set $k := k * F^{-1/p}$.

3. Compute $\delta\boldsymbol{\alpha}$, the direction of motion, and test for convergence, as in the basic algorithms.
4. Compute $\mathbf{d} = A\delta\boldsymbol{\alpha}$, and find λ to minimize

$$\sum_{i=1}^{m} (k(r_i - \lambda d_i))^p.$$

5. If $\alpha + \lambda\delta\alpha \in R$, then set $\gamma = \lambda$. Otherwise, choose γ to be the maximum value such that $\alpha + \gamma\delta\alpha \in R$.

Remark. This procedure is closely related to the application of the basic algorithms to the original problem (1.6). The computed value of k at each step gives the inverse of the current value of the norm.

Except for the case $p = 2$ (when an analytical solution is available), the minimum in step 4 was obtained in the following simple manner. A direct search procedure, starting from $\lambda = 0$, was used to obtain three values of λ spanning the minimum. A quadratic was then fitted through these 3 points, and the value of λ giving the minimum of this quadratic accepted as the true value.

The algorithm proved to be effective for problems involving large values of p, provided that reasonable care was taken in deciding when to accept convergence. Essentially, this amounts to deciding when to accept $\delta\alpha = 0$ in step 3. In addition to the obvious criterion

$$||\delta\alpha||_\infty < T,$$

where T is some prescribed tolerance, it was found necessary to take into account the step length γ, and the sensitivity of the solution by also testing to see whether

$$||\alpha_i - \alpha_{i+1}||_\infty < T\ ||\alpha_i||_\infty,$$

where α_i and α_{i+1} are two consecutive values of α, or

$$|1/k_i - 1/k_{i+1}| < T,$$

where k_i and k_{i+1} are two consecutive values of k. (For all the examples of this paper, $T = 10^{-6}$.) This latter test was only invoked if it was found that the convergence was ultimately slow, with the directions of progress confined to lie in the intersection of a fixed set of hyperplanes $r_i = 0$. Both algorithms tended to suffer from this defect for larger problems, with the projected gradient method slightly worse. This reflects the slow ultimate convergence generally associated with steepest descent type methods, and a better algorithm in this case is that of [13].

Our first example is a simple problem which possesses an analytical solution.

Example 1.

$$A = \begin{bmatrix} 1.0 & 0.5 \\ -1.0 & 2.0 \\ 1.0 & -1.0 \end{bmatrix}, \qquad b = \begin{bmatrix} 1.0 \\ 0.0 \\ 0.5 \end{bmatrix}.$$

It is readily verified from Lemma 6 that the problem with this data is solved by

$$\alpha_1^{(p)} = \frac{8(5/3)^{1/(p-1)} + 5}{10(5/3)^{1/(p-1)} + 6},$$

$$\alpha_2^{(p)} = 2 - 2\alpha_1^{(p)},$$

for $1 < p < \infty$.

The solution was first obtained for $p = 2$ using the initial approximation $\alpha = (0, 0)^T$. Solutions were then obtained for an increasing sequence of values of p, using the previous solution as initial approximation. For $p = 10$, the value of the objective

TABLE 1

p	α_1	α_2	$\|\|r\|\|_p$
2	0.808 823	0.382 353	0.085 749
4	0.811 200	0.377 599	0.073 556
10	0.812 060	0.375 880	0.066 752
100	0.812 460	0.375 081	0.062 915
500	0.812 492	0.375 016	0.062 583

Solutions to Example 1

function of (3.1) at the minimum is $O(10^{-12})$, and it is interesting to note that without the scaling introduced above, the solution for $p = 4$ was accepted for $p = 10$ and for all higher values of p.

In Table 1, we give solutions for values of p up to $p = 500$. The value of k was initially set to 1.0 for $p = 2$, and to the current value for other values of p. Note that the final value of k for each p gives the inverse of the minimum value of the norm. The calculated coefficients were in fact correct to one figure in the 8th decimal place. The exact solution for $p = \infty$ is

$$\alpha_1 = 0.8125, \qquad \alpha_2 = 0.375, \qquad ||\mathbf{r}||_\infty = 0.0625.$$

We now present two examples derived by discretizing continuous approximation problems.

Example 2. Let

$$r(x) = x^{1/2} - (\alpha_1 + \alpha_2 x + \alpha_3 x^2), \qquad 1 \leqq x \leqq 2,$$

and define

$$r_i = r(x_i), \qquad i = 1, 2, \cdots, m.$$

TABLE 2

p	α_1	α_2	α_3	$\|\|r\|\|_p$
2	0.444 697	0.626 623	-0.071 320	0.003 113
4	0.445 189	0.625 842	-0.071 031	0.002 047
10	0.445 284	0.625 691	-0.070 975	0.001 659
40	0.445 077	0.626 020	-0.071 097	0.001 537
100	0.444 982	0.626 170	-0.071 153	0.001 523
500	0.444 926	0.626 261	-0.071 186	0.001 516
∞	0.444 911	0.626 283	-0.071 194	0.001 514

Solutions to Example 2

In Table 2, we present the results of computing one-sided approximations on the set of points

$$x_i = 1 + 0.1(i - 1), \qquad i = 1, 2, \cdots, 11.$$

The initial approximation for $p = 2$ was $\boldsymbol{\alpha} = (0, 0, 0)^T$.

Example 3. Let

$$r(x) = b(x) - \sum_{i=1}^{5} \alpha_i x^{i-1}, \qquad 0 \leqq x \leqq 2,$$

where

$$\begin{aligned} b(x) &= e^x, && 0 \leqq x \leqq 1, \\ &= e^{-x} + e - e^{-1}, && 1 \leqq x \leqq 2. \end{aligned}$$

In Table 3, we present the results of computing one-sided approximations on the set of $m = 21$ points

$$x_i = 0.1(i - 1), \qquad i = 1, 2, \cdots, 21.$$

The initial approximation for $p = 2$ was taken to be $\boldsymbol{\alpha} = (0, 0, 0, 0, 0)^T$. (In many cases, however, it may be worthwhile calculating the solution for $p = 1$ (or $p = \infty$) to use as a good initial approximation for any other value of p.)

Remark. Because of the presence of a constant in the approximating functions of Examples 2 and 3, the one-sided Chebyshev approximations are just shifted two-sided Chebyshev approximations (Phillips [15]).

Finally, we present some examples of the use of one-sided approximations to obtain approximate numerical solutions to certain types of ordinary differential equations, with error bounds. The derivation of the error bounds is based on the differential operators involved satisfying certain monotone properties. The use of such properties to obtain approximate solutions with error bounds has been considered, for example, by Collatz [4], [5] and Rosen [18], [19].

We illustrate the principle involved for the problem of finding the function u in some bounded domain D, where u satisfies

$$L[u] = f \quad \text{in } D$$

TABLE 3

p	α_1	α_2	α_3	α_4	α_5	$\|r\|_p$
2	1.000 00	-0.306 09	4.925 44	-4.132 23	0.965 88	0.402 344
4	1.000 00	-0.642 49	5.814 48	-4.835 91	1.137 51	0.270 738
10	1.000 00	-0.985 94	6.714 76	-5.536 86	1.302 90	0.230 744
20	1.000 00	-1.101 63	7.020 71	-5.777 62	1.359 71	0.221 012
40	1.000 00	-1.136 45	7.104 44	-5.831 70	1.367 62	0.216 331
∞	1.000 00	-1.165 57	7.176 62	-5.881 70	1.376 62	0.212 302

Solutions to Example 3

and takes prescribed values on ∂D, the boundary of D. Here, L is a linear differential operator, and we say that this operator is *monotone* if the conditions $L[v] \leqq L[w]$ in D, and $v \geqq w$ on ∂D imply that $v \geqq w$ on $D \cup \partial D$.

For example, consider the ordinary differential equation boundary value problem

$$L[y(x)] = f(x), \qquad a \leqq x \leqq b, \tag{4.1}$$

subject to

$$y(a) = \gamma_1, \qquad y(b) = \gamma_2, \tag{4.2}$$

where L is a linear differential operator. Now suppose $z_1(x)$ is such that

$$\begin{aligned} &L[z_1] \leqq f(x), \qquad a \leqq x \leqq b,\\ &z_1(a) \geqq \gamma_1,\\ &z_1(b) \geqq \gamma_2, \end{aligned}$$

and $z_2(x)$ is such that

$$\begin{aligned} &L[z_2] \geqq f(x), \qquad a \leqq x \leqq b,\\ &z_2(a) \leqq \gamma_1,\\ &z_2(b) \leqq \gamma_2. \end{aligned}$$

Then if L is monotone in the sense described above, it follows that in $[a, b]$

$$z_2(x) \leqq y(x) \leqq z_1(x). \tag{4.3}$$

For example, a suitable L is given by

$$L[y] = y'' + g(x)y' + h(x)y \tag{4.4}$$

where $f(x)$, $g(x)$ and $h(x)$ are bounded in $[a, b]$, and $h(x) \leqq 0$ in $[a, b]$ (Protter and Weinberger [16]).

Our method of solution is as follows. Let

$$\phi(\alpha, x) = \sum_{i=1}^{n} \alpha_i \phi_i(x), \qquad \phi_i(x) \in C^2[a, b], \tag{4.5}$$

be an approximation to $y(x)$ defined by Eqs. (4.1), (4.4) and (4.2), and let x_j, $j = 1, 2, \cdots, m$, be a set of points such that $a = x_1 < x_2 < \cdots < x_m = b$.

Let

$$r(x) = f(x) - L[\phi(\alpha, x)]$$

and assume that there exists at least one α satisfying $\phi(\alpha, a) = \gamma_1$, $\phi(\alpha, b) = \gamma_2$ such that $r(x) \geqq 0$, $a \leqq x \leqq b$. Then a suitable $z_1(x)$, which will be optimum with respect to the form of the approximation $\phi(\alpha, x)$, can be obtained by choosing α to solve the problem

$$\begin{aligned} &\text{minimize} \quad \|r(x)\|_p \qquad (p \geqq 1)\\ &\text{subject to} \quad r(x) \geqq 0, \qquad a \leqq x \leqq b, \end{aligned}$$

and the additional constraints

$$\phi(\alpha, a) = \gamma_1, \qquad \phi(\alpha, b) = \gamma_2. \tag{4.6}$$

Approximations to this optimum $z_1(x)$ can be obtained by discretizing the problem on the set of points x_j, $j = 1, 2, \cdots, m$, and solving the linear discrete one-sided problem defined by

$$r_i = r(x_i), \qquad i = 1, 2, \cdots, m,$$

$$b_i = f(x_i), \qquad i = 1, 2, \cdots, m,$$

and

$$A_{ij} = L[\phi_j(x_i)], \qquad i = 1, 2, \cdots, m; \, j = 1, 2, \cdots, n.$$

The constraints (4.6) may be included as two extra inequality constraints, which may be forced to hold with equality.

A similar approach is used to approximate the optimum $z_2(x)$, the only difference being that we define

$$r(x) = -f(x) + L[\phi(\alpha, x)].$$

A monotone property, similar to the one defined above, can be shown to apply to initial value problems with the operator (4.4), where the initial conditions are $y(a) = \delta_1$, $y'(a) = \delta_2$. Then, if we define $z_1(x)$ such that

$$L[z_1] \geqq f(x), \qquad a \leqq x \leqq b,$$

$$z_1(a) \geqq \delta_1,$$

$$z_1'(a) \geqq \delta_2,$$

and $z_2(x)$ such that

$$L[z_2] \leqq f(x), \qquad a \leqq x \leqq b,$$

$$z_2(a) \leqq \delta_1,$$

$$z_2'(a) \leqq \delta_2,$$

then, under the same conditions on $f(x)$, $g(x)$ and $h(x)$ as before, we have $z_2(x) \leqq y(x) \leqq z_1(x)$ (Protter and Weinberger [16]).

Suppose, therefore, that we obtain approximations $\phi(\beta, x)$ and $\phi(\gamma, x)$ of the form (4.5) to the optimum $z_1(x)$ and $z_2(x)$ respectively. Then, assuming that the monotone property is unaffected by the discretization, we have

$$\phi(\gamma, x_i) \leqq y(x_i) \leqq \phi(\beta, x_i), \qquad i = 1, 2, \cdots, m.$$

Then it is clear that

$$|y(x_i) - \phi(\mathbf{w}, x_i)| \leqq \phi(\mathbf{v}, x_i), \qquad i = 1, 2, \cdots, m, \tag{4.7}$$

where

$$\mathbf{w} = (\beta + \gamma)/2, \qquad \mathbf{v} = (\beta - \gamma)/2.$$

Example 4.

$$y'' + \frac{2x}{1 + x^2} y' - \frac{2}{1 + x^2} y = \frac{4}{1 + x^2}, \qquad 0 \leqq x \leqq 1, \tag{4.8}$$

TABLE 4

x_i	$p = 1$		$p = 2$		$p = 10$	
	$e(x_i)$	$\phi(v, x_i)$	$e(x_i)$	$\phi(v, x_i)$	$e(x_i)$	$\phi(v, x_i)$
0.08	2.4 E-4	1.2 E-3	2.8 E-4	1.3 E-3	2.2 E-4	1.4 E-3
0.16	4.9 E-4	2.0 E-3	4.8 E-4	2.3 E-3	3.6 E-4	2.5 E-3
0.24	5.7 E-4	2.6 E-3	4.6 E-4	3.0 E-3	2.4 E-4	3.3 E-3
0.32	4.6 E-4	3.1 E-3	2.4 E-4	3.5 E-3	8.1 E-5	3.9 E-3
0.40	2.9 E-4	3.3 E-3	3.5 E-5	3.8 E-3	4.4 E-4	4.3 E-3
0.48	2.1 E-4	3.4 E-3	2.0 E-4	3.9 E-3	6.7 E-4	4.4 E-3
0.56	2.9 E-4	3.4 E-3	1.6 E-4	3.8 E-3	6.6 E-4	4.3 E-3
0.64	5.0 E-4	3.3 E-3	5.6 E-5	3.5 E-3	4.1 E-4	3.9 E-3
0.72	7.3 E-4	2.9 E-3	3.4 E-4	3.1 E-3	5.7 E-5	3.4 E-3
0.80	8.1 E-4	2.4 E-3	5.1 E-4	2.5 E-3	2.3 E-4	2.6 E-3
0.88	6.2 E-4	1.6 E-3	4.3 E-4	1.7 E-3	2.9 E-4	1.7 E-3
0.96	2.0 E-4	6.2 E-4	1.4 E-4	6.2 E-4	1.1 E-4	6.2 E-4

Errors and Error Bounds for Example 4:$n = 6$

TABLE 5

x_i	$n = 6$		$n = 9$	
	$e(x_i)$	$\phi(v, x_i)$	$e(x_i)$	$\phi(v, x_i)$
0.08	2.2 E-4	1.4 E-3	5.9 E-7	1.6 E-5
0.16	3.6 E-4	2.5 E-3	1.0 E-6	2.8 E-5
0.24	2.4 E-4	3.3 E-3	1.4 E-6	3.8 E-5
0.32	8.8 E-5	3.9 E-3	6.9 E-7	4.4 E-5
0.40	4.6 E-4	4.3 E-3	2.7 E-6	4.8 E-5
0.48	6.9 E-4	4.4 E-3	2.0 E-6	4.9 E-5
0.56	6.8 E-4	4.3 E-3	3.8 E-7	4.8 E-5
0.64	4.4 E-4	3.9 E-3	1.4 E-6	4.4 E-5
0.72	8.1 E-5	3.4 E-3	2.9 E-7	3.8 E-5
0.80	2.1 E-4	2.6 E-3	2.2 E-6	3.0 E-5
0.88	2.7 E-4	1.7 E-3	1.4 E-6	1.9 E-5
0.96	1.0 E-4	6.2 E-4	6.4 E-8	7.0 E-6

Errors and Error Bounds for Example 4:$p = \infty$

$$y(0) = y(1) = 1. \tag{4.9}$$

This has the exact solution

$$y(x) = 1 - 3\pi x/4 + 3x \tan^{-1} x.$$

Approximations $\phi(\alpha, x) = \sum_{i=1}^{n} \alpha_i x^{i-1}$ were obtained for $y(x)$ in the manner described above for a number of values of p up to $p = 500$. In Tables 4 and 5, we present some actual errors and computed error bounds derived from Eq. (4.7) for approximations on 26 equispaced points in [0, 1] using the L_1, L_2, L_{10} and L_∞ norms. Note that, for $p = 1$ and $p = \infty$, the constraints (4.6) are readily incorporated into the linear programming formulations of these problems. The true error is denoted by

$$e(x_i) = |y(x_i) - \phi(\mathbf{w}, x_i)|, \qquad i = 1, 2, \cdots, m.$$

Example 5.

$$y'' - \frac{2x}{1+x} y' - \frac{1-x}{1+x} y = \frac{2e^x}{1+x},$$

$$y(0) = y'(0) = 1.$$

This has the exact solution

$$y(x) = e^x(x + 1/(1+x)).$$

Polynomial approximations and error bounds were again obtained on 26 equispaced points in [0, 1]. In Tables 6 and 7, we give results similar to those in Tables 4 and 5.

5. Conclusion. We have considered methods for calculating linear discrete one-sided L_p approximations to given data, for all values of p satisfying $1 \leqq p \leqq \infty$. The methods have been illustrated by examples, including applications to the numerical solution of ordinary differential equations, with error bounds. It is clear

TABLE 6

x_i	$p = 1$		$p = 2$		$p = 10$	
	$e(x_i)$	$\phi(v, x_i)$	$e(x_i)$	$\phi(v, x_i)$	$e(x_i)$	$\phi(v, x_i)$
0.04	1.0 E-5	1.5 E-5	8.5 E-6	1.3 E-5	5.1 E-6	9.2 E-6
0.12	2.7 E-5	1.2 E-4	1.9 E-5	1.1 E-4	2.6 E-6	8.2 E-5
0.20	2.5 E-5	3.3 E-4	3.5 E-5	2.9 E-4	7.6 E-5	2.3 E-4
0.28	1.5 E-4	6.2 E-4	1.5 E-4	5.4 E-4	2.0 E-4	4.5 E-4
0.36	3.2 E-4	9.8 E-4	2.9 E-4	8.7 E-4	3.2 E-4	7.5 E-4
0.44	4.9 E-4	1.4 E-3	4.1 E-4	1.3 E-3	4.1 E-4	1.1 E-3
0.52	6.5 E-4	1.9 E-3	5.0 E-4	1.7 E-3	4.4 E-4	1.6 E-3
0.60	7.8 E-4	2.5 E-3	5.5 E-4	2.3 E-3	4.1 E-4	2.2 E-3
0.68	9.2 E-4	3.1 E-3	5.8 E-4	2.9 E-3	3.5 E-4	2.9 E-3
0.76	1.1 E-3	3.8 E-3	6.3 E-4	3.7 E-3	3.1 E-4	3.7 E-3
0.84	1.3 E-3	4.6 E-3	7.4 E-4	4.5 E-3	3.1 E-4	4.6 E-3
0.92	1.6 E-3	5.5 E-3	9.2 E-4	5.5 E-3	3.8 E-4	5.7 E-3
1.0	1.9 E-3	6.6 E-3	1.1 E-3	6.7 E-3	4.9 E-4	6.9 E-3

Errors and Error Bounds for Example 5: $n = 6$

TABLE 7

x_i	$n = 6$		$n = 9$	
	$e(x_i)$	$\phi(v, x_i)$	$e(x_i)$	$\phi(v, x_i)$
0.04	4.7 E-6	8.7 E-6	3.0 E-9	6.9 E-8
0.12	4.9 E-6	7.9 E-5	2.8 E-7	6.2 E-7
0.20	7.9 E-5	2.2 E-4	3.4 E-7	1.7 E-6
0.28	2.0 E-4	4.4 E-4	1.8 E-8	3.5 E-6
0.36	3.2 E-4	7.4 E-4	1.2 E-7	5.8 E-6
0.44	4.0 E-4	1.1 E-3	2.2 E-7	8.8 E-6
0.52	4.1 E-4	1.6 E-3	7.6 E-7	1.3 E-5
0.60	3.6 E-4	2.2 E-3	9.6 E-7	1.7 E-5
0.68	2.9 E-4	2.9 E-3	6.8 E-7	2.3 E-5
0.76	2.2 E-4	3.7 E-3	3.2 E-7	2.9 E-5
0.84	2.0 E-4	4.6 E-3	3.8 E-7	3.6 E-5
0.92	2.5 E-4	5.7 E-3	7.4 E-7	4.5 E-5
1.0	3.3 E-4	7.0 E-3	7.5 E-7	5.5 E-5

Errors and Error Bounds for Example 5: $p = \infty$

that the general approach can be a powerful one, and may be applied to other types of operator equation, for example, to elliptic partial differential equations and to integral equations. The author hopes to consider this in more detail in a further paper.

Department of Mathematics
University of Dundee
Dundee, Scotland

1. I. BARRODALE & A. YOUNG, "Algorithms for best L_1 and L_∞ linear approximations on a discrete set," *Numer. Math.*, v. 8, 1966, pp. 295–306. MR **33** #5096.

2. I. BARRODALE & F. D. K. ROBERTS, *Applications of Mathematical Programming to L_p Approximation*, Dept. of Mathematics Report #32, May 1970; *Nonlinear Programming*, Academic Press, New York, 1970, pp. 447–464.

3. R. BOJANIC & R. DE VORE, "On polynomials of best one-sided approximation," *Enseignment Math.*, (2) v. 12, 1966, pp. 139–164. MR **35** #4647.

4. L. COLLATZ, "Approximation in partial differential equations," in *On Numerical Approximation* (Proc. Sympos., Madison, 1958), Publ. no. 1, Math. Res. Center, U.S. Army, University of Wisconsin, Univ. of Wisconsin Press, Madison, Wis., 1959, pp. 413–422. MR **21** #2361.

5. L. COLLATZ, "Monotonicity and related methods in nonlinear differential equations problems," in *Numerical Solutions of Nonlinear Differential Equations*, Wiley, New York, 1966, pp. 65–87. MR**35** # 1223.

6. R. DE VORE, "One-sided approximation of functions," *J. Approximation Theory*, v. 1, 1968, pp. 11–25. MR**37** #5584.

7. A. V. FIACCO & G. P. MCCORMICK, *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York, 1968. MR **39** #5152.

8. R. Fletcher, J. A. Grant & M. D. Hebden, "The calculation of linear best L_p approximations," *Comput. J.*, v. 14, 1971, pp. 276–279.

9. R. Fletcher, "Minimizing general functions subject to linear constraints," in *Numercial Methods for Non-linear Optimisation,* F. A. Lootsma (Editor), Academic Press, London, New York, 1972, pp. 279–296.

10. G. Hadley, *Linear Programming*, Addison-Wesley Ser. in Indust. Management., Addison-Wesley, Reading, Mass., 1962. MR **24** #B1669.

11. J. T. Lewis, "Computation of best one-sided L_1 approximation," *Math Comp.*, v. 24, 1970, pp. 529–536. MR **42** #8656.

12. G. Marsaglia, "One-sided approximations by linear combinations of functions," in *Approximation Theory*, Academic Press, London, 1970, pp. 233–242. MR **42** #1307.

13. B. A. Murtagh & R. W. H. Sargent, "A constrained minimisation method with quadratic convergence," in *Optimization* (Sympos., Univ. of Keele, Keele, 1968), Academic Press, London, 1969, pp. 215–245, MR **44** #1442.

14. M. R. Osborne & G. A. Watson, "On the best linear Chebyshev approximation," *Comput. J.*, v. 10, 1967, pp. 172–177. MR **36** #1892.

15. D. L. Phillips, "A note on best one-sided approximations," *Comm. ACM,* v. 14, 1971, pp. 598–600.

16. M. H. Protter & H. F. Weinberger, *Maximum Principles in Differential Equations,* Prentice-Hall, Englewood Cliffs, N.J., 1967. MR **36** #2935.

J. B. Rosen, "The gradient projection method for nonlinear programming, I: Linear constraints," *J. Soc. Indust. Appl. Math.*, v. 8, 1960, pp. 181–217. MR **22** #3601.

18. J. B. Rosen, "Approximate computational solution of nonlinear parabolic partial differential equations by linear programming," in *Numerical Solutions of Nonlinear Differential Equations* (Proc. Adv. Sympos., Madison, Wis., 1966), Wiley, New York, 1966, pp. 265–296. MR **34** #7048.

19. J. B. Rosen, "Approximate solution and error bounds for quasi-linear elliptic boundary value problems," *SIAM J. Numer. Anal.,* v. 7, 1970, pp. 80–103. MR **41** #9452.

20. G. A. Watson, "On the best linear one-sided Chebyshev approximation," *J. Approximation Theory*, v. 7, 1973, pp. 48–58.

21. P. Wolfe, "Methods of nonlinear programming," in *Recent Advances in Mathematical Programming*, McGraw-Hill, New York, 1963, pp. 67–86. MR **27** #5617.

22. P. Wolfe, "Methods of nonlinear programming," in *Nonlinear Programming* (NATO Summer School, Menton, 1964), North-Holland, Amsterdam, 1967, pp. 97–131. MR **35** #7697.

23. R. H. Bartels & G. H. Golub, "Stable numerical methods for obtaining the Chebyshev solution to an overdetermined system of equations," *Comm. ACM,* v. 11, 1968, pp. 401–406.

24. I. Barrodale & F. D. K. Roberts, *An Improved Algorithm for Discrete l_1 Linear Approximation*, MRC Technical Summary Report #1172, January 1972.