JOURNAL OF APPROXIMATION THEORY **39**, 374–385 (1983)

The Convergence of the Best Discrete Linear L_p Approximation as $p \to 1$

JÜRGEN FISCHER

Mathematisches Institut A der Universität Stuttgart, Pfaffenwaldring 57, D-7000 Stuttgart 80, West Germany

Communicated by G. Meinardus

Received February 8, 1982

It is well known that the best discrete linear L_p approximation converges to a special best Chebyshev approximation as $p \to \infty$. In this paper it is shown that the corresponding result for the case $p \to 1$ is also true. Furthermore, the special best L_1 approximation obtained as the limit is characterized as the unique solution of a nonlinear programming problem on the set of all L_1 solutions.

1. INTRODUCTION

For a given $m \times n$ matrix A (with $m > n$) and $y \in \mathbb{R}^m$ the discrete linear L_p approximation problem can be stated as minimizing over $\mathbb{R}^n$

$$\|Ax - y\|_p^p = \sum_{j=1}^{m} |a_j'x - y_j|^p, \tag{1.1}$$

where $A' = (a_1, ..., a_m)$, $a_j \in \mathbb{R}^n$ and $y' = (y_1, ..., y_m)$. Under the general assumption rank $A = n$, the above problem has a unique solution $x(p)$ for $1 < p < \infty$. To exclude trivial considerations, we furthermore assume that $y \notin \{Ax | x \subset \mathbb{R}^n\}$.

For the two limiting cases $p = 1$ (L_1 problem) and $p = \infty$ (Chebyshev problem), a vector minimizing (1.1) is in general not unique. In 1963 it was shown by Descloux [3] that $\lim_{p \to \infty} x(p) = x(\infty)$ exists, even if the Chebyshev solution fails to be unique. Moreover, the so-called "strict Chebyshev solution" $x(\infty)$ can be characterized in a certain sense as the "best of the best" Chebyshev approximations (see also p. 239 ff. in [9] for an extensive discussion).

In this paper the corresponding result for $p \to 1$ is derived; the basic idea is to use appropriate dual formulations of the L_p and L_1 problems.

0021-9045/83 $3.00

Furthermore, the special L_1 solution $\lim_{p\to 1} x(p)$ is shown to be the unique solution of an appropriate nonlinear programming problem on the set of all L_1 solutions and thus can be computed numerically.

In the following $p \to 1$ is always used in the sense of $p \to 1+$. A' is the transpose of the matrix A, and $\|\cdot\|_p$ denotes the L_p norm (for $1 \leqslant p \leqslant \infty$) defined in (1.1). For reference we state

LEMMA 1.1. (i) *$x(p)$ is bounded for $1 < p < \infty$.*

(ii) *Every cluster point of $x(p)$, $p \to 1$, is an L_1 solution.*

Proof. For $v \in \mathbb{R}^m$ and $1 \leqslant p \leqslant \infty$ we have

$$\|v\|_\infty \leqslant \|v\|_p \leqslant \|v\|_1. \tag{1.2}$$

Let $r(p) = Ax(p) - y$ denote the vector of residuals, with $r(\infty) = Ax(\infty) - y$. From (1.2) and the optimality of $x(p)$ we obtain that

$$\|r(p)\|_\infty \leqslant \|r(p)\|_p \leqslant \|r(\infty)\|_p \leqslant \|r(\infty)\|_1$$

for $p > 1$, and thus $r(p)$ is bounded. (i) now follows with $x(p) = (A'A)^{-1} A'(r(p) + y)$.

With an optimal L_1 solution $\tilde{x}$ we have $\|r(p)\|_p \leqslant \|A\tilde{x} - y\|_p \leqslant \|A\tilde{x} - y\|_1$ for $p > 1$ and therefore $\limsup_{p\to 1} \|r(p)\|_p \leqslant \|A\tilde{x} - y\|_1$, which implies (ii).

2. DUALITY RELATIONSHIPS

With the new variables $r_j = a_j'x - y_j$, $j = 1,..., m$, the original problem of minimizing (1.1) is transformed into the constrained problem

$$\min_{(x,r)\in \mathbb{R}^{n+m}} \|r\|_p^p \qquad \text{s.t.} \quad Ax - r = y. \tag{2.1}$$

Using the Lagrangian

$$L(x, r; u) = \|r\|_p^p + u'(Ax - r - y) \tag{2.2}$$

the primal problem (2.1) may be written as

$$\inf_{(x,r)} \sup_u L(x, r; u),$$

and its dual problem is then given by

$$\sup_u \inf_{(x,r)} L(x, r; u).$$

In the special case of (2.2) we have

$$\inf_{(x,r)} L(x, r; u) = \inf_{r} \{\|r\|_p^p - u'r\} + \inf_{x} (A'u)' x - u'y,$$

and by elementary calculations we obtain

$$\inf_{r} \{\|r\|_p^p - u'r\} = -(1/q)(1-1/q)^{q-1} \|u\|_q^q,$$

where q is related to p via the equation $1/p + 1/q = 1$ and the inf is attained for

$$r_j = (1 - 1/q)^{q-1} |u_j|^{q-1} \operatorname{sgn}(u_j), \qquad j = 1,..., m. \tag{2.3}$$

Thus we have

$$\begin{aligned}\inf_{(x,r)} L(x, r; u) &= -(1/q)(1-1/q)^{q-1} \|u\|_q^q - y'u \text{ if } A'u = 0,\\ &= -\infty \qquad\qquad \text{otherwise,}\end{aligned}$$

and therefore the dual problem can be finally written as

$$\min_{u \in \mathbb{R}^m} \{(1/q)(1-1/q)^{q-1} \|u\|_q^q + y'u\} \quad \text{s.t. } A'u = 0. \tag{2.4}$$

Since, for $q > 1$, the objective function in (2.4) is strictly convex and tends to $+\infty$ if $\|u\|_q \to \infty$, problem (2.4) has a unique optimal solution $u(q)$ for every $q > 1$.

From standard duality theory (see Chapter 8 in [6]) and Eq. (2.3) we obtain the following relationship between $x(p)$ and $u(q)$ for $1 < p < \infty$:

$$Ax(p) - y = (1 - 1/q)^{q-1} \begin{pmatrix} |u_1(q)|^{q-1} \operatorname{sgn}(u_1(q)) \\ \vdots \\ |u_m(q)|^{q-1} \operatorname{sgn}(u_m(q)) \end{pmatrix}. \tag{2.5}$$

Similarly, for $p = 1$, (1.1) can be formulated as the linear problem

$$\min_{(x,r) \in \mathbb{R}^{n+m}} \sum_{j=1}^{m} r_j \qquad \text{s.t.} \quad -r_j \leqslant a_j'x - y_j \leqslant r_j, j = 1,..., m,$$

or, equivalently,

$$\min_{(x,r)} (0', e') \begin{pmatrix} x \\ r \end{pmatrix} \qquad \text{s.t.} \begin{pmatrix} A & -I \\ -A & -I \end{pmatrix} \begin{pmatrix} x \\ r \end{pmatrix} \leqslant \begin{pmatrix} y \\ -y \end{pmatrix}, \tag{2.6}$$

where $e' = (1,..., 1)$ and I is the $m \times m$ identity matrix. The dual problem is

$$\min_{(u_1,u_2)\in\mathbb{R}^{2m}} (y', -y')\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \quad \text{s.t.} \quad \begin{pmatrix} A' & -A' \\ -I & -I \end{pmatrix}\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} = \begin{pmatrix} 0 \\ -e \end{pmatrix},$$

$$\begin{pmatrix} u_1 \\ u_2 \end{pmatrix} \geqslant 0, \tag{2.7}$$

and with $u = u_1 - u_2$ we obtain the simplified form

$$\min_{u\in\mathbb{R}^m} y'u \qquad \text{s.t.} \quad A'u = 0, -e \leqslant u \leqslant e. \tag{2.8}$$

Here the inequality constraints are equivalent to $\|u\|_\infty \leqslant 1$, and u_1, u_2 can be computed from u by $u_1 = (e+u)/2$, $u_2 = (e-u)/2$.

3. Some Results Concerning $u(q)$

In this section we show that $\lim_{q\to\infty} u(q)$ exists. Let $w_1 > 0$ denote the optimal value of the L_1 problem. Then the optimal value of problem (2.8) is equal to $-w_1 < 0$.

LEMMA 3.1. (i) *Every cluster point $\tilde{u}$ of $u(q)$, $q \to \infty$, satisfies $\|\tilde{u}\|_\infty \leqslant 1$.*

(ii) $\lim_{q\to\infty} y'u(q) = -w_1$.

Proof. (i) Let $\tilde{u}$ be a cluster point of $u(q)$ satisfying $\|\tilde{u}\|_\infty > 1$. Then, for a sequence (q_k) with $\lim_{k\to\infty} q_k = +\infty$ and $\lim_{k\to\infty} u(q_k) = \tilde{u}$, we have

$$\lim_{k\to\infty} \left\{(1/q_k)(1-1/q_k)^{q_k-1} \sum_{j=1}^{m} |u_j(q_k)|^{q_k} + y'u(q_k)\right\} = +\infty,$$

which contradicts the fact that the optimal value of problem (2.4) is always negative.

(ii) Multiplying (2.5) by $u(q)$ we obtain

$$(1-1/q)^{q-1} \|u(q)\|_q^q = -y'u(q),$$

and therefore the optimal value of problem (2.4) is equal to $(1-1/q)\,y'u(q)$. Since every optimal solution $\bar{u}$ of problem (2.8) is feasible for (2.4), we have

$$(1-1/q)\,y'u(q) \leqslant (1/q)(1-1/q)^{q-1} \|\bar{u}\|_q^q + y'\bar{u}$$

and thus

$$\limsup_{q\to\infty} y'u(q) \leqslant y'\bar{u} = -w_1.$$

On the other hand, since every cluster point of $u(q)$ is feasible for problem (2.8), we obtain

$$\liminf_{q \to \infty} y'u(q) \geqslant -w_1.$$

Next we consider the following modification of problem (2.4):

$$\min_{u \in \mathbb{R}^m} \|u\|_q^q \qquad \text{s.t.} \quad A'u = 0, y'u = -w_1. \tag{3.1}$$

LEMMA 3.2. *The unique optimal solution of problem* (3.1) *for* $q > 1$ *is given by*

$$\hat{u}(q) = u(q)/\tau(q), \qquad \text{with} \quad \tau(q) = -y'u(q)/w_1.$$

Proof. The uniqueness follows from the strict convexity of the objective function. Obviously, $\hat{u}(q)$ is feasible for problem (3.1), and with $u(q) = \tau(q)\,\hat{u}(q)$ we obtain from (2.5)

$$Ax(p) - y = (1 - 1/q)^{q-1} (\tau(q))^{q-1} \begin{pmatrix} |\hat{u}_1(q)|^{q-1} \operatorname{sgn}(\hat{u}_1(q)) \\ \vdots \\ |\hat{u}_m(q)|^{q-1} \operatorname{sgn}(\hat{u}_m(q)) \end{pmatrix}.$$

But this equation implies that the gradient of the objective function of (3.1) at $\hat{u}(q)$ is a linear combination of the gradients of the constraints, and thus $\hat{u}(q)$ satisfies the optimality conditions for problem (3.1).

THEOREM 3.3. (i) $\lim_{q \to \infty} u(q) = u(\infty)$ *exists and is an optimal solution of problem* (2.8).

(ii) *If* $J = \{j \mid |u_j(\infty)| = 1\}$ *and* $\varepsilon_j = u_j(\infty)$ *for* $j \in J$, *then the set of all optimal solutions of* (2.8) *is given by*

$$\{u \in \mathbb{R}^m \mid A'u = 0, \|u\|_\infty \leqslant 1, u_j = \varepsilon_j \text{ for } j \in J\}.$$

Proof. Problem (3.1) consists in finding the point of minimal L_q norm on the linear manifold $\{u \mid A'u = 0, y'u = -w_1\}$. The results of Descloux [3] then imply that $\lim_{q \to \infty} \hat{u}(q) = \hat{u}(\infty)$ exists and is equal to the strict Chebyshev solution of the problem

$$\min_{u \in \mathbb{R}^m} \|u\|_\infty \qquad \text{s.t.} \quad A'u = 0, y'u = -w_1. \tag{3.2}$$

Lemma 3.1 implies that $\lim_{q \to \infty} \tau(q) = 1$, and therefore $\lim_{q \to \infty} u(q) = u(\infty)$ exists and is equal to $\hat{u}(\infty)$; furthermore, $u(\infty)$ is an optimal solution of problem (2.8). Since every optimal solution $\bar{u}$ of (2.8) satisfies $\|\bar{u}\|_\infty = 1$, we have $\|u(\infty)\|_\infty = 1$, and the set of optimal solutions of (2.8) coincides with that of problem (3.2). The properties of the strict Chebyshev solution then imply that every optimal solution $\bar{u}$ of (2.8) satisfies $\bar{u}_j = \varepsilon_j$ for $j \in J$. The proof is now completed by applying Lemma A.1 to the pair of problems (2.6), (2.7) and using the connection between (2.7) and (2.8).

4. Existence and Characterization of $\lim_{p \to 1} x(p)$

LEMMA 4.1. (i) *The set of all optimal solutions of the L_1 problem is given by*

$$M_1 = \{x \in \mathbb{R}^n \mid a_j' x = y_j, j \notin J;\ \varepsilon_j a_j' x \geqslant \varepsilon_j y_j, j \in J\},$$

and there is an $\bar{x} \in M_1$ such that $\varepsilon_j a_j' \bar{x} > \varepsilon_j y_j$ for all $j \in J$.

(ii) *The set M_1 is bounded.*

Proof. (i) follows from Lemma A.1 (applied to (2.6) and (2.7)) and Theorem 3.3(ii).

(ii) For $x \in M_1$ we have

$$w_1 = \sum_{j \in J} (\varepsilon_j a_j' x - \varepsilon_j y_j).$$

Since every term in this sum is nonnegative, we obtain

$$\varepsilon_j y_j \leqslant \varepsilon_j a_j' x \leqslant w_1 + \varepsilon_j y_j \qquad \text{for} \quad j \in J,$$

and therefore $\max_{j=1,\ldots,m} |a_j' x|$ is bounded on M_1. The result then follows from rank $A = n$.

Lemma 4.1(i) together with Theorem 6.5 in [10] implies that the relative interior of M_1 is given by

$$\operatorname{ri} M_1 = \{x \in M_1 \mid \varepsilon_j a_j' x > \varepsilon_j y_j \text{ for } j \in J\}.$$

From (2.5) it follows that, for sufficiently large q (and therefore for $1 < p \leqslant \bar{p}$ with a $\bar{p} > 1$), $\varepsilon_j a_j' x(p) - \varepsilon_j y_j > 0$ holds for all $j \in J$. Now, by Lemma 1.1(i) and Lemma 4.1(ii), there is a $c_1 > 0$ such that, for all $j \in J$,

$$0 < \varepsilon_j a_j' x \quad - \varepsilon_j y_j \leqslant c_1 \qquad \text{for all} \quad x \in \operatorname{ri} M_1,$$

and

$$0 < \varepsilon_j a_j' x(p) - \varepsilon_j y_j \leqslant c_1 \qquad \text{for all} \quad 1 < p \leqslant \bar{p}.$$

By Taylor's theorem there is a $c_2 > 0$ such that, for $0 < t \leqslant c_1$ and $1 \leqslant p \leqslant \bar{p}$,

$$|t^p - t - (p-1)\, t \ln t| \leqslant c_2(p-1)^2$$

holds.

Therefore, for $x \in \operatorname{ri} M_1$ we obtain

$$\begin{aligned}\sum_{j=1}^{m} |a_j' x - y_j|^p &= \sum_{j \in J} (\varepsilon_j a_j' x - \varepsilon_j y_j)^p \\ &= \sum_{j \in J} (\varepsilon_j a_j' x - \varepsilon_j y_j) + (p-1) \sum_{j \in J} (\varepsilon_j a_j' x - \varepsilon_j y_j) \\ &\quad \ln (\varepsilon_j a_j' x - \varepsilon_j y_j) + O((p-1)^2), \qquad p \to 1.\end{aligned}$$

Here the first sum is equal to w_1, and the function

$$f(x) = \sum_{j \in J} (\varepsilon_j a_j' x - \varepsilon_j y_j) \ln(\varepsilon_j a_j' x - \varepsilon_j y_j)$$

may be extended continuously onto M_1, if $t \ln t$ is interpreted as 0 for $t = 0$.

LEMMA 4.2. *f is strictly convex on M_1, and the problem*

$$\min_{x \in M_1} f(x) \tag{4.1}$$

has a unique optimal solution $x^* \in \operatorname{ri} M_1$.

Proof. For $x \in \operatorname{ri} M_1$ we have

$$\nabla^2 f(x) = \sum_{j \in J} (\varepsilon_j a_j' x - \varepsilon_j y_j)^{-1} a_j a_j'.$$

Assume that M_1 contains more than one point (otherwise the result is trivial) and let $s \neq 0$ be such that $a_j' s = 0$ for $j \notin J$. Then, since rank $A = n$, there is a $j_0 \in J$ such that $a_{j_0}' s \neq 0$. This implies that the restriction of $\nabla^2 f(x)$ to the subspace orthogonal to all a_j, $j \notin J$, is positive definite for all $x \in \operatorname{ri} M_1$ and thus yields the strict convexity of f.

It remains to show that the minimum of f on M_1 is not attained on the relative boundary $M_1 \backslash \operatorname{ri} M_1$. But this follows from the fact that, for every $x \in M_1 \backslash \operatorname{ri} M_1$ and every $\bar{x} \in \operatorname{ri} M_1$, the directional derivative of f at x in the direction $\bar{x} - x$ is equal to $-\infty$.

Now we are ready to prove

THEOREM 4.3. $\lim_{p \to 1} x(p) = x^*$.

Proof. Since, in general, $x(p) \notin M_1$ for $p > 1$, f is regarded in the following as being defined and continuous on $\{x \in \mathbb{R}^n \mid \varepsilon_j a_j' x \geqslant \varepsilon_j y_j, j \in J\}$. For $j \notin J$ we have $|u_j(\infty)| < 1$, and from (2.5) and $1/(q-1) = p-1$ it follows that

$$|a_j' x(p) - y_j| = (1 - 1/q)^{q-1} |u_j(q)|^{q-1} = O((p-1)^2), \ p \to 1.$$

This yields

$$\begin{aligned} \|Ax(p) - y\|_p^p &= \sum_{j \notin J} |a_j' x(p) - y_j|^p + \sum_{j \in J} (\varepsilon_j a_j' x(p) - \varepsilon_j y_j)^p \\ &= \sum_{j \in J} (\varepsilon_j a_j' x(p) - \varepsilon_j y_j) + (p-1) f(x(p)) \\ &\quad + O((p-1)^2), \qquad p \to 1. \end{aligned}$$

On the other hand,

$$\begin{aligned} \|Ax^* - y\|_p^p = \sum_{j \in J} (\varepsilon_j a_j' x^* - \varepsilon_j y_j) + (p-1) f(x^*) \\ + O((p-1)^2), \qquad p \to 1, \end{aligned}$$

and therefore

$$\begin{aligned} &\|Ax(p) - y\|_p^p - \|Ax^* - y\|_p^p \\ &= \left(\sum_{j \in J} \varepsilon_j a_j \right)' (x(p) - x^*) + (p-1)(f(x(p)) - f(x^*)) \\ &\quad + O((p-1)^2), \qquad p \to 1. \end{aligned}$$

From $A' u(\infty) = 0$ we obtain that

$$\begin{aligned} \left(\sum_{j \in J} \varepsilon_j a_j \right)' (x(p) - x^*) &= \left(-\sum_{j \notin J} u_j(\infty) a_j \right)' (x(p) - x^*) \\ &= -\sum_{j \notin J} u_j(\infty)(a_j' x(p) - y_j) \\ &= O((p-1)^2), \qquad p \to 1, \end{aligned}$$

and thus

$$\begin{aligned} \|Ax(p) - y\|_p^p - \|Ax^* - y\|_p^p = (p-1)(f(x(p)) - f(x^*)) \\ + O((p-1)^2), \qquad p \to 1. \end{aligned} \tag{4.2}$$

From the optimality of $x(p)$ it follows that the left-hand side of (4.2) is non-positive for $p > 1$.

Now let $\tilde{x} \in M_1$ be a cluster point of $x(p)$, $p \to 1$, and (p_k) a sequence with $\lim_{k\to\infty} p_k = 1$ and $\lim_{k\to\infty} x(p_k) = \tilde{x}$. Then $f(\tilde{x}) \geqslant f(x^*)$, and (4.2) implies that $\lim_{k\to\infty} f(x(p_k)) = f(\tilde{x}) = f(x^*)$ and therefore $\tilde{x} = x^*$.

5. Some Remarks and Examples

From Lemma 4.1(i) and Theorem 3.3.(ii) it follows immediately

LEMMA 5.1. (i) *The L_1 problem has a unique optimal solution if and only if* span $\{a_j \mid j \notin J\} = \mathbb{R}^n$.

(ii) *The dual problem* (2.8) *has a unique optimal solution if and only if the set $\{a_j \mid j \notin J\}$ is linearly independent.*

Let $|J|$ denote the cardinality of the set J. Then we obtain

LEMMA 5.2. *Let A satisfy the Haar condition* (*that is, every $n \times n$ submatrix of A is nonsingular*). *Then*

(i) *the L_1 problem has a unique optimal solution if and only if* $|J| \leqslant m - n$,

(ii) *the problem* (2.8) *has a unique optimal solution if and only if* $|J| \geqslant m - n$.

To determine $\lim_{p\to 1} x(p)$ if the optimal L_1 solution is not unique we have to solve problem (4.1); since this is essentially a strictly convex minimization problem under linear equality constraints, it can be solved easily by existing efficient algorithms. But to be able to formulate problem (4.1) we need $u(\infty)$ or, at least, the set J and the ε_j, $j \in J$. $u(\infty)$ can be determined by applying to problem (3.2) an algorithm computing the strict Chebyshev solution (see [4] or [1]). Alternatively, the set J may be identified by solving the L_p problem for a value of p "sufficiently" close to 1; an algorithm for this problem is described in [5].

If A satisfies the Haar condition, then Lemma 5.2 shows that, if the optimal L_1 solution is not unique, $u(\infty)$ is the unique optimal solution of problem (2.8) and thus can be computed simply by solving this linear problem.

Finally, two examples are discussed, both having several L_1 solutions. The first one appears in [2, p. 44]. Here $n = 2$, $m = 6$, and though the Haar condition is violated, $u(\infty) = (1, -1, -1, 0, -1, 1)'$ is unique. The set M_1 of all optimal L_1 solutions can be described by $2x_1 + 4x_2 = 11.1$ and $1.77 \leqslant x_1 \leqslant 2.51666....$ By eliminating one variable problem (4.1) can be reduced to a one-dimensional minimization problem. The solution

$x^* = (2.08802984,\ 1.73098508)'$ differs considerably from the value reported at the end of [8].

The second example is taken from [7] and has

$$A = \begin{pmatrix} 1 & 2 \\ 2 & 3 \\ 3 & 5 \\ 6 & 10 \end{pmatrix}, \qquad y = \begin{pmatrix} 6 \\ 9 \\ 14 \\ 24 \end{pmatrix}.$$

Again $u(\infty) = (-1, -1, -1, 1)'$ is unique, and M_1 is defined by the four inequalities

$$\begin{aligned} -x_1 - \ \ 2x_2 &\geqslant \ -6, \\ -2x_1 - \ \ 3x_2 &\geqslant \ -9, \\ -3x_1 - \ \ 5x_2 &\geqslant -14, \\ 6x_1 + 10x_2 &\geqslant \ \ 24. \end{aligned}$$

The two-dimensional problem (4.1) has the optimal solution

$$x^* = (1.18241272,\ 1.81758728)',$$

which again differs from the value computed in [7].

Appendix

For the linear programming problem

$$\min_{x \in \mathbb{R}^n} c'x \qquad \text{s.t.} \quad a_j' x \leqslant b_j, j = 1,..., m, \tag{A.1}$$

we denote the set of all optimal solutions by M_P and assume that $M_P \neq \varnothing$. Then, with $A' = (a_1,..., a_m)$ and $b = (b_1,..., b_m)'$, the dual problem is given by

$$\min_{u \in \mathbb{R}^m} b'u \qquad \text{s.t.} \quad A'u = -c,\ u \geqslant 0, \tag{A.2}$$

with M_D denoting the set of all optimal solutions.

LEMMA A.1. *There is a set $K \subset \{1,..., m\}$ such that*:

(1) $M_P = \{x \in \mathbb{R}^n \mid a_j' x = b_j, j \in K;\ a_j' x \leqslant b_j, j \notin K\}$,

and there is an $\bar{x} \in M_P$ satisfying $a_j' \bar{x} < b_j$ for all $j \notin K$;

(2) $M_D = \{u \in \mathbb{R}^m | A'u = -c, u \geqslant 0; u_j = 0 \ \textit{for}\ j \notin K\}$,

and there is a $\hat{u} \in M_D$ *satisfying* $\hat{u}_j > 0$ *for all* $j \in K$.

Proof. Define $K = \{j | a_j' x = b_j \text{ for all } x \in M_p\}$.

(1) The definition of K implies that, for every $j \notin K$, there is an $x^{(j)} \in M_p$ such that $a_j' x^{(j)} < b_j$. Let k denote the cardinality of K, and define

$$\bar{x} = (m-k)^{-1} \sum_{j \notin K} x^{(j)}$$

(if $k = m$ choose any $\bar{x} \in M_p$). Then $\bar{x} \in M_p$, and $a_j' \bar{x} < b_j$ holds for every $j \notin K$.

Let $C = \{y \in \mathbb{R}^n | y = \sum_{j \in K} u_j a_j, u_j \geqslant 0\}$ denote the convex cone generated by the vectors a_j, $j \in K$. The optimality conditions for $\bar{x}$ then imply that $-c \in C$. Therefore, the objective function of (A.1) is constant on $\{x | a_j' x = b_j, j \in K\}$, and M_p is given by the formula in part (1) of the lemma.

(2) Let $\bar{u} \in M_D$ be a vector of Lagrange multipliers corresponding to $\bar{x}$; that is,

$$-c = \sum_{j \in K} \bar{u}_j a_j \qquad \text{with} \quad \bar{u}_j \geqslant 0.$$

For an arbitrary $\tilde{u} \in M_D$ let $\tilde{x}$ be an associate optimal solution of (A.1). Then

$$-c = \sum_{j \in K} \bar{u}_j a_j = \sum_{j=1}^{m} \tilde{u}_j a_j,$$

and therefore

$$\sum_{j \notin K} \tilde{u}_j a_j = \sum_{j \in K} (\bar{u}_j - \tilde{u}_j) a_j.$$

This gives

$$\sum_{j \notin K} \tilde{u}_j a_j'(\bar{x} - \tilde{x}) = \sum_{j \in K} (\bar{u}_j - \tilde{u}_j) a_j'(\bar{x} - \tilde{x}) = 0. \tag{A.3}$$

But, for $j \notin K$, $\tilde{u}_j > 0$ implies that

$$a_j'(\bar{x} - \tilde{x}) = a_j' \bar{x} - a_j' \tilde{x} < b_j - a_j' \tilde{x} = 0,$$

and thus (A.3) cannot be true unless $\tilde{u}_j = 0$ holds for all $j \notin K$.

Furthermore, $-c \in \operatorname{ri} C$. Otherwise there is an $s \in \operatorname{span}\{a_j | j \in K\}$, $s \neq 0$, such that $c's = 0$ and $a_j' s \leqslant 0$ for $j \in K$. But then $a_j' s < 0$ for at least one

$j \in K$, and $\bar{x} + \sigma s \in M_p$ holds for sufficiently small $\sigma > 0$, in contradiction to part (1) of the lemma. Now, by Theorem 6.9 in [10] it follows that

$$\operatorname{ri} C = \left\{ y \in \mathbb{R}^n \mid y = \sum_{j \in K} u_j a_j,\ u_j > 0 \text{ for all } j \in K \right\},$$

and this implies the existence of $\hat{u}$.

References

1. N. N. Abdelmalek, Computing the strict Chebyshev solution of overdetermined linear equations, *Math. Comp.* **31** (1977), 974–983.
2. E. W. Cheney, "Introduction to Approximation Theory," McGraw-Hill, New York, 1966.
3. J. Descloux, Approximations in L^p and Chebyshev approximations, *J. SIAM* **11** (1963), 1017–1026.
4. C. S. Duris and M. G. Temple, A finite step algorithm for determining the "strict" Chebyshev solution to $Ax = b$, *SIAM J. Numer. Anal.* **10** (1973), 690–699.
5. J. Fischer, An algorithm for discrete linear L_p approximation, *Numer. Math.* **38** (1981), 129–139.
6. O. L. Mangasarian, "Nonlinear Programming," McGraw–Hill, New York, 1969.
7. G. Merle and H. Späth, Computational experience with discrete L_p-approximation, *Computing* **12** (1974), 315–321.
8. R. W. Owens, An algorithm for best approximate solutions of $Ax = b$ with a smooth strictly convex norm, *Numer. Math.* **29** (1977), 83–91.
9. J. R. Rice, "The Approximation of Functions," Vol. II, Addison–Wesley, Reading, Mass., 1969.
10. R. T. Rockafellar, "Convex Analysis," Princeton Univ. Press, Princeton, N. J., 1970.