JOURNAL OF APPROXIMATION THEORY 17, 107-114 (1976)

Algorithms for Solving the Dual Problem for $Av = b$ with Varying Norms*

MICHAEL S. BOROWSKY

Naval Surface Weapons Center Dahlgren Laboratory, Dahlgren, Virginia

Communicated by E. W. Cheney

Received April 2, 1974

1. INTRODUCTION

Let V be a linear space, X a normed linear space, and $A: V \to X$ a linear transformation from V into X. Also, let $b \in X \sim R(A)$ where $R(A)$ denotes the range of A. The equation $Av = b$ thus has no solution and we shall call the problem of finding a $\hat{v}$ in V such that

$$\| b - A\hat{v} \| \leq \| b - Av \|$$

for all v in V the *primary problem*. A solution of the primary problem shall be called a *best approximate solution* of $Av = b$. Phelps [6] has shown that if X is a Banach space having a predual (i.e., there exists a normed linear space Y such that the dual space of Y, $Y^* = X$) and if $R(A)$ is weak* closed then the primary problem has a solution. In fact (see [3]), if X is *strictly convex*, i.e.,

$$\| \tfrac{1}{2}(x + y) \| < 1 \quad \text{if} \quad \| x \| = \| y \| = 1 \text{ and } x \neq y$$

then the primary problem has at most one solution. The *dual problem*, which always has a solution (see [2, 4]) consists of finding a bounded linear functional $\hat{f}$ in

$$R(A)^{\perp} = \{ f \in X^* : f(x) = 0 \text{ for all } x \in R(A) \}$$

such that $\| \hat{f} \| = 1$ and

$$\hat{f}(b) = \max_{\substack{f \in R(A)^{\perp} \\ \| f \| \leq 1}} f(b).$$

* Much of this work was a Ph.D. dissertation at Drexel University, Philadelphia, Pennsylvania under the direction of Professor Howard Anton. The author also wishes to thank Professors Charles S. Duris and Robert Busby for their valuable assistance.

Since in certain instances the solution of the primary problem follows from the solution of the dual problem we shall investigate algorithms for solving the dual problem. In fact if $(\| \cdot \|_n)$ is a sequence of strictly convex norms defined on a normed linear space X and $\| \cdot \|$ is a norm on X such that $\lim_{n\to\infty} \| a \|_n = \| a \|$ for all a in X, then solutions of the dual problem with respect to $\| \cdot \|$ are generated as limit points of sequences of approximate solutions of the dual problems with respect to $\| \cdot \|_n$.

2. Dual Vectors

If X is a normed linear space and $x \in X$ then for $f \in X^*$ we write $f(x) = (x/f)$. The *dual norm* $\| \cdot \|'$ of a norm $\| \cdot \|$ on X is defined to be the usual norm on X^*; i.e., if $f \in X^*$ then

$$\| f \|' = \sup_{\|x\|=1} |(x/f)|.$$

Furthermore, if $f \in X^* \sim \{0\}$ and $x \in X$, then x is called a $\| \cdot \|$-dual vector for f if $\| x \| = 1$ and

$$(x/f) = \max_{\substack{z \in X \\ \|z\|=1}} |(z/f)| = \| f \|'$$

Let X be a Banach space and let ϕ: $X \to X^{**}$ be the canonical embedding of X in X^{**}, i.e., $\phi(x) = \hat{x}$ where $\hat{x}(f) = f(x)$ for all f in X^*. This mapping enables us to identify X with a subspace of X^{**} and will be used in the following theorem.

Theorem 1. (i) *If X is a Banach space having a predual M, then each $f \in M \sim \{0\}$ has a dual vector in X.*

(ii) *If X is a strictly convex Banach space having a predual M, then each $f \in M \sim \{0\}$ has a unique dual vector in X.*

(iii) *If A: $V \to X$ is a linear transformation from a linear space V into a normed linear space X with strictly convex dual and $b \in X \sim R(A)$ then the dual problem has a unique solution.*

(iv) *Let V be a linear space, X a normed linear space, A: $V \to X$ a linear transformation, and assume $b \in X \sim R(A)$.*

(a) *If the primary problem has a solution and if $\hat{f}$ is a solution of the dual problem, then there exists a dual vector $\hat{f}'$ for $\hat{f}$ such that*

$$Av = b - (b/\hat{f})\hat{f}'$$

is consistent.

(b) *If $\hat{f}$ is a solution of the dual problem and $\hat{f}'$ is any dual vector for $\hat{f}$ such that*

$$Av = b - (b/\hat{f})\hat{f}'$$

is consistent, then every solution is a solution of the primary problem.

Proof. See [1].

3. The Algorithms

THEOREM 2. *Let V be a linear space, X a finite-dimensional linear space with a sequence $(\|\cdot\|_n)$ of strictly convex norms and a norm $\|\cdot\|$. Assume that for each a in X $\lim_{n\to\infty} \| a \|_n = \| a \|$, and let $A: V \to X$ be a 1–1 linear transformation such that $b \in X \sim R(A)$. Then the algorithm below generates a sequence $(\hat{f}_n)$ with the following properties:*

(1) *$(\hat{f}_n)$ has at least one limit point.*

(2) *Every limit point of $(\hat{f}_n)$ is a solution of the $\|\cdot\|$-dual problem.*

(3) *If $\hat{f}_n'$ is the $\|\cdot\|_n$-dual vector to $\hat{f}_n$, then $(b - (b/\hat{f}_n)\hat{f}_n')$ has a $\|\cdot\|$ limit point and every $\|\cdot\|$ limit point is of the form $b - (b/\hat{f})\hat{f}'$ where $\hat{f}$ is a solution of the $\|\cdot\|$-dual problem and $\hat{f}'$ is a $\|\cdot\|$-dual vector to $\hat{f}$. Furthermore $Av = b - (b/\hat{f})\hat{f}'$ has a unique solution which is a best $\|\cdot\|$-approximate solution to $Av = b$.*

Step 0. *Select a fixed basis $B = \{g_1, \ldots, g_k\}$ for $\{R(A) \cup b\}^\perp$ and define $F: X \to \{R(A) \cup b\}^\perp$ by $F(x) = \sum_{j=1}^{k} \overline{(x/g_j)}\, g_j$.*

Step 1. *Set $i = 0$, $n = 1$ and choose $f_0 \in R(A)^\perp$ so that $\| f_0 \|_n' = 1$ and $(b/f_0) > 0$.*

Step 2. *Compute the $\|\cdot\|_n$-dual vector of f_i, call it $f_i'^n$.*

Step 3. *Compute $h_i = F(f_i'^n)$.*

Step 4. *If $\| h_i \|' \leqslant 1/n$ let $\hat{f}_n = f_i$, and to to (7). If not, go to (5).*

Step 5. *Find α_i in C (the complex numbers) such that $\| f_i - \alpha_i h_i \|_n' \leqslant \| f_i - \lambda h_i \|_n'$ for all λ in C.*

Step 6. *Let $f_{i+1} = (f_i - \alpha_i h_i)/(\| f_i - \alpha_i h_i \|_n')$, $i = i + 1$ and go to (2).*

Step 7. *Let $\hat{f}_n' = f_i'^n$.*

Step 8. *Let $i = 0$, $f_i = \hat{f}_n/\| \hat{f}_n \|_{n+1}'$, $n = n + 1$, and go to (2).*

Proof (1). Since $\dim(X^*) < \infty$ there is a constant $k > 0$ such that $k \| \hat{f}_n \|' \leqslant \| \hat{f}_n \|_n' = 1$ for all n sufficiently large. To see that the same k works for all large n we observe that given $0 < \epsilon < 1$ there is an $N > 0$ so that for $n \geqslant N$, $(1 - \epsilon) \| a \|' \leqslant \| a \|_n' \leqslant (1 + \epsilon) \| a \|'$ for all $a \in X^*$ (see [5, p. 104]).

Letting $k = (1 - \epsilon)$ yields the results. Therefore $\|\hat{f}_n\|' \leqslant 1/k$ for all sufficiently large n, and $(\hat{f}_n)$ has a $\|\cdot\|'$ limit point.

Proof (2). Let $\hat{f}$ be a $\|\cdot\|'$ limit point of $\hat{f}_n$. Then there is a subsequence, $(\hat{f}_{n_j})$, such that $\lim_{j\to\infty} \hat{f}_{n_j} = \hat{f}$. We show that $(\hat{f}'_{n_j})$ has a $\|\cdot\|$ limit point and that every limit point is a $\|\cdot\|$-dual vector of $\hat{f}$. Since $\|\hat{f}'_{n_j}\|_{n_j} = 1$ for all j, by going to a subsequence if necessary, there is a z in X such that $z = \lim_{j\to\infty} \hat{f}'_{n_j}$.

Since $(\hat{f}'_{n_j}/\hat{f}_{n_j}) = 1$ it follows that

$$\lim_{j\to\infty} (\hat{f}'_{n_j}/\hat{f}_{n_j}) = (z/\hat{f}) = 1.$$

But

$$\|\hat{f}'\| = \|z\| = 1$$

(see [5]). Hence z is a $\|\cdot\|$-dual vector to $\hat{f}$; let $z = \hat{f}'$.

Since F is continuous and $\|F(\hat{f}'_{n_j})\|' \leqslant 1/n_j$,

$$\lim_{j\to\infty} F(\hat{f}'_{n_j}) = F(\hat{f}') = 0.$$

But $F(\hat{f}') = 0$ implies $\hat{f}$ is a solution of the $\|\cdot\|$-dual problem This follows from the observation that $F(\hat{f}') = 0$ implies $\hat{f}' \epsilon \operatorname{lin} \{R(A) \cup b\}$

Proof (3). Let $B = \operatorname{lin}\{R(A) \cup b\}$. Now $\hat{f}' \in B$, which imp ies that there is an $\alpha \in C$ and an $x_0 \in V$ such that

$$\alpha b - Ax_0 = (b/\hat{f})\hat{f}'.$$

Since $\hat{f} \in R(A)^\perp$,

$$(\alpha b - Ax_0/\hat{f}) = \alpha(b/\hat{f})$$

and

$$\begin{aligned}(\alpha b - Ax_0/\hat{f}) &= (b/\hat{f})(\hat{f}'/\hat{f}) \\ &= (b/\hat{f}).\end{aligned}$$

Therefore $\alpha = 1$ and

$$Ax_0 = b - (b/\hat{f})\hat{f}'.$$

Now for any $x \in V$,

$$\begin{aligned}\|b - Ax\| &= \|b - Ax\|\,\|\hat{f}\|' \\ &\geqslant |(b - Ax/\hat{f})| \\ &= |(b/\hat{f})| \\ &= \|b - Ax_0\|\end{aligned}$$

and x_0 is a best $\|\cdot\|$-approximate solution to $Av = b$.

The fact that $\| h_i \|' \leqslant 1/n$ is eventually satisfied follows from [1, pp. 11–18] (see Appendix).

In the special case where A is an $m \times n$ matrix the previous algorithm can be modified to obtain a sequence whose limit points are solutions of the primary problem.

THEOREM 3. *Let $Ax = b$ be an overdetermined system of m equations in n unknowns with rank $(A) = n$. Let $(\| \cdot \|_n)$ be a sequence of strictly convex norms on C^m, and suppose there is a norm, $\| \cdot \|$, on C^m such that for each a in C^m, $\lim_{n\to\infty} \| a \|_n = \| a \|$. Then the algorithm below generates sequences $(\hat{f}_n)$ and $(\hat{x}_n)$ having the following properties:*

(1) *$(\hat{f}_n)$ has at least one limit point.*

(2) *Every limit point of $(\hat{f}_n)$ is a solution to the dual problem with respect to $\| \cdot \|$.*

(3) *$(\hat{x}_n)$ has at least one limit point.*

(4) *Every limit point of $(\hat{x}_n)$ is a solution to the primary problem with respect to $\| \cdot \|$.*

Steps 0–7. *Same as* Steps 0–7 *in Theorem* 2.

Step 8. *Find the best l^2 approximate solution, $\hat{x}_n$, of $Ax = b - (b/\hat{f}_n)\hat{f}_n'$.*

Step 9. *If $\| b - A\hat{x}_n \| - (1/\| f_n \|)(b/f_n) < \epsilon$ for some predetermined $\epsilon > 0$, take $\hat{x}_n = x^*$, a $\| \cdot \|$-best approximate solution of $Ax = b$, and $f_n/\| f_n \|'$ a solution of the $\| \cdot \|$-dual problem. If not, go to* (10).

Step 10. *Let $i = 0, f_i = \hat{f}_n/\| \hat{f}_n \|'_{n+1}$ $n = n + 1$, and go to* (2).

Proof. The first two results follow from Theorem 2. Now

$$\hat{x}_n = (A^T A)^{-1} A^T (b - (b/\hat{f}_n)\hat{f}_n')$$

and by part (3) of Theorem 2, it follows that by going to a subsequence if necessary $\lim_{n\to\infty} (b - (b/\hat{f}_n)\hat{f}_n') = b - (b/\hat{f})\hat{f}'$ where $\hat{f}$ is a solution of the $\| \cdot \|$-dual problem and $\hat{f}'$ is a $\| \cdot \|$-dual vector to $\hat{f}$. Let x^{**} be the solution to $Ax = b - (b/\hat{f})\hat{f}'$. Then x^{**} is a best $\| \cdot \|$-approximate solution of $Ax = b$.

Let $\alpha_n = b - (b/\hat{f}_n)\hat{f}_n'$ and $\alpha = b - (b/\hat{f})\hat{f}'$. Then,

$$\begin{aligned} \| \hat{x}_n - x^{**} \| &= \|(A^T A)^{-1} A^T \alpha_n - (A^T A)^{-1} A^T \alpha \| \\ &= \|(A^T A)^{-1} A^T (\alpha_n - \alpha)\| \\ &\leqslant \|(A^T A)^{-1} A^T \| \cdot \| \alpha_n - \alpha \|. \end{aligned}$$

Therefore,

$$\lim_{n\to\infty} \| \hat{x}_n - x^{**} \| \leqslant \|(A^T A)^{-1} A^T \| \lim_{n\to\infty} \| \alpha_n - \alpha \| = 0.$$

With additional assumptions, the algorithm in Theorem 2 can be extended to the case where X has infinite dimension. This is the content of the following theorem.

THEOREM 4. *Let V be a linear space, X a Banach space with strictly convex norms $(\| \cdot \|_n)$ and a norm $\| \cdot \|$ such that for each a in X, $\| a \| \leqslant \| a \|_n$ $n = 1, 2,\ldots$ and $\lim_{n\to\infty} \| a \|_n = \| a \|$. Furthermore, suppose X has a separable predual Y and $A\colon V \to X$ is a 1–1 linear transformation such that $R(A)$ is weak* closed, $R(A)^\perp \subset Y$, $\dim(X/R(A)) < \infty$, and $b \in X \sim R(A)$. Then the algorithm of Theorem 2 generates a sequence $(\hat{f}_n)$ with the following properties:*

(1) *$(\hat{f}_n)$ has a $\| \cdot \|'$ limit point.*

(2) *Every $\| \cdot \|'$ limit point of $(\hat{f}_n)$ is a solution to the $\| \cdot \|$-dual problem.*

(3) *If $\hat{f}_n'$ is the $\| \cdot \|_n$-dual vector to $\hat{f}_n$, then $(b - (b/\hat{f}_n)\hat{f}_n')$ has a weak* limit point with respect to $\| \cdot \|$ and every weak* limit point is of the form $b - (b/\hat{f})\hat{f}'$, where $\hat{f}$ is a solution to the $\| \cdot \|$-dual problem and $\hat{f}'$ is a $\| \cdot \|$-dual vector to $\hat{f}$. Furthermore $Av = b - (b/\hat{f})\hat{f}'$ has a unique solution which is a best $\| \cdot \|$-approximate solution to $Av = b$.*

Proof (1). Since $\dim(R(A)^\perp) < \infty$ there is a $k > 0$ so that for sufficiently large n, $\| \hat{f}_n \|' \leqslant k \| \hat{f}_n \|_n' = k$. Therefore there is a subsqeuence $(\hat{f}_{n_i})$ and an $\hat{f} \in R(A)^\perp$ such that $\lim_{i\to\infty} \| \hat{f}_{n_i} - \hat{f} \|' = 0$.

Proof (2). Let $\hat{f}$ be a $\| \cdot \|'$ limit point of $(\hat{f}_n)$. Then by passing to a subsequence if necessary, we have $\lim_{n\to\infty} \| \hat{f}_n - \hat{f} \|' = 0$. Since $\| \hat{f}_n' \| \leqslant \| \hat{f}_n' \|_n = 1$ $n = 1, 2,\ldots$, by passing to a subsequence if necessary there is a $z \in X$, $\| z \| \leqslant 1$, such that $\hat{f}_n' \to^{w^*} z$. We show that z is a $\| \cdot \|$-dual vector to $\hat{f}$. Now $0 \leqslant |(z/\hat{f}) - 1| \leqslant |(z/\hat{f}) - (\hat{f}_n'/\hat{f})| + |(\hat{f}_n'/\hat{f}) - (\hat{f}_n'/\hat{f}_n)| \leqslant |(z/\hat{f}) - (\hat{f}_n'/\hat{f})| + \| \hat{f}_n - \hat{f} \|_n \| \hat{f}_n' \|_n = |((z - \hat{f}_n')/\hat{f})| + \| \hat{f}_n - \hat{f} \|_n'$. Since $\dim(R(A)^\perp) < \infty$ there is a constant $L > 0$ so that $\| \hat{f}_n - \hat{f} \|_n' \leqslant L \| \hat{f}_n - \hat{f} \|'$, for n sufficiently large.

Therefore $0 \leqslant |(z/\hat{f}) - 1| \leqslant |((z - \hat{f}_n')/\hat{f})| + L \| \hat{f}_n - \hat{f} \|'$. It follows that $0 \leqslant |(z/\hat{f}) - 1| \leqslant \lim_{n\to\infty} |((z - \hat{f}_n')/\hat{f})| + L \lim_{n\to\infty} \| \hat{f}_n - \hat{f} \|' = 0$ since $\hat{f}_n' \to^{w^*} z$ and $\hat{f} \in R(A)^\perp \subset Y$. Hence $(z/\hat{f}) = 1$. But $\| \hat{f} \|' = 1$ and $\| z \| \leqslant 1$ implies that $\| z \| = 1$, and therefore z is a $\| \cdot \|$-dual vector, $\hat{f}'$, to $\hat{f}$. Since F is weak* continuous, $\lim_{n\to\infty} F(\hat{f}_n') = F(\hat{f}')$ and $\hat{f}$ is a solution to the $\| \cdot \|$-dual problem.

Proof (3). This follows in a similar manner as the proof of Theorem 2.

Appendix

To show that $\| h_i \|' \leqslant 1/n$ is eventually satisfied in the algorithm we show that for fixed n

$$\lim_{i\to\infty} h_i = \lim_{i\to\infty} F(f_i'^n) = 0.$$

Consider therefore the algorithm in Theorem 2 with n fixed, Steps (7), (8) eliminated, and Step (4) modified to read

Step (4). If $h_i = 0$, stop since f_i is a solution of the dual problem. If $h_i \neq 0$, go to (5).

Assume that the modified algorithm does not terminate and let $\rho_i = (b/f_i)$ and $\rho = (b/f)$ where f is a solution of the dual problem. Since $\rho_i \leqslant \rho$, $i = 1, 2, \ldots$ and

$$\rho_{i+1} - \rho_i = (b/f_{i+1}) - (b/f_i) = (1/\| f_i - \alpha_i h_i \|_n' - 1)(b/f_i) > 0,$$

it follows that (ρ_i) is a monotone sequence bounded above by ρ (observe that if $h_i = F(f_i'^n) \neq 0$, then there is a scalar α_i such that $0 < \| f_i - \alpha_i h_i \|_n' = \min_\lambda \| f_i - \lambda h_i \|_n' < 1$). Thus there is a real number σ such that $\lim_{i\to\infty} \rho_i = \sigma \leqslant \rho$.

Suppose $\sigma < \rho$. Since $\| f_i \|_n' = 1$, $i = 1, 2, \ldots$ there is a subsequence (f_{i_m}) converging to some f in $R(A)^\perp$ which implies $f_{i_m}' \to f'$. Therefore

$$\lim_{m\to\infty} \left\| \sum_{j=1}^{k} (f_{i_m}'/g_j)\, g_j - \sum_{j=1}^{k} (f'/g_j)\, g_j \right\|_n' = \lim_{m\to\infty} \| h_{i_m} - h \|_n' = 0$$

where $h = F(f')$. Now $h \neq 0$, for $h = 0$ implies f is a solution of the dual problem (if $f \in R(A)^\perp$ with $\| f \|_n' = 1$ and $(b/f) > 0$ then f is a solution of the dual problem if and only if $F(f') = 0$) and

$$\sigma = \lim_{m\to\infty} \rho_{i_m} = \lim_{m\to\infty} (b/f_{i_m}) = (b/f) = \rho$$

which is a contradiction. Therefore $\| h \|_n' = d > 0$. Since $h_{i_m} \to h$, it follows that $\| h_{i_m} \|_n' > d/2$ for m sufficiently large and

$$1 > \| f_{i_m} - \alpha_{i_m} h_{i_m} \|_n' > | \alpha_{i_m} | \, \| h_{i_m} \|_n' - 1.$$

Therefore $0 < | \alpha_{i_m} | < 4/d$ for m sufficiently large and by passing to a subsequence if necessary we have that there exists an $\hat{\alpha}$ so that $\lim_{m\to\infty} \alpha_{i_m} = \hat{\alpha}$. Now

$$1 = \lim_{m\to\infty} (\rho_{i_m}/\rho_{i_{m+1}}) = \lim_{m\to\infty} \| f_{i_m} - \alpha_{i_m} h_{i_m} \|_n' \leqslant \lim_{m\to\infty} \| f_{i_m} - \lambda h_{i_m} \|_n'$$

which contradicts the fact that $\min_\lambda \| f - \lambda h \|_n' < 1$. Hence $\sigma = \rho$.

If (f_{i_m}) is a convergent subsequence of (f_i) and if $\hat{f} = \lim_{m\to\infty} f_{i_m}$ then $\|\hat{f}\|_n' = 1$ and $(b/\hat{f}) = \rho$. Therefore $\hat{f}$ is a solution of the dual problem. If the algorithm terminates at the ith step then $F(f_i') = 0$ which implies that f_i is a solution of the dual problem.

REFERENCES

1. M. S. BOROWSKY, Best Approximate Solutions of $Av = b$. Ph.D. dissertation, Drexel University, June 1973.
2. R. C. BUCK, Applications of Duality in Approximation Theory, Approximation of Functions (1965), 27–42.
3. E. W. CHENEY, "Introduction to Approximation Theory," McGraw–Hill, New York, 1966.
4. C. S. DURIS, An Algorithm for Solving Overdetermined Linear Equations in the l^p Sense, Drexel University, Mathematics Report No. 70-10, 1970.
5. B. R. KRIPKE, Best Approximation With Respect to Nearby Norms, *Numer. Math.* **6** (1964), 103–105.
6. R. R. PHELPS, Uniqueness of Hahn-Banach Extensions and Unique Best Approximations, *Trans. Amer. Math. Soc.* **95** (1960), 238–255.