JOURNAL OF APPROXIMATION THEORY 16, 245–250 (1976)

On Minimum Norm and Best Approximate Solutions of $Av = b$ in Normed Spaces

HOWARD ANTON AND C. S. DURIS

Department of Mathematics, Drexel University, Philadelphia, Pennsylvania 19104

Communicated by E. W. Cheney

Received April 30, 1974

INTRODUCTION

Let V and X be real vector spaces, $A: V \to X$ a linear transformation, and b a fixed vector in X. We shall be interested in the operator equation

$$Av = b. \tag{1}$$

If this equation has no solutions, we call it *inconsistent* (INC); otherwise, we call it *consistent* (CON). If (1) is inconsistent and X is normed, then a vector $\hat{v}$ satisfying

$$\| b - A\hat{v} \| \leqslant \| b - Av \|,$$

for all $v \in V$ will be called a *best approximate solution* of $Av = b$. If (1) is consistent and V is normed, then a vector $\hat{v}$ satisfying

$$A\hat{v} = b,$$

and

$$\| \hat{v} \| \leqslant \| v \|$$

for all v satisfying (1) will be called a *minimum norm solution* of $Av = b$. We will call the problem of finding best approximate solutions, problem (INC) and the problem of finding minimum norm solutions, problem (CON).

In addition to eigenvalue problems in integral equations, and classical finite-dimensional least squares problems, inconsistent operator equations arise directly from physical problems such as the integral equation formulation of the interior Neumann problem for the Laplacian on a simply connected region with smooth boundary (see, e.g., [1, 2]). Similarly, problem (CON) arises directly from physical problems; for example, many minimum energy or minimum effort optimal control problems reduce to finding minimum norm solutions of consistent operator equations (see, e.g., [3, 4]). We will show that under appropriate conditions, a (CON) problem can be transformed into an (INC) problem and conversely. This result is of interest because it makes it possible to solve (CON) problems using algorithms for (INC) problems (and conversely).

In [5], we formulated a dual maximization problem whose solution can be used to solve problem (INC) for certain operators. We also gave an algorithm for solving the dual problem under the assumption that A has a finite-codimensional range. In this paper, we formulate a dual maximization problem whose solution can be used to solve problem (CON) for certain operators. We also generalize the algorithm given in [5] so it applies to a wide class of maximization problems, including the dual maximization problem (CON).

1. Preliminaries

We will use the following notation: X^* will denote the dual space of a normed linear space X and $S_X = \{x \in X: \| x \| = 1\}$. If $f \in X^*$ and $x \in X$, then $\langle x, f \rangle$ will denote $f(x)$. If W is a subset of X, then $W^\perp = \{f \in X^*: \langle x, f \rangle = 0$ for all $x \in W\}$. We will call a subspace W of X *proximinal* if each vector in X has a closest vector in W. We will also need the following well-known consequence of the Hahn–Banach theorem (see, e.g., [6, 7]).

THEOREM 1.1. *If W is a subspace of a real normed linear space X, and x is a vector in X, then the following "duality" relation holds*:

$$\inf\{\| x - w \|: w \in W\} = \max\{\langle x, f \rangle : f \in W^\perp, \| f \| = 1\}. \tag{2}$$

We will denote the kernel and range of a linear transformation A, respectively, by $\ker(A)$ and $\mathrm{R}(A)$.

DEFINITION. If $f \in X^* - \{0\}$, then we call $x \in S_X$ a *dual vector* for f if $\langle x, f \rangle = \| f \|$.

The following standard results can be found, e.g., in [8].

THEOREM 1.2. *If X is reflexive and $f \in X^* - \{0\}$, then there is at least one dual vector for f. If, in addition, X is rotund, then the dual vector is unique.*

2. Relationships between Problems (CON) and (INC)

Assume $A: V \to X$ has finite-dimensional kernel with basis $\{v_1, v_2, \dots, v_n\}$ and define $A_1: V^* \to \ker(A)$ by

$$A_1(f) = \sum_{i=1}^{n} \langle v_i, f \rangle v_i .$$

THEOREM 2.1. *If $b \neq 0$ and $\bar{v}$ is any fixed solution of $Av = b$, then v is a minimum norm solution of $Av = b$ if and only if $v = \bar{v} - A_1\hat{v}^*$, where $\hat{v}^*$ is a best approximate solution of $A_1v^* = \bar{v}$.*

Remark. To paraphrase this result, the minimum norm solutions of $Av = b$ are precisely the error vectors for the best approximate solutions of $A_1v^* = \bar{v}$.

Proof of 2.1. For $v^* \in V^*$, the operator equation

$$A_1v^* = \bar{v}$$

is inconsistent; otherwise, $A\bar{v} = AA_1v^* = 0$ contradicts the assumption $0 \neq b$.

If

$$v = \sum_{i=1}^{n} k_i v_i$$

is any vector in $\ker(A)$, then $v = A_1(g)$, where $g \in V^*$ is chosen so that $\langle v_i, g\rangle = k_i$. Thus, $R(A_1) = \ker(A)$. It follows that the solutions of $Av = b$ are precisely the vectors of the form

$$v = \bar{v} - A_1v^*,$$

where $v^* \in V^*$. Thus, v is a minimum norm solution of $Av = b$ if and only if

$$v = \bar{v} - A_1\hat{v}^*,$$

where $\hat{v}^*$ is a best approximate solution of $A_1v^* = \bar{v}$.

Theorem 2.1 shows that under appropriate conditions, a (CON) problem can be transformed into an (INC) problem. We now show that under appropriate conditions, an (INC) problem can be transformed into a (CON) problem.

Assume $A: V \to X$ has a closed finite-codimensional range; let $\{x_1^*, x_2^*, \ldots, x_m^*\}$ be a basis for $R(A)^\perp$ and define $A_0: X \to X^*$ by

$$A_0(x) = \sum_{i=1}^{m} \langle x, x_i^*\rangle\, x_i^*.$$

Further, let $b_0 = A_0 b$.

THEOREM 2.2. *A vector v is a best approximate solution of $Av = b$ if and only if $b - Av$ is a minimum norm solution of $A_0x = b_0$.*

Proof. Since $R(A) = \ker(A_0)$ and since b is a solution of $A_0x = b_0$, it follows that the solutions of $A_0x = b_0$ are the vectors of the form

$$x = b - Av, \qquad v \in V.$$

Thus, v is a best approximate solution of $Av = b$ if and only if $x = b - Av$ is a minimum norm solution of $A_0 x = b_0$.

Remark. It follows from this theorem that a best approximate solution of $Av = b$ can be obtained by solving $A_0 x = b_0$ for a minimum norm solution $\hat{x}$ and then solving $Av = b - \hat{x}$ exactly.

3. A Duality Theorem for Problem (CON)

In this section, we prove a theorem relating the solution of problem (CON) to the solution of a dual maximization problem. This result is analogous to the duality theorem [5, Theorem 1.2] for problem (INC).

THEOREM 3.1. *Let V be a rotund and reflexive Banach space, X a vector space, and $A: V \to X$ a linear transformation whose kernel is a proximinal subspace of V. Let $\bar{v}$ be any solution of $Av = b$. If $\hat{f}$ is any solution of the dual problem*

$$\langle \bar{v}, \hat{f} \rangle = \max\{\langle \bar{v}, f \rangle : f \in \ker(A)^{\perp}, \|f\| = 1\},$$

and if $\hat{f}^*$ is the dual vector for $\hat{f}$, then $\langle \bar{v}, \hat{f} \rangle \hat{f}^*$ is a minimum norm solution of $Av = b$.

Proof. Let

$$\hat{v} = \langle \bar{v}, \hat{f} \rangle \hat{f}^*. \tag{3}$$

Then,

$$\|\hat{v}\| = |\langle \bar{v}, \hat{f} \rangle| \, \|\hat{f}^*\| = \langle \bar{v}, \hat{f} \rangle. \tag{4}$$

If we denote the value in (4) by ρ, it follows from (2) and (4) that

$$\rho = \|\hat{v}\| = \inf\{\|\bar{v} - e\| : e \in \ker(A)\}. \tag{5}$$

If v is any solution of $Av = b$, then $\bar{v} - v \in \ker(A)$ so that from (5)

$$\|v\| = \|\bar{v} - (\bar{v} - v)\| \geqslant \rho = \|\hat{v}\|. \tag{6}$$

Thus, if $\hat{v}$ is a solution of $Av = b$, it must be a minimum norm solution, by (6). To complete the proof, we shall show

$$\bar{v} - \hat{v} \in \ker(A),$$

from which it follows that $\hat{v}$ is a solution of $Av = b$.

Let $g \in \ker(A)^{\perp}$. Then,

$$\langle \bar{v} - \langle \bar{v}, \hat{f} \rangle \hat{f}^*, g \rangle = \langle \bar{v}, g \rangle - \rho \langle \hat{f}^*, g \rangle. \tag{7}$$

Since $\ker(A)$ is proximinal, there exists $\hat{e} \in \ker(A)$ such that

$$\| \bar{v} - \hat{e} \| = \rho.$$

Thus, from (7),

$$\langle \bar{v} - \langle \bar{v}, \hat{f} \rangle \hat{f}^*, g \rangle = \langle \bar{v} - \hat{e}, g \rangle - \rho \langle \hat{f}^*, g \rangle.$$

Letting $w = (1/\rho)(\bar{v} - \hat{e})$, this can be rewritten as

$$\langle \bar{v} - \langle \bar{v}, \hat{f} \rangle \hat{f}^*, g \rangle = \rho \langle w, g \rangle - \rho \langle \hat{f}^*, g \rangle. \tag{8}$$

Taking $g = \hat{f}$ in (8), it follows that

$$0 = \rho \langle w, \hat{f} \rangle - \rho \langle \hat{f}^*, \hat{f} \rangle,$$

and consequently,

$$\langle w, \hat{f} \rangle = \| \hat{f} \|.$$

Since $\| w \| = 1$, it follows from the uniqueness of the dual vector that

$$w = \hat{f}^*.$$

It now follows from (8) that

$$\langle \bar{v} - \langle \bar{v}, \hat{f} \rangle \hat{f}^*, g \rangle = 0$$

for all $g \in \ker(A)^\perp$. Since $\ker(A)$ is proximinal, it is closed, so that

$$\bar{v} - \langle \bar{v}, \hat{f} \rangle \hat{f}^* \in \ker(A);$$

that is

$$\bar{v} - \hat{v} \in \ker(A).$$

4. An Algorithm for Obtaining Maximizing Functionals

In this section, we give a modification of the algorithm stated in [5, Theorem 2.1]. Under appropriate conditions, this modified algorithm can be used to find the maximizing functional $\hat{f}$ in (2).

THEOREM 4.1. *If X and X^* are uniformly rotund and W has finite codimension, then the following algorithm yields the maximizing functional $\hat{f}$ of* (2) *in a finite number of steps or else yields a sequence (f_i) that converges strongly to the maximizing functional.*

Step 1. Select a fixed basis $B = \{w_1, w_2, \ldots, w_n\}$ for $\{W \cup x\}^\perp$ and let $F: X \to \{W \cup x\}^\perp$ be defined by

$$F(z) = \sum_{i=1}^{n} \langle z, w_i \rangle\, w_i .$$

Step 2. Choose $f_0 \in W^\perp$ such that $\|f_0\| = 1$ and $\langle x, f_0 \rangle > 0$.

Step 3. Set $i = 0$.

Step 4. Compute $h_i = F(f_i^*)$.

Step 5. If $h_i = 0$, then stop since f_i is the maximizing functional. If $h_i \neq 0$, then go to Step (6).

Step 6. Determine α_i such that

$$\|f_i - \alpha_i h_i\| \leqslant \|f_i - \lambda h_i\|, \qquad \text{for all } \lambda.$$

Step 7. Set $f_{i+1} = (f_i - \alpha_i h_i)/\|f_i - \alpha_i h_i\|$, increase i by 1 and return to Step 4.

The proof is an obvious modification of the proof in [5]; we omit the details.

Remark. The algorithm in Theorem 4.1 together with Theorem 3.1 can be used to solve problem (CON) directly. However, the algorithm in Theorem 4.1 requires that $\ker(A)^\perp$ have finite dimension. Thus, if A is a Fredholm operator (e.g., $A = \lambda I - U$, where U is compact) then $\ker(A)$ also has finite dimension, (see [8]) which implies that the domain of A is finite dimensional.

References

1. W. J. Kammerer and M. Z. Nashed, Iterative methods for best approximate solutions of linear integral equations of the first and second kinds, *J. Math. Anal. Appl.* **40** (1972), 547–572.
2. P. R. Garabedian, "Partial Differential Equations," Wiley, New York, 1964.
3. W. A. Porter, "Modern Foundations of Systems Engineering," Macmillan, New York, 1966.
4. J. A. Cadzow, Algorithms for minimum effort control problems, *IEEE Trans. Automatic Control*, February (1971).
5. Howard Anton and C. S. Duris, On an algorithm for best approximate solutions to $Av = b$ in normed linear spaces, *J. Approximation Theory*, **8**, (1973), 133–141.
6. R. C. Buck, Applications of duality in approximation theory, *in* "Approximation of Functions," pp. 27–42. Elsevier, Amsterdam, 1965.
7. Ivan Singer, "Best Approximation in Normed Linear Spaces by Elements of Linear Subspaces," Springer-Verlag, Berlin, 1970.
8. S. Lang, "Analysis II," Addison Wesley, Reading, MA. 1969.