JOURNAL OF APPROXIMATION THEORY 51, 11–18 (1987)

The Method of Weighting and Approximation of Restricted Pseudosolutions

G. D. CALLON*

Department of Mathematics, Miami University-Middletown, Middletown, Ohio 45042, U.S.A.

AND

C. W. GROETSCH

Department of Mathematical Sciences, University of Cincinnati, Cincinnati, Ohio 45221, U.S.A.

Communicated by Oved Shisha

Received April 4, 1984; revised August 5, 1985

1. INTRODUCTION

Suppose $K: H_1 \to H_2$ is a bounded linear operator and H_1, H_2 are Hilbert spaces. A vector x is said to be a *least squares solution* of the equation

$$Kx = g \tag{1}$$

if

$$\|Kx - g\| = \inf\{\|Ky - g\| : y \in H_1\}$$

(we denote the inner product and associated norm in each of the Hilbert spaces by $(\cdot, \cdot)$ and $\|\cdot\|$, respectively). A least squares solution of (1) exists if and only if $g \in R(K) + R(K)^\perp$, where $R(K)$ is the range of K. Moreover, the set of least squares solutions is closed and convex and hence contains an element of smallest norm. The operator $K^\dagger: \mathscr{D}(K^\dagger) = R(K) + R(K)^\perp \to H_1$, which associates with g the minimal norm least squares solution $K^\dagger g$ of (1) is called the Moore–Penrose pseudoinverse of K and the set of all least squares solutions is $K^\dagger g + N(K)$, where $N(K)$ is the nullspace of K (see e.g. [4]). Therefore $K^\dagger$ provides a general means of uniquely solving least squares problems associated with (1) and many ways

* The work of this author was supported by the Charles Phelps Taft Foundation.

0021-9045/87 $3.00

of approximating $K^{\dagger}g$ are known [4]. In particular, the so-called Tikhonov approximations converge to $K^{\dagger}g$, i.e.,

$$(K^*K+\alpha I)^{-1}K^*g \to K^{\dagger}g \qquad \text{as} \qquad \alpha \to 0 \ (\alpha > 0),$$

where K^* is the adjoint of K. In the case of matrices the convergence of the Tikhonov approximations was first established by den Broeder and Charnes [1].

In this note we are concerned with a restricted version of (1) and with a particular method of approximating a restricted pseudosolution. Specifically, suppose $L: H_1 \to H_3$ is another bounded linear operator and we wish to satisfy (at least in a generalized sense) the equation

$$Lx = h \tag{2}$$

in addition to (1). In the special case when $h=0$, the problem of finding the minimal norm least squares solution of (1) subject to (2) was studied by Minamide and Nakamura [8] (see also [6, 5]) and they called the generalized solution operator the restricted pseudoinverse.

We are interested in finding a least squares solution of (1) among all least squares solutions of (2). That is, among all least squares solutions of (2), we wish to find a vector x which minimizes $\|Kx-g\|$. We will call this vector x the *restricted pseudosolution* of (1)–(2). To be more specific, we wish to find $x \in L^{\dagger}h + N(L)$ such that

$$\|Kx-g\| = \inf\{\|Ku-g\| : u \in L^{\dagger}h + N(L)\}.$$

For matrices this problem has been studied extensively by Eldén [2, 3], and in the Hilbert space setting by Morozov [9, 10], and Oganesjan and Starostenko [12]. Let $\bar{K} = K \mid N(L)$ be the restriction of K to $N(L)$, then writing

$$x = L^{\dagger}h + z, \qquad z \in N(L),$$

the above conditions are equivalent to

$$\|\bar{K}z - (g - KL^{\dagger}h)\| = \inf\{\|\bar{K}w - (g - KL^{\dagger}h)\| : w \in N(L)\}.$$

In Lemma 2.1 we show that at least one such z exists for any $g \in H_2$, and hence a pseudosolution of the restricted problem exists if and only if $h \in \mathscr{D}(L^{\dagger})$ and is unique if and only if $N(K) \cap N(L) = \{0\}$ (see also [3]). In this case, the unique restricted pseudosolution is given by

$$\begin{aligned} x &= L^{\dagger}h + \bar{K}^{\dagger}(g - KL^{\dagger}h) \\ &= (I - \bar{K}^{\dagger}K)\, L^{\dagger}h + \bar{K}^{\dagger}g \end{aligned} \tag{3}$$

(See also [8, Theorem 3.1].)

We will study a method, which in the finite dimensional case is called the method of weighting [3, 13], for approximating the restricted pseudosolution x. The key to our analysis is a change in the underlying geometry in the space H_1. Once this is done, some well-known facts about Tikhonov's method can be applied in the new structure to give a new proof of the convergence of the method of weighting in Hilbert space. This enables us to clearly see the relationship between the method of weighting and Tikhonov regularization and provides more insight into the structure of restricted pseudosolutions. We also note the correspondence between the restricted pseudoinverse and the concept of L-pseudoinversion [10].

2. Preliminary Results

We assume throughout that $N(K) \cap N(L) = \{0\}$ and that the product transformation $(K, L)\colon H_1 \to H_2 \times H_3$ has closed range. These assumptions are equivalent to Morozov's assumptions that K and L are jointly closed and complementary [10]. Define a new inner product $[\cdot, \cdot]$ on H_1 by

$$[z, w] = (Kz, Kw) + (Lz, Lw),$$

and denote the associated norm by $|\cdot|$ (Locker and Prenter [7] call this the "star" inner product). Then H_1 is a Hilbert space under the inner product $[\cdot, \cdot]$ and we will denote this Hilbert space by $\mathscr{H}_1$. Since (K, L) has closed range we have [4, Theorem 1.2.3]

$$|x|^2 = \|Kx\|^2 + \|Lx\|^2 \geq m \|x\|^2$$

for some $m > 0$ and all $x \in H_1$. Therefore convergence in $\mathscr{H}_1$ implies convergence in H_1. In order to distinguish the underlying space, we will denote the operator L when acting on $\mathscr{H}_1$ by $\mathscr{L}$, i.e., $\mathscr{L}\colon \mathscr{H}_1 \to H_3$ is given by $\mathscr{L}y = Ly$. Note that $\mathscr{L}^\dagger$ and $L^\dagger$ are distinct operators, although they have a common domain.

Our first result relates (3) to $\mathscr{L}^\dagger h$. First we note that $R(\bar{K})$ is closed and hence $\mathscr{D}(\bar{K}^\dagger) = H_2$ [4].

LEMMA 2.1. *$R(\bar{K})$ is closed.*

Proof. Suppose $\bar{K}x_n = y_n \to y \in H_2$. Then $(K, L)\, x_n \to (y, 0) \in H_2 \times H_3$ and since (K, L) has closed range there is an $x \in H_1$ with $(K, L)\, x = (y, 0)$. Therefore $x \in N(L)$ and $y = \bar{K}x$. ∎

THEOREM 2.2. *If $h \in \mathscr{D}(L^\dagger)$, then $\mathscr{L}^\dagger h = (I - \bar{K}^\dagger K)\, L^\dagger h$.*

Proof. Let $w = L^\dagger h - \bar{K}^\dagger K L^\dagger h$. Then (since $R(\bar{K}^\dagger) \subseteq N(L)$), $Lw = LL^\dagger h$ and hence w is a least squares solution of Eq. (2). We must show that $|w|$ is minimal among all least squares solutions of (2), or equivalently that $[w, v] = 0$ for all $v \in N(L)$. If $v \in N(L)$, then if $\bar{Q}$ is the orthogonal projection of H_2 onto $R(\bar{K})$,

$$\begin{aligned}[w, v] &= (Kw, Kv) \\ &= (K(I - \bar{K}^{\dagger}K)\, L^{\dagger}h, \bar{K}v) \\ &= ((I - \bar{Q})\, KL^{\dagger}h, \bar{K}v) = 0\end{aligned}$$

since $R(I - \bar{Q}) = R(\bar{K})^{\perp}$.

Therefore w is the least squares solution of (2) with minimal $|\cdot|$-norm, i.e., $w = \mathscr{L}^{\dagger}h$. ∎

The restricted pseudosolution (3) can now be expressed as

$$x = \mathscr{L}^{\dagger}h + \bar{K}^{\dagger}g. \tag{4}$$

Of course, $\bar{K}^{\dagger}$ is the restricted pseudoinverse operator of Minamide and Nakamura [8]. The following result sheds more light on the vector $\mathscr{L}^{\dagger}h$.

PROPOSITION 2.3. *If $h \in \mathscr{D}(L^{\dagger})$, then $\mathscr{L}^{\dagger}h$ is the unique least squares solution of* (2) *satisfying $\|K\mathscr{L}^{\dagger}h\| \leqslant \|Ky\|$ for any least squares solution y of* (2).

Proof. If y is a least squares solution of (2), then $y = \mathscr{L}^{\dagger}h + \eta$, for some $\eta \in N(\mathscr{L}) = N(L)$. Therefore, since $\mathscr{L}^{\dagger}h$ is the minimal $|\cdot|$-norm least squares solution of (2),

$$\begin{aligned}\|K\mathscr{L}^{\dagger}h\|^2 + \|L\mathscr{L}^{\dagger}h\|^2 &= |\mathscr{L}^{\dagger}h|^2 \\ &\leqslant |y|^2 = \|Ky\|^2 + \|L(\mathscr{L}^{\dagger}h + \eta)\|^2 \\ &= \|Ky\|^2 + \|L\mathscr{L}^{\dagger}h\|^2,\end{aligned}$$

and hence $\|K\mathscr{L}^{\dagger}h\| \leqslant \|Ky\|$, with equality only if $y = \mathscr{L}^{\dagger}h$. ∎

Proposition 2.3 shows that $\mathscr{L}^{\dagger}h$ is Eldén's $L_{IK}^{+}h$ [2] and Oganesjan and Starostenko's x_{h0} [12]. The following characterization of $\bar{K}^{\dagger}g$ is of independent interest and will prove useful in the approximation scheme of the next section.

THEOREM 2.4. *If $g \in \mathscr{D}(K^{\dagger})$, then $\bar{K}^{\dagger}g = \mathscr{P}K^{\dagger}g$, where $\mathscr{P}$ is the orthogonal projector of $\mathscr{H}_1$* onto $N(L)$.

Proof. We emphasize that the projector $\mathscr{P}$ is orthogonal with respect to the $[\cdot, \cdot]$-inner product.

Let $u=K^{\dagger}g$ and $\bar{u}=\bar{K}^{\dagger}g$. Then u and $\bar{u}$ satisfy the respective equations [4]

$$Ku=Qg \qquad \text{and} \qquad K\bar{u}=\bar{Q}g,$$

where Q and $\bar{Q}$ are the orthogonal projections of H_2 onto $\overline{R(K)}$ and $R(\bar{K})$, respectively.

Therefore for any $z\in N(L)$,

$$\begin{aligned}[u-\bar{u}, z] &= (Ku-K\bar{u}, Kz)\\ &= (g, QKz)-(g, \bar{Q}Kz)\\ &= (g, QKz-\bar{Q}Kz)=0\end{aligned}$$

since

$$\bar{Q}Kz=\bar{Q}\bar{K}z=\bar{K}z=Kz=QKz,$$

i.e., $u-\bar{u}$ is $[\cdot, \cdot]$-orthogonal to $N(L)$. Since $\bar{u}\in N(L)$, it follows that

$$\bar{K}^{\dagger}g=\bar{u}=\mathscr{P}\bar{u}=\mathscr{P}u=\mathscr{P}K^{\dagger}g. \quad \blacksquare$$

The next result, which relates the adjoints of the operators L and $\mathscr{L}$, will be used in the sequel.

LEMMA 2.5. $L^*=(K^*K+L^*L)\,\mathscr{L}^*$.

Proof. For $u\in H_3$ and $w\in H_1$, we have

$$\begin{aligned}(L^*u, w) &= (u, Lw)=[\mathscr{L}^*u, w]\\ &= (K\mathscr{L}^*u, Kw)+(L\mathscr{L}^*u, Lw)\\ &= ((K^*K+L^*L)\,\mathscr{L}^*u, w). \quad \blacksquare\end{aligned}$$

3. THE METHOD OF WEIGHTING

The method of weighting for approximating the restricted pseudosolution x of (1)–(2) consists of finding the minimizer x_ε of the functional

$$F(u)=\|Lu-h\|^2+\varepsilon\,\|Ku-g\|^2 \qquad (\varepsilon>0) \tag{5}$$

It follows that the operators $L^*L+\varepsilon K^*K$ and $\varepsilon^{-1}L^*L+K^*K$ are invertible [11] and

$$\begin{aligned}x_\varepsilon &= (L^*L+\varepsilon K^*K)^{-1}L^*h+(\varepsilon^{-1}L^*L+K^*K)^{-1}K^*g\\ &= \qquad\qquad y_\varepsilon \qquad\qquad + \qquad\qquad z_\varepsilon \qquad\qquad .\end{aligned} \tag{6}$$

Morozov calls x the K-pseudosolution of Eq. (2) and has proved that $x_\varepsilon \to x$ as $\varepsilon \to 0$. We offer an alternate proof which illustrates the relationship between the method of weighting and Tikhonov's method.

We have seen in (4) that the restricted pseudosolution may be written $x = \mathscr{L}^\dagger h + \bar{K}^\dagger g$. We know that the Tikhonov approximations $(\mathscr{L}^*\mathscr{L} + \alpha I)^{-1}\mathscr{L}^* h$ converge in $\mathscr{H}_1$ to $\mathscr{L}^\dagger h$ as $\alpha \to 0$. Since the product transformation (K, L) has a trivial nullspace and closed range it follows that these approximations also converge to $\mathscr{L}^\dagger h$ in the Hilbert space H_1 (see e.g. [4, Theorem 1.2.3]). Lemma 2.5 allows us to relate these approximations to the method of weighting:

$$\begin{aligned}(\mathscr{L}^*\mathscr{L} + \alpha I)^{-1}\mathscr{L}^* h &= ((K^*K + L^*L)^{-1}L^*L + \alpha I)^{-1}(K^*K + L^*L)^{-1}L^*h \\ &= (1-\varepsilon)(\varepsilon K^*K + L^*L)^{-1}L^*h \\ &= (1-\varepsilon)\, y_\varepsilon,\end{aligned}$$

where $\varepsilon = \alpha/(1+\alpha)$. Therefore the convergence $y_\varepsilon \to \mathscr{L}^\dagger h$ as $\varepsilon \to 0$ follows from the well known theory for Tikhonov's method.

Concerning the second term in the method of weighting, note that z_ε is the unique minimizer of the functional

$$\|Lz\|^2 + \varepsilon \|Kz - g\|^2.$$

But, if $g \in \mathscr{D}(K^\dagger)$ and Q is the orthogonal projector of H_2 onto $\overline{R(K)}$, then, since $KK^\dagger g = Qg$,

$$\begin{aligned}\|Lz\|^2 + \varepsilon\|Kz - g\|^2 &= \|Lz\|^2 + \varepsilon\|Kz - Qg\|^2 + \varepsilon\|Qg - g\|^2 \\ &= \|Lz\|^2 + \varepsilon\|K(z - K^\dagger g)\|^2 + \varepsilon\|Qg - g\|^2 \\ &= \|L(K^\dagger g - u)\|^2 + \varepsilon\|Ku\|^2 + \varepsilon\|Qg - g\|^2 \\ &= \|Lu - LK^\dagger g\|^2 + \varepsilon\|Ku\|^2 + \varepsilon\|Qg - g\|^2,\end{aligned}$$

where $u = K^\dagger g - z$. Therefore $z_\varepsilon = K^\dagger g - u_\varepsilon$, where u_ε minimizes

$$\|Lu - LK^\dagger g\|^2 + \varepsilon\|Ku\|^2,$$

that is,

$$\begin{aligned}u_\varepsilon &= (L^*L + \varepsilon K^*K)^{-1}L^*LK^\dagger g \\ &= \frac{1}{1-\varepsilon}\left(\mathscr{L}^*\mathscr{L} + \frac{\varepsilon}{1-\varepsilon} I\right)^{-1}\mathscr{L}^*\mathscr{L}K^\dagger g,\end{aligned}$$

by Lemma 2.5. Hence,

$$(1-\varepsilon)\, z_\varepsilon = (1-\varepsilon)\, K^\dagger g - \left(\mathscr{L}^*\mathscr{L} + \frac{\varepsilon}{1-\varepsilon} I\right)^{-1}\mathscr{L}^*\mathscr{L}K^\dagger g. \tag{7}$$

Now, since $\mathscr{L}^{\dagger}\mathscr{L} = I - \mathscr{P}$ (e.g., [4]), we have by the convergence of the Tikhonov approximations

$$\left(\mathscr{L}^*\mathscr{L} + \frac{\varepsilon}{1-\varepsilon} I\right)^{-1} \mathscr{L}^*\mathscr{L}K^{\dagger}g \to (I - \mathscr{P})\, K^{\dagger}g$$

as $\varepsilon \to 0$. Therefore, by (7) and Theorem 2.4,

$$\begin{aligned} z_{\varepsilon} &\to K^{\dagger}g - (I - \mathscr{P})\, K^{\dagger}g \\ &= \mathscr{P}K^{\dagger}g = \bar{K}^{\dagger}g \end{aligned}$$

as $\varepsilon \to 0$.

Finally, we note that in fact $z_{\varepsilon} \to \bar{K}^{\dagger}g$ as $\varepsilon \to 0$ for *any* $g \in H_2$. Indeed, if $g \in H_2$ and $\tilde{g} \in \mathscr{D}(K^{\dagger})$, then

$$\begin{aligned} \|z_{\varepsilon} - \bar{K}^{\dagger}g\| \leqslant{} & \|(\varepsilon^{-1}L^*L + K^*K)^{-1}K^*(g - \tilde{g})\| \\ & + \|(\varepsilon^{-1}L^*L + K^*K)^{-1}K^*\tilde{g} - \bar{K}^{\dagger}\tilde{g}\| + \|\bar{K}^{\dagger}\tilde{g} - \bar{K}^{\dagger}g\|. \end{aligned}$$

We have shown that the middle term on the right converges to zero as $\varepsilon \to 0$. Since $\mathscr{D}(K^{\dagger})$ is dense in H_2, $\bar{K}^{\dagger}$ is bounded, and $(\varepsilon^{-1}L^*L + K^*K)^{-1}$ is uniformly bounded, it follows that $z_{\varepsilon} \to \bar{K}^{\dagger}g$ as $\varepsilon \to 0$.

By using well-known results on Tikhonov's method, we have therefore established the convergence of the method of weighting in Hilbert space. (The original proof given by Morozov [9, Theorem 3, p. 7] uses entirely different techniques). We summarize the main result in the following theorem.

THEOREM 3.1. *Suppose K: $H_1 \to H_2$ and L:$H_1 \to H_3$ are bounded linear operators and the product operator (K, L): $H_1 \to H_2 \times H_3$ has trivial nullspace and closed range. If $h \in \mathscr{D}(L^{+})$ and $g \in H_2$, then the minimizer x_{ε} of* (5) *converges as $\varepsilon \to 0$ to the restricted pseudosolution of* (1)–(2).

The ideas developed by Locker and Prenter [7] can be used to establish convergence of the method of weighting in the case in which K is an unbounded densely defined closed linear operator with closed range such that

$$\|Lx\| \geqslant \beta\|x\| \qquad \text{for} \quad x \in \| N(K) \qquad (\beta > 0).$$

In this case Lemma 2.1 and Theorem 2.4 above remain valid. Locker and Prenter show that $(\mathscr{D}(K), [\cdot, \cdot])$ is a Hilbert space, which we may denote $\mathscr{H}_1$. If we understand by L the operator L restricted to $\mathscr{D}(K)$ (with the usual inner product) and denote by $\mathscr{L}$ the operator L acting on $\mathscr{H}_1$, then

Theorem 2.2 and the decomposition (4) hold. It then follows [7, Lemma 4.1] that

$$\mathscr{L}^* = (K^*K + L^*L)^{-1} L^*$$

and the convergence of the method for $h \in \mathscr{D}(L^\dagger)$, $g \in \mathscr{D}(K^\dagger)$ can be established as above.

REFERENCES

1. G. G. DEN BROEDER AND A. CHARNES, "Contributions to the Theory of Generalized Inverses for Matrices," Purdue University, Lafayette, Indiana, 1957.
2. L. ELDÉN, "A note on weighted pseudoinverses with applications to the regularization of Fredholm integral equations of the first kind," Department of Mathematics, Linköping University, Report LiTH-MAT-R-1975-11, 1975.
3. L. ELDÉN, A weighted pseudoinverse, generalized singular values, and constrained least squares problems, *BIT* **22** (1982), 487–502.
4. C. W. GROETSCH, "Generalized Inverses of Linear Operators: Representation and Approximation," Dekker, New York, 1977.
5. J. HARTUNG, A note on restricted pseudoinverses, *SIAM J. Math. Anal.* **10** (1979), 266–273.
6. R. B. HOLMES, "A Course in Optimization and Best Approximation," Lecture Notes in Mathematics, Vol. 257, Springer-Verlag, Berlin/Heidelberg/New York, 1972.
7. J. LOCKER AND P. M. PRENTER, Regularization with differential operators I: general theory, *J. Math. Anal. Appl.* **74** (1980), 504–529.
8. N. MINAMIDE AND K. NAKAMURA, A restricted pseudoinverse and its application to constrained minima, *SIAM J. Appl. Math.* **19** (1970), 167–177.
9. V. A. MOROZOV, "Methods for Solving Incorrectly Posed Problems," Springer-Verlag, New York, 1984.
10. V. A. MOROZOV, *L*-pseudoinversion and its properties, *Soviet Math. Dokl.* **18** (1977), No. 2, 335–339.
11. M. Z. NASHED, Approximate regularized solutions to improperly posed linear integral and operator equations, *in* "Constructive and Computational Methods for Differential and Integral Equations" (D. Colton and R. P. Gilbert, Eds.), Lecture Notes in Mathematics, Vol. 430, Springer-Verlag, Berlin/Heidelberg/New York, 1974.
12. S. M. OGANESJAN AND V. I. STAROSTENKO, Some properties of an *L*-pseudosolution and of an element minimizing a generalized Tikhonov functional, *Soviet Math. Dokl.* **25** (1982), No. 2, 383–385.
13. C. VANLOAN, A generalized SVD analysis of some weighting methods for equality constrained least squares, *in* "Matrix Pencils" (B. Kagstrom and O. Ruhe, Eds.), Lecture Notes in Mathematics, Vol. 973, Springer-Verlag, Berlin/Heidelberg/New York, 1983.