JOURNAL OF APPROXIMATION THEORY 33, 214–223 (1981)

# Optimal Approximation and the Method of Least Squares

F. J. DELVOS AND WALTER SCHEMPP

*Lehrstuhl für Mathematik I, University of Siegen, Hölderlinstrasse 3, D-5900 Siegen 21, Federal Republic of Germany*

*Communicated by Oved Shisha*

Received September 11, 1980

IN MEMORY OF PROFESSOR ARTHUR SARD

## 1. INTRODUCTION

It is well known that the roots of the theory of generalized inverses (see, e.g., Bjerhammar [3], Albert [1]) are in the calculus of observations (or "theory of errors"). For instance, the variational characterization of the generalized inverse relies on the (direct) method of least squares (cf. Groetsch [8]). On the other hand, the generalized inverse of linear operators is closely related to the functional analytic theory of splines (Laurent [11], Delvos [5]). Aside from the direct method there exists a more intrinsic (and less well-known) approach to the theory of least squares that may be traced back to C. F. Gauss [6] (also see Grossmann [9]). In the recent literature this approach is known as the Gauss–Markov theorem (Albert [1], Beutler and Root [2]) and Neyman–David theorem (Linnik [12]). In the present paper we apply the theory of generalized inverses of linear operators to show that the minimal norm least-squares solution admits a characterization as the unique Čebyšev center (Holmes [10]) of a certain bounded set of "admissible" solutions of the given linear operator equation.

This geometric interpretation is revealed to be closely related to the notion of optimal approximation in the functional analytic theory of splines (Sard [15]). Our goal is to display these connections which we have not seen elsewhere in the literature.

## 2. GENERALIZED INVERSES AND LEAST-SQUARES SOLUTIONS

In this section the fundamentals of generalized inverses for continuous linear operators with closed range in Hilbert spaces and their applications to the method of least squares are summarized.



0021-9045/81/110214-10$02.00/0

Let $H_1, H_2$ denote two complex Hilbert spaces and let $T \in B(H_1, H_2)$ be a continuous linear operator of $H_1$ into $H_2$ such that the range $R(T)$ of $T$ is a closed vector subspace of $H_2$. Then we have (cf. Groetsch [8])

$$R(T) = N(T^*)^{\perp}, \tag{1}$$

where the right hand side denotes the orthogonal complement (in $H_2$) of the kernel of the adjoint $T^*: H_2 \to H_1$ of $T$. It should be observed that the continuous linear operator $T^* \in B(H_2, H_1)$ has also a closed range and satisfies the condition

$$R(T^*) = N(T)^{\perp}. \tag{2}$$

Linear operators with closed range are called normally solvable (Petryshyn [13]). Any normally solvable continuous linear operator $T \in B(H_1, H_2)$ admits a generalized inverse

$$T^+ \in B(H_2, H_1). \tag{3}$$

The operator $T^+$ is uniquely determined by the four Penrose equations that read as follows:

$$TT^+T = T, \tag{$4_1$}$$

$$T^+TT^+ = T^+, \tag{$4_2$}$$

$$(TT^+)^* = TT^+, \tag{$4_3$}$$

$$(T^+T)^* = T^+T. \tag{$4_4$}$$

For this and some other equivalent definitions of the generalized inverse, the reader is referred to the monograph by Groetsch [8].

Let us now describe the least-squares solution of the linear operator equation

$$Tx = b \qquad (b \in H_2) \tag{5}$$

in terms of the generalized inverse $T^+$ of $T$.

A vector $x \in H_1$ is called a least-squares solution of the equation (5) if and if it satisfies the condition

$$\|Tx - b\| \leqslant \|Tu - b\| \qquad (u \in H_1). \tag{6}$$

A least-squares solution $x_0 \in H_1$ of (5) is of minimal norm if the conditions

$$\|Tx_0 - b\| = \|Tx - b\| \qquad (x \in H_1, x \neq x_0) \tag{7}$$

imply the strict inequality

$$\|x_0\| < \|x\|. \tag{8}$$

The following result is well-known. For the proof see, for instance, Holmes [10].

THEOREM 1. *Let $T \in B(H_1, H_2)$ denote a continuous linear normally solvable operator. For any $b \in H_2$, the element of $H_1$ given according to*

$$x_0 = T^+ b \tag{9}$$

*represents the unique least-squares solution of the linear operator equation* (5) *of minimal norm.*

## 3. Least-Squares Solutions as Čebyšev Centers

Aside from the Hilbert spaces $H_1, H_2$ let $H_3$ denote a third complex Hilbert space. Let the continuous linear operator $F \in B(H_1, H_3)$ be fixed.

Consider the linear operator equation

$$Tx + \delta b = b \qquad (x \in H_1), \tag{10}$$

where $x \in H_1$ is supposed to denote the "exact" solution and $\delta b \in H_2$ is the error of the observation "measured by $T$." The aim is to approximate the "true value" $Fx \in H_3$ in terms of the measured observation $Gb$, where $G \in B(H_2, H_3)$ belongs to a certain set $C(r, F)$ of operators that will be specified now.

To this end, suppose that $Gb \in H_3$ is an unbiased approximation of $Fx$, i.e., that $\delta b = 0$ implies $Gb = Fx$. Then we have by (10)

$$F = GT. \tag{11}$$

Moreover, let $V \in B(H_2)$ denote a fixed non-negative selfadjoint operator on $H_2$, i.e., a continuous operator of the space $H_2$ into itself such that

$$V = V^*, \qquad V \geqslant 0. \tag{12}$$

Suppose that there exists a real number $r > 0$ such that

$$C(r, F) = \{G \in B(H_2, H_3) \mid F = GT, \operatorname{Tr}(GVG^*) \leqslant r^2\} \neq \varnothing. \tag{13}$$

In (13), Tr denotes the trace, the main properties of which may be found in Sard [14] and Schatten [17]. Now we are in a position to define the set $A(r, F) \subseteq H_3$ of admissible approximations of $Fx$ according to

$$A(r, F) = \{Gb \mid G \in C(r, F)\}. \tag{14}$$

We shall suppose that $A(r,F)$ forms a bounded subset of $H_3$. This hypothesis is quite natural. Denoting by $P_{R(T)}$ the orthogonal projector on the (closed) vector subspace $R(T)$ of $H_2$, the additional hypothesis

$$P_{R(T)}V = VP_{R(T)} \tag{15}$$

will be made. The invariance property (15) admits a statistical interpretation: If $V$ is considered as a variance operator of the observation error $\delta b$ (Sard [14, 16]) then (15) says that the "error" in $R(T)^{\perp}$ is uncorrelated with the error in $R(T)$ (see Beutler and Root [2]). For the rôle that (15) plays in the functional analytic spline theory, see also Delvos [5].

For any element $z \in H_3$ the error incurred with respect to the (bounded) set (14) is given by

$$E(z) = \sup_{w \in A(r,F)} \|z - w\| = \sup_{G \in C(r,F)} \|z - Gb\|. \tag{16}$$

A Čebyšev center (cf. Holmes [10]) of $A(r,F)$ is an element $z_0 \in H_3$ which best represents the set $A(r,F)$, i.e., an element $z_0 \in H_3$ such that

$$E(z_0) \leqslant E(z) \qquad (z \in H_3) \tag{17}$$

holds. In this case, $E(z_0)$ is called the Čebyšev radius of the set $A(r,F)$. In other words: The Čebyšev center $z_0$ has a minimal maximal error. The notion of minimal maximal error is fundamental in the theory of optimal approximation. See Golomb and Weinberger [7] for the scalar case (i.e., $F \in H_1'$ is a continuous linear form on $H_1$) and Sard [15] for the general (non-scalar) case.

The special structure of the set (14) allows an application of the generalized hypercircle method (Synge [18], Davis [4], Sard [15], Golomb and Weinberger [7], Holmes [10]) in order to determine the Čebyšev center of $A(r,F)$, the set of admissible approximations of $Fx$.

Retain the preceding notations. The following result will be basic for our approach.

LEMMA 1. *Let the operator $G \in B(H_2, H_3)$ belong to the set $C(r,F)$ for a suitable number $r > 0$, i.e., suppose*

$$F = GT, \qquad \mathrm{Tr}(GVG^*) < +\infty. \tag{18}$$

*Moreover, suppose that the operator*

$$G_0 = FT^+ \tag{19}$$

*satisfies* $\mathrm{Tr}(G_0 V G_0^*) < +\infty$. *Then we have*

$$F = G_0 T \tag{20}$$

*and the equality of the traces*

$$\mathrm{Tr}(GVG^*) = \mathrm{Tr}(G_0 V G_0^*) + \mathrm{Tr}((G - G_0)\, V(G - G_0)^*) \tag{21}$$

*obtains.*

*Proof.* From (18) we conclude $N(T) \subseteq N(F)$. Hence the inclusion

$$N(F)^\perp \subseteq N(T)^\perp \tag{22}$$

holds for the orthogonal complements in the Hilbert space $H_1$. Taking into account that $T^+ T = P_{R(T^+)} = P_{R(T^*)} = P_{N(T)^\perp}$ holds, we conclude from (19), $F = F P_{N(F)^\perp}$ and (22) that

$$\begin{aligned} G_0 T &= F T^+ T \\ &= F P_{N(F)^\perp} P_{N(T)^\perp} \\ &= F P_{N(F)^\perp} \\ &= F, \end{aligned} \tag{23}$$

as (20) asserts. Since $GV^{1/2}$ and $G_0 V^{1/2}$ are Hilbert–Schmidt operators of $H_2$ into $H_3$, the operators $(G - G_0)\, V(G - G_0)^*$, $G_0 V(G - G_0)^*$, $(G - G_0)\, V G_0^*$ all have finite traces (Sard [14]). It follows for the left hand side of (21)

$$\begin{aligned} \mathrm{Tr}(GVG^*) = \mathrm{Tr}(G_0 V G_0^*) & \\ &+ \mathrm{Tr}((G - G_0)\, V(G - G_0)^*) \\ &+ \mathrm{Tr}((G - G_0)\, V G_0^*) \\ &+ \mathrm{Tr}(G_0 V(G - G_0)^*). \end{aligned} \tag{24}$$

Since $V$ is a selfadjoint operator, the equality (21) is established if we have proved that

$$G_0 V(G - G_0)^* = 0 \tag{25}$$

holds. Indeed, we obtain by (19), $(4_2)$, $TT^+ = P_{R(T)}$, (15), $(4_3)$, (18) and (19):

$$\begin{aligned}
G_0 V(G-G_0)^* &= FT^+ V(G-G_0)^* \\
&= FT^+ TT^+ V(G-G_0)^* \\
&= FT^+ P_{R(T)} V(G-G_0)^* \\
&= FT^+ VP_{R(T)}(G-G_0)^* \\
&= FT^+ VTT^+ (G-G_0)^* \\
&= FT^+ V(TT^+)^*(G-G_0)^* \\
&= FT^+ (GTT^+ - G_0 TT^+) \\
&= FT^+ (FT^+ - FT^+) \\
&= 0
\end{aligned} \tag{26}$$

This completes the proof. ∎

*Remark* 1. Under the hypotheses of Lemma 1, equality (21) shows that for any $r > 0$ we have

$$\operatorname{Tr}(G_0 V G_0^*) = \inf_{G \in C(r,F)} \operatorname{Tr}(GVG^*). \tag{27}$$

Thus, using the terminology of Sard [15], the operator $G_0$ given by (19) is an abstract spline with respect to the observation $G \rightsquigarrow GT$ and the coobservation $G \rightsquigarrow GV^{1/2}$.

The next lemma provides an expression of the translate of the "hypercircle" $C(r, F)$ as defined by (13).

LEMMA 2. *If the number $\rho > 0$ is defined according to*

$$\rho^2 = r^2 - \operatorname{Tr}(G_0 V G_0^*), \tag{28}$$

*where $G_0 = FT^+$ satisfies $\operatorname{Tr}(G_0 V G_0^*) < +\infty$, then we have*

$$C(r, F) = G_0 + C(\rho, 0). \tag{29}$$

*Proof.* Let $K \in C(\rho, 0)$, i.e., $KT = 0$ and $\operatorname{Tr}(KVK^*) \leqslant \rho^2$. An application of Lemma 1 yields $F = (G_0 + K)T$ and

$$\begin{aligned}
\operatorname{Tr}((G_0+K)V(G_0+K)^*) &= \operatorname{Tr}(G_0 V G_0^*) + \operatorname{Tr}(KVK^*) \\
&\leqslant \operatorname{Tr}(G_0 V G_0^*) + \rho^2 \\
&= r^2.
\end{aligned} \tag{30}$$

Consequently, $(G_0 + K) \in C(r, F)$ and therefore $G_0 + C(\rho, 0) \subseteq C(r, F)$. On the other hand, for a given $G \in C(r, F)$ define $K = G - G_0$. Then we have $KT = 0$ and (21) implies

$$\begin{aligned} \operatorname{Tr}(KVK^*) &= \operatorname{Tr}(GVG^*) - \operatorname{Tr}(G_0 VG_0^*) \\ &\leqslant r^2 - \operatorname{Tr}(G_0 VG_0^*) \\ &= \rho^2. \end{aligned} \tag{31}$$

This shows $K \in C(\rho, 0)$, i.e., the inclusion $C(r, F) \subseteq G_0 + C(\rho, 0)$ holds and the identity (29) is established. ∎

We are now in a position to determine the (unique) Čebyšev center $z_0 \in H_3$ of the set $A(r, F) \subseteq H_3$ of admissible approximations of $Fx$. Our proof makes use of the techniques developed by Sard [15].

THEOREM 2. *Let $x_0 = T^{+} b$. Then the unique Čebyšev center of the set $A(r, F)$ (see (14)) is given by*

$$z_0 = G_0 b = F x_0. \tag{32}$$

*Proof.* We have to establish the inequality

$$E(z_0) < E(z) \qquad (z \in H_3, z \neq z_0). \tag{33}$$

An application of Lemma 2 yields (cf. (16))

$$\begin{aligned} E(z)^2 &= \sup_{G \in C(r, F)} \|z - Gb\|^2 \\ &= \sup_{K \in C(\rho, 0)} \|z - G_0 b - Kb\|^2 \\ &= \sup_{K \in C(\rho, 0)} \max(\|z - z_0 - Kb\|^2, \|z - z_0 + Kb\|^2) \\ &\geqslant \sup_{K \in C(\rho, 0)} \frac{1}{2}(\|z - z_0 - Kb\|^2 + \|z - z_0 + Kb\|^2) \\ &= \sup_{K \in C(\rho, 0)} (\|Kb\|^2 + \|z - z_0\|^2) \\ &= \|z - z_0\|^2 + \sup_{K \in C(\rho, 0)} \|Kb\|^2. \end{aligned} \tag{34}$$

From (34) we conclude

$$E(z)^2 \geqslant \|z - z_0\|^2 + E(z_0)^2 \tag{35}$$

From (35) inequality (33) follows. ∎

Let us consider the special case

$$\begin{aligned} H_3 &= \mathbb{C}, \\ V &= \mathrm{id}_{H_2}, \end{aligned} \tag{36}$$

$$F \in H_1' \text{ such that } N(T) \subseteq N(F) \text{ (cf. (11))}.$$

Then we have $F^*(1) \in R(F^*) \subseteq R(T^*)$. Furthermore

$$\mathrm{Tr}(GG^*) = \|G\|^2 = (G^*(1) | G^*(1)). \tag{37}$$

In particular, there exists a number $r > 0$ such that $C(r, F) \neq \varnothing$.

An application of Theorem 2 yields the following infinite dimensional version of the Neyman-David theorem (cf. Linnik [12]):

COROLLARY 1. *Let hypotheses* (36) *be satisfied and choose* $r > 0$ *so that* $C(r, F) \neq \varnothing$ *holds. Then*

$$z_0 = Fx_0 = FT^+ b \tag{38}$$

*represents the unique complex number so that the strict inequality*

$$\begin{aligned} &\sup\{|Fx_0 - Gb| \mid GT = F, \|G\| \leqslant r\} \\ &\quad < \sup\{|z - Gb| \mid GT = F, \|G\| \leqslant r\} \end{aligned} \tag{39}$$

*holds for all* $z \in \mathbb{C}$, $z \neq Fx_0$.

COROLLARY 2. *Let* $\rho > 0$ *so that* $\rho^2 = r^2 - \|G_0\|^2$. *In the present case, the Čebyšev radius of the set* $A(r, F)$ *is given according to*

$$E(z_0) = \rho \|Tx_0 - b\|. \tag{40}$$

*Proof.* In view of (16), (32), and (13) we have

$$E(z_0) = \sup\{|G_0 b - Gb| \mid F = GT, \|G\| \leqslant r\}. \tag{41}$$

By Lemma 2 we may write $G = G_0 + K$, where $K \in C(\rho, 0)$. Thus, we obtain

$$\begin{aligned} E(z_0) &= \sup\{|Kb| \mid KT = 0, \|K\| \leqslant \rho\} \\ &= \sup\{(b | K^*(1)) \mid T^*K^*(1) = 0, \|K^*(1)\| \leqslant \rho\}. \end{aligned} \tag{42}$$

Taking into account that $N(T^*) = N(T^+)$ holds, we have

$$\begin{aligned} E(z_0) &= \sup\{(b | K^*(1) - TT^+ K^*(1)) \mid T^*K^*(1) = 0, \|K^*(1)\| \leqslant \rho\} \\ &= \sup\{(b | P_{R(T)^\perp} K^*(1)) \mid T^*K^*(1) = 0, \|K^*(1)\| \leqslant \rho\} \\ &= \sup\{(P_{R(T)^\perp} b | K^*(1)) \mid T^*K^*(1) = 0, \|K^*(1)\| \leqslant \rho\}. \end{aligned} \tag{43}$$

Since $P_{R(T)^\perp} b = P_{N(T^*)} b$ belongs to $N(T^*)$, the identity

$$E(z_0) = \rho \|P_{R(T)^\perp} b\| = \rho \|TT^+ b - b\| \tag{44}$$

follows. If we observe (9), the equality (40) obtains. ∎

In a second example, choose

$$\begin{aligned} H_1 &= H_3, \\ F &= P_{N(T)^\perp}. \end{aligned} \tag{45}$$

Let $V \in B(H_2)$ satisfy the conditions (12), (13) and

$$\mathrm{Tr}(V) < +\infty. \tag{46}$$

In view of $T^+ T = P_{N(T)^\perp}$ there exists a number $r > 0$ so that the hypercircle

$$C(r, P_{N(T)^\perp}) = \{G \in B(H_2, H_1) \mid GT = P_{N(T)^\perp}, \mathrm{Tr}(GVG^*) \leqslant r^2\} \tag{47}$$

is not empty; $C(r, P_{N(T)^\perp})$ is the set of inner inverses $G$ (i.e., $TGT = T$) with bounded variance $\mathrm{Tr}(GVG^*) \leqslant r^2$. See Sard [14, 16]. We assume that the set $A(r, P_{N(T)^\perp})$ of admissible approximations is bounded. Since in this example $x_0 = Fx_0$, we obtain immediately from Theorem 2 the following result:

COROLLARY 3. *The minimal norm least-squares solution $x_0 = T^+ b$ of the equation $Tx + \delta b = b$ is the unique Čebyšev center of the set $A(r, P_{N(T)^\perp})$ of admissible approximations $Gb$ to the "true" solution $x$*, i.e.,

$$\sup_{G \in C(r, P_{N(T)^\perp})} \|x_0 - Gb\| < \sup_{G \in C(r, P_{N(T)^\perp})} \|z - Gb\| \ (z \neq x_0). \tag{48}$$

*Remark* 2. An important example occurs when the space $H_2$ is finite dimensional. Then condition (41) holds trivially and Corollary 2 is true for positive definite variance matrices $V$ that satisfy the invariance property (15). The choice $V = \sigma \mathrm{id}_{H_2}$ is of particular importance.

*Note added in proof.* The authors have been informed that in a recent note by C. W. Groetsch, "Generalized splines and generalized inverses," *Numer. Funct. Anal. Optim.* 2 (1980), 93–97, a common functional analytic framework for splines and generalized inverses has been given. In this connection they also refer to the forthcoming paper of M. Tasche, "A unified approach to interpolation methods," *J. Integral Equations*, in press.

## REFERENCES

1. A. ALBERT, "Regression and the Moore–Penrose Pseudoinverses," Academic Press, New York, 1972.

2. F. J. BEUTLER AND W. L. ROOT, The operator pseudoinverse in control and systems identification, *in* "Generalized Inverses and Applications" (M. Z. Nashed, Ed.), Academic Press, New York, 1976.
3. A. BJERHAMMAR, "Theory of Errors and Generalized Inverse Matrices," Elsevier, Amsterdam, 1973.
4. P. J. DAVIS, "Interpolation and Approximation," Ginn (Blaisdell), London, 1963.
5. F. J. DELVOS, Splines and pseudoinverses, *RAIRO Anal. Numer.* **12** (1978), 313–324.
6. C. F. GAUSS, "Abhandlungen zur Methode der kleinsten Quadrate" (A. Börsch and P. Simon, Eds.), Berlin, 1889.
7. M. GOLOMB AND H. F. WEINBERGER, Optimal approximation and error bounds, *in* "On Numerical Approximation" (R. E. Langer, Ed.), Madison, Univ. of Wisconsin Press, 1959.
8. C. W. GROETSCH, "Generalized Inverses of Linear Operators," Dekker, New York, 1977.
9. W. GROSSMANN, "Grundzüge der Ausgleichsrechnung," 3. Aufl., Springer-Verlag, Berlin/Heidelberg/New York, 1969.
10. R. B. HOLMES, "A Course on Optimization and Best Approximation," Springer-Verlag, Berlin/New York, 1972.
11. P.-J. LAURENT, "Approximation et optimisation," Herrmann, Paris, 1972.
12. J. W. LINNIK, "Die Methode der kleinsten Quadrate in moderner Darstellung," Deutscher Verlag der Wissenschaften, Berlin, 1961.
13. W. PETRYSHYN, On generalized inverses and the uniform convergence of $(I - \beta K)^n$ with applications to iterative methods, *J. Math. Anal. Appl.* **18** (1967), 417–439.
14. A. SARD, "Linear Approximation," Amer. Math. Soc., Providence, R. I., 1963.
15. A. SARD, Approximation based on nonscalar observations, *J. Approx. Theory* **8** (1973), 315–334.
16. A. SARD, Approximation and variance, *Trans. Amer. Math. Soc.* **73** (1952), 428–446.
17. R. SCHATTEN, "Norm Ideals of Completely Continuous Operators," Ergebnisse der Mathematik und ihrer Grenzgebiete 27, Springer-Verlag, Berlin/Göttingen/Heidelberg, 1960.
18. J. L. SYNGE, "The Hypercircle in Mathematical Physics," Cambridge Univ. Press, London, 1957.