387

No/

MSC-ED-R-67-59

MSC INTERNAL NOTE

PSEUDOINVERSION FOR OPERATOR'S DEFINED ON FINITE DIMENSIONAL HILBERT SPACES

By

John L. Engvall

National Aeronautics and Space Administration
Manned Spacecraft Center
Houston, Texas

September 1967

70090

MSC INTERNAL NOTE NO. ED-R-67-59

PSEUDOINVERSION FOR OPERATOR'S DEFINED ON FINITE DIMENSIONAL HILBERT SPACES

Prepared by: John L. Engvall

Approved by: H. P. Decell, Jr.
Chief, Theory and Analysis Office

Approved by: Eugene H. Brock
Chief, Computation and Analysis Division

National Aeronautics and Space Administration
Manned Spacecraft Center
Houston, Texas

CONTENTS

INTRODUCTION

This paper presents an algorithm for computing a pseudo-inverse for any linear operator, $A : R_n \rightarrow R_m$, where R_n and R_m are Hilbert spaces of degree N and M, respectively. Problems involving linear operators that are "ill-conditioned" or "difficult to invert" are being encountered in ever increasing numbers. Computer programs using standard inversion techniques often provide no solution to these problems and no information indicating an alternate approach to the problem.

The algorithm to be discussed provides a method for minimizing $||Ax - b||$ for a fixed vector b in R_m. The inner products for R_n and R_m are not restricted to the usual inner product for Euclidean complex space. The notation for the inner product of two vectors x and y will be (x,y) and the only norm used will be $||x|| = \sqrt{(x,x)}$. Although the inner product on R_n may be different from the one on R_m, there will be no special notation to differentiate these inner products since the space to which the vectors x and y belong will dictate the inner product to be used. Computer programs for minimizing $||Ax - b||$ for any norm other than the Euclidean norm are essentially nonexistent at MSC. It should be pointed out that standard inversion procedures are more desirable for inversion of square matrices, and sometimes for solving the least squares problem, provided that one is inverting "nicely-behaved" operators.

AN ALGORITHM FOR PSEUDOINVERSION

Let $A : R_n \to R_m$ be a linear operator. Let $e_1, e_2, \cdots, e_n$ denote unit orthogonal vectors in R_n and the vectors $A(e_i)$ in R_m will be denoted by b_i. The subspace of R_m spanned by the vectors $b_1, b_2, \cdots b_n$ will be referred to as $R(A)$ or the range of the operator A.

If y is an element of R_m, then there exists a unique vector x belonging to the range of A such that

$$||y - x|| = \inf_{m \varepsilon R(A)} ||y - m|| \tag{1}$$

The purpose of this paper is to characterize all vectors, $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, such that

$$x = \sum_{i=1}^{n} \alpha_i \, b_i = \sum_{i=1}^{n} \alpha_i \, A(e_i) = A(\alpha) \tag{2}$$

and to provide an algorithm for calculating some of these vectors.

If an orthogonal basis for $R(A)$ is known, then a representation for x can be obtained by

$$x = P(y) \tag{3}$$

where $P : R_m \to R(A)$ is the orthogonal projection operator onto the range of A. Hence the collection of all vectors, α, satisfying (1) are precisely those such that

$$P(y) = A(\alpha) \tag{4}$$

Using the Gram-Schmidt process, a set of vectors, $z_1, z_2, \cdots, z_n$, which span $R(A)$ can be represented in terms of the b_i as follows. Let $z_1 = b_1$ and for $j = 2,3,\cdots,n$. Let

$$z_j = b_j - \sum_{i=}^{j-1} \frac{(b_j, z_i)}{||z_i||^2} z_i \tag{5}$$

where the sum is taken over the z_i such that $||z_i|| \neq 0$.

Define $Q = \{z_i : ||z_i|| \neq 0\}$. Then Q is a collection of orthogonal vectors which span $R(A)$ and the projection operator can be defined by

$$P(y) = \sum_{\substack{i=1 \\ ||z_i|| \neq 0}}^{n} \frac{(y,\ z_i)}{||z_i||^2}\, z_i \tag{6}$$

If a set of vectors, $c_1, c_2, \cdots, c_n$, can be defined such that they span R_n and $A(c_i) = z_i$, then any vector, $\alpha = (\alpha_1, \alpha_2, \cdots, \alpha_n)$, such that

$$\alpha_i = \frac{(y,\ z_i)}{||z_i||^2} \tag{7}$$

for those z_i such that $||z_i|| \neq 0$ satisfies

$$A\left(\sum_{i=1}^{n} \alpha_i\ C_i\right) = \sum_{i=1}^{n} \alpha_i\ z_i$$

$$= \sum_{\substack{i=1 \\ ||z_i|| \neq 0}}^{n} \frac{(y,\ z_i)}{||z_i||^2}\, z_i + \Theta$$

$$= P(y) \tag{8}$$

Moreover, this will completely characterize all vectors, c , such that $A(c) = P(y)$. Suppose there is a vector, $\hat{c}$, such that $A(\hat{c}) = P(y)$; then there is a unique set of scalars $(\alpha_1, \alpha_2, \cdots \alpha_n)$ such that $\hat{c} = \sum_{i=1}^{n} \alpha_i c_i$ (since the vectors c_i are linearly independent). The vectors, $z_i \in Q$, are orthogonal and span $R(A)$, thus the coefficients of the z_i are such that

$$\sum_{i=1}^{n} a_i z_i = P(y) \quad z_i \in Q \tag{9}$$

are unique. Thus,

$$\alpha_i = (y, z_i)/||z_i||^2 \tag{10}$$

for every z_i in Q .

If the vectors, $c_1, c_2, \cdots, c_n$, can be defined, then the collection of all vectors β such that $A(\beta) = P(y)$ is given by

$$\beta = \left\{ \sum_{i=1}^{n} \alpha_i c_i \right. \tag{11}$$

the α_i satisfy (10) above $\Big\}$.

The vectors, $c_1, c_2, \cdots, c_n$, are defined by the following recurrence relation.

Initialize $c_1 = e_1$ then for $j = 2,3,\cdots,n$ let

$$c_j = e_j - \sum_{\substack{i=1 \\ ||z_i|| \neq 0}}^{j-1} \frac{(b_j,\ z_i)}{||z_i||^2}\, c_i \tag{12}$$

It remains to be shown that $A(c_i) = z_i$, and that the vectors c_i are linearly independent. First, note that $A(c_1) = A(e_1) = b_1 = z_1$. Then, for $j = 2,3,\cdots,n$

$$A(c_j) = A(e_j) - \sum_{\substack{i=1 \\ ||z_i|| \neq 0}}^{j-1} \frac{(b_j,\ z_i)}{||z_i||^2}\, A(c_i) \tag{13}$$

$$A(c_j) = b_j - \sum_{\substack{i=1 \\ ||z_i|| \neq 0}}^{j-1} \frac{(b_j,\ z_i)}{||z_i||^2}\, A(c_j) \tag{14}$$

Thus

$$A(c_j) = z_j \quad \text{for} \quad j = 2,3,\cdots,n$$

One can quickly observe that the vectors c_j are linearly independent by expanding (12), and noting that if $\sum_{i=1}^{k} a_i c_i = \Theta$, $a_k \neq 0$, then there exists scalars $d_1, d_2, \cdots, d_{k-1}$ such that

$$\sum_{i=1}^{k-1} d_i e_i + e_k = \Theta$$

which is a contradiction.

Consider now the operator, $\hat{A}$, defined as follows. For any $q \neq \Theta$ in $R(A)$, q has a unique representation as a linear combination of the vectors in Q . If

$$q = \sum_{z_i \varepsilon Q} a_i z_i$$

define

$$\hat{A}(q) = \sum a_i c_i$$

and

$$\hat{A}(\Theta) = \Theta$$

If $H(A)$ is used to denote the subspace of R_n spanned by the vectors c_i such that $z_i \neq \Theta$, then $\hat{A} : R(A) \to H(A)$ is a one-to-one onto map. Moreover, for any x in $H(A)$, $\hat{A}A(x) = x$, and for any q in $R(A)$, $A\hat{A}(q) = q$. We cannot say that $A\hat{A} = \hat{A}A$ since their domains are different. We cannot say that the matrix representation of either $\hat{A}A$ or $A\hat{A}$ is the identity map since $\hat{A}A : R_n \to H(A)$ is the identity only on $H(A)$ and $A\hat{A} : R(A) \to R_m$, but $R(A) \subset R_m$ may not be generated by the e_i in R_m.

The following operator does provide a convenient representation for a solution to (1). For $y \in R_m$, define $A^+(y) = \hat{A}\, P(y)$, then $\alpha = A^+(y)$ is a solution to (1). This follows immediately since $P(y)$ belongs to the range of A. Thus, $A(\alpha) = A\hat{A}\, P(y) = P(y)$. If $N(A)$ denotes the space spanned by the c_i such that $z_i = \Theta$, then $E = \{g = A^+(y) + \mu : \mu \in N(A)\}$ is the collection of all vectors satisfying (2).

For those who are familiar with the generalized inverse, it should be pointed out that A^+ satisfies the

Penrose equations, $A^+ AA^+ = A^+$ and $AA^+ A = A$, but not the other Penrose equations. Moreover, $A^+ y$ satisfies (1), but it is not necessarily the vector having minimal norm which satisfies this equation.

APPLICATIONS TO MATRIX INVERSION

For computational purposes, the following modification of this algorithm is more practical. Choose a tolerance ε such that if $||x|| < \varepsilon$ for any vector x, $x = \theta$. Normalize all vectors b_i and let

$$d_i = b_i/||b_i|| \tag{15}$$

Obtain orthogonal unit vectors $\hat{z}$ by the following recurrence relation.

Initialize $\hat{\hat{z}}_1 = \hat{z}_1 = z_1 = d_1$. For $j = 2,3,\cdots,n$ let

$$z_j = d_j - \sum_{\substack{i=1 \\ ||z_i|| \neq 0}}^{j-1} (d_j, \hat{\hat{z}}_i)\, \hat{\hat{z}}_i \tag{16}$$

Use Wilkinson's modified Gram-Schmidt

$$\hat{z}_j = z_j - \sum_{i=1}^{j-1} (z_j, \hat{\hat{z}}_i) \hat{\hat{z}}_i \qquad (17)$$

$$||z_i|| \neq 0$$

If $||z_j|| < \varepsilon$, set $\hat{\hat{z}}_j = \Theta$. Otherwise, set

$$\hat{\hat{z}}_j = \hat{z}_j / ||\hat{z}_j||$$

Notice that ε is used as a relative error tolerance since all vectors are normalized.

All other calculations are carried out as before and we have $c_1 = e_1$ and for $j = 2,3,\cdots,n$

$$c_j = \frac{\frac{1}{||b_j||} e_j - \sum_{i=1}^{j-1} (d_j, \hat{\hat{z}}_i) c_i}{||\hat{z}_j||} \qquad (18)$$

If the vectors $\hat{\hat{z}}_i$ are stored by columns in a matrix ZH and the c_i are stored in a matrix ZI by columns. Then

$$A^+ = ZI(ZH^T) \qquad (19)$$

and

$$P(y) = ZH(ZH^T y) \tag{20}$$

For square nonsingular matrices, $A^+ = A^{-1}$.

Moreover, it would be advisable to use standard matrix inversion algorithms for matrices that are not ill-conditioned.

For matrices that are ill-conditioned, this algorithm provides the following advantages:

1. The reason for the error return "ill-conditioned matrix" is accompanied by reason as to why A^{-1} is difficult to compute. For example, if $||\hat{z}_5|| = 10^{-6}$, then there exists scalars g_1, g_2, g_3, g_4 such that

$$\sum_{i=1}^{4} g_i \, b_i + b_5 = \varepsilon \tag{21}$$

where

$$||\varepsilon|| \leq 10^{-6} \, ||b_5||$$

2. Although the matrix is ill-conditioned, a solution x is computed such that $\delta = ||Ax - b|| \neq 0$ is a

minimum, where $x = A^{+} y$. Even though $\delta = ||Ax - b|| \neq 0$, the value for δ is often quite acceptable.

3. The range of A is characterized by unit orthogonal vectors. Using these, the value for $\inf||Ax - y||$ can be computed for any number of values for y without calculating A^{+}, that is $\delta = ||ZH(ZH^{T} y) - y||$.

4. The null space, $N(A)$, is characterized by linearly independent vectors.

If the computation of $A^{+} y$ is the primary objective, then the error check for equation (17) could be augmented or replaced by the alternate check

$$\text{if } \left[\frac{(y, \hat{\hat{z}}_j)}{\sum_{i=1}^{j-1} |(y, z_i)|^2} \right] < \varepsilon \quad \text{set } \hat{\hat{z}}_j = \Theta$$

APPLICATION TO "LEAST SQUARES" PROBLEMS

The problem of minimizing $||Aq - y||$, where $A = \{a_{ij}\} = f_j(x_i)$ for $i = 1,2,\cdots,m$ and $j = 1,2,\cdots,n$, occurs frequently in applied mathematics. If $||Aq - y||^2 = (Aq - y)^* (Aq - y)$ then this problem is usually solved using orthogonal polynomials, or in the case of general functions, by the solution to the normal equations

$$q = (A^* A)^{-1} A^* y$$

provided $A^* A$ is nonsingular.

Polynomial approximations are insufficient for solving a large class of problems. For this class of problems and for polynomial problems where nonconsecutive coefficients must be zero, the algorithm described in this paper has the following advantages:

1. The matrix $A^* A$ is often "ill-conditioned" when A is not. The matrix $A^* A$ is never used, thus avoiding excessive rounding error in some instances.

2. A solution is always obtained even when A is not of full rank, that is, when $A^* A$ is singular.

3. If k significant decimal places are desired in the solution, then set $\varepsilon = 10^{-k}$ and all functions which individually do not affect the first k significant digits of the solution will be discarded. This will permit more accurate orthogonalization and consequently a more accurate estimation of the coefficients for the more significant functions. This also allows more functions to be used for fitting and is sometimes the deciding factor in obtaining an acceptable value for $||Aq - y||$.

4. The value for $\min||Aq - y||$ can be computed without computing the solution $q = A^+ y$. Thus, if the value is not acceptable, no time need be used to compute A^+ , that is,
$\min||Aq - y|| = ||ZH\ ZH^T y - y||$.

5. The process described in 3 and 4 can be carried out sequentially with respect to the functions f_k . That is to say that after J columns of the ZH matrix have been computed, if $\delta = ||ZH\ ZH^T y - y||$ is sufficiently small, then the solution, $A^+ = ZI(ZH^T)$, can be computed using only that portion of the matrices available and setting

$q_i = 0$ for $i > J$. If δ is not acceptable, then the process can be continued until the number of functions is exhausted or an acceptable value is obtained.

It should be pointed out, as in the case of square matrices, that standard procedures for solving the least squares problem require less computer time and fewer storage locations than the algorithm discussed. In this sense, the standard procedures are much more desirable provided an acceptable solution is obtained.

CONCLUSION

The fact that other algorithms are often more desirable must be emphasized strongly. There are highly recommended methods based upon extremely well planned numerical analysis using less computer storage and less execution time. The Gram-Schmidt orthogonalization is known to be an unstable process and the accuracy of the process has not been optimized in the procedure presented in this paper. This algorithm does have the advantage that all quantities computed have an obvious relation to the original operator. This permits one to determine the reasons for the difficulty in inverting an operator or in computing A^+ . For "least squares" or minimum norm type problems, the method has the advantages that 1) solutions can be computed sequentially with respect to the approximating functions, 2) the norm of the error vector can be determined without computing A^+ , and 3) functions affecting the solution in "insignificant" decimal places can be automatically discarded. Problems requiring only two or three significant decimal places occur quite often in problems related to spaceflight. Discarding functions in this manner can also improve the behavior of the solution at points between the values of the independent variable used for the regression analysis.