

# Extending Pseudo Inverses for Matrices to Linear Operators in Hilbert Space

M. A. Murray-Lasso \*<sup>1</sup>

<sup>1</sup>Facultad de Ingeniería, Universidad Nacional Autónoma de México (UNAM).  
Circuito Escolar s/n, Ciudad Universitaria, Coyoacán  
Mexico City, Mexico 04510  
\*mamurraylasso@yahoo.com

## ABSTRACT

In this paper formulas derived by the author for calculating the pseudo inverse of any matrix are generalized to linear operators in Hilbert space. The pseudo inverse is seldom required unless there are many right side vectors, which become known at different times. The minimum square solution of functional equations is also presented for a single right-side vector. Some definitions and theorems of functional analysis are included. An application to a simple minimum energy optimal control problem is presented in detail.

Keywords pseudo inverse operators, minimum norm optimization, linear operators, Hilbert space, discretization.

## RESUMEN

En este artículo se generalizan fórmulas, deducidas por el autor, para el cálculo de la pseudo inversa de cualquier matriz, a operadores lineales en espacios de Hilbert. La pseudo inversa raramente se necesita a menos que muchos vectores del lado derecho se presenten en diferentes tiempos. La solución de mínimos cuadrados de ecuaciones funcionales se presenta también. Se presenta en detalle una aplicación a la solución de un problema sencillo de control óptimo de mínima energía.

## 1. Introduction

The solution of systems of linear equations occupies a central place in applied mathematics. Even though systems with  $n$  equations and  $n$  variables are common in practice, in many cases one has to solve systems with more or less equations than unknowns, that is overdetermined or undetermined systems. Depending on the ranks of the coefficient matrix and the augmented matrix one can know if there is a solution and if the solution is unique. When no solution exists what can be done is to search for the vector which gives the residue with least Euclidean norm.

When there is an infinity of solutions, of all of them one can look for the shortest one in Euclidean norm. When there is only one solution it can be obtained by multiplying the right side of the system by the inverse matrix. When there are no solutions or many solutions what can be done is to multiply the right side by the pseudo inverse of the matrix

and the shortest solution with least residue is obtained.

When the concept of pseudo inverse of a matrix is extended to spaces with infinite dimensions, the concepts of rank of a matrix and augmented matrix lose their applicability. However, other theorems applicable to both, spaces of finite and infinite dimensions, are valid. Thus, a linear functional equation  $Lx = y$  with  $y$  known, has a solution if  $y$  is in the range of the linear operator  $L$ .

The solution is unique if the only solution to the homogeneous equation  $Lx = 0$  is the zero function. If we work carefully, much can be done to apply theorems which are true for finite dimensional spaces to Hilbert spaces of infinite dimensions. In this article, a generalization of a theorem (see Murray-Lasso [1, 2]) for the calculation of pseudo inverse matrices to linear operators in Hilbert space, is given.

## 2. Expressions for the pseudo inverse of a matrix

The decomposition of a rectangular matrix  $A$  in singular values is given by Equation (1):

$$A = U \Lambda V^T \tag{1}$$

where  $A$  is any  $m \times n$  matrix of real numbers with rank  $k$ ,  $U$  is a matrix whose columns are the  $m$  orthonormal characteristic vectors associated with the non-zero characteristic values of the symmetric matrix  $AA^T$  (in this article the superindex  $T$  represents the transposed matrix) and the matrix  $V$  is formed with the orthonormal characteristic vectors associated with the non-zero characteristic values of the symmetric matrix  $A^T A$ . Because of the symmetry of  $AA^T$  and  $A^T A$  the corresponding characteristic values are real and positive and their square roots (the singular values) can be taken positive. The non-zero characteristic values coincide when properly ordered. Even in the case of repeated characteristic values, the characteristic vectors can be manipulated to form an orthonormal system. The vectors of  $U$  and  $V$  should be ordered so that they coincide with the common non-zero characteristic values of the matrices  $AA^T$  and  $A^T A$ . The matrix  $\Lambda$  is a diagonal matrix with positive numbers called the singular values of  $A$  and can be written in the form shown in Equation (2):

$$\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_2 & \dots & 0 \\ 0 & 0 & \dots & 0 \\ 0 & 0 & \dots & \lambda_k \end{bmatrix} \tag{2}$$

Generally, the singular values are ordered as shown in Expression (3)

$$\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n \tag{3}$$

The pseudo inverse  $A^+$  of the matrix  $A$  can, by the definition, be given by Equation (4)

$$A^+ = V \Lambda^{-1} U^T \tag{4}$$

(See Lanczos [6], Forsythe [12], Noble [13], Dahlquist and Björk [14].)

Now, if we sacrifice the property that the matrix  $\Lambda$  be diagonal and instead of the semiorthonormal matrices  $V$  and  $U$  we place matrices whose columns generate the same spaces, Murray-Lasso [1] has shown that the pseudo inverse  $A^+$  can be written as Equation (5):

$$A^+ = Q \Phi^{-1} P^T \tag{5}$$

where the non-singular matrix  $\Phi = P^T A Q$  and the matrices  $P$  and  $Q$  are matrices whose columns form a basis for the spaces generated by the columns of  $A$  and  $A^T$ .

Murray-Lasso [2] established that the pseudo inverse  $A^+$  can be calculated using Equation (6) in the following manner:

$$A^+ = QX \tag{6}$$

where the matrix  $X$  is any solution of the matrix Equation (7):

$$P^T A Q X = P^T \tag{7}$$

Here  $P$  is a matrix whose columns include a set of linearly independent vectors and their linear combinations capable of generating the space of the columns of  $A$ ;  $Q$  is a matrix whose columns include a set of linearly independent vectors and their combinations capable of generating the space of the columns of  $A^T$ .

As a particular case, we can take  $P = A$  and  $Q = A^T$ . Hence, we can write Equations (6) and (7) as Equations (8) and (9):

$$A^+ = A^T X \tag{8}$$

where  $X$  is any matrix solution of Equation (9):

$$A^T A A^T X = A^T \tag{9}$$

Generally, in the solution of a problem of least-square-minimum-norm, it is laborious to calculate the pseudo inverse if it is only used to calculate the solution for a single right side. It is much less work to solve the problem directly for the given right side. The equations corresponding to (8) and (9) for a single right side can be written as Equation (10):

$$\mathbf{x}_{\min} = \mathbf{A}^T \mathbf{z} \text{ where } \mathbf{z} \text{ is any solution of } \mathbf{A}^T \mathbf{A} \mathbf{A}^T \mathbf{z} = \mathbf{A}^T \mathbf{y} \quad (10)$$

### L<sub>2</sub> Space

One of the Hilbert spaces we will be using is the L<sub>2</sub> space of real functions in the interval [a, b] over the field of real numbers. In this space, the elements are real functions quadratically integrable in the interval [a, b], which is to say that that any element **f** in L<sub>2</sub>. satisfies Equation (11):

$$\int_a^b |f(x)|^2 dx < \infty \quad (11)$$

The inner product of this space is defined by Equation (12):

$$(f, g) = \int_a^b f(x)g(x)dx \quad (12)$$

and the norm of a vector is defined through the inner product as Equation (13):

$$\|f\| = \sqrt{(f, f)} \quad (13)$$

The metric, that is the distance between two points (vectors), is defined through the norm as Equation (14):

$$\rho(f, g) = \|f - g\| \quad (14)$$

### 3. Linear Operators

If E and F are linear vector spaces over the same field  $\Phi$ , we will denote with L(E, F) the set consisting of all linear transformations from E to F. This set, whose elements we will call Linear Operators, is a vector space over  $\Phi$  if we define for  $S, T \in L(E, F)$  and  $\alpha \in \Phi$  the operations of addition and multiplication by a scalar with the properties given by Equations (15), and (16):

$$\text{Addition: } (S + T)x = Sx + Tx, \quad x \in E \quad (15)$$

$$\begin{aligned} &\text{Multiplication by a scalar:} \\ &(\alpha S)x = \alpha(Sx), \quad x \in E, \quad \alpha \in \Phi \end{aligned} \quad (16)$$

If E = F, we write L(E) instead of L(E, E) and we call the elements of L(E) Linear Operators over E. If  $T \in L(E, F)$ , we call E the dominion of T and we call F the co-dominion of T. We call the set  $N(T) = \{x \in E: Tx = 0\}$  the null space of T. We call the set  $R(T) = \{y \in F: y = Tx \text{ for some } x \in E\}$  the range of T. The set N(T) is a subspace of E, and the set R(T) is a subspace of F. L(E, T) is *one to-one* iff (if and only if)  $N(T) = \{0\}$  and is *over* iff  $R(T) = F$ . The null transformation  $0 \in L(E, F)$  is that given by  $0x = 0$  for all  $x \in E$ . The null space of the 0 transformation is  $N(0) = E$ , while its range  $R(0) = 0$ . The identity operator  $I_E \in L(E)$  is given by  $I_E x = x$  for all  $x \in E$ . Additionally  $N(I_E) = \{0\}$  and  $R(I_E) = F$ . All that was said in this section about linear operators holds for any linear vector space, including Banach and Hilbert spaces. (Canavati Ayub [5]).

### 4. Operator Algebra

Linear Operators in Hilbert space form a vector space. The product of bounded operators mapping a Hilbert space to itself is defined by Equation (17):

$$(AB)x = A(Bx), \text{ for any } x \in H \quad (17)$$

We can also define the sum of operators in Hilbert space through Equation (18):

$$(A + B)x = Ax + Bx, \text{ for any } x \in H \quad (18)$$

Likewise, the multiplication of an operator by a scalar k is defined by Equation (19):

$$(kA)x = k(Ax) \quad (19)$$

The operator product, sum, and the image of the multiplication of a vector by a scalar are all vectors that belong to a Hilbert space when the operators are bounded.

The norm  $\|A\|$  of an operator A in a Hilbert space is defined by Equation (20):

$$\|A\| = \sup_{\|x\| \leq 1} \|Ax\|, \text{ or equivalently } \|A\| = \sup_{\|x\| \neq 0} \frac{\|Ax\|}{\|x\|} \quad (20)$$

Operators with a finite norm are said to be *bounded* (See Kolmogorov and Fomin [7]). Sometimes, mappings between a Hilbert space and one of its subspaces are considered.

**5. Adjoint Operators in Hilbert Space**

To each linear operator  $T$  in a real Hilbert space  $H$  corresponds another operator  $T^*$  also in Hilbert space known as its *adjoint*. This adjoint operator is defined in terms of the inner product by Equation (21):

$$(Tx, y) = (x, T^*y) \text{ or } (Tx, y) - (x, T^*y) = 0, \text{ for any pair } x, y \in H \tag{21}$$

In real Euclidean spaces of finite dimensions, operators are represented with matrices and the adjoint operator is represented with the transpose of that matrix. Many of the properties of transposed matrices are inherited by adjoint operators in Hilbert spaces. For example, the adjoint operator is also linear, the norms of the operator and its adjoint are equal, that is,  $\|A\| = \|A^*\|$ ; this means that if the operator is bounded, the adjoint operator is also bounded and both map a finite sphere into a finite sphere. Additionally, if an operator is continuous, so is the adjoint. When an operator and its adjoint are equal, the operator is called *self adjoint*; such operators have properties very similar to those of symmetric matrices. (See Vulikh [4])

**6. Determination of Adjoints for Some Particular Cases**

In the solution of certain linear differential and Fredholm integral equations like Equation (22) appear frequently:

$$Kx = y(s) = \int_a^b K(s, t)x(t)dt \tag{22}$$

It involves an operator which we will call  $K$  because Equation (22) can be written  $y = Kx$ . Because of the properties of integrals, the operator  $K$  is linear. We will assume that  $K(s, t)$  is a continuous function of two variables defined in the square  $a \leq s, t \leq b$  and that  $K$  belongs to the function space  $L_2$  defined in the interval  $[a, b]$ . In this case the adjoint operator has as its kernel  $K(t, s)$  (notice the variables have been interchanged in

the kernel) and expression (22) is converted to Equation (23):

$$K^*x = y(s) = \int_a^b K(t, s)x(t)dt \tag{23}$$

In case  $K(s, t) = K(t, s)$ , we are dealing with a self-adjoint operator  $K$ . Note the similarity with matrices. If we consider integration as a generalized summation and the two variables of  $K(s, t)$  as continuous, instead of discrete, indices, whereas for a matrix to find the adjoint (transpose) we interchange the indices, for an integral operator we interchange the variables of  $K(s, t)$ .

In many problems in engineering and other disciplines, the models appear as ordinary or partial linear differential equations. The differential equations go together with initial or boundary conditions which are an integral part of the model. To find the adjoint operators of the differential operators, one can apply Equation (21). There are other methods for determining the adjoint operator of a differential operator. One of them involves discretizing the derivatives by means of finite differences. We obtain a matrix which is then transposed and observed when we make the interval of discretization tend to zero. We can deduce which operator the transposed matrix tends to. One must be careful with the boundary conditions because, depending on them, the discretized matrices may be different. When the matrix resulting from the discretization is symmetric then the differential operator is self-adjoint. A second method is to convert the differential model into an integral model and apply the method described above. In this case the Green function of the differential system turn to be very useful. (Lanczos [6])

We now work a simple case to illustrate the application of Equation (21) to obtain the adjoint of a linear ordinary differential operator. Consider the differential operator corresponding to a unit mass moving due to a force that varies with time  $x''(t) = B(t)$ , in the time interval  $[0, 1]$ . The initial conditions are  $x(0) = x'(0) = 0$ . We will call  $D = \frac{d^2}{dt^2}$  the formal part of the differential operator, and  $D^*$  its adjoint. We begin by determining the

differential part of the adjoint operator. Using Equation (21)  $xy'' - yD^*x = 0$ . We notice that if we assume that  $D^* = \frac{d^2}{dt^2}$  the following equation is true:

$$\int_0^1 [x''(t)y(t) - x(t)D^*y(t)]dt = \int_0^1 [x''(t)y(t) - x(t)y''(t)]dt = \int_0^1 \frac{d}{dt} (xy' - yx')dt = [xy' - yx']_0^1 \equiv 0$$

The term in brackets depends only on the boundary conditions which if it becomes zero it justifies the assumption that  $D^*$  is the second derivative with respect to time. We conclude that the adjoint operator which operates on  $y$  is  $\frac{d^2}{dt^2}$ .

We now investigate the boundary condition of the right side. Given the initial conditions on  $x$ , the contribution of the adjoint in the lower limit is zero, while in the upper limit where  $t = 1$ , we have

$$x(1)y'(1) - y(1)x'(1) = 0$$

Since nothing was specified about  $x(t)$  and  $x'(t)$  for  $t = 1$ , the equation forces  $y(1) = y'(1) = 0$ . Therefore, the adjoint operator sought when applied to  $y$  is

$$\frac{d^2}{dt^2} y(t), y(1) = y'(1) = 0$$

Nothing was said about the right side because it does not belong to the operator.

(Analogously, in an  $n$ -dimensional Euclidean space, the transpose only has to do with the coefficient matrix and has nothing to do with the right-hand side vector. (Lanczos [6].)

An important case which appears in dynamic systems ruled by linear differential equations with constant coefficients is that of vector Equation (24):

$$\mathbf{x}' = \mathbf{Ax} + \mathbf{Bu} \tag{24}$$

where  $\mathbf{x}$  is a vector with  $n$  functions of time representing the state of the system,  $\mathbf{A}$  is an  $n \times n$  constant matrix,  $\mathbf{B}$  is an  $n \times m$  constant matrix and  $\mathbf{u}$  is a vector of functions of time for the control. The

adjoint system corresponding to  $\mathbf{x}' = \mathbf{Ax}$ ,  $\mathbf{x}(0) = \mathbf{0}$  defined in the time interval  $[0, T]$ , is  $\mathbf{z}' = -\mathbf{A}^*\mathbf{z}$ ,  $\mathbf{z}(T) = \mathbf{0}$  defined in the time interval  $[0, T]$ . The two system can be put in integral form respectively as follows :

$$\int_0^T e^{\mathbf{A}(T-t)}[.]dt, \quad \int_0^T e^{-\mathbf{A}^*(T-t)}[.]dt$$

where in the places marked  $[.]$  go the right side functions  $\mathbf{Bu}(t)$  and  $\mathbf{u}^T(t)\mathbf{B}^T$ . It should be noted that we are working with the inverses of the differential operators.

Lanczos [6] gives formulas to find the adjoint operators corresponding to linear differential operators that vary with time. For the operator given in Equation (25):

$$Dv(x) = \sum_{k=0}^r p_k(x)v^{(k)}(x) \tag{25}$$

He gives, for the  $k$ -th term of the operator  $D$ , the  $k$ -th term of the adjoint operator as shown in Equation (26):

$$D^*u(x) = (-1)^k \frac{d^k}{dx^k} [p_k(x)u(x)]. \tag{26}$$

when the coefficients  $p_k(x)$  are constant, each term of the differential operator consists of the formal part of the term including the constant coefficient multiplied by a constant whose value is either 1 or -1 depending of whether the order of the devivative is even or odd, respectively. When we apply this rule, we

deduce that the operator  $a \frac{d^2}{dx^2}$  has itself for adjoint in the formal part, the boundary conditions are still pending.

### 7. Theorems on norm minimization in Hilbert space

**Theorem 1:** Let  $G$  and  $H$  be Hilbert spaces, and let  $A$  be a bounded operator mapping  $G$  in  $H$  with a closed range in  $H$ . Then the vector  $\mathbf{x}$  of minimum norm that satisfies  $A\mathbf{x} = \mathbf{y}$  is given by  $\mathbf{x} = A^*\mathbf{z}$  wher  $\mathbf{z}$  is any solution of  $AA^*\mathbf{z} = \mathbf{y}$ .

**Corollary 1:** Under the assumptions of the previous theorem, if the operator  $AA^*$  is invertible, the optimal solution takes the form  $x = A^*(AA^*)^{-1}y$ .

**Theorem 2:** Let  $A$  be a linear bounded operator acting in two Hilbert spaces. The following are true;

1.  $[R(A)]^\perp = N(A^*)$ , 2.  $R(A) = [N(A^*)]^\perp$ , 3.  $[R(A^*)]^\perp = N(A)$ , 4.  $R(A^*) = [N(A)]^\perp$ .

The lines above the symbols denote the closure of the corresponding subspace.

(For the proofs of these theorems see Luenberger [11])

If we assume both  $R(A)$  and  $R(A^*)$  are closed, Theorem 2 justifies the diagram of Figure 1.

In the diagram in Figure 1, the domain of  $A$  and the co-domain of  $A^*$  are made to coincide because they

are isomorphic spaces. Both the domain of  $A$  and the co-domain of  $A^*$  are the direct sum ( $\oplus$ ) of the null space of  $A$  and the range of  $A^*$ . Likewise, the domain of  $A^*$  and co-domain of  $A$  are the direct sum of the null space of  $A^*$  and the range of  $A$ . The symbol " $\perp$ " denotes the orthogonal space of the symbol preceding it.

We give a geometric illustration in a three-dimensional Euclidean space of the problem of finding the minimum norm solution corresponding to Theorem 1. Any solution of a linear problem  $Ax = y$  for  $y \in R(A)$  is given by the sum of any vector  $x_0$  satisfying the equation and any vector in the null space of  $A$ . If the null space of  $A$  consists only of the zero vector, the solution is unique, else there exists an infinity of solutions forming a linear manifold given by the sum of a vector  $x_0$  plus the subspace  $N(A)$ . Of all the vectors in the linear manifold there exists one (and only one) which is the shortest, and the given vector is orthogonal to  $N(A)$ .

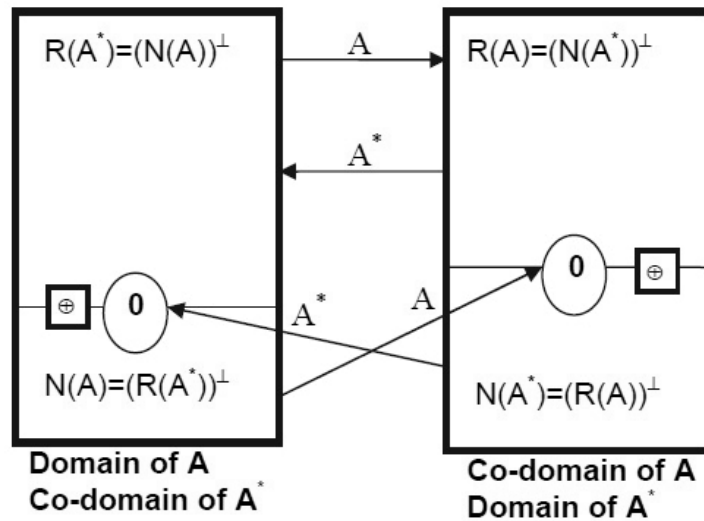


Figure 1. Schematic Diagram of the mappings associated with an operator  $A$  and its adjoint  $A^*$ .

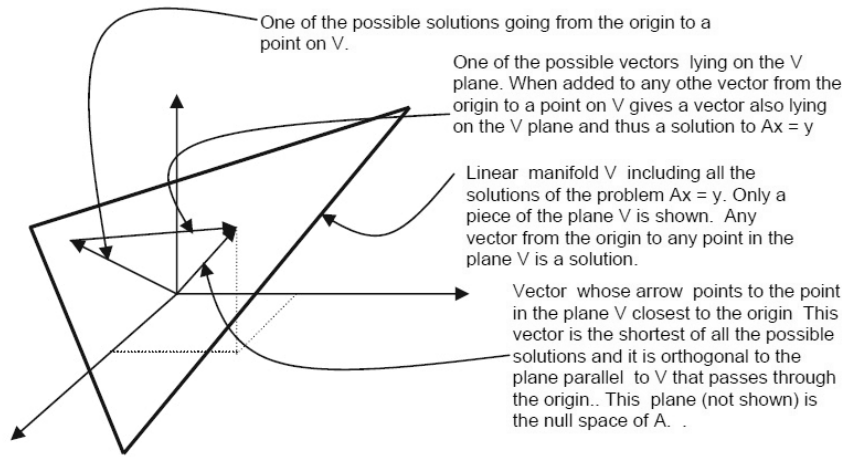


Figure 2. Diagram showing the connection of all solutions of a linear system  $Ax = y$  and any one solution and a linear manifold parallel to the null space of  $A$ . The shortest solution is orthogonal to the plane parallel to  $V$  passing through the origin (null space of  $A$ .)

In theorem 1, we assume that the equation  $Ax = y$  has a solution because  $y$  is in the range of  $A$ . If we assume that  $y$  is not in the range of  $A$ , then the equation has no solution. However, in many cases it is of interest to obtain the vector  $x$  which when operated by  $A$  is closest to  $y$ . The approximation is measured by the Euclidean distance, which is to say that we want to find  $x_m$  such that

$$x_m = \min_x \|Ax - y\| \tag{27}$$

To solve this problem, we invoke the following theorem:

**Theorem 3:** (For proof see Luenberger [11].) Let  $H$  be a Hilbert space and  $M$  a closed subspace of  $H$ . Corresponding to any vector  $x \in H$ , there exists a unique vector  $m_0 \in M$  such that  $\|x - m_0\| \leq \|x - m\|$  for all  $m \in M$ . Furthermore, a necessary and sufficient condition that  $m_0 \in M$  be the unique minimizing vector is that  $x - m_0$  be orthogonal to  $M$ . (See illustration in Figure 3)

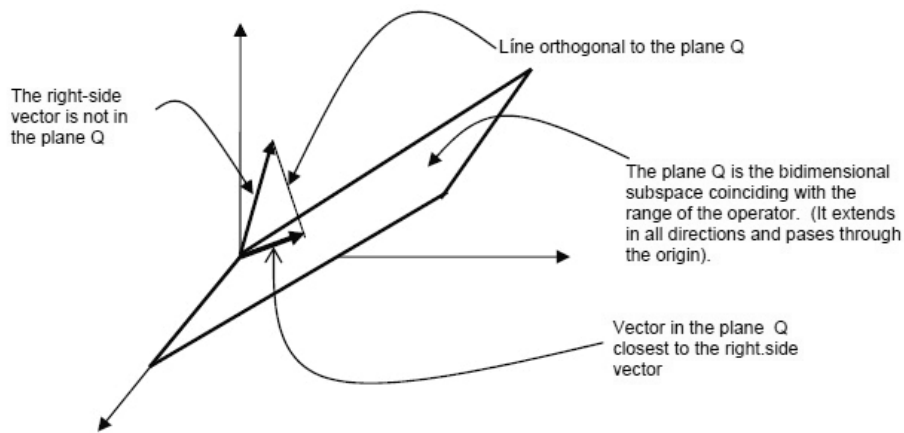


Figure 3. Diagram illustrating that when the right-side vector of  $Ax = y$  is not in the range of  $A$  there is no solution, but the best approximation in Euclidean norm to a solution is the one that takes as the right-side vector the orthogonal projection of  $y$  on the range.

The use which we will give to theorem 3 is that when in the equation  $Ax = y$  the vector  $y$  is not in the range of  $A$ , we will orthogonally project the vector  $y$  to the range of  $A$  and we will use said projection as the right side. Therefore, the best approximation, in the sense indicated, to the solution of  $Ax = y$  corresponds to the solution of  $A^*Ax = A^*y$ .

The vector  $y$  is not in the range of  $A$  and therefore has a component in the space orthogonal to the range of  $A$ . According to Figure 1, this orthogonal complement is the null space of  $A^*$ , in other words, the operator  $A^*$  sends to the zero vector any component in the null space of  $A^*$  leaving the resulting vector completely on the range of  $A$ . It also leaves the resulting vector in the range of  $A^*$  since this resulting vector is being premultiplied by  $A^*$  and is therefore in the image space of  $A^*$ . The operation of  $A^*$  on the left side leaves this side also in the range of  $A^*$ .

If the null space of  $A$  is the zero vector, the equation  $A^*Ax = A^*y$  will have a unique solution. On the other hand, in case  $A$  has a null space different from the zero vector, there will be an infinity of solutions. The shortest of these solutions is the one we are looking for. To get a solution in the most general case which involves the possibility that both  $A$  and  $A^*$  have null spaces different from the zero vector, we start by assuming that the right side of  $Ax = y$  is in the range of  $A$ ; we then apply Theorem 1 to obtain the shortest solution  $x_m = A^*z$  where  $z$  is any solution to  $AA^*z = y$ . We now admit the possibility that  $y$  is not in the range of  $A$ , we take care of this contingency by premultiplying both sides of the last equation by  $A^*$  and we obtain Equation (28):

$$A^*AA^*z = A^*y, \quad x_m = A^*z \tag{28}$$

We justify this move by noting that  $AA^*$  has the same range as  $A$ . To prove this we note that  $AA^*$  has the same null space as  $A^*$ , since if  $x \in N(A^*)$  which means  $A^*x = 0$  then so is  $AA^*x = 0$ . Conversely, if  $AA^*x = 0$  then  $(x, AA^*x) = (A^*x, A^*x) = \|A^*x\|^2 = 0$  which implies  $A^*x = 0$ . Now, according to Figure 1, the range of an operator is the orthogonal complement of the null space of the adjoint, we can deduce that the ranges of  $A$  and  $AA^*$  are the same.

When we compare Equation (28) with Equation (10), we can see that Equation (28) is an extension of the solution to problems of minimum-squares-minimum-norm in  $n$ -dimensional Euclidean space to the solution of the same problem in Hilbert space.

**8. Illustrative example. Application to optimal control**

To illustrate the theory presented, we work a simple example in the area of optimal control. Consider a direct current motor controlled by its magnetic field. Originally, the motor is at rest in the angle  $x = 0$ . We want to activate the motor with a time varying voltage  $u(t)$  and take it to the angle  $x = 1$  (in radians) in a time  $T = 1$  and to leave it there with zero angular velocity using minimum energy. The corresponding differential equation of motion is given by Gille [8] and Athans and Falb [9] and is

$$\frac{d^2}{dt^2} x(t) = u(t), \quad x(0) = 0, \quad \frac{d}{dt} x(0) = 0$$

where  $x(t)$  is the angular position of the rotor and  $u(t)$  is the field voltage which is related to the total energy given to the motor through the control. This total energy, which we wish to minimize, is given by

$$E = \int_0^1 [u(t)]^2 dt.$$

We begin by writing the system equation in canonical state variable form. We define  $x_1(t) = x(t)$ ,  $x_2(t) = x_1'(t)$ , from which  $x_2'(t) = x_1''(t)$  and  $x_1''(t) = u(t)$ . If we write this in matrix form we have

$$\mathbf{x}'(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{b}u(t), \quad \mathbf{x}(0) = \mathbf{0};$$

$$\begin{bmatrix} x_1' \\ x_2' \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(t), \quad \begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

The vectors and matrices  $\mathbf{A}$ ,  $\mathbf{b}$ ,  $\mathbf{x}$  are easily identified. The solution to the vector equation is

$$\mathbf{x}(t) = \int_0^t e^{\mathbf{A}(t-z)} \mathbf{b}u(z) dz$$

To impose the condition that in  $t = 1$ , the rotor reaches the position  $x = 1$  with angular velocity zero, we write



$$\mathbf{x}(1) = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \text{ which is translated into}$$

$$\int_0^1 e^{A(1-z)} \mathbf{b}u(z) dz = \begin{bmatrix} 1 \\ 0 \end{bmatrix}$$

We calculate the exponential  $e^{At}$  using the following instructions of the program *Mathematica* (Wolfram [10]):

```
A={{0,1},{0,0}}
{{0, 1}, {0, 0}}
MatrixForm[MatrixExp[A t]]
1 t
0 1
```

We are somewhat skeptic that the result is so simple. To check it, we recall the definition of the exponential of a matrix  $\exp(\mathbf{A}t) = \mathbf{I} + \mathbf{A}t / 1! + \mathbf{A}^2t^2 / 2! + \dots$ . Using the matrix  $\mathbf{A}$ , we find that  $\mathbf{A}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}$ , therefore all the higher powers above the first are zero and  $\exp(\mathbf{A}t) = \mathbf{I} + \mathbf{A}t$  which explains why the result is so simple.

Mathematically, our problem is to find  $u(t)$  of minimum norm subject to the condition that

$$\int_0^1 e^{A(1-z)} \mathbf{b}u(z) dz = \begin{bmatrix} 0 \\ 1 \end{bmatrix}. \text{ The condition can be}$$

represented as a functional equation  $Au = c$ , where the operator  $A$  maps the space  $L_2[0, 1]$  into a subspace of two dimensions generated by the functions  $w_1(t) = 1$ ,  $w_2(t) = t$ . Since all spaces of finite dimensions are closed, we can apply Theorem 1. The operator is given by

$$A = \int_0^1 e^{A(t-z)} \mathbf{b}[\cdot] dz. \text{ To find } u(t) \text{ of minimum norm, we must solve the functional equation}$$

$AA^*v = A^*c$  and the minimum norm solution will be given by  $u = A^*v$ . We now find the adjoint operator  $A^*$ . We have pointed above that if an operator is given by  $\int_a^b K(s,t)[\cdot] dt$ , where  $K(s,t)$  is a scalar function of two variables, its adjoint, which maps an  $n$ -dimensional subspace into a Hilbert space  $L_2[0, 1]$ , is given by the kernel of the

integral with the variables  $s, t$  interchanged. In the case that  $K(s, t)$  is a matrix, besides the interchange of variables the matrix must be transposed. If this is taken into consideration we have

$$A = \mathbf{b}^T e^{-A^T(1-z)} [\cdot]$$

The interchange of  $t$  and  $z$  is achieved by placing a minus sign in front of the exponent of  $A$ . Notice that  $A^*$  is a matrix of functions. The product  $AA^*v$  and  $A^*c$  are

$$AA^*v = \left[ \int_0^1 e^{A(1-z)} \mathbf{b}b^T e^{-A(1-z)} dz \right] v, \quad u(z) = A^*c$$

$$= \mathbf{b}^T e^{-A^T(1-z)} c$$

$$\mathbf{b}b^T \text{ is } \begin{bmatrix} 0 \\ 1 \end{bmatrix} \begin{bmatrix} 0 & 1 \end{bmatrix} = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}. \quad e^{A(1-z)} = \begin{bmatrix} 1 & 1-z \\ 0 & 1 \end{bmatrix},$$

$$e^{-A^T(1-z)} = \begin{bmatrix} 1 & 0 \\ z-1 & 1 \end{bmatrix}. \text{ therefore, carrying out the}$$

product of the functions under the integral sign we get

$$AA^* = \int_0^1 \begin{bmatrix} (1-z)(z-1) & 1-z \\ z-1 & 1 \end{bmatrix} dz$$

Introducing the integrand of the previous equation into *Mathematica* and solving the integral we have

```
T = {{{(1 - t) (t - 1), 1 - t},{t - 1, 1}}
{{{(1 - t) (-1 + t), 1 - t}, {-1 + t, 1}}
Q=Integrate[T,{t,0,1}]
General::intinit: Loading integration packages.
1 1 1
{{(-), -}, {-(-), 1}}
3 2 2
```

We now invert Q:

```
R=MatrixForm[Inverse[Q]]
-12 6
-6 4
```

We multiply the inverse by  $c$  to obtain  $v$ :

```
v=Inverse[Q].{1, 0}
{-12, -6}
```

Finally, we operate  $\mathbf{A}^*$  on  $\mathbf{v}$  to obtain  $u$ :

$$u(t) = \mathbf{A}^* \mathbf{v} = \mathbf{b}^T e^{-\mathbf{A}^T(1-t)} \mathbf{v} = \begin{bmatrix} 0 & 1 \end{bmatrix} \begin{bmatrix} 1 & 0 \\ t-1 & 1 \end{bmatrix} \begin{bmatrix} -12 \\ -6 \end{bmatrix} = 6 - 12t$$

If we integrate the function  $u(t)$  from 0 to 1, we obtain the angular velocity  $v(t) = x'(t) = 6t - 6t^2$ . It is apparent that for  $t = 1$ ,  $v(0) = 0$ . We now integrate  $v(t)$  from 0 to 1 and get the angular position  $x(t) = 3t^2 - 2t^3$ . Again for  $t = 0$ ,  $x(0) = 0$

and for  $t = 1$ ,  $x(1) = 1$ . Therefore, all the boundary conditions hold for this problem.

In Figures 4, 5 and 6, the functions  $u(t)$ ,  $v(t)$  and  $x(t)$  are shown respectively. Figure 4 shows the control function, Figure 5 the angular velocity function, and Figure 6 the angular position of the rotor. It is evident from the graphs that the boundary conditions are fulfilled. The theory tells us that there is no function in  $L^2[0, 1]$  with a smaller total energy delivered to the motor measured by  $E = \int_0^1 [u(t)]^2 dt$ .

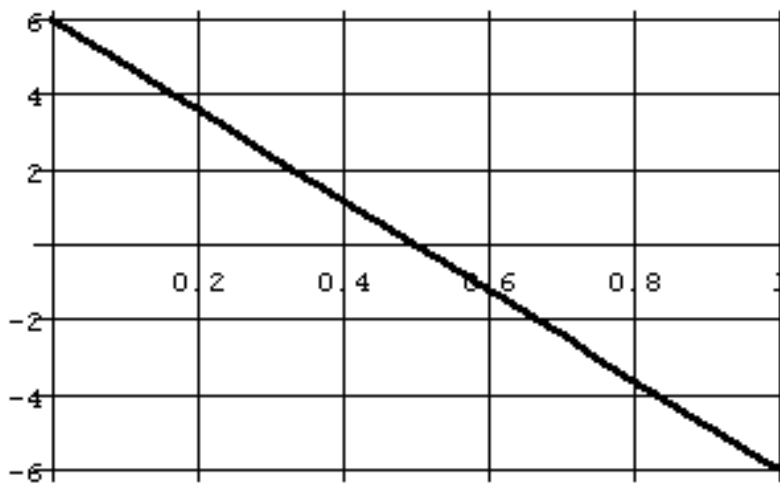


Figure 4. Graph of the optimal voltage  $u(t)$ .

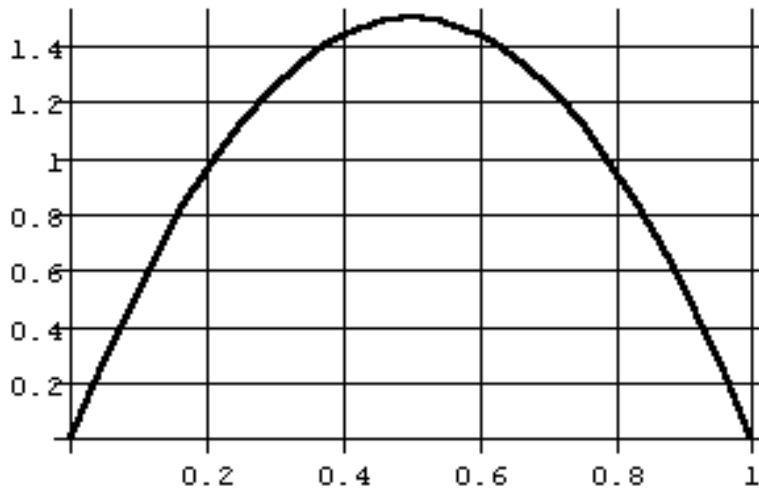


Figure 5. Graph of the optimal angular velocity  $x'(t)$ .

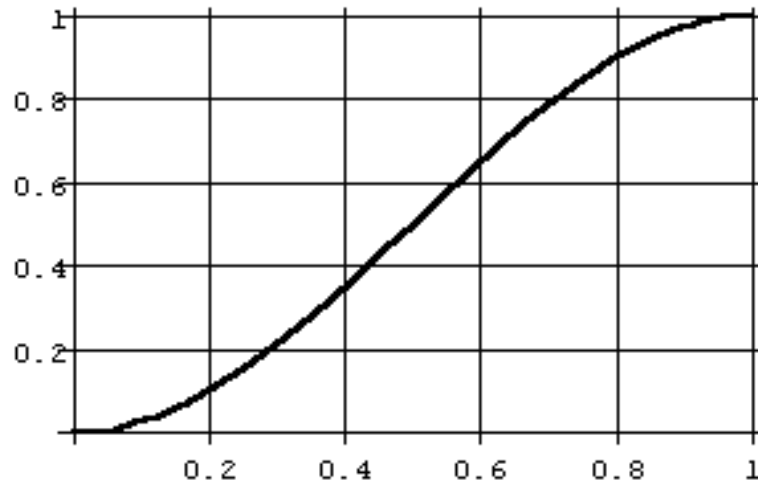


Figure 6. Graph of the optimal angular position  $x(t)$ .

There are many other functions  $u(t)$  which can take the rotor from angle  $x = 0$  to the angle  $x = 1$  starting and ending with zero angular velocity. An example is shown in Figure 7. On Figures 8 and 9, the angular velocity  $v(t) = x'(t)$  and the angular position  $x(t)$  are shown. Nevertheless,  $u(t)$  in Figure 4 is the function in  $L_2[0, 1]$  that minimizes  $\int_0^1 [u(t)]^2 dt$ . The function has norm  $\|u\| = 12$ . While the norm of the

discontinuous function of Figure 7 is  $\|u\| = 16$ . An example of a continuous function which does the same job at a different cost is shown in Figure 10. The corresponding angular velocity and angular position are shown in Figures 11 and 12. The integral of the square of the function from 0 to 1 of the function in Figure 10 is  $2\pi^2 = 19.7392$  which is larger than the two others considered.

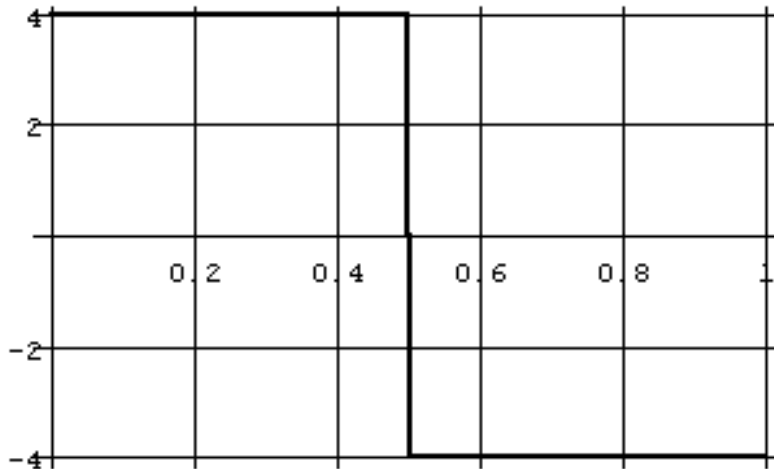


Figure 7. Another  $u(t)$  which takes the rotor from rest in  $x = 0$  to rest in  $x = 1$  in one second.

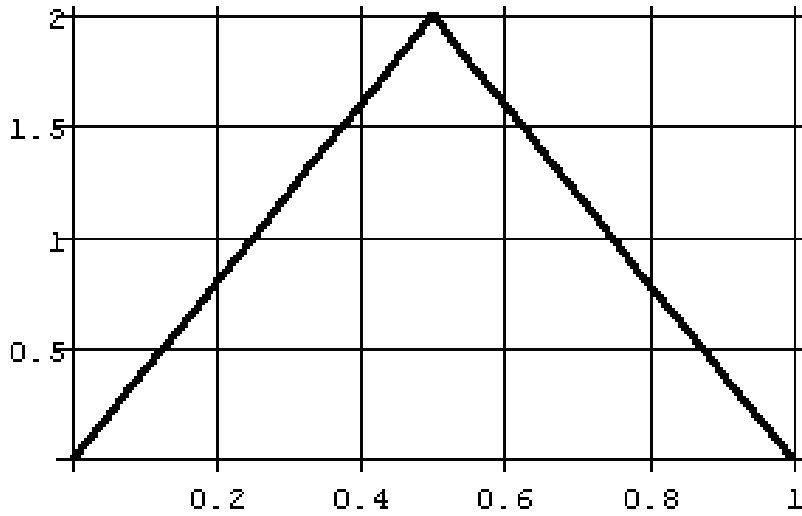


Figure 8. Angular velocity corresponding to  $u(t)$  in Figure 7. Note the zero velocity in both time extremes.

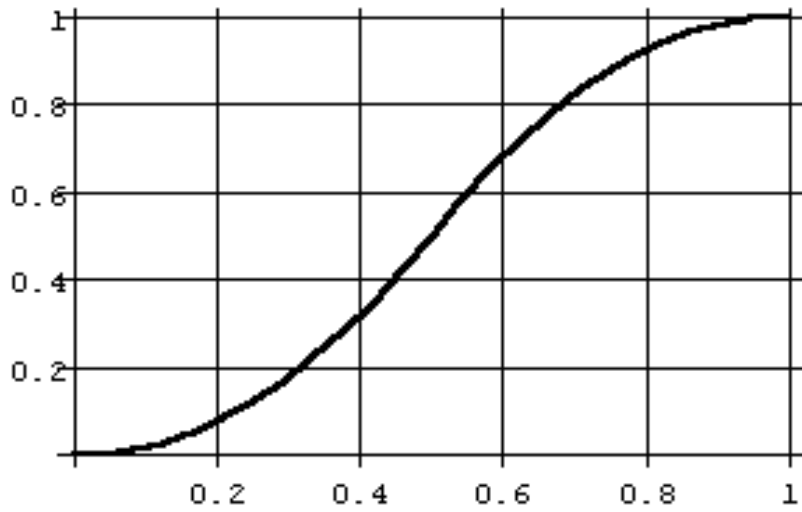


Figure 9. Angular position  $x(t)$  corresponding to  $u(t)$  in Figure 7. Note the function has zero derivative in the extremes.

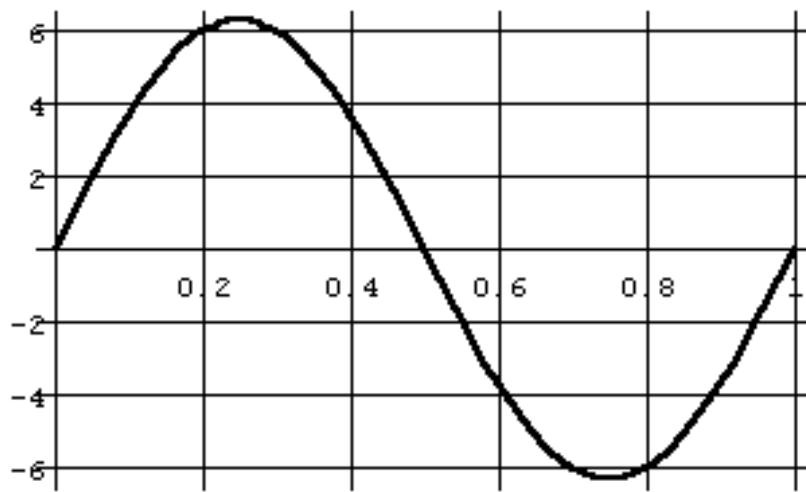


Figure 10. Another  $u(t)$  which takes the rotor from rest in  $x = 0$  to rest in  $x = 1$  in one second.

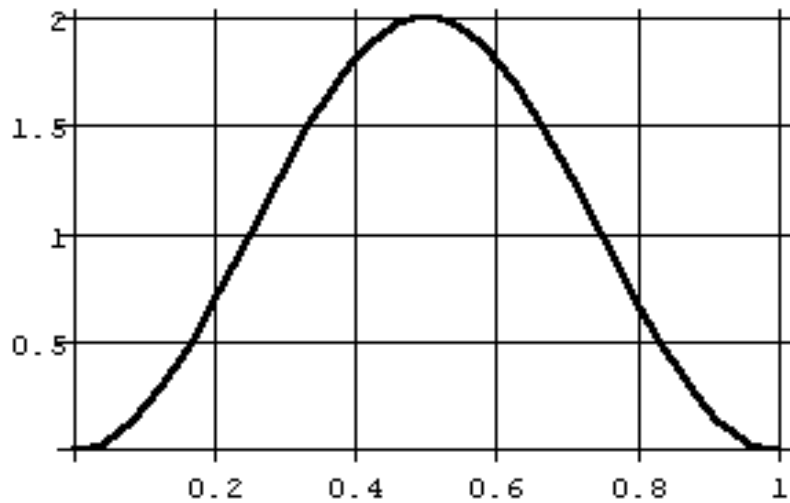


Figure 11. Angular velocity corresponding to  $u(t)$  of Figure 10. Note the zero velocity on both extremes.

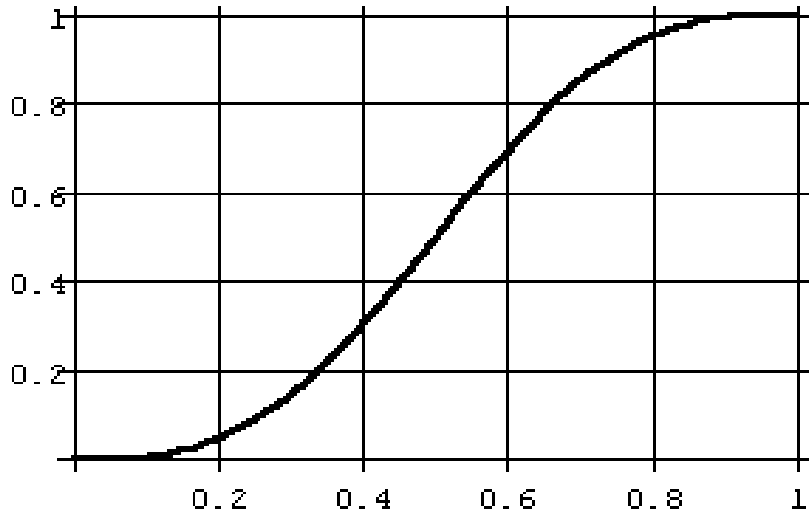


Figure 12. Angular position related to the function in Figure 10. Note the zero derivative in both extremes.

## 9. Conclusion

In this paper a generalization of an expression for the calculation of the pseudo inverse of finite matrices to Hilbert space is given. The expression remains invariant as long as we interpret the finite dimensional vectors as vectors in Hilbert space, matrices as linear operators in Hilbert space and the transposed matrices as adjoint operators. Models which frequently occur in applications involve linear differential operators of finite order. In this case, the differential operator spans a finite dimensional subspace of the Hilbert space and the problem involves a mixture of matrices and integral operators. A simple example of this situation is solved in the article. Some basic definitions and results regarding Hilbert spaces are given in the paper to clarify the notation. As an illustration, a simple minimal-norm optimal control problem is solved in detail.

## References

- [1] Murray – Lasso, M. A. (2007): “Linear Vector Space Derivation of New Expressions for the Pseudo Inverse of Rectangular Matrices”, *Journal of Applied Research and Technology*, Vol. 5, No. 3, pp. 150 – 159.
- [2] Murray-Lasso, M. A. (2008): “Alternative Methods of Calculation of a Pseudo Inverse of a Non – Full Rank Matrix”. *Journal of Applied Research and Technology*, Vol. 6, No. 3, pp. 170 – 183.
- [3] Dunford, N. and Schwartz, J. T. (1957): *Linear Operators, Part I: General Theory*. Interscience Publishers, Inc., New York.
- [4] Vulikh, V. Z. (1963): *Functional Analysis for Scientists and Technologists*. Pergamon Press, Oxford.
- [5] Canavati Ayub, J. A. (1998): *Introducción al Análisis Funcional*. Fondo de Cultura Económica, México.
- [6] Lanczos, C. (1961): *Linear Differential Operators*. D. Van Nostrand Company Limited, London
- [7] Kolmogorov A. N. and Fomin S. V. (1975): *Elements of the Theory of Functional Analysis*. Graylock Press. Rochester, NY.

[8] Gille, J. C. , et al. (1959): Feedback Control Systems, Mc Graw Hill Book Company, New York.

[9] Athans, M. and Falb, P. L. (1966): Optimal Control. Mc Graw – Hill Book Company, New York.

[10] Wolfram, S. (1991): Mathematica. Addison – Wesley Publishing Company, Reading, MA.

[11] Luenberger, D. G. (1969): Optimization by Vector Space Methods. John Wiley & Sons, Inc., New York.

[12] Forsythe, G. E. and C. B. Moller (1967): Computer Solution of Linear Algebraic Systems. Prentice\_hall, Inc. Englewood Cliffs, NJ.

[13] Noble, B. (1969): Applied Linear Algebra. Prentice-Hall, Inc. Englewood Cliffs, NJ.[14] Dahlquist G. and A. Björck (1974): Numerical Methods. Prentice-Hall, Inc. Englewood Cliffs, NJ.

### **Acknowledgement**

The author is indebted to Mario Rodriguez-Green, mathematician, who helped him with bibliographic material and reviewed the manuscript.

**Author's Biography****Marco Antonio Murray-Lasso**

Dr. Murray-Lasso is a mechanical-electrical engineer from UNAM and holds a master of science degree as well as a doctor of science degree, both from MIT. He has held different positions abroad, namely adjunct professor of mathematics at Newark College of Engineering, associate professor of systems engineering at Case Western Reserve University, member of the technical staff at Bell Telephone Laboratories, consultant with NASA (circuit design for space missions). In Mexico, Dr. Murray-Lasso is also a National Researcher, professor and researcher at UNAM (50 years), founding president of Academia Nacional de Ingeniería (National Academy of Engineering). Also, he is or has been international president of Consejo Internacional de Academias de Ingeniería y Ciencias Tecnológicas (International Council of Academies of Engineering and Technological Sciences), CAETS; international director of the International Society for Technology in Education (ISTE); president of Academia Mexicana de Ciencia de Sistemas (Mexican Academy of Systems Science), AMCS; founding president of the Honor Council of Academia Mexicana de Ciencias, Artes, Tecnología y Humanidades (Mexican Academy of Sciences, Arts, Technology and Humanities), AMCATH; coordinator engineering, informatics and general secretary of Academia Mexicana de Ciencias (Mexican Academy of Sciences), AMC; founding president of Sociedad Mexicana de Computación en la Educación (Mexican Society for Computing in Education), SOMECE; honor academician at Academia de Ingeniería y Academia de Ciencia en Sistemas (Academy of Engineering and Academy of Systems Science); member of Academia de Tecnología y Academia de Informática (Academy of Technology and Academy of Informatics); consultant with Banco Nacional de México, PEMEX, Departamento de Comunicaciones y Transportes de México (Department of Communications and Transportation of Mexico), Conacyt, Stanford Research Institute, University of Texas, Universidad de Guanajuato (university of Guanajuato), Mexican Senate and UNESCO. Additionally, Dr. Murray-Lasso has been counselor with MIT Educational Council (for 30 years), author or coauthor of 300 technical articles and 12 books. Finally, he appears in *Who is Who in Science and Engineering* and *Who is Who in the World*.