# ALTERNATIVE METHODS OF CALCULATION OF THE PSEUDO INVERSE OF A NON FULL-RANK MATRIX

M. A. Murray-Lasso

Facultad de Ingeniería UNAM

ABSTRACT

The calculation of the pseudo inverse of a matrix is intimately related to the singular value decomposition which applies to any matrix be it singular or not and square or not. The matrices involved in the singular value decomposition of a matrix A are formed with the orthogonal eigen vectors of the symmetric matrices $A^TA$ and $AA^T$ associated with their nonzero eigenvalues which forms a diagonal matrix. If instead of using the eigenvectors, which are difficult to calculate, we use any set of vectors that span the same spaces, which are easier to obtain, we can get simpler expressions for calculating the pseudoinverse, although the diagonal matrix of eigenvalues is filled. All numerical work to obtain the pseudo inverse whose components are rational numbers when the original matrix is also rational reduces to elementary row operations. We can, thus, generalize the least-squares/ minimum-length normal equations for full-rank matrices and solve said problems and obtain the pseudo inverse in terms of A and $A^{T.}$ without solving any eigen problems or factoring matrices.

KEY WORDS: pseudo inverse, singular values, normal equations, least-squares, minimum-length.

RESUMEN

El cálculo de la seudo inversa de una matriz está íntimamente relacionado con la descomposición de valores singulares aplicable a cualquier matriz, singular o no y cuadrada o no. Las matrices involucradas en la descomposición en valores singulares de una matriz A están formadas con los vectores característicos ortogonales de las matrices simétricas $A^TA$ y $AA^T$ asociados con los valores característicos no nulos, los cuales forman una matriz diagonal. Si, en lugar de usar los vectores característicos, los cuales son difíciles de calcular, se usa cualquier conjunto de vectores que generan los mismos espacios, que son más fáciles de obtener, se pueden obtener expresiones más simples para el cálculo de la seudo inversa, no obstante que la matriz diagonal se llena. Todo el trabajo numérico se reduce a operaciones elementales de filas obteniéndose seudo inversas con componentes racionales cuando la matriz original tiene componentes racionales. De esta manera podemos generalizar las ecuaciones normales de mínimos cuadrados / longitud mínima de matrices de rango completo, resolver el problema y obtener la seudo inversa en términos de A y $A^T$ sin resolver problemas de vectores característicos o factorizar matrices.

PALABRAS CLAVE: seudo inversa, valores singulares, ecuaciones normales, mínimos cuadrados, longitud mínima.

## 1. INTRODUCTION

The pseudo inverse of an $m \times n$ rectangular matrix, where $m$ and $n$ are any natural numbers, is a generalization of the inverse of a square matrix and may be used to solve systems of simultaneous linear equations of any sort. In the case in which the system has a unique solution, the result obtained with the pseudo inverse coincides with the one obtained with the standard inverse. In the case in which there are many (an infinity) of solutions, the pseudo inverse obtains the shortest solution in the euclidean sense. In the case there is no solution, the pseudo inverse obtains a vector which has minimum residue and of all the ones that have the given minimum residue obtains the shortest.

When the rank of the matrix of coefficients **A** of the system of equations is equal to the minimum of the number of rows or columns the matrix is said to be of *full rank.* In such cases there are simple formulas to calculate the pseudo inverse $\mathbf{A}^+$, namely

Rank of **A** is equal to the number of rows: $\mathbf{A}^+ = \mathbf{A}^T(\mathbf{A}\mathbf{A}^T)^{-1}$

Rank of **A** is equal to the number of columns: $\mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T$

When the rank of the matrix is neither equal to the number of rows nor of the columns, the calculation of the pseudo inverse is more involved. The best-known manner of calculation (see Dahlquist and Björk [1]) obtains the singular value decomposition of matrix **A** and exchanges the nonzero singular values for their reciprocal, leaving the zero singular values untouched. This, however, requires the solution of the eigenvalue and eigenvector problem of the symmetric matrices $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$. Golub and Reinsch [2] give an alternative method together with listings of computer programs for obtaining the singular value decomposition. Other methods require factoring the matrix **A** into factors of full rank and apply to each factor one of the formulas above. Noble [3] provides a method for doing the factoring. Murray – Lasso [4] gives a method which involves inverting a matrix obtained by pre and post multiplying matrix **A** on the right and the left with matrices formed with bases for the spaces spanned by the columns and rows of matrix **A** and post and pre multiplying the resulting inverse with the same matrices.

To solve a shortest-length / minimum-square linear equation problem in general it is not necessary to compute the pseudo inverse explicitly since it is more efficient to multiply the right-side vector by the succesive matrices to its left and the part of inverting a matrix can be obviated by factoring the matrix to be inverted in triangular factors **LU** (known as the LU-decomposition) and applying a forward and backward substitution process to obtain the solution for each different right side. In the case of large matrices, this process requires a smaller number of operations than obtaining the inverse and multiplying it by matrices on the right and left and finally multiplying the resultant matrix by the right side vector.

In this paper, we present a method for the calculation of the pseudo inverse which is based on the same ideas as that of [4] but which does not require the previous analysis of the ranks and determination of the bases of the spaces spanned by the rows and columns of **A,** nor the calculation of a standard inverse, but relies on the row-reduction to echelon form of a matrix with a number of rows equal to the smaller of $m$ or $n$. Only $k$ columns need to be processed, since the matrix has rank $k$ and as soon as the $k$-th column is processed, zeros will appear in all succeeding rows and the computation can be stopped. To process one right-side vector, the matrix to be reduced has $n + 1$ columns; if the explicit pseudo inverse is desired, the number of columns is $2n$. The most labor intensive part of the whole process is the multiplication of the matrices.

## 2. THE SPACES OF THE ROWS AND COLUMNS OF A

The space spanned by the columns of an $m \times n$ matrix **A** is its *range.* The range is a subspace of the co-domain of **A**. The dimension of the range is equal to the rank of **A,** which is also the number of linearly independent columns of **A**. The orthogonal complement of the range space is the null space of $\mathbf{A}^T$. The sum of the dimensions of the range of **A** and the null space of $\mathbf{A}^T$ is equal to the number of columns of **A**. The space spanned by the rows of **A,** which is the same space as that spanned by the columns of $\mathbf{A}^T$, is the range space of $\mathbf{A}^T$, which is a subspace of the domain of **A**. The rank of $\mathbf{A}^T$ is equal to the rank of **A**, since both matrices have the same determinants of different orders, therefore, both, the ranges of **A** and $\mathbf{A}^T$ have the same dimension. The orthogonal complement of the range of $\mathbf{A}^T$ is the null space of **A**. This information is condensed schematically in Figure 1.

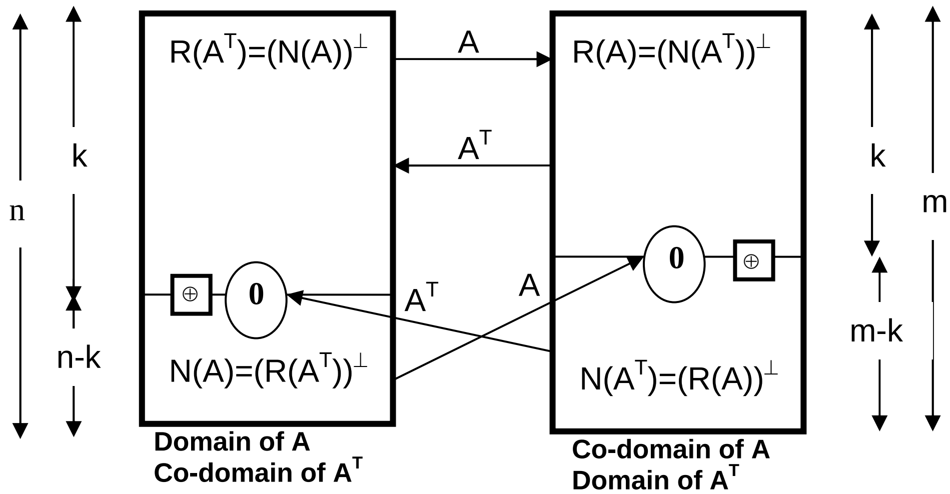


*Figure 1.*

In Figure 1, R and N stand for the range and null space of its argument. The superscript $^{\perp}$ stands for the orthogonal complement of the subspace it refers to. The two pieces separated by a thin line of the domain and co-domain of **A**,when direct summed (⊕) give the whole space in question. The doble arrows on the extreme left and right of the figure are the dimensions of the subspaces inquestion. The **0's** in the ovals are the zero vectors of the spaces. The arrows denote the direction of the mapping and the symbols next to the arrows refer to the operators doing the mapping. $\mathbf{A}^T$ corresponds to the adjoint of operator **A,** which is represented by the transpose of the matrix representing **A.**

Although there are four spaces involved, namely: the domain of **A**, the domain of $\mathbf{A}^T$, the co-domain of **A** and the co-domain of $\mathbf{A}^T$, we have assumed the co-domain of **A** and the domain of $\mathbf{A}^T$ is the same space, and the domain of **A** and the co-domain of $\mathbf{A}^T$ is also the same, since the dimensions coincide. Some authors (Zadeh and Desoer [5]) reduce, without loss of generality, all spaces to one, by adding zeros to the matrix and the shorter vectors to make the matrix square and all vectors equally long. When this approach is used, some facts such as the assertion that the nullity of **A** and $\mathbf{A}^T$ are equal are true. For calculation purposes, which is our aim, the padding with zeros is cumbersome, hence we will not take this approach. However we quote many facts from Zadeh and Desoer [5] because most of them apply to both approaches.

When treating the problem

$$\mathbf{Ax} = \mathbf{b}$$

where **A** is an $m \times n$ matrix of known rational numbers, **x** is an *n*-vector of unknowns, **b** is an *m*-vector of known rational numbers, and we wish to find the unknown vector **x**, we will assume that **A** represents a linear operator mapping an *n*-space to an *m*-space and that the vectors and operator are represented with respect to so called natural orthonormal bases represented by [1, 0, 0, ... , 0], [0, 1, 0, ... , 0], ..., [0, 0, 0, ..., 1], with the proper number of components to correspond to the dimensionality of the space in question. We assume all spaces are Euclidean, thus, the inner product of two vectors $\{x_i\}$ and $\{y_i\}$, i = 1, 2, ..., *r,* is given by

$$\mathbf{x.y} = x_1y_1 + x_2y_2 + x_3y_3 + ... + x_ry_r$$

and two vectors are orthogonal if their inner product is zero. The Euclidean length $||\mathbf{x}||$ of a vector **x** is equal to the positive square root of the inner product of the vector with itself.

$$\|\mathbf{x}\| = \sqrt{\mathbf{x}.\mathbf{x}} = \sqrt{x_1^2 + x_2^2 + ... + x_n^2}$$

If the length of a vector is zero, the vector must be the zero vector [0 0 ... 0].

## 3. THE SPACES OF $A^TA$ AND $AA^T$

There is a close connection between the spaces of **A** and $\mathbf{A}^T$ and the ones of $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$. First, the null space of **A** is the same as the null space of $\mathbf{A}^T\mathbf{A}$. This can be shown by noticing that any vector **x** which satisfies $\mathbf{Ax} = \mathbf{0}$ also satisfies $\mathbf{A}^T\mathbf{Ax} = \mathbf{0}$, that is any vector in the null space of **A** is in the null space of $\mathbf{A}^T\mathbf{A}$. Likewise, for any vector y that satisfies $\mathbf{A}^T\mathbf{Ay} = \mathbf{0}$, also satisfies $\mathbf{Ay} = \mathbf{0}$, since we must have $\mathbf{x}^T\mathbf{A}^T\mathbf{Ay} = \mathbf{0}$ which can be written $(\mathbf{Ay})^T.\mathbf{Ay} = \|\mathbf{Ay}\|^2 = \mathbf{0}$ which implies $\mathbf{Ay} = \mathbf{0}$. Hence, all the vectors in both null spaces are the same. In a similar fashion, it can be shown that the null spaces of $\mathbf{AA}^T$ and $\mathbf{A}^T$ are the same. From the diagram of Fig. 1, since the range of a matrix is the orthogonal complement of the null space of its transpose, we can deduce that the ranges of **A** and $\mathbf{AA}^T$ are the same and also that the ranges of $\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{A}$ are the same.

## 4. DEDUCTION OF FORMULAS FOR THE CALCULATION OF THE PSEUDO INVERSE

The singular value decomposition of a rectangular matrix can be expressed as

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

where **A** is any $m\times n$ matrix of rank $k$ (we assume $k < m < n$.) **U** is a matrix whose columns are the $m$ orthonormal eigenvectors of the $m\times m$ symmetric matrix $\mathbf{AA}^T$, (the sperscript $^T$ denotes the transpose of the matrix), and the matrix **V** is formed with the $n$ orthonormalized eigenvectors of the symmetric matrix $\mathbf{A}^T\mathbf{A}$. The non-zero eigenvalues of both matrices are equal.in number and value and the order of the eigenvectors must be such that both the i-th column of **U** and the i-th column of **V** correspond to the same eigenvalue. The $m\times n$ matrix $\Lambda$ is of the form

$$\Lambda = \begin{bmatrix} \Lambda_1 & \mathbf{0}_1 \\ \mathbf{0}_2 & \mathbf{0}_3 \end{bmatrix} \tag{2}$$

where $\Lambda_1$ is a $k\times k$ diagonal matrix whose $k$ diagonal elements are the real positive square roots of the comon eigenvalues of $\mathbf{AA}^T$ and $\mathbf{A}^T\mathbf{A}$ ordered in the same order as the eigenvectors of **U** and **V.** The matrices $\mathbf{0}_1$, $\mathbf{0}_2$, and $\mathbf{0}_3$ are zero matrices of orders $k\times(n-k)$, $(m-k)\times k$, and $(m-k)\times(n-k)$, respectively. $\mathbf{0}_3$ is associated with the eigenvalue zero, which must appear in both $\mathbf{A}^T\mathbf{A}$ and $\mathbf{AA}^T$ if they both have rank $k < m, n$. (Lanczos [6], pp. 120 –124.) The pseudoinverse $\mathbf{A}^+$ of **A** is

$$\mathbf{A}^+ = \mathbf{V}\Lambda^+\mathbf{U}^T \tag{3}$$

where the pseudo inverse $\Lambda^+$ is of the same form as Matrix $\Lambda$ and the non-zero diagonal elements of $\Lambda_1$ are replaced by their reciprocals. (Dahlquist and Björk [1],p.144.) The calculation of eigenvectors involves difficulties, more so when there are repeated eigenvalues (as it often happens with the zero eigenvalue) and the corresponding eigenvectors have to be orthogonalized. In some of the methods for calculating eigenvalues and eigenvectors, irrational numbers are introduced, whereas the exact pseudoinverse of a matrix of fractional numbers also has fractional numbers.

Equation (1) can be simplified if we take out the eigenvectors corresponding to the eigenvalue zero. In such a case equation (1) becomes

$$\mathbf{A} = \mathbf{U}_p \Lambda_1 \mathbf{V}_p^T \qquad (4)$$

where $\mathbf{U}_p$ and $\mathbf{V}_p$ are semiorthonormal matrices whose columns are the eigenvectors which correspond to the non-zero eigenvalues of $\mathbf{A}^T\mathbf{A}$ and $\mathbf{A}\mathbf{A}^T$ , respectively, ordered as before, that is, in such a way that the i-th eigenvector of each matrix corresponds to the same eigenvalue. The pseudo inverse of **A** is now

$$\mathbf{A}^+ = \mathbf{V}_p \Lambda_1^{-1} \mathbf{U}_p^T \qquad (5)$$

If in equation (5) we sacrifice the condition that $\Lambda_1$ be diagonal, we can replace the work of solving eigenproblems for matrix inversion of a square non-singular matrix of the same size as $\Lambda_1$ which is a considerable saving. All that is necessary is to replace matrix $\mathbf{U}_p$ with a matrix **P** with a minimum number of columns (which does not have to be orthonormal) that spans the same space as **U,** which is the same space as that spanned by $\mathbf{A}\mathbf{A}^T$, and replace $\mathbf{V}_p$ with a matrix **Q** with a minimum number of columns (which does not have to be orthonormal) that spans the same space as **V** which is the same space as that spanned by $\mathbf{A}^T\mathbf{A}$.

It is easy to get a set of vectors that spans the same space as $\mathbf{A}^T\mathbf{A}$, all we have to do is find a set of $k$ linearly independent vectors of $\mathbf{A}^T\mathbf{A}$, where $k$ is the rank of $\mathbf{A}^T\mathbf{A}$. Similarly for $\mathbf{A}\mathbf{A}^T$. A possible method is to column reduce the matrices which are equivalent to row-reducing them since they are symmetric, and then taking the non-zero rows of the reduced matrix as columns of the matrices **P** and **Q** that will replace them.

From Equation (1) we can isolate $\Lambda$ by premultiplying both sides by $\mathbf{U}^T$ and postmultiplying by **V**. Since matrices **U** and **V** are orthonormal, their transposes are their inverses and we obtain

$$\Lambda = \mathbf{U}^T\mathbf{A}\mathbf{V} \qquad (6)$$

and in a similar fashion, because of the zeros in $\Lambda$, we can arrive at

$$\Lambda_1 = \mathbf{U}_p^T\mathbf{A}\mathbf{V}_p \qquad (7)$$

replacing $\mathbf{U}_p$ with **P** and $\mathbf{V}_p$ with **Q,** which are not formed with eigenvectors but span the same space as the corresponding eigenvectors, Equation (7) becomes

$$\Phi = \mathbf{P}^T\mathbf{A}\mathbf{Q} \qquad (8)$$

where the $k \times k$ matrix $\Phi$ is not diagonal, but it has rank k and is therefore non singular and can be inverted. Thus, the equation that corresponds to Equation (5), which gives the pseudoinverse of **A,** is

$$\mathbf{A}^+ = \mathbf{Q}\Phi^{-1}\mathbf{P}^T \qquad (9)$$

where $\mathbf{A}^+$ is an $n \times m$ (the dimensions of $\mathbf{A}^T$) matrix of rank $k$, **Q** is an $n \times k$ matrix of rank $k$, and $\mathbf{P}^T$ is a $k \times m$ matrix of rank $k$.

## 5. ILLUSTRATIVE NUMERICAL EXAMPLE

Let us take the $6 \times 4$ matrix **A** representing an overdetermined system, (taken from Noble [3], p 145)

$$\mathbf{A} = \begin{bmatrix} -1 & 0 & 1 & 2 \\ -1 & 1 & 0 & -1 \\ 0 & -1 & 1 & 3 \\ 0 & 1 & -1 & -3 \\ 1 & -1 & 0 & 1 \\ 1 & 0 & -1 & -2 \end{bmatrix}$$

The symmetric matrix $\mathbf{A}^T\mathbf{A}$ is

$$\mathbf{A}^T\mathbf{A} = \begin{bmatrix} 4 & -2 & -2 & -2 \\ -2 & 4 & -2 & -8 \\ -2 & -2 & 4 & 10 \\ -2 & -8 & 10 & 28 \end{bmatrix}$$

Row reducing the previous matrix, we obtain the two columns of **Q** from the non-zero rows of the row reduced matrix giving

$$\mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \\ -2 & -3 \end{bmatrix}$$

and doing the same with $\mathbf{A}\mathbf{A}^T$ we get

$$\mathbf{A}\mathbf{A}^T = \begin{bmatrix} 6 & -1 & 7 & -7 & 1 & -6 \\ -1 & 3 & -4 & 4 & -3 & 1 \\ 7 & -4 & 11 & -11 & 4 & -7 \\ -7 & 4 & -11 & 11 & -4 & 7 \\ 1 & -3 & 4 & -4 & 3 & -1 \\ -6 & 1 & -7 & 7 & -1 & 6 \end{bmatrix}$$

and after row-reducing it, we obtain for **P**

$$\mathbf{P} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 1 & -1 \\ -1 & 1 \\ 0 & -1 \\ -1 & 0 \end{bmatrix}$$

Using Equation (8), we obtain

$$\Phi = \mathbf{P}^T\mathbf{AQ} = \begin{bmatrix} -26 & -36 \\ 16 & 30 \end{bmatrix}$$

From which

$$\Phi^{-1} = \begin{bmatrix} -5/34 & -3/17 \\ 4/51 & 13/102 \end{bmatrix}$$

We now apply Equation (9)

$$\mathbf{A}^{+} = \mathbf{Q}\Phi^{-1}\mathbf{P}^T = \begin{bmatrix} -\frac{5}{34} & -\frac{3}{17} & \frac{1}{34} & -\frac{1}{34} & \frac{3}{17} & \frac{5}{34} \\ \frac{4}{51} & \frac{13}{102} & -\frac{5}{102} & \frac{5}{102} & -\frac{13}{102} & -\frac{4}{51} \\ \frac{7}{102} & \frac{5}{102} & \frac{1}{51} & -\frac{1}{51} & -\frac{5}{102} & -\frac{7}{102} \\ \frac{1}{17} & -\frac{1}{34} & \frac{3}{34} & -\frac{3}{34} & \frac{1}{34} & -\frac{1}{17} \end{bmatrix}$$

We notice that the exact pseudo inverse of **A** has only rational numbers. This result coincides with the one obtained from the program *Matematica* through the function PseudoInverse[A]. (Wolfram [7], p.850.)

## 6. SOME EXTENSIONS

Matrices **P** and **Q** of Equations (8) and (9) are formed with columns that represent linearly indpendent vectors that span the ranges of matrices $\mathbf{AA}^T$ and $\mathbf{A}^T\mathbf{A}$, respectively. We can save some effort if we note that an arbitrary vector $\eta$ which lies in the range of **A**, can be represented by $\eta = \mathbf{Ay}$ where **y** is a vector in the domain of **A.** Now $\eta$ is a vector in the co-domain of **A**, which is the direct sum of N($A^T$), the null space of $\mathbf{A}^T$**,** and R(A), the range of **A** The rank *k* of **A,** which we assume is less than *m, n* (the number of rows and columns of **A),** implies that **A** has a null space which is orthogonal to R($A^T$). When we multiply $\eta$ by $\mathbf{A}^T$, that is $\mathbf{A}^T\mathbf{Ax}$**,** the resultant vector, which in general belongs to both N($A^T$) and R(A), is completely in the range of **A**, because the component that belongs to N($A^T$) has been annihilated. The same result is achieved by a **P** which contains only linearly independent vectors that span the column space of $\mathbf{A}^T\mathbf{A}$ or that of **A**. A dual argument can be given for insuring that a vector lie completely in the range of $\mathbf{A}^T$. Either we premultiply an arbitrary vector by $\mathbf{AA}^T$ or we multiply it by a matrix **Q** whose columns are linearly indpendent vectors that span the column space of $\mathbf{AA}^T$ or that of $\mathbf{A}^T$. A useful conclusion of this discussion is that we need not find the products $\mathbf{A}^T\mathbf{A}$ and $\mathbf{AA}^T$ to search for linearly independent vectors that span their ranges. It is sufficient to find *k* linearly independent vectors among the columns of **A** to form the columns of **P**; and it is sufficient to find *k* linearly independent *rows* of **A** (equivalent to finding linearly independent columns of $\mathbf{A}^T$).to form the *columns* of **Q**. This result was obtained by Murray-Lasso [4] using a different line of reasoning. There, using the same example, he row-reduced **A** and $\mathbf{A}^T$ to obtain the matrices

$$\mathbf{P}^T = \begin{bmatrix} 1 & 0 & 1 & -1 & 0 & -1 \\ 0 & 1 & -1 & 1 & -1 & 0 \end{bmatrix} \qquad \mathbf{Q} = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ -1 & -1 \\ -2 & -3 \end{bmatrix}$$

which when used in Equations (8) and (9) give identical results for the pseudo inverse of **A.**

In the theory of Linear Operators in Infinite Dimensional Spaces it may not always be clear what corresponds to the linear independent rows and columns and, in some cases, we have to deal with the whole linear operators. It is therefore convenient to have formulas for the calculation of the pseudo inverse in which only the operators **A** and $\mathbf{A}^T$ appear and not pieces of them. When the rank *k* of matrix **A** coincides with the minimum of *m*, *n* (the dimensions of **A**) the expressions for the pseudo inverse are

$$\mathbf{A}^+ = \mathbf{A}^T(\mathbf{AA}^T)^{-1} \quad \text{and} \quad \mathbf{A}^+ = (\mathbf{A}^T\mathbf{A})^{-1}\mathbf{A}^T \qquad (10)$$

(Noble [3], pp.142-143.) This is the kind of formulas we are seeking to attempt to apply to more geneal linear operators.

Using the same ideas that we used for deducing Equations (8) and (9), we postulate the following equation with **A** an $m\times n$ matrix of rank $k < m, n.$ and $\mathbf{A}^T$ playing the role of **Q**. and **A** playing the role of **P**.

$$(\mathbf{A}^T\mathbf{AA}^T)\,\xi = \mathbf{A}^T\mathbf{b}$$

What we have done is lifting the restriction of matrices **P = A** and $\mathbf{Q} = \mathbf{A}^T$ to have rank *k* equal to one of their dimensions. This leaves matrices **P** and **Q** with a null space and, hence, the solutions are not unique. Matrix ($\mathbf{A}^T\mathbf{AA}^T$), which would correspond in Equation (8) to $\Phi$ which has to be inverted, is not a $k\times k$ matrix of rank *k* which can be inverted. Instead, it is an $n\times m$ matrix of rank *k* necessarily singular. Instead of inverting it, what we can do is to obtain one of the solutions by the method of row-reducing it. Because the right side is in the range of **A**, which is also the range of $\mathbf{A}^T\mathbf{AA}^T$, the system of equations has an infinity of solutions. Any one of the solutions when pre-multiplied by $\mathbf{A}^T$ to return to the original variable **x** will result in the same vector, since the portion of the answer that is in the null space of $\mathbf{A}^T$ will be annihilated by $\mathbf{A}^T$. An illustrative example will clarify this:

## 7. AN ILLUSTRATIVE NUMERICAL EXAMPLE OF THE EXTENDED CASE

Let us take again matrix **A** used in the previous example and consider the problem **Ax** = **b.** Let **b =**Column [1 2 3 4 5 6]. Vector **b** is not in the range of **A**, (we can ascertain this by forming the augmented matrix [**A | b**], row- reduce it to normal echelon form and verify that the right side becomes one of the unit columns, showing that the augmented matrix has rank $k + 1$ so the system has no solution.) By multiplying both sides of the equation by $\mathbf{A}^{T,}$ we assure that both sides of the equation are in the range of **A** and, therefore, there is a solution. If we solved the problem as it stands, we would obtain an infinity of solutions, all having the property that they minimize ||Ax – b|| in the euclidean sense. To obtain a solution that is in the range of $\mathbf{A}^T$ which would be orthogonal to N(A), and, therefore, the shortest of all the solutions, we make $\mathbf{x} = \mathbf{A}^T\xi$; the vector $\xi$ is the set of coefficients of the columns of $\mathbf{A}^T$ which are forcing the vector **x** to lie in R($A^T$). The products $\mathbf{A}^T\mathbf{A}\mathbf{A}^T$ and $\mathbf{A}^T\mathbf{b}$ are

$$\mathbf{A}^T\mathbf{A}\mathbf{A}^T = \begin{bmatrix} -10 & -4 & -6 & 6 & 4 & 10 \\ -16 & 14 & -30 & 30 & -14 & 16 \\ 26 & -10 & 36 & -36 & 10 & -26 \\ 68 & -34 & 102 & -102 & 34 & -68 \end{bmatrix} \qquad \mathbf{A}^T\mathbf{b} = \begin{bmatrix} 8 \\ -2 \\ -6 \\ -10 \end{bmatrix}$$

We should point out that $\mathbf{A}^T\mathbf{b}$ does not have the appearance of being the projection of vector **b** which is in a six-dimensional space on the range of **A.** What happens is that **b** is represented with respect to the column vectors of matrix **A** which has four columns, therefore, there are four coefficients. Because the vectors are not linearly independent, the representation is not unique. We shoud not worry, however, because we have subjected the left side vector **Ax** to the same transformation, making the equality between both sides valid. **.** Recall that the rightmost $\mathbf{A}^T$ has to do with the fact that we have made the transformation $\mathbf{x} = \mathbf{A}^T\xi$**.** When we augment the matrix with the vector and row-reduce the augmented matrix, we obtain

$$[\mathbf{A}^T\mathbf{A}\mathbf{A}^T \mid \mathbf{A}^T\mathbf{b}]_{\text{row reduced}} = \begin{bmatrix} 1 & 0 & 1 & -1 & 0 & -1 & -26/51 \\ 0 & 1 & -1 & 1 & -1 & 0 & -37/51 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

From the row-reduced matrix we obtain an obvious solution (which we will call “the natural solution.”)

$$\xi = \text{Column } [-26/51 \quad -37/51 \quad 0 \quad 0 \quad 0 \quad 0]$$

There are many other solutions given by $\xi$ = Column [$\xi_1$ $\xi_2$ $\xi_3$ $\xi_4$ $\xi_5$ $\xi_6$] with the components $\xi_i$ satisfying

$$\xi_1 = -26/51 - \xi_3 + \xi_4 + \xi_6$$
$$\xi_2 = 13/102 + \xi_3 - \xi_4 + \xi_5$$

where the values of $\xi_3$, $\xi_4$, $\xi_5$ and $\xi_6$ are arbitrary. We see that the solution for $\xi$ is undetermined, however, when we return to the original variable **x,** all the solutions collapse into one. For example, if we take the natural solution $\xi_n$ (all the arbitrary components are zero), we have

$$\mathbf{x}_n = \mathbf{A}^T \xi_n = \text{Column } [\ 21/17 \quad -37/51 \quad -26/51 \quad -5/17\ ]$$

Let us now take another solution $\xi_m$ by choosing $\xi_3 = 1$, $\xi_4 = -1$, $\xi_5 = 0$, $\xi_6 = 2$. The corresponding vector is

$$\xi_m = \text{Column } [\ -26/51 \quad 65/51 \quad 1 \quad -1 \quad 0 \quad 2]$$

and the corresponding $\mathbf{x}_m$ is

$$\mathbf{x}_m = \mathbf{A}^T \xi_m = \text{Column } [\ 21/17 \quad -37/51 \quad -26/51 \quad -5/17]$$

which coincides with the previous answer. The same thing would happen with any solution to the linear system that we choose. The reason that the answers are the same is that the general solution obtained from the row reduction consists of a particular solution plus any solution in the null space of the matrix that was row reduced (which coincides with the null space of **A.**) When any of those solutions is multiplied by $\mathbf{A}^T$, the resulting vector will have its portion in the null space of **A** annihilated (including the portion in the null space of **A** that the particular solution may have.) Thus, the result is a vector in the orthogonal complement of N(A) which is $R(A^T)$. The vector is unique and the shortest of all the solutions. (See Zadeh and Desoer [5], Theorem 6, p. C.15.)

Our conclusion is that it is not necessary to have matrices **P** and **Q** in Equations (8) and (9) that have linearly independent vectors as long as, instead of inverting matrix $\Phi$, what we do is obtaining any solution given by matrix $\mathrm{X}$ to the following problem with many right sides

$$(\mathbf{A}^T\mathbf{A}\mathbf{A}^T)\mathrm{X} = \mathbf{A}^T, \qquad \mathbf{A}^+ = \mathbf{A}^T\mathrm{X} \tag{11}$$

The right hand side of the equation on the left is obtained by letting the **b** vector adopt all the values of the columns of the unit matrix so that the resulting $\mathrm{X}$ when multiplied by matrix $\mathbf{A}^T$ produces the pseudo inverse. Equations (11) is the formula we were seeking for calculating the pseudo inverse of an arbitrary $m \times n$ matrix containing only operators **A** and $\mathbf{A}^T$.

The first of the equations (11) can be solved by adjoining to the matrix $\mathbf{A}^T\mathbf{A}\mathbf{A}^T$ the matrix $\mathbf{A}^T$; row-reducing the adjoined matrix, considering only the natural solution (making zero all the variables not belonging to unit columns) and, according to the second equation (11) , multiplying the resulting matrix (adding the zero rows that are necessary to really have a solution, not only copying the result in the reduced matrix $\mathbf{A}^T$ that was adjoined) by $\mathbf{A}^T$.

## 8. NUMERICAL EXAMPLE ILLUSTRATING THE USE OF EQUATION (11)

A numerical example using the same matrix **A** follows: From a previous example we already have matrix $\mathbf{A}^T\mathbf{A}\mathbf{A}^T$. When matrix $\mathbf{A}^T$ is adjoined on the right, we have matrix

$$[\ \mathbf{A}^T\mathbf{A}\mathbf{A}^T \,|\, \mathbf{A}^T\ ] = \left[\begin{array}{cccccc:cccccc} -10 & -4 & -6 & 6 & 4 & 10 & -1 & -1 & 0 & 0 & 1 & 1 \\ -16 & 14 & -30 & 30 & -14 & 16 & 0 & 1 & -1 & 1 & -1 & 0 \\ 26 & -10 & 36 & -36 & 10 & -26 & 1 & 0 & 1 & -1 & 0 & -1 \\ 68 & -34 & 102 & -102 & 34 & -68 & 2 & -1 & 3 & -3 & 1 & -2 \end{array}\right]$$

When this matrix is row-reduced, we obtain

$$[\mathbf{A}^T\mathbf{A}\mathbf{A}^T \mid \mathbf{A}^T]_{row-reduced} = \begin{bmatrix} 1 & 0 & 1 & -1 & 0 & -1 & \frac{7}{12} & \frac{5}{102} & \frac{1}{51} & -\frac{1}{51} & -\frac{5}{102} & -\frac{7}{102} \\ 0 & 1 & -1 & 1 & -1 & 0 & \frac{4}{51} & \frac{13}{102} & -\frac{5}{102} & \frac{5}{102} & -\frac{13}{102} & -\frac{4}{51} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

from which we obtain $\mathrm{X}$**,** adding two rows of zeros to the last 6 columns of the last matrix.

$$\mathrm{X} = \begin{bmatrix} \frac{7}{102} & \frac{5}{102} & \frac{1}{51} & -\frac{1}{51} & -\frac{5}{102} & -\frac{7}{102} \\ \frac{4}{51} & \frac{13}{102} & -\frac{5}{102} & \frac{5}{102} & -\frac{13}{102} & -\frac{4}{51} \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and when $\mathrm{X}$ is pre-multiplied by $\mathbf{A}^T$ ,we get the pseudo inverse $\mathbf{A}^+$, which coincides with the one obtained before.

## 9. MAIN RESULT

From the previous discussion we can make the following statement:

*If matrix **P** is formed with columns which include a set of linear independent vectors spanning the space of the columns of **A**, and matrix **Q** is formed in the same way but with columns that span the space of the rows of **A**, then the pseudo inverse $\mathbf{A}^+$ of **A** can be calculated by*

$\mathbf{A}^+ = \mathbf{Q}X$, *where matrix $X$ is any solution of the matrix equation* $\mathbf{P}^T\mathbf{A}\mathbf{Q}X=\mathbf{P}^T$ (12)

## 10. NUMERICAL EXAMPLE APPLYING THE MAIN RESULT

A simple example illustrates the statement. We again use matrix **A** employed before. The first two columns of **A** are independent as are the first two rows, if we include them in **P** and **Q** and add the 4$^{th}$, column and row, we get the following **P** and **Q** matrices

$$\mathbf{P} = \begin{bmatrix} -1 & 0 & 2 \\ -1 & 1 & -1 \\ 0 & -1 & 3 \\ 0 & 1 & -3 \\ 1 & -1 & 1 \\ 1 & 0 & -2 \end{bmatrix}, \qquad \mathbf{Q} = \begin{bmatrix} -1 & -1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & -1 \\ 2 & -1 & -3 \end{bmatrix}$$

Solving the equation $(\mathbf{P}^T\mathbf{A}\mathbf{Q})\mathrm{X} = \mathbf{P}^T$ for **X** and using the "natural solution" obtained by row-reduction we get

$$\mathrm{X}_{\text{natural}} = \begin{bmatrix} \frac{7}{102} & \frac{5}{102} & \frac{1}{51} & -\frac{1}{51} & -\frac{5}{102} & -\frac{7}{102} \\ \frac{4}{51} & \frac{13}{102} & -\frac{5}{102} & \frac{5}{102} & -\frac{13}{102} & -\frac{4}{51} \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

The product $\mathbf{Q}\mathrm{X}_{\text{natural}}$ gives the pseudo inverse $\mathbf{A}^{+}$, which coincides with the ones obtained before.

## 11. MATHEMATICA PROGRAMS

We now supply *Mathematica* programs to calculate the least-square/minimum- length solution when $k < n < m$ for a single right side and for the calculation of the pseudo inverse $\mathbf{A}^{+}$.( For the case $k < m < n$, one can calculate the pseudo inverse of the transpose and transpose the result)

For a matrix **A** and right side vector **b** (which we assume have already been entered into *Mathematica* in the arrays a and b, respectively. The solution is stored in the array x.)

```
at = Transpose[a]; x = at.LinearSolve[at.a.at],at.b]
```

For the calculation of the pseudo inverse which after execution is displayed and stored in the array psinv (we assume matrix A has been input; this is not shown.)

```
at=Transpose[a];
M=at.a.at;
DD={};Do[DD=Join[DD,Join[M[[i]],at[[i]]]],{i,1,Dimensions[a]
[[2]]}];
DD=Partition[DD,2 Dimensions[a][[1]]];
DDrr=RowReduce[DD];
p=DDrr[[Range[1,Dimensions[a][[2]]],Range[Dimensions[a][[1]]+1,2 Dimensions[a][[1]]]]];
p0=Join[p,Table[0,{Dimensions[a][[1]]-Dimensions[a][[2]]},
{Dimensions[a][[1]]}]];
psinv=at.p0
```

The explanation to the listing is as follows: The first line calculates the transpose of **A** and calls it at. In the next line, product $\mathbf{A}^{T}\mathbf{A}\mathbf{A}^{T}$ is calculated and called M. The next two lines do the job of setting up the partitioned matrix $[\mathbf{A}^{T}\mathbf{A}\mathbf{A}^{T} \mid \mathbf{A}^{T}]$ by extending each row of M with the corresponding row of at, and then declaring that a new matrix with rows with a number of componentes twice the number of columns of $\mathbf{A}^{T}\mathbf{A}\mathbf{A}^{T}$ (which is the same as twice the number of columns of $\mathbf{A}^{T}$) components is to be called DD. In the next line DD is row-reduced and called DDrr. In the next line, we take the piece of DDrr including rows 1 to the number of rows of A, and columns from 1 + the number of rows of **A** to twice the number of rows of **A** and call that matrix p. In the next line, we append to the bottom of p a matrix with a number of rows equal to the difference between the number of rows and columns of **A** and a number of columns of zeros equal to the number of rows of **A**, and call it p0. In the last line, we define the matrix psinv with the product of matrices $\mathbf{A}^{T}$ and p0. No output is displayed for all the lines that end with a ";". But because the last line does not end in a ";" it displays the output of the last operation and, thus, displays the pseudo inverse of **A**.

It should be mentioned that *Mathematica* has a primitive function PseudoInverse, thus the listings shown are for the purpose of formally documenting the method given in the paper.

## 12. CONCLUSION

We have presented additional methods of calculating the Moore – Penrose Pseudo Inverse of a general matrix, (Moore [8], Penrose [9]), particularly useful when the matrix is not full rank. The methods presented are extensions of the ones presented by Murray – Lasso [4] and the normal equations (Davis [9]); the last of which is valid for full rank matrices. The author hopes the methods presented can be generalized to linear operators in infinite dimensional spaces (Lanczos [10]) where in some cases it is not easy to distinguish rows and columns but the linear operator and its adjoint may be available. Several numerical examples are provided to clarify the discussion and *Mathematica* programs are provided.

## 13. REFERENCES

[1] Dahlquist, G. & A. Bjork (1974): Numerical Methods. Prentice – Hall, Inc., Englewood Cliffs, NJ.

[2] Golub, G. H. & C. Reinsch (1971): “Singular Value Decomposition and Least Square Solutions.” Contribution 1/10 in Wilkinson, J. H. & C. Reinsch: Handbook for Automatic Computation. Volume II, Linear Algebra. Springer – Verlag, New York, pp. 134 – 151.

[3] Noble, B. (1969): Applied Linear Algebra. Prentice – Hall, Inc., Englewood Cliffs, NJ.

[4] 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. .

[5] Zadeh, L. & C. Desoer (1963): Linear System Theory: The State Space Approach. Mc Graw – Hill Book Company, Inc., New York

[6] Lanczos, C. (1961): Linear Differential Operators. D. Van Nostrand Company Limited, Londres.

[7] Wolfram, S. (1991): Mathematica: A System for Doing Mathematics by Computer. Second Edition. Addison – Wesley Publishing Company, Inc., Redwood City, CA.

[8] Moore, E. H. (1920): “On the Reciprocal of the General Algebraic Matrix.” Bulletin of the American Mathematical Society, Vol. 26, pp. 394 – 395.

[9] Penrose, R. (1955): “A Generalized Inverse for Matrices.” Proceedings of the Cambridge Philosophical Society, Vol. 51, pp. 406 – 413.

[10] Davis, P. J. (1963): Interpolation and Approximation. Blaisdell Publishing Company, New York.