

Applications of the Moore-Penrose Generalized Inverse to Linear Systems of Algebraic Equations

Asmaa Mohammed Kanan^{1,*}, Asma Ali Elbeleze², Afaf Abubaker²

¹Mathematics Department, Science Faculty, Sabratha University, Sabratha, Libya

²Mathematics Department, Science Faculty, Zawia University, Zawia, Libya

Email address:

asmaakanan20@gmail.com (A. M. Kanan)

*Corresponding author

To cite this article:

Asmaa Mohammed Kanan, Asma Ali Elbeleze, Afaf Abubaker. Applications of the Moore-Penrose Generalized Inverse to Linear Systems of Algebraic Equations. *American Journal of Applied Mathematics*. Vol. 7, No. 6, 2019, pp. 152-156. doi: 10.11648/j.ajam.20190706.11

Received: October 4, 2019; Accepted: October 30, 2019; Published: November 18, 2019

Abstract: In this work, we consider linear systems of algebraic equations. These systems are studied utilizing the theory of the Moore-Penrose generalized inverse or shortly (MPGI) of matrices. Some important algorithms and theorems for computation the MPGI of matrices are given. The singular value decomposition (SVD) of a matrix has a very important role in computation the MPGI, hence it is useful to study the solutions of over- and under-determined linear systems. We use the MPGI of matrices to solve linear systems of algebraic equations when the coefficients matrix is singular or rectangular. The relationship between the MPGI and the minimal least squares solutions to the linear system is expressed by theorem. The solution of the linear system using the MPGI is often an approximate unique solution, but for some cases we can get an exact unique solution. We treat the linear algebraic system as an algebraic equation with coefficients matrix A (square or rectangular) with complex entries. A closed form for solution of linear system of algebraic equations is given when the coefficients matrix is of full rank or is not of full rank, singular square matrix or non-square matrix. The results are taken from the works mentioned in the references. A few examples including linear systems with coefficients matrix of full rank and not of full rank are provided to show our studding.

Keywords: Linear Algebraic Systems, MPGI, Rank, SVD, Least Squares Solutions

1. Introduction

The system of equations

$$Ax = b; A \in \mathbb{C}^{m \times n}, x \in \mathbb{C}^n, b \in \mathbb{C}^m \quad (1)$$

has been exploited in several contexts. It is consistent, i.e., has a solution for x if and only if b is in the range of A , if b is not in the range of A then $b - Ax$ is nonzero for all $x \in \mathbb{C}^n$, so we will go to find an approximate solution of (1), i.e., we want to find a vector x making $b - Ax$ closest to zero. If the coefficients matrix A is square, that means, the number of equations is equal to the number of unknowns, in this a case A either is of full rank or is not of full rank. If A is square and of full rank, that means, A is invertible, so there exists unique solution of the system $Ax = b$. If A is an arbitrary matrix in $\mathbb{C}^{m \times n}$, then x is a vector with n components and b is a vector with m components. If m is greater than n , that means, there more equations than unknowns, in this a case (1)

is called over-determined system and has no solution. Conversely, if n is greater than m , that means, there more unknowns than equations, in this a case (1) is called under-determined system and has infinite number of solution. Even though $m > n$ or $n > m$ the linear system (1) still has a natural unique solution, it's called "least squares solution" [1-5]. The concept of the Moore-Penrose generalized inverse of matrices has been explained in many references [1, 6-12]. We will use the MPGI to solve the linear systems of algebraic equations $Ax = b$ with coefficients matrix A [13-16].

A solution utilizing the MPGI is the minimal least squares solution. When b belongs to the range of A ($b \in R(A)$) then the notions of solutions using the MPGI and least squares solution coincide.

In this paper, A^* is the transposed conjugate complex of A . If A is a real matrix then A^* reduces to the transposed A^T . The trace of A is denoted by $Tr(A)$, the zero matrix is denoted by O . I always denotes an identity matrix.

2. Significance of the Study

The concept of the Moore-Penrose generalized inverse of matrix has been exploited recently in many contexts. It presents solutions for many singular problems in linear algebra, especially for linear singular algebraic systems. The contribution of this study is that:

- 1) Explaining the relationship between the MPGI and the least squares solutions.
- 2) Distinguish the importance of using the MPGI to solve the linear systems of an algebraic equations.
- 3) It gives clue to depth study of the MPGI to get an exact solutions instead an approximate solutions.
- 4) It invites other researchers interested to find other forms for the MPGI and using them to solve the linear algebraic systems, and hence getting unique exact solutions instead approximate solutions.

3. The Moore-Penrose Generalized Inverse

In this section, we introduction some of important definitions, algorithms, and theorems.

Definitions 3.1:

If $A \in \mathbb{C}^{m \times n}$, then $A^\dagger \in \mathbb{C}^{n \times m}$ is unique, and it is called the Moore-Penrose generalized inverse of A if it satisfies the following condition:

- 1) $AA^\dagger A = A$
- 2) $A^\dagger AA^\dagger = A^\dagger$
- 3) $(AA^\dagger)^* = AA^\dagger$
- 4) $(A^\dagger A)^* = A^\dagger A$.

The next theorem is useful to compute A^\dagger when A has a full rank factorization.

Theorem 3.1:

If $A = BC$ where $A \in \mathbb{C}^{m \times n}$, $B \in \mathbb{C}^{m \times r}$, $C \in \mathbb{C}^{r \times n}$, and $r = \text{rank}(A) = \text{rank}(B) = \text{rank}(C)$, then

$$A^\dagger = C^*(CC^*)^{-1}(B^*B)^{-1}B^*.$$

Theorem 3.2:

If $A \in \mathbb{C}^{m \times n}$ and $\text{rank}(A) = 1$, then

$$A^\dagger = \frac{1}{\alpha} A^*$$

where

$$\alpha = \text{Tr}(A^*A) = \sum_{i,j} |a_{ij}|^2.$$

Definitions 3.2:

A matrix $E \in \mathbb{C}^{m \times n}$ which has rank r is said to be in row echelon form if E is of the form

$$E = \begin{bmatrix} C_{r \times n} \\ \dots \\ O_{(m-r) \times n} \end{bmatrix},$$

where the elements c_{ij} of $C(=C_{r \times n})$ satisfy the following conditions:

- 1) $c_{ij} = 0$ when $i > j$.
- 2) The first non-zero entry in each row of C is 1.
- 3) If $c_{ij} = 1$ is the first non-zero entry of the i th row then the j th column of C is the unit vector e_i whose only non-zero entry is in the i th position.

Definitions 3.3:

A matrix $G \in \mathbb{C}^{n \times n}$ is said to be in hermite echelon form if its elements g_{ij} satisfies the following conditions:

- 1) G is upper triangular (i.e., $g_{ij} = 0$ when $i > j$).
- 2) g_{ii} is either 0 or 1.
- 3) If $g_{ii} = 0$ then $g_{ik} = 0$ for every k , $1 \leq k \leq n$.
- 4) If $g_{ii} = 1$ then $g_{ki} = 0$ for every $k \neq i$.

Note that, for $A \in \mathbb{C}^{n \times n}$ can always be row reduced to a hermite form G_A by using elementary row operations on A .

Algorithm 3.1:

To obtain the MPGI of a square matrix:

- 1) Row reduce A^* to its hermite form G_{A^*} (or its hermite echelon form G_{A^*}).
- 2) Select the distinguished columns of A^* (they are the columns which correspond the columns e_1, e_2, \dots , and e_r in G_{A^*} which are the unit vectors, where $r = \text{rank}(A^*)$). Label these columns v_1, v_2, \dots, v_r and place them as columns in a matrix L .
- 3) Form the matrix AL .
- 4) Form $I - G_{A^*}$ and select the non-zero columns from this matrix. Label these columns w_1, w_2, \dots, w_{n-r} .
- 5) Place the columns of AL and the w_i as columns in a matrix $M = [AL : w_1 : w_2 : \dots : w_{n-r}]$ and compute M^{-1} .
- 6) Place the first r rows of M^{-1} (in the same order as they in M^{-1}) in a matrix called R .
- 7) Compute A^\dagger as $A^\dagger = LR$.

Note that, it is easy to use this algorithm for non-square matrices. That is by adding zero rows or zero columns to construct a square matrix, after that find MPGI for the result matrix $[A : O]$, where

$$[A : O]^\dagger = \begin{bmatrix} A^\dagger \\ \dots \\ O^* \end{bmatrix}, \begin{bmatrix} A^\dagger \\ \dots \\ O^* \end{bmatrix}^\dagger = [A^\dagger : O^*].$$

Algorithm 3.2:

To obtain the full rank factorization and the MPGI for $A \in \mathbb{C}^{m \times n}$:

- 1) Reduce A to row echelon form E_A .
- 2) Select the distinguished columns of A (they are the columns which correspond the columns e_1, e_2, \dots, e_r in E_A) and place them as the columns in a matrix B in the same order as they appear in A .
- 3) Select the non-zero rows form E_A and place them as rows in a matrix C in the same order as they appear in E_A .
- 4) Compute $(CC^*)^{-1}$ and $(B^*B)^{-1}$.
- 5) Compute A^\dagger as $A^\dagger = C^*(CC^*)^{-1}(B^*B)^{-1}B^*$.

4. Least Squares Solutions

The name "least squares" comes from the definition of the

Euclidean norm [1] which is the square root of a sum of squares. That is, the Euclidean norm of $u = [u_1, \dots, u_p]^* \in \mathbb{C}^p$ is denoted by $\|u\|$, where

$$\|u\| = \left(\sum_{i=1}^p |u_i|^2 \right)^{\frac{1}{2}} = (u^* u)^{\frac{1}{2}}.$$

We consider the problem of finding solution w to (1). If (1) is inconsistent, we have to find w that makes $Aw - b$ as small as possible.

Definitions 4.1:

Suppose that $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^m$. Then a vector $w \in \mathbb{C}^n$ is called least squares solution to (1), if $\|Aw - b\| \leq \|Av - b\|$ for all $v \in \mathbb{C}^n$. A vector w is called minimal least squares solution to (1) if w is least squares solution to (1) and $\|w\| < \|u\|$ for all other least squares solutions u .

Next, we have an important theorem. It relates minimal least squares solutions of (1) and the MPGI of A , it also answers our question: what kind of answer is $A^\dagger b$?

Theorem 4.1:

Suppose that $A \in \mathbb{C}^{m \times n}$ and $b \in \mathbb{C}^m$. Then $A^\dagger b$ is the minimal least squares solution to $Ax = b$.

Note that, the minimal least squares solutions $A^\dagger b$ is called the approximate solution to (1) [3]. If $X \in \mathbb{C}^{n \times m}$, $Ax = b$ has a solution $x = Xb$ is the minimal least squares solution then $X = A^\dagger$.

5. Solutions of Linear Algebraic System

This section deals with the linear algebraic systems and their solution techniques. The solution of these systems is the unknown vector $x \in \mathbb{C}^n$ which satisfies those systems, but in some cases, we can not find such x , so we search for the $w \in \mathbb{C}^n$ that makes Aw is closest to b . In other words, we search for the $w \in \mathbb{C}^n$ that makes $\|b - Aw\|$ as small as possible [2].

5.1. Linear Systems with Coefficients Matrix Is of Full Rank

In this kind of linear systems, we study three cases:

- 1) $A \in \mathbb{C}^{n \times n}$ is square matrix of full rank [17], that is $\text{rank}(A) = m = n$. In this a case, the unique solution of $Ax = b$ is

$$x = A^\dagger b, A^\dagger = A^{-1} \quad (2)$$

- 2) $A \in \mathbb{C}^{m \times n}$ is non-square (rectangular) matrix, $\text{rank}(A) = n < m$, this means, A is of full column rank, so it has left inverse A^\dagger satisfies $A^\dagger A \simeq I$, $A A^\dagger \neq I$. In this a case there is unique solution to (1), it is the least squares solution which is on the form:

$$x = A^\dagger b, A^\dagger = (A^* A)^{-1} A^* \quad (3)$$

- 3) $A \in \mathbb{C}^{m \times n}$ is also non-square matrix but $\text{rank}(A) = m < n$, this means, A is of full row rank and has right inverse satisfies $AA^\dagger \simeq I$, $A^\dagger A \neq I$. In this a case there is unique solution to (1) is the least squares solution, it is:

$$x = A^\dagger b, A^\dagger = A^* (AA^*)^{-1} \quad (4)$$

5.2. Linear Systems with Coefficients Matrix Is Not of Full Rank

When the coefficients matrix A is not of full rank (square or non-square), we can not use the last forms (in subsection 5.1) of A^\dagger . There are other forms for A^\dagger but the solution still $x = A^\dagger b$. These forms are:

A. The first form:

$$A^\dagger = V \Sigma^\dagger U^*$$

We use this form when A has a singular value decomposition [5, 20], that is $A = U \Sigma V^*$, where $U \in \mathbb{C}^{m \times m}$ and $V \in \mathbb{C}^{n \times n}$ are unitary matrices [18], $\Sigma \in \mathbb{C}^{m \times n}$ is a diagonal matrix. Note that, this method is slow.

B. The second form:

$$A^\dagger = \frac{1}{\alpha} A^*, \alpha = \text{Tr}(A^* A)$$

This form uses when $\text{rank}(A) = 1$, $A \in \mathbb{C}^{m \times n}$.

C. The third form:

$$A^\dagger = C^* (CC^*)^{-1} (B^* B)^{-1} B^*$$

There are some conditions to use this form, they are:

- 1) $A = BC$, where $B \in \mathbb{C}^{m \times r}$ and $C \in \mathbb{C}^{r \times n}$.
- 2) $\text{rank}(A) = \text{rank}(B) = \text{rank}(C)$.

Since, we have the equation

$$Ax = b$$

multiply by A^\dagger :

$$A^\dagger Ax = A^\dagger b$$

from the last equation, we get a new linear equation

$$A'x = b'$$

when we solve this equation we will get multiple solutions using arbitrary parameter (parameters), if we choose that parameter to equal zero, then we will get an exact solution. In other words, x satisfies $\|Ax - b\| = 0$ or $Ax = b$ [13]. Note that, the solution we get using the third form is an exact solution.

6. Numerical Examples

In this section, we give some examples to show our work.

Example 6.1 consider the system

$$x_1 + 3x_2 = 17$$

$$5x_1 + 7x_2 = 19$$

$$11x_1 + 13x_2 = 23$$

We can write this system as in (1), where

$$A = \begin{pmatrix} 1 & 3 \\ 5 & 7 \\ 11 & 13 \end{pmatrix} \in \mathbb{C}^{3 \times 2}, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, b = \begin{pmatrix} 17 \\ 19 \\ 23 \end{pmatrix}.$$

Since $\text{rank}(A) = 2 = n$, then A is of full column rank, in this case A has a left inverse is

$$A^\dagger = (A^*A)^{-1}A^* \\ = \begin{pmatrix} -0.5197 & -0.2169 & 0.2373 \\ 0.4277 & 0.2041 & -0.1313 \end{pmatrix}.$$

We know that the solution for this a case is on the form

$$x = A^\dagger b$$

that is

$$x \simeq \begin{pmatrix} -7.5 \\ 8.13 \end{pmatrix}.$$

Note that, such solution is an approximate solution, it is the least squares solution.

Example 6.2 Let us consider the system

$$x_1 - 2x_2 = 5$$

$$-3x_1 + 6x_2 = -15$$

Using matrix-vector notation, this system is written $x = Ab$, where

$$A = \begin{pmatrix} 1 & -2 \\ -3 & 6 \end{pmatrix} \in \mathbb{C}^{2 \times 2}, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, b = \begin{pmatrix} 5 \\ -15 \end{pmatrix}$$

Note that, $\text{rank}(A) = 1$, A is not of full rank, so we can use one of forms of A^\dagger explained in subsection 5. 2, the solution will be the same for any form. Let us use the first form

$$A^\dagger = V \Sigma^\dagger U^*$$

The eigenvalues of matrix $A^\dagger A$ [18] are $\lambda_1 = 50$ and $\lambda_2 = 0$. The singular values of matrix A are $\sigma_1 = 5\sqrt{2}$ and $\sigma_2 = 0$, hence

$$\Sigma = \begin{pmatrix} 5\sqrt{2} & 0 \\ 0 & 0 \end{pmatrix}, V = \begin{pmatrix} \frac{-1}{\sqrt{5}} & \frac{2}{\sqrt{5}} \\ \frac{2}{\sqrt{5}} & \frac{1}{\sqrt{5}} \end{pmatrix}, U = \begin{pmatrix} \frac{-1}{\sqrt{10}} & 0 \\ \frac{3}{\sqrt{10}} & 0 \end{pmatrix}.$$

Note that $A = U\Sigma V^*$, hence $A^\dagger = V \Sigma^\dagger U^*$. That is

$$A^\dagger = \frac{1}{50} \begin{pmatrix} 1 & -3 \\ -2 & 6 \end{pmatrix}.$$

The solution to the given system is

$$x = \begin{pmatrix} \frac{1}{50} & \frac{-3}{50} \\ \frac{-2}{50} & \frac{6}{50} \end{pmatrix} \begin{pmatrix} 5 \\ -15 \end{pmatrix} = \begin{pmatrix} 1 \\ -2 \end{pmatrix}.$$

Note that, the solution using the SVD is an exact solution.

Example 6.3 In this example we will show how to find the solution of algebraic system using the third form and get exact solution. Let us consider the following system:

$$2x_1 + 2x_2 = 2$$

$$x_1 + x_2 = 1$$

$$3x_1 + 3x_2 = 3$$

As before you can note that

$$A = \begin{pmatrix} 2 & 2 \\ 1 & 1 \\ 3 & 3 \end{pmatrix} \in \mathbb{C}^{3 \times 2}, x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix}, b = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}.$$

Since $\text{rank}(A) = 1$, then A is not of full rank, we use

$$A^\dagger = C^*(CC^*)^{-1}(B^*B)^{-1}B^*, A = BC$$

and use algorithm 3.2 to obtain:

$$B = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix}, C = (1 \quad 1)$$

So

$$A^\dagger = \frac{1}{28} \begin{pmatrix} 2 & 1 & 3 \\ 2 & 1 & 3 \end{pmatrix}.$$

We have the system

$$\begin{pmatrix} 2 & 2 \\ 1 & 1 \\ 3 & 3 \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 2 \\ 1 \\ 3 \end{pmatrix},$$

Multiplying the last system on the left by A^\dagger gives, the corresponding system $A^\dagger Ax = A^\dagger b$, which reduces to

$$\begin{pmatrix} \frac{14}{28} & \frac{14}{28} \\ \frac{14}{28} & \frac{14}{28} \\ \frac{14}{28} & \frac{14}{28} \end{pmatrix} \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} \frac{14}{28} \\ \frac{14}{28} \\ \frac{14}{28} \end{pmatrix},$$

this is a new system is denoted by $A'x = b'$, the solution to such system is

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1-t \\ t \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix} + t \begin{pmatrix} -1 \\ 1 \end{pmatrix},$$

where t is an arbitrary parameter. The solution $x = (x_1 \ x_2)^T$ for which Ax is equal to b is by taking the parameter $t = 0$, i.e.,

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \end{pmatrix},$$

Note that, the way we used in this example gives an exact solution.

7. Conclusion

Our study shows the importance of using the MPGI of the matrices for solving the linear system of algebraic equations $Ax = b$. The relationship between the MPGI and the least squares solutions was given. Our study also explains how to solve the linear algebraic system using the MPGI, in other words, which the form of the MPGI we can choose, where we proposed of the some forms for the MPGI of a matrix A when

A is of full rank and is not of full rank. We explained how to use the third form to obtain an exact solution. An approximate solution was often used.

References

- [1] S. L. Campbell, C. D. Meyer, Jr., Generalized Inverses of Linear Transformations, Pitman, London, 1979.
- [2] W. W. Hager, Applied Numerical Linear Algebra, Prentice Hall, USA, 1988.
- [3] A. Ben-Israel, T. N. E. Greville, Generalized Inverses: Theory and Applications, Springer-Verlage, New York (2003).
- [4] G. H. Golub, Least squares, Singular values and matrix approximations, Aplikace matematiky, 13 (1968) 44-51.
- [5] G. H. Golub, C. Reinsch, Singular value decomposition and least squares solutions, Numer. Math., 14 (1970) 403-420.
- [6] G. Bahadur-Thapa, P. Lam-Estrada, J. Lopez-Bonilla, On the Moore-Penrose Generalized Inverse Matrix, World Scientific News, 95 (2018) 100-110.
- [7] M. Zuhair Nashed (Ed.), Generalized inverses and applications, Academic Press, New York, (1976).
- [8] R. Penrose, A generalized inverse for matrices, Proc. Camb. Phil. Soc., 51 (1955) 406-413.
- [9] H. Ozden, A Note On The Use Of Generalized Inverse Of Matrices In Statistic, Istanbul Univ. Fen Fak. Mat. Der., 49 (1990) 39-43.
- [10] A. Ben-Israel, Generalized Inverses of Matrices and Their Applications, Springer, (1980) 154-186.
- [11] C. E. Langenhop, On Generalized Inverse of Matrices, SIAM J. Appl. Math., 15 (1967) 1239-1246.
- [12] C. R. Rao, S. K. Mitra, Generalized Inverse of a Matrix and Its Applications, New York: Wiley, (1971) 601-620.
- [13] A. M. Kanan, Solving the Systems of Linear Equations Using the Moore-Penrose Generalized Inverse, Journal Massarat Elmeiya, Part 1 (2) (2017) 3-8.
- [14] J. Lopez-Bonilla, R. Lopez-vazquez, S. Vidal-Beltran, Moore-Penrose's inverse and solutions of linear systems, World Scientific News, 101 (2018) 246-252.
- [15] Abu-Saman, A. M., Solution of Linearly-Dependent Equations by Generalized Inverse of Matrices, Int. J. Sci. Emerging Tech, 4 (2012) 138-142.
- [16] J. Z. Hearon, Generalized inverses and solutions of linear systems, Journal of research of notional Bureau of standards-B. Mathematical Sciences, 72B (1968) 303-308.
- [17] D. H. Griffel, Linear Algebra and its Applications, vol. 1, first course.
- [18] F. Chatelin, Eigenvalues of Matrices, WILEY, France, 1807.
- [19] T. N. E. Greville, The Pseudoinverse of a rectangular singular matrix and its application to the solution of systems of linear equations, SIAM Rev., 1 (1960) 38-43.
- [20] G. H. Golub, W. Kahan, Calculating the singular values and pseudo-inverse of a matrix, SIAM J. Numer. Anal., 2 (B) (1965) 205-224.