1 Introduction

The inverse of a matrix A can only exist if A is nonsingular. This is an important theorem in linear algebra, one learned in an introductory course. In recent years, needs have been felt in numerous areas of applied mathematics for some kind of inverse like matrix of a matrix that is singular or even rectangular. To fulfill this need, mathematicians discovered that even if a matrix was not invertible, there is still either a left or right sided inverse of that matrix. A matrix $A \in \mathbb{C}^{m\times n}$ is left invertible (right invertible) so that there is a matrix $L(R) \in \mathbb{C}^{n\times m}$ so that

$$LA = I_n \qquad\qquad (AR = I_m)\,.$$

This property, where every matrix has some inverse-like matrix, is what gave way to the defining of the generalized inverse.

The generalized inverse has uses in areas such as inconsistent systems of least squares, properties dealing with eigenvalues and eigenvectors, and even statistics. Though the generalized inverse is generally not used, as it is supplanted through various restrictions to create various different generalized inverses for specific purposes, it is the foundation for any pseudoinverse. Arguably the most important generalized inverses is the Moore-Penrose inverse, or pseudoinverse, founded by two mathematicians, E.H. Moore in 1920 and Roger Penrose in 1955. Just as the generalized inverse the pseudoinverse allows mathematicians to construct an inverse like matrix for any matrix, but the pseudoinverse also yields a unique matrix. The pseudoinverse is what is so important, for example, when solving for inconsistent least square systems as it is constructed in a way that gives the minimum norm and therefore the closest solution.

2 Generalized Inverse

If A is any matrix, there is a generalized inverse, A^- such that,

$$AA^-A = A.$$

As mentioned before, this equation is extrapolated from the conjecture that any matrix has at least a one sided inverse. If we assume that A^- is equal to either L or R we see that

$$ALA = A(LA) = AI = A \qquad\qquad ARA = (AR)A = IA = A$$

If A is a $n\times m$ matrix though, A^- is then a $m\times n$ matrix, and the resultant identity matrix either has its rank equal to the columns or rows of A.It is obvious to point out as well that when $m = n$ and when $rank(A) = n$ then $A^- = A^{-1}$. There are other properties, some trivial, some interesting, but the most important part of the generalized inverse though, is that A^- is not unique.

3 Moore-Penrose Inverse

Definition 1. *If $A \in \mathbb{M}_{n,m}$, then there exists a unique $A^+ \in \mathbb{M}_{m,n}$ that satisfies the four Penrose conditions:*

1. $AA^+A = A$
2. $A^+AA^+ = A^+$
3. $A^+A = (A^+A)^*$ *Hermitian*
4. $AA^+ = (AA^+)^*$ *Hermitian*

Where M^* is the conjugate transpose of matrix M.

If A is nonsingular, it is clear that $A^+ = A^{-1}$ trivially satises the four equations. Since the pseudoinverse is known to be unique, which we prove shortly, it follows that the pseudoinverse of a nonsingular matrix is the same as the ordinary inverse.

Theorem 3.1. *For any $A \in \mathbb{C}_{n,m}$ there exists a $A^+ \in \mathbb{C}_{m,n}$ that satisfies the Penrose conditions.*

Proof. The proof of this existence theorem is lengthy and is not included here, but can be taken as conjecture. A version of the proof can be found in *Generalized Inverses: Theory and Applications* □

Theorem 3.2. *For a matrix $A \in \mathbb{M}_{n,m}$, then there exists a unique $A^+ \in \mathbb{M}_{m,n}$*

Proof. Suppose that there are two matrices, B and C that satisfy the four penrose conditions (1,2,3,4) so that

$$\begin{aligned} B &= BAB && (2) \\ &= (A^*B^*)B && (4) \\ &= (A^*C^*A^*)B^*B && (1) \\ &= (CA)(A^*B^*B) && (4) \\ &= CAB && (2) \end{aligned}$$

and

$$\begin{aligned} C &= CAC && (2) \\ &= C(C^*A^*) && (3) \\ &= CC^*(A^*B^*A^*) && (1) \\ &= (CA)(B) && (3) \\ &= CAB && (2). \end{aligned}$$

Therefore $B = C$. □

5 Pseudoinverse In Least Squares

The pseudoinverse is most often used to solve least squares systems using the equation $A\vec{x} = \vec{b}$. When $\vec{b}$ is in the range of A, there is at least one or more solutions to the system. If $\vec{b}$ is not in the range of A, then there are no solutions to the system, but it is still desirable to to find a $\vec{x_0}$ that is closest to a solution. The residual vector is a key component to solve these systems, and is given as $\vec{r} = A\vec{x} - \vec{b}$.

Definition 3. *The norm of a vector is written as* $||\vec{a}||$ *such that* $||\vec{a}|| = \sqrt{\vec{a}^2}$.

Definition 4. *A least squares solution to a system is a vector such that*

$$||\vec{r_0}|| = ||A\vec{x_0} - \vec{b}|| \leq ||A\vec{x} - \vec{b}||$$

The unique least squares solution is given when the $\vec{x_0}$ creates a minimum in the norm of the residual vector.

Theorem 5.1. $\vec{x_0} = A^+\vec{b}$ *is the best approximate solution of* $A\vec{x} = \vec{b}$.

Proof. For any $x \in \mathbb{C}^m$,

$$A\vec{x} - \vec{b} = A(\vec{x} - A^+\vec{b}) + (I - AA^+)(-\vec{b})$$

where $I - AA^+$ is an orthogonal projector onto $N(A^*)$, which by *corollary 1* of *theorem 3.3* we know is also a projector onto $N(A^+)$, then the summation on the right hand side is of orthogonal vectors. Using Pythagorean theorem with the norm, we can deduce that

$$\begin{aligned} ||A\vec{x} - \vec{b}||^2 &= ||A(\vec{x} - A^+\vec{b})||^2 + ||(I - AA^+)(-\vec{b})||^2 \\ &= ||A(\vec{x} - \vec{x_0})||^2 + ||A\vec{x_0} - \vec{b}||^2 \\ &\geq ||A\vec{x_0} - \vec{b}||^2. \end{aligned}$$

Now we can say that the norm of the residual vector is at its minimum when $\vec{x} = \vec{x_0}$. □

This theorem allows us to affirm that $A^+\vec{b}$ is either the unique least squares solution or is the least squares solution of minimum norm.

Example 3.

$$A = \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} \qquad \vec{b} = \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix}$$

Using the corollary 2 of theorem 3.3, $\vec{x_0} = A^+\vec{b} = (A^*A)^{-1}A^*\vec{b}$.

$$\vec{x_0} = A^+\vec{b} \qquad A^+ = \begin{bmatrix} 1/9 & 4/9 & -1/9 \\ 2/9 & -1/9 & 5/18 \end{bmatrix}$$

$$\begin{aligned} \vec{x_0} &= \begin{bmatrix} 1/9 & 4/9 & -1/9 \\ 2/9 & -1/9 & 5/18 \end{bmatrix} \begin{pmatrix} 1 \\ 2 \\ 3 \end{pmatrix} \\ &= \begin{pmatrix} 2/3 \\ 5/6 \end{pmatrix} \\ A\vec{x_0} &= \begin{bmatrix} 1 & 2 \\ 2 & 0 \\ 0 & 2 \end{bmatrix} \begin{pmatrix} 2/3 \\ 5/6 \end{pmatrix} \\ &= \begin{pmatrix} 7/3 \\ 4/3 \\ 5/3 \end{pmatrix} \end{aligned}$$

$$||\vec{r}|| = \begin{pmatrix} 7/3 & 4/3 & 5/3 \end{pmatrix} \begin{pmatrix} 7/3 \\ 4/3 \\ 5/3 \end{pmatrix} = \begin{pmatrix} 10 \end{pmatrix}$$