Generalized Inverses

Kashvi Srivastava

August 2018

Abstract

The following report contains the basic terminologies and concepts related to Generalized Inverses. Proven and unproven results from texts in the field have been restated and verified in layman language. The report also aims at giving an insight into the computational aspects and theorems regarding various types of generalized inverses. Conclusive proofs have been provided to some problems stated in the texts, 'Linear Algebra and Linear Models by Ravindra B. Bapat', 'Linear Algebra by A.Ramachandara Rao and P. Bhimasankaran' and 'Generalized Inverses by Adi Ben-Israel and Thomas N.E. Greville'.

Contents

1. History

Coining of the concept of *generalized inverses* dates back to 1903 when a particular generalized inverse, then known as pseudo-inverse, of an initial operator was given by **Fredholm**. Further, in 1912, **Hurwitz** characterized the class of all pseudo inverses and used the finite dimensionality of the null spaces of the Fredholm operators to give a simple algebraic construction. Generalized inverses of differential and integral operators led to the concept of generalized inverses of matrices.

E.H. Moore defined a 'general reciprocal' (unique inverse) for every finite matrix (square or rectangular) in his publication in 1920. Extensions of these ideas were made for matrices by **Siegel** in 1937 and for operators by **Tseng** in 1933. But no systematic study of the subject was made until 1955 when **Penrose** redefined the Moore inverse in a different way. These properties were also recognized by **Bjerhammar**, who rediscovered Moore's inverse. Thus, Penrose sharpened and extended Bjerhammar's results on linear systems and the important research publications therefore led to what is known as Moore-Penrose Inverse.

About the same time, **Rao** gave a method of computing pseudo inverses of singular matrices and applied it to normal equations to express the variances of estimators. These pseudo inverses defined by Rao did not satisfy all the restrictions imposed by Moore and Penrose. It was therefore different from the Moore-Penrose inverse. In a later paper, Rao showed that, to deal with problems of linear equations, a less defined inverse compared to Moore-Penrose inverse, is sufficient. Such an inverse was called generalized inverse. This g-inverse is not unique and thus opens the scope of extensive study in matrix algebra.

Since 1955, the subject experienced huge research and developments. The major contributors are **Greville, Ben-Israel and Charnes, Chipman, Scroggs and Odell**. Moreover, **Bose** mentions the use of g-inverse in his work , "Analysis of Variance". **Bott and Duffin** defined constrained inverse of a square matrix. Further applications were considered by **Chernoff, Mitra** and **Bhimasankaram**. The subject still provides a platform for research and findings. The applications of generalized inverses in statistics and mathematics have been extensively expanding ever since.

2. Elementary Properties

We are well-versed with the concept of inverse of a square matrix. Such a unique inverse exists if the matrix is non-singluar. Non-singularity of a square matrix A can be derived from the computation of determinant of the matrix i.e., $|A| = 0$ or the rank of the matrix i.e., $\rho(A) = n$ where n is the order of the matrix A. Many equivalent statements can derive the non-singularity of a square matirx.
Hence, unique inverse of such a matrix A is defined as A^{-1} such that $AA^{-1} = A^{-1}A = I_n$.

However, if a matrix is singular or non-square, a unique inverse does not exist. In this case, the concept of generalized inverses is implemented. A $m \times n$ matrix A has a *generalized inverse* G of order $n \times m$ such that $AGA = A$. G is also refered to as A^{-}. Let G be a *g-inverse* of a non-singular square matrix A, then G is a unique inverse such that $G = A^{-1}$. Every matrix has a *g-inverse* and each matrix has infinitely many *g-inverses*. Further, we see some important properties of *generalized inverses.*

2.1. Let A be an $m \times n$ matrix and $A^T = A$, that is, A is symmetric. Let A be defined over a field with characteristic $\neq 0$. Then, A has a symmetric *g-inverse.*

Proof. Let G be a *g-inverse* of A.
$\implies AGA = A$

Taking transpose on both sides,

$$\begin{aligned} (A(GA))^T &= A^T \\ (GA)^T A^T &= A^T \\ A^T G^T A^T &= A^T \\ AG^T A &= A \end{aligned}$$

$\therefore G^T$ is a *g-inverse* of A.

Now, we can verify that $\frac{1}{2}(G + G^T)$ is also a *g-inverse* of A.

$$\begin{aligned} A\frac{1}{2}(G+G^T)A &= \frac{1}{2}(A(G+G^T))A \\ &= \frac{1}{2}(AG + AG^T)A \\ &= \frac{1}{2}(AGA + AG^TA) \\ &= \frac{1}{2}(A + A) \\ &= A \end{aligned}$$

Hence, a symmetric *g-inverse* exists. □

2.2. Let A be an $m \times n$ matrix and G_1 and G_2 be two *g-inverses* of A. Let F be the field over which A is defined. Then, $\alpha(G_1) + (1-\alpha)(G_2)$ is also a *g-inverse* of A $\forall$ $\alpha \in F$.

Proof. Given, $AG_1A = A$ and $AG_2A = A$.
Now,

$$\begin{aligned} A(\alpha(G_1) + (1-\alpha)(G_2))A &= \alpha(AG_1A) + (1-\alpha)(AG_2A) \\ &= \alpha(A) + (1-\alpha)(A) \\ &= \alpha(A) + A - \alpha(A) \\ &= A. \end{aligned}$$

Hence, $\alpha(G_1) + (1-\alpha)(G_2)$ is also a *g-inverse* of A. □

2.3. Let A be an $m \times n$ matrix and G be a *g-inverse* of A. Then, AG is idempotent and $\rho(A) = \rho(AG)$.

Proof. Since matrix multiplication is associative.

$$\begin{aligned}(AG)(AG) &= (AGA)(G)\\ &= AG\end{aligned}$$

$\therefore AG$ is idempotent.

$$\begin{aligned}\text{Now, let }\ & x \in col(AB)\\ & \implies x = ABy \text{ for some column matrix } y\\ & \because By = z \text{ for some column matrix } z\\ & \therefore x = Az\\ & \therefore x \in col(A)\\ & col(AB) \subseteq col(A)\end{aligned}$$

Taking dimensions of both sides, we get $\rho(AB) \leq \rho(A)$.
Similarly, this is true for B as $\rho(AB) \leq \rho(B)$.
Thus, combining both

$$\rho(AG) \leq min(\rho(A), \rho(G)). \tag{2.1}$$

Now, $\rho(A) = \rho(AGA) = \rho((AG)A) \leq \rho(AG)$.
Combining this with (2.1), we get,

$$\rho(A) = \rho(AG). \tag{2.2}$$

□

2.4. Let A be an $m \times n$ matrix and G be a *g-inverse* of A. Then, GA is idempotent and $\rho(A) = \rho(GA)$.

Proof. On similar lines as the previous proof,

$$\begin{aligned}(GA)(GA) &= (G)(AGA)\\ &= GA\end{aligned}$$

$\therefore GA$ is idempotent.
Now, $\rho(A) = \rho(AGA) = \rho(A(GA)) \leq \rho(GA)$.
Combining this with (2.1), we get,

$$\rho(A) = \rho(GA).$$

□

2.5. Let A be a matrix and G be a *g-inverse*.Then,

$$\rho(A) \leq \rho(G)$$

Proof. By any of the above proofs, we see,
$\rho(A) = \rho(AGA) = \rho((AG)A) \leq \rho(AG) \leq \rho(G)$.

$$\therefore \rho(A) \leq \rho(G) \tag{2.3}$$

□

2.6. Let A be an $m \times n$ matrix and G be a *g-inverse* of A. The class of *g-inverses* of A is given by

$$G + U - GAUAG \tag{2.4}$$

where U is an arbitrary matrix.

Proof. The proof is completed in two steps:

(i) Prove $G + U - GAUAG$ is a *g-inverse* of matrix A.

(ii) Let H be a *g-inverse* . Prove that H is of the form $G + U - GAUAG$.

$$\begin{aligned} A(G + U - GAUAG)A &= (AG + AU - AGAUAG)A \\ &= (AG + AU - AUAG)A \\ &= AGA + AUA - AUAGA \\ &= A + AUA - AUA \\ &= A. \end{aligned}$$

Hence, (i) is proved.

Now, H is a *g-inverse* of $A \implies AHA = A$.

Substituting $U = H - G$ in (2.4),

$$\begin{aligned} G + U - GAUAG &= G + H - G - GA(H - G)AG \\ &= H - G(A(H - G)A)G \\ &= H - G(AHA - AGA)G \\ &= H - G(A - A)G \\ &= G. \end{aligned}$$

Hence, (ii) is proved.

□

2.7. Let A be an $m \times n$ matrix and G be a *g-inverse* of A. The class of *g-inverses* of A is also given by

$$G + V(I - AG) + (I - GA)W \tag{2.5}$$

where V and W are arbitrary matrices.

Proof. On similar lines as the previous proof, this proof is completed in two steps:

(i) Prove $G + V(I - AG) + (I - AG)W$ is a *g-inverse* of matrix A.

(ii) Let H be a *g-inverse* . Prove that H is of the form $G + V(I - AG) + (I - GA)W$.

$$\begin{aligned} A(G + V(I - AG) + (I - GA)W)A &= (AG + AV(I - AG) + A(I - GA)W)A \\ &= (AG + AV(I - AG) + (A - AGA)W)A \\ &= (AG + AV(I - AG) + (0 - 0)W)A \\ &= (AGA + AV(I - AG)A) \\ &= (A + AV(A - AGA) \\ &= (A + AV(0 - 0) \\ &= A \end{aligned}$$

Hence, (i) is proved.

Now, H is a *g-inverse* of $A \implies AHA = A$.

Substituting $V = H - G$ and $W = HAG$ in (2.5),

$$\begin{aligned} G + V(I - AG) + (I - GA)W &= G + (H - G)(I - AG) + (I - GA)(HAG) \\ &= G + (H - G)(I - AG) + (HAG - GAHAG) \\ &= G + (H - G)(I - AG) + (HAG - GAG) \\ &= G + H - G - (H - G)(AG) + HAG - GAG \\ &= H - HAG + GAG + HAG - GAG \\ &= H \end{aligned}$$

Hence, (ii) is proved. $\square$

$\because U, V$ and W are arbitrary matrices, H can take infinitely many values. Hence, each matrix A has infinitely many *g-inverses* possible.

2.8. Reflexive Property

Let A be a matrix and G be a *g-inverse* of A. G is said to be reflexive if A is also a *g-inverse* of G,that is,

$$AGA = A \text{ and } GAG = G.$$

Further, G is a reflexive *g-inverse* iff $\rho(G) = \rho(A)$

Proof. Combining (2.2) and (2.1), we get,
$\rho(A) = \rho(AG) = \rho(GA) \because G$ is also a *g-inverse* a A
We exchange A and G in the above equation,
$\rho(G) = \rho(GA) = \rho(AG)$.
Equating both,

$$\rho(A) = \rho(G) \tag{2.6}$$

Thus, we can see equality of (2.3) is given by (2.6) only if G is a reflexive *g-inverse*.
Conversely, given, $\rho(A) = \rho(G), AGA = A$.
As stated in 2.3, we know,
$col(AB) \subseteq col(A)$
$\therefore col(GA) \subseteq col(G)$
$\implies G = GAX$ for some X
Now,

$$\begin{aligned} GAG &= GAGAX \\ &= G(AGA)X \\ &= GAX \\ &= G \end{aligned}$$

Hence, G is reflexive. □

2.9. Let A be a matrix and G_1 and G_2 be two *g-inverses* of A. Then, G_1AG_2 is a reflexive *g-inverse* of A.

Proof. Given, $AG_1A = A$ and $AG_2A = A$.
Now,

$$\begin{aligned} A(G_1AG_2)A &= (AG_1A)G_2A \\ &= (A)G_2A \\ &= A \end{aligned}$$

Hence, G_1AG_2 is a *g-inverse* of A.
Also,

$$\begin{aligned} (G_1AG_2)A(G_1AG_2) &= G_1(AG_2A)(G_1AG_2) \\ &= G_1A(G_1AG_2) \\ &= G_1(AG_1A)G_2 \\ &= G_1AG_2 \end{aligned}$$

Hence, A is a *g-inverse* of G_1AG_2.
Combining both results,
G_1AG_2 is a reflexive *g-inverse* of A. □

3. Basic Computation

Generalized inverse of an $m \times n$ matrix A can be calculated by various methods.

3.1. *G-inverse* of a zero matrix
Every matrix is a *g-inverse* of a null matrix.
$0G0 = 0$ is satisfied by every G.

3.2. Let A be an $m \times n$ matrix of the form,

$$A = \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix}$$

where,
I_r = Identity matrix of order $r \times r$
0_1 = Null matrix of order $r \times (n-r)$
0_2 = Null matrix of order $(n-r) \times r$
0_3 = Null matrix of order $(m-r) \times (n-r)$
Then, *g-inverse* of A is an $n \times m$ matrix which can be represented in the form,

$$G = \begin{bmatrix} I_r & B \\ C & D \end{bmatrix}$$

where,
B, C and D are arbitrary matrices of orders $r \times (m-r), (n-r) \times r$ and $(n-r) \times (m-r)$ respectively.

Proof. This can be verified as follows:

$$\begin{aligned}\begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix} \begin{bmatrix} I_r & B \\ C & D \end{bmatrix} \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix} \\ &= \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix} \begin{bmatrix} I_r & 0_1 \\ C & 0_4 \end{bmatrix} \\ &= \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix}\end{aligned}$$

where,
$0_4 =$ Null matrix of order $(n-r) \times (n-r)$
$\therefore AGA = A$
Hence, G is *g-inverse* of A. □

3.3. Let A be an $m \times n$ matrix such that $\rho(A) = r$ and A is represented in the form,

$$A = P \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix} Q$$

where,
P and Q are non-singular matrices.
Then, *g-inverse* of A is an $n \times m$ matrix which can be represented in the form,

$$G = Q^{-1} \begin{bmatrix} I_r & B \\ C & D \end{bmatrix} P^{-1}$$

All notations are defined as in the previous proof.

Proof. This can be verified as follows:

$$\begin{aligned}P \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix} QQ^{-1} \begin{bmatrix} I_r & B \\ C & D \end{bmatrix} P^{-1}P \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix} Q \\ &= P \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix} I \begin{bmatrix} I_r & B \\ C & D \end{bmatrix} I \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix} Q \\ &= P \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix} \begin{bmatrix} I_r & B \\ C & D \end{bmatrix} \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix} Q \\ &= P \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix} Q\end{aligned}$$

$\because$ By the previous proof,we know,

$$\begin{bmatrix} I_r & B \\ C & D \end{bmatrix} \text{ is a } \textit{g-inverse} \text{ of } \begin{bmatrix} I_r & 0_1 \\ 0_2 & 0_3 \end{bmatrix}$$

$\therefore AGA = A$.
Hence, G is *g-inverse* of A. □

3.4. Let A be an $m \times n$ matrix of the form,

$$A = \begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix}$$

where,
I_r = Identity matrix of order $r \times r$
B = Arbitrary matrix of order $r \times (n-r)$
0_1 = Null matrix of order $(n-r) \times r$
0_2 = Null matrix of order $(m-r) \times (n-r)$
Then, *g-inverse* of A is an $n \times m$ matrix which can be represented in the form,

$$G = \begin{bmatrix} I_r & 0_3 \\ 0_1 & C \end{bmatrix}$$

where,
0_3 = Null matrix of order $r \times (n-r)$
C = Arbitrary matrix of order $(m-r) \times (n-r)$

Proof. This can be verified as follows:

$$\begin{aligned} \begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix} \begin{bmatrix} I_r & 0_3 \\ 0_1 & C \end{bmatrix} \begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix} & \\ &= \begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix} \begin{bmatrix} I_r & B \\ 0_1 & 0_4 \end{bmatrix} \\ &= \begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix} \end{aligned}$$

$\therefore AGA = A$.
Hence, G is a *g-inverse* of A.

□

3.5. Let A be an $m \times n$ matrix of the form,

$$A = P \begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix} Q$$

where,
P and Q are non-singular matrices.
Then, *g-inverse* of A is an $n \times m$ matrix which can be represented in the form,

$$G = Q^{-1}\begin{bmatrix} I_r & 0_3 \\ 0_1 & C \end{bmatrix} P^{-1}$$

All notations are defined as in the previous proof.

Proof. This can be verified as follows:

$$\begin{aligned} &P\begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix} QQ^{-1}\begin{bmatrix} I_r & 0_3 \\ 0_1 & C \end{bmatrix} P^{-1}P\begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix} Q \\ &\qquad = P\begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix} I \begin{bmatrix} I_r & 0_3 \\ 0_1 & C \end{bmatrix} I \begin{bmatrix} I_r & B \\ 0_1 & C \end{bmatrix} Q \\ &\qquad = P\begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix}\begin{bmatrix} I_r & 0_3 \\ 0_1 & C \end{bmatrix}\begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix} Q \\ &\qquad = P\begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix} Q \end{aligned}$$

$\because$ By the previous proof,we know,

$$\begin{bmatrix} I_r & 0_3 \\ 0_1 & C \end{bmatrix} \text{ is a } \textit{g-inverse} \text{ of } \begin{bmatrix} I_r & B \\ 0_1 & 0_2 \end{bmatrix}$$

$\therefore AGA = A$.
Hence, G is *g-inverse* of A. □

3.6. Let A be an $m \times n$ matrix such that $\rho(A) = r$ and A is partitioned in the form,

$$A = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$

$\because \rho(A) = r, \exists$ a non-singular submatrix of order r.
By permutations, this can be changed to non-singular leading submatrix.
$\therefore$ Let A_{11} is non-singular.
Then, *g-inverse* of A is an $n \times m$ matrix which can be represented in the form,

$$G = \begin{bmatrix} (A_{11})^{-1} & 0_1 \\ 0_2 & 0_3 \end{bmatrix}$$

where,
0_1 = Null matrix of order $r \times (n-r)$
0_2 = Null matrix of order $(n-r) \times r$
0_3 = Null matrix of order $(m-r) \times (n-r)$

Proof. This can be verified as follows:

$$\begin{aligned}
&\begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} (A_{11})^{-1} & 0_1 \\ 0_2 & 0_3 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \\
&= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} I_r & (A_{11})^{-1}A_{12} \\ 0_2 & 0_3 \end{bmatrix} \\
&= \begin{bmatrix} A_{11} & A_{11}(A_{11})^{-1}A_{12} \\ A_{21} & A_{21}(A_{11})^{-1}A_{12} \end{bmatrix} \\
&= \begin{bmatrix} A_{11} & IA_{12} \\ A_{21} & A_{21}(A_{11})^{-1}A_{12} \end{bmatrix} \\
&= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}
\end{aligned}$$

The last step is explained as $\rho(A) = r$.
$\therefore A_{12} = A_{11}X$ for some matrix X.
$\implies X = (A_{11})^{-1}A_{12}$.
and $A_{22} = A_{21}X$ for the same matrix X.
$\therefore A_{22} = A_{21}(A_{11})^{-1}A_{12}$
$\therefore AGA = A$
Hence, G is *g-inverse* of A. □

Before learning the next method of computing *generalized inverses*, we learn the concept of **Left and Right Inverses**.

3.7. Let A be an $m \times r$ matrix such that $\rho(A) = r$,that is, A is a full column rank matrix. Then, $\exists$ a matrix A_l of order $r \times m$ called the *left-inverse* of A such that

$$A_l A = I_r$$

Proof. A is a full column rank matrix.
$\therefore$ Columns of A are linearly independent.

$$A = \begin{bmatrix} A_1 & A_2 & \dots & A_r \end{bmatrix}$$

$\therefore \alpha = \{(A_1), (A_2), ..., (A_r)\}$ forms a basis of S such that $S \subseteq R^m$ with $\rho(S) = r$.

We extend this basis to a basis of R^m.
$\beta = \{(A_1), (A_2), ..., (A_r), (X_1), ..., (X_{m-r})\}$

Let, a matrix X be defined as

$$X = \begin{bmatrix} X_1 & X_2 & \dots & X_{m-r} \end{bmatrix}$$

$\therefore [A : X]$ is a non-singular matrix with rank $= m$.
$\implies \exists$ a unique inverse and let this be partitioned as:

$$\begin{bmatrix} A_l \\ Y \end{bmatrix}$$

$$\because \begin{bmatrix} A_l \\ Y \end{bmatrix} \begin{bmatrix} A & X \end{bmatrix} = I$$

$$\begin{bmatrix} A_l A & A_l X \\ YA & YX \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & I_{m-r} \end{bmatrix}$$

Equating both sides, we get,

$$A_l A = I_r$$

□

3.8. Let A be an $r \times n$ matrix such that $\rho(A) = r$,that is, A is a full row rank matrix. Then, $\exists$ a matrix A_r of order $n \times r$ called the *right-inverse* of A such that

$$AA_r = I_r$$

Proof. A is a full row rank matrix.
$\therefore$ Columns of A are linearly independent.

$$A = \begin{bmatrix} A_1 \\ A_2 \\ \vdots \\ A_r \end{bmatrix}$$

$\therefore \alpha = \{(A_1), (A_2), ..., (A_r)\}$ forms a basis of S such that $S \subseteq R^n$ with $\rho(S) = r$.

We extend this basis to a basis of R^n.
$\beta = \{(A_1), (A_2), ..., (A_r), (X_1), ..., (X_{n-r})\}$

Let, a matrix X be defined as

$$X = \begin{bmatrix} X_1 \\ X_2 \\ \vdots \\ X_{n-r} \end{bmatrix}$$

$$\therefore \begin{bmatrix} A \\ X \end{bmatrix} \text{ is a non-singular matrix with rank} = \text{n}$$

$\implies \exists$ a unique inverse and let this be partitioned as:

$$\begin{bmatrix} A_r & Y \end{bmatrix}$$

$$\because \begin{bmatrix} A \\ X \end{bmatrix} \begin{bmatrix} A_r & Y \end{bmatrix} = I$$

$$\begin{bmatrix} AA_r & AY \\ XA_r & XY \end{bmatrix} = \begin{bmatrix} I_r & 0 \\ 0 & I_{n-r} \end{bmatrix}$$

Equating both sides, we get,

$$AA_r = I_r$$

□

After learning about these two important classes of inverses, we learn a method of computing *g-inverse* of a matrix by its **rank-factorization**.

3.9. Let A be an $m \times n$ matrix with $\rho(A) = r$. Then, $\exists$ two matrices , P and Q of orders $m \times r$ and $r \times n$ respectively such that $\rho(P) = \rho(Q) = r$ and $A = PQ$. This decomposition is known as the rank factorization of A.

3.10. Let A be an $m \times n$ matrix with $\rho(A) = r$ and the rank factorization of A is given as $A = PQ$ for a full column rank matrix P and a full row rank matrix Q.
Then, *g-inverse* of A is an $n \times m$ matrix such that,

$$G = Q_r P_l$$

where,
Q_r is the right inverse of Q and
P_l is the left inverse of P.

Proof. This can be verified as follows:

$$\begin{aligned} AGA &= PQQ_rP_lPQ \\ &= P(QQ_r)(P_lP)Q \\ &= P(I_r)(I_r)Q \\ &= PQ \\ &= A \end{aligned}$$

Hence, G is a *g-inverse* of A. $\square$

4. Linear Equations

Generalized inverses are used to effectively solve system of equations.
We first learn about the consistency and inconsistency of a given system of linear equations.
Given a system of equations,

$$\begin{aligned} a_{11}x_1 + a_{12}x_2 + \ldots + a_{1n}x_n &= b_1 \\ a_{21}x_1 + a_{22}x_2 + \ldots + a_{2n}x_n &= b_2 \\ \vdots &= \vdots \\ a_{m1}x_1 + a_{m2}x_2 + \ldots + a_{mn}x_n &= b_n \end{aligned} \tag{4.1}$$

This can be rewritten in the form of matrices as

$$Ax = b$$

where,

$$A = \begin{bmatrix} a_{11} & a_{12} & \ldots & a_{1n} \\ a_{21} & a_{22} & \ldots & a_{2n} \\ \vdots & \vdots & & \vdots \\ a_{m1} & a_{m2} & \ldots & a_{mn} \end{bmatrix}$$

$$x = \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix}$$

$$b = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_m \end{bmatrix}$$

The above defined system is a **general linear system**.
If $b =$ Null matrix of order $m \times 1$, then, the system of equations is known as a **homogenous system of equations**.
Next, we find the augmented matrix,

$$\begin{aligned} X &= \begin{bmatrix} A & b \end{bmatrix} \\ &= \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} & b_1 \\ a_{21} & a_{22} & \dots & a_{2n} & b_2 \\ \vdots & \vdots & & \vdots & \vdots \\ a_{m1} & a_{m2} & \dots & a_{mn} & b_m \end{bmatrix} \end{aligned}$$

A system of equations is said to be **consistent** if $\rho(A) = \rho(X)$.

$$\implies dim(col(A)) = dim(col(X)) \tag{4.2}$$

Now, let $p \in col(X)$
$\therefore [A : b]q = p$ for some column matrix q.
$\therefore p \in col(A)$
Hence, $col(X) \subseteq col(A)$.
Combining this result with (4.2), we get,
$col(X) = col(A)$
Hence, obviously $b \in col(A)$.
We have the following result.

4.1. The system of linear equations $Ax = b$ is consistent iff $b \in col(A)$.

4.2. Let A be an $m \times n$ matrix. Then, G is a *g-inverse* of A iff Gb is a solution to $Ax = b$ when the system is consistent, or simply, when $b \in col(A)$.

Proof. If Gb is a solution to $Ax = b$,

$$AGb = b$$

$\because b \in col(A)$
$b = Ay$ for some column matrix y.

Substituting this in the above equation,
$AGAy = Ay$

Now, in particular,
Let $y = [1, 0, 0, \ldots]$,
$\therefore$ first column of AGA = first column of A.

Now, let $y = [0, 1, 0, \ldots]$,
$\therefore$ second column of AGA = second column of A.

and so on.

Thus, $AGA = A$.
$\therefore G$ is a *g-inverse* of A.

Conversely,
$AGA = A$ and $b \in col(A)$
Multiplying y on both sides,
$AGAy = Ay$
$\implies AGb = b$
$\therefore Gb$ is a solution to consistent $Ax = b$. □

Before discussing the class of solutions of a system of linear equations in terms of *generalized inverse*, we see,

4.3. The set of all solutions,S, of a consistent system $Ax = b$ is given by

$$S = u + Null(A)$$

where u is one of the solutions to $Ax = b$.

Proof. Let $v \in S$
$\therefore Av = b$
Also, $Au = b$.
Subtracting both, $A(v - u) = 0$
$\therefore v - u \in Null(A)$
$v \in u + Null(A)$.
$\implies S \subseteq u + Null(A)$

Now, let $v \in u + Null(A)$
$v - u \in Null(A)$
$\therefore A(v - u) = 0$ $Av = Au$

$Av = b$
$\therefore v \in S$
$\implies u + Null(A) \subseteq S$

Combining both results,

$$S = u + Null(A)$$

□

4.4. Let A be an $m \times n$ matrix and G be a *g-inverse* of A. Then, a general solution of $Ax = 0$ is $(I - GA)z$
where, z is an arbitrary vector.

Proof.

$$\begin{aligned} A(I - GA)z &= (A - AGA)z \\ &= (A - A)z \\ &= 0 \end{aligned}$$

Hence, $(I - GA)z$ is a solution of $Ax = 0$.
Also, Let u be a solution to $Ax = 0$.
It can easily be shown that u is of the form $(I - GA)z$ by substituing $z = u$.

$$\begin{aligned} (I - GA) &= (I - GA)u \\ &= u - GAu \\ &= u - G(0) \\ &= u \end{aligned}$$

Combining the two results,
$(I - GA)z$ is the class of solutions of $Ax = 0$.
Also, $(I - GA)z \in Null(A)$.

□

4.5. Let A be an $m \times n$ matrix and G be a *g-inverse* of A. Then, a general solution of a consistent system of linear equarions $Ax = b$ is
$Gb + (I - GA)z$
where, z is an arbitrary vector.

Proof. By 4.2, we know Gb is a solution to consisten $Ax = b$.
Let $u = Gb$.
Also, by 4.3, we know $S \subseteq u + Null(A)$.
Lastly, by the previous proof, $(I - GA)z \in Null(A)$.
Combining the above statements, a general solution v is such that,
$v \in S$
$\implies v = Gb + (I - GA)z$ for some arbitrary matrix z.

□

4.6. Let A be an $m \times n$ matrix and G be a *g-inverse* of A. Then, a general solution of a consistent system of linear equarions $y^T A = c^T$ is
$c^T G + w^T (I - GA)$
where, w is an arbitrary vector.

The class of solutions to a consistent system of linear equations is used to derive some important special types of *generalized inverses* namely, *Minimum Norm inverse* and *least Squares Inverse.*

5. Minimum Norm Inverse

Let A be an $m \times n$ matrix. An $n \times m$ matrix G is said to be a *minimum norm g-inverse* of A if,

$$AGA = A$$
$$(GA)^T = GA$$

Now, **Euclidean Norm** of a column matrix is given by

$$||x|| = \sqrt{x^T x} \tag{5.1}$$

5.1. Let A be an $m \times n$ matrix. An $n \times m$ matrix is a *minimum norm g-inverse* of A iff $x = Gb$ is a solution of minimum norm to a consistent system of equations $Ax = b$.

Proof. By 4.5, the class of solutions is given by $Gb + (I - GA)z$ where z is arbitrary.
Given, $AGA = A$ and $(GA)^T = GA$.
Now,

$$\begin{aligned} ||Gb + (I - GA)z||^2 &= (Gb + (I - GA)z)^T (Gb + (I - GA)z) \\ &= ((Gb)^T + ((I - GA)z)^T)(Gb + (I - GA)z) \\ &= (Gb)^T (Gb) + ((I - GA)z)^T (I - GA)z \\ &+ (Gb)^T (I - GA)z + ((I - GA)z)^T Gb \\ &= ||Gb||^2 + ||(I - GA)z||^2 + 2(Gb)^T (I - GA)z \\ &= ||Gb||^2 + ||(I - GA)z||^2 + 2b^T G^T (I - GA)z \end{aligned} \tag{5.2}$$

$\because (Gb)^T(I - GA)z$ is 1×1
$\therefore (Gb)^T(I - GA)z = ((I - GA)z)^T Gb.$

Now, the system $Ax = b$ is consistent, $\implies b \in col(A)$.
$\therefore b = Ay$ for some column matrix y.

$$\begin{aligned}
b^TG^T(I - GA)z &= (Ay)^TG^T(I - GA)z \\
&= y^TA^TG^T(I - GA)z \\
&= y^T(GA)^T(I - GA)z \\
&= y^T(GA)(I - GA)z \\
&= y^T(GA - GAGA)z \\
&= y^T(GA - GA)z \\
&= 0
\end{aligned}$$

Substituting this result in 5.2, we get
$||Gb + (I - GA)z||^2 = ||Gb||^2 + ||(I - GA)z||^2$
$||Gb||^2 \leq ||Gb + (I - GA)z||^2$
$||Gb|| \leq ||Gb + (I - GA)z||$
$\therefore Gb$ is the solution with minimum norm.

Now, conversely, Gb is a solution of $Ax = b$ with minimum norm.
$||Gb|| \leq ||Gb + (I - GA)z||$
is true for all z.
$||Gb||^2 \leq ||Gb + (I - GA)z||^2$
$||Gb||^2 \leq ||Gb||^2 + ||(I - GA)z||^2 + 2b^TG^T(I - GA)z$
$||(I - GA)z||^2 + 2b^TG^T(I - GA)z \geq 0$

Now, suppose, $2b^TG^T(I - GA)z$ is initially negative.
Replacing y^T by αy^T where $\alpha > 0$ and $|\alpha|$ is sufficiently large.
$\therefore ||(I - GA)z||^2 + 2b^TG^T(I - GA)(\alpha z) < 0$ which is a contradiction.

Similarly, now suppose, $2b^TG^T(I - GA)z$ is initially positive.
Replacing y^T by βy^T where $\beta < 0$ and $|\alpha|$ is sufficiently large.
$\therefore ||(I - GA)z||^2 + 2b^TG^T(I - GA)(\beta z) < 0$ which is a contradiction.

$$\therefore 2b^TG^T(I - GA)z = 0 \text{ for all } z.$$

$$\begin{aligned}
A^TG^T(I-GA) &= 0 \\
A^TG^T - A^TG^TGA &= 0 \\
A^TG^T &= A^TG^TGA \\
(GA)^T &= (GA)^TGA \\
(GA) &= (GA)^T(GA^T)^T \\
&= (GA)^TGA
\end{aligned} \tag{5.3}$$

$\therefore (GA)^T = GA$
Thus, G is a *minimum norm g-inverse* of A. □

Let A be an $m \times n$ matrix. An $n \times m$ matrix G is said to the *minimum norm N-g-inverse* of A if,

$$AGA = A$$
$$(GA)^TN = N(GA)$$

N-norm of a column matrix is given by

$$\sqrt{x^TNx}$$

where N is a positive-definite matrix.

5.2. Let A be an $m \times n$ matrix. An $n \times m$ matrix G is a *minimum N-norm g-inverse* of A iff $x = Gb$ is a solution of minimum N-norm to a consistent system of equations $Ax = b$.

Proof. On similar lines as the previous proof, the class of solutions is given by $Gb + (I - GA)z$ where z is arbitrary.
Given, $AGA = A$ and $(GA)^TN = NGA$.

Now,

$$\begin{aligned}
(Gb+(I-GA)z)^TN(Gb+(I-GA)z) &= ((Gb)^T + ((I-GA)z)^T)N(Gb+(I-GA)z) \\
&= (Gb)^TN(Gb) + ((I-GA)z)^TN((I-GA)z) \\
&\quad + (Gb)^TN(I-GA)z + ((I-GA)z)^TN(Gb)
\end{aligned} \tag{5.4}$$

Now, since,

$$\begin{aligned}(Gb)^T N(I-GA)z &= ((Gb)^T N(I-GA)z)^T \\ &= ((I-GA)z)^T((Gb)^T N)^T \\ &= ((I-GA)z)^T N^T(Gb) \\ &= ((I-GA)z)^T N(Gb)\end{aligned}$$

Substituting this back,
$(Gb+(I-GA)z)^T N(Gb+(I-GA)z) = (Gb)^T N(Gb)+((I-GA)z)^T N((I-GA)z) + 2(Gb)^T N(I-GA)z$

Now, the system $Ax = b$ is consistent, $\implies b \in col(A)$.
$\therefore b = Ay$ for some column matrix y.

$$\begin{aligned}(Gb)^T N(I-GA)z &= (Ay)^T G^T N(I-GA)z \\ &= y^T A^T G^T N(I-GA)z \\ &= y^T(GA)^T N(I-GA)z \\ &= y^T NGA(I-GA)z \\ &= y^T N(GA-GAGA)z \\ &= y^T N(GA-GA)z \\ &= 0\end{aligned}$$

Substituting this in (5.4),
$(Gb+(I-GA)z)^T N(Gb+(I-GA)z) = (Gb)^T N(Gb)+((I-GA)z)^T N((I-GA)z)$
$(Gb)^T N(Gb) \leq (Gb+(I-GA)z)^T N(Gb+(I-GA)z)$
$\sqrt{(Gb)^T N(Gb)} \leq \sqrt{(Gb+(I-GA)z)^T N(Gb+(I-GA)z)}$
$\therefore Gb$ is the solution with minimum norm.

Now, conversely, Gb is a solution of $Ax = b$ with minimum N-norm.

On similar lines as the above proof, we state that
$(Gb)^T N(I-GA)z = 0$
$\because b = Ay$ for some column matrix y.

$$
\begin{aligned}
(Gb)^T N(I - GA)z &= 0 \\
y^T A^T G^T N(I - GA)z &= 0 \\
(GA)^T N(I - GA)z &= 0 \\
(GA)^T N &= (GA)^T N(GA) \\
N^T(GA) &= (N(GA))^T (GA)^T \\
N(GA) &= (GA)^T N(GA)
\end{aligned}
$$

$\therefore (GA)^T N = NGA$

Thus, G is a *minimum N-norm g-inverse* of A. □

5.3. Let A be an $m \times n$ matrix and G be a *minimum norm g-inverse* of A. The class of all *minimum norm g-inverses* is given by X such that $XA = GA$.

Proof. Given, $AGA = A$ and $(GA)^T = GA$.
Now, $AXA = A(XA) = A(GA) = A$.
Hence, X is a *g-inverse* of A.

Also,$(XA)^T = (GA)^T = GA = XA$.
Combining both results, X is a *minimum norm g-inverse*.

□

6. Least Squares Inverse

Let A be an $m \times n$ matrix. An $n \times m$ matrix G is said to be a *least squares g-inverse* of A if,

$$AGA = A$$
$$(AG)^T = AG$$

6.1. Let A be an $m \times n$ matrix. An $n \times m$ matrix is a *least squares g-inverse* of A iff $x = Gb$ is a solution to a consistent system of equations $Ax = b$ such that $||AGb - y||$ is minimum out of all the solutions, that is,

$$||AGb - b|| \leq ||Ax - b||$$

Proof. The class of all soltuions given by 4.5 is true for all z.
Thus, let $x = Gb + w$ for some w where G is the *least squares g-inverse* of A. Now,

$$\begin{aligned} ||Ax - b||^2 &= ||A(Gb + w) - b||^2 \\ &= ||AGb - b + Aw||^2 \\ &= ||AGb - b||^2 + ||Aw||^2 + (Aw)^T(AGb - b) \\ &+ (AGb - b)^T(Aw) \end{aligned} \tag{6.1}$$

In this, $(Aw)^T(AGb - b)$ is a scalar.
$\therefore (Aw)^T(AGb - b) = (AGb - b)^T(Aw)$
$b \in col(A) \implies b = Ay$ for some y.

$$\begin{aligned} w^T A^T (AGb - b) &= w^T A^T (AGAy - Ay) \\ &= w^T A^T (AG - I)Ay \\ &= w^T (A^T AG - A^T)Ay \\ &= w^T (A^T (AG)^T - A^T)Ay \\ &= w^T (A^T G^T A^T - A^T)Ay \\ &= w^T (A^T - A^T)Ay \\ &= 0 \end{aligned}$$

Substituting this in (6.1), we get
$||Ax - b||^2 = ||AGb - b||^2 + ||Aw||^2$
$||AGb - b||^2 \leq ||Ax - b||^2$
$||AGb - b|| \leq ||Ax - b||$
$\therefore$ Gb is a least squares solution.

Now, conversely, Gb is a solution of $Ax = b$ such that
$||AGb - b|| \leq ||Ax - b||$ is true for all solutions x.
Substituting $b = Ax$ in this,
$||AGAx - Ax|| \leq ||Ax - Ax||$
$AGAx - Ax = 0$
$AGA = A$
Let $x = Gb + w$ for some w.
$||AGb - b|| \leq ||AGb - b + Aw||$
$||AGb - b||^2 \leq ||AGb - b + Aw||^2$
Using (6.1)
$||AGb - b||^2 \leq ||AGb - b||^2 + ||Aw||^2 + (Aw)^T(AGb - b) + 2(Aw)^T(AGb - b)$
$||Aw||^2 + (Aw)^T(AGb - b) + 2(Aw)^T(AGb - b) \geq 0$

Now, suppose $w^T A^T (AGb - b)$ is initially negative.
Replacing w^T by αw^T where $\alpha > 0$ and $|\alpha|$ is sufficiently large.
$\therefore$ $||Aw||^2 + (Aw)^T(AGb - b) + 2(Aw)^T(AGb - b) < 0$ which is a contradiction.

Similarly, now suppose, $w^T A^T (AGb - b)$ is initially positive.
Replacing w^T by βw^T where $\beta < 0$ and $|\alpha|$ is sufficiently large.
$\therefore$ $||Aw||^2 + (Aw)^T(AGb - b) + 2(Aw)^T(AGb - b) < 0$ which is a contradiction.

$$\therefore w^T A^T (AGb - b) = 0$$

$$
\begin{aligned}
A^T(AGb - b) &= 0 \\
A^TAGb - A^Tb &= 0 \\
A^T &= A^TAG \\
G^TA^T &= G^TA^TAG \\
(AG)^T &= (AG)^TAG \\
AG &= (AG)^T((AG)^T)^T \\
AG &= (AG)^TAG
\end{aligned}
$$

$\therefore (AG)^T = AG$.
Thus, G is a *least squares g-inverse* of A. □

Let A be an $m \times n$ matrix. An $n \times n$ matrix G is saide to be its *M-least squares g-inverse* if,

$$
\begin{gathered}
AGA = A \\
(AG)^TM = M(AG)
\end{gathered}
$$

where, M is a postive-definite matrix.

6.2. Let A be an $m \times n$ matrix. An $n \times m$ matrix G is a *M-least squares g-inverse* of A iff $x = Gb$ is a solution to a consistent system of equations $Ax = b$ such that $(AGb - b)^TM(AGb - b) \leq (Ax - b)^TM(Ax - b)$ for all solutions x.

Proof. On similar lines as the previous proof, the class of solutions is given by $x = Gb + w$ where w is arbitrary.
Given, $AGA = A$ and $(AG)^TM = MAG$.
Now,

$$
\begin{aligned}
(Ax - b)^TM(Ax - b) &= (AGb - b + Aw)^TM(AGb - b + Aw) \\
&= ((AGb - b)^T + (Aw)^T)M(AGb - b + Aw) \\
&= (AGb - b)^TM(AGb - b) + (Aw)^TM(Aw) \\
&\quad + (AGb - b)^TM(Aw) + (Aw)^TM(AGb - b)
\end{aligned} \tag{6.2}
$$

Now, since,

$$
\begin{aligned}
(Aw)^TM(AGb - b) &= ((Aw)^TM(AGb - b))^T \\
&= (AGb - b)^T((Aw)^TM)^T \\
&= (AGb - b)^TM^T(Aw) \\
&= (AGb - b)^TMAw
\end{aligned}
$$

Substituting this back,
$(Ax-b)^TM(Ax-b) =$
$(AGb-b)^TM(AGb-b)+(Aw)^TM(Aw)+2(Aw)^TM(AGb-b)$

Now, the system $Ax = b$ is consistent $\implies b \in col(A)$.
$\therefore b = Ay$ for some column matrix y.

$$\begin{aligned}(Aw)^TM(AGb-b) &= (Aw)^TM(AGAy-Ay)\\ &= (Aw)^TM(Ay-Ay)\\ &= 0\end{aligned}$$

$\therefore (Ax-b)^TM(Ax-b) = (AGb-b)^TM(AGb-b)+(Aw)^TM(Aw)$
$(AGb-b)^TM(AGb-b) \leq (Ax-b)^TM(Ax-b)$
$\therefore Gb$ is the M-least squares solution.

Now, conversely, Gb is a M-least squares solution of $Ax = b$.
On similar lines as the above proof, we state that
$(Aw)^TM(AGb-b) = 0$

$$\begin{aligned}(Aw)^TM(AGb-b) &= 0\\ w^TA^TM(AGb-b) &= 0\\ A^TMb &= A^TMAGb\\ G^TA^TMb &= G^TA^TMAGb\\ (AG)^TM &= (AG)^TM(AG)\\ M^T(AG) &= (AG)^T(M(AG)^T)^T\\ M(AG) &= (AG)^TM^T(AG)\\ M(AG) &= (AG)^TM(AG)\end{aligned}$$

$\therefore MAG = (AG)^TM$
Thus, G is a *M-least squares g-inverse* of A. □

6.3. Let A be an $m \times n$ matrix and G be a *least squares g-inverse* of A. The class of all *least squares g-inverses* is given by X such that $AX = AG$.

Proof. Given, $AGA = A$ and $(AG)^T = AG$.
Now, $AXA = (AX)A = (AG)A = A$.
Hence, X is a *g-inverse* of A.

Also,$(AX)^T = (AG)^T = AG = AX$.
Combining both results, X is a *least squares g-inverse.*

□

7. Moore Penrose Inverse

Let A be an $m \times n$ matrix. An $n \times m$ matrix G is said to the *Moore-Penrose inverse* of A if,

$$AGA = A \tag{7.1}$$

$$GAG = G \tag{7.2}$$

$$(AG)^T = AG = G^T A^T \tag{7.3}$$

$$(GA)^T = GA = A^T G^T \tag{7.4}$$

A *Moore-Penrose inverse* is also denoted by A^+.

7.1. Let A be an $m \times n$ matrix. Then, an $n \times m$ matrix G is its *Moore-Penrose inverse* iff $x = Gb$ is a minimum norm and least squares solution to a consistent system of linear equations $Ax = b$.

Proof. Combining (7.1) and (7.4) and proceeding as in 5.1, we get $x = Gb$ as the minimum norm solution.
Combining (7.1) and (7.3) and proceeding as in 6.1, we get $x = Gb$ as the least squares solution.
Following the converse steps in both 5.1 and 6.1, we conclude that *Moore-Penrose inverse* has the properties of both *minimum norm g-inverse* and *least squares g-inverse.*

□

7.2. A *Moore-Penrose inverse* G of a matrix A is unique.

Proof. On the contrary, let G_1 and G_2 be two Moore-Penrose inverses of A. Thus, (7.1) to (7.4) are all true for both G_1 and G_2.

$$\begin{aligned}
G_1 &= G_1AG_1 && \text{Using (7.2)}\\
&= G_1G_1^TA^T && \text{Using (7.3)}\\
&= G_1G_1^TA^TG_2^TA^T && \text{Using (7.1)}\\
&= G_1G_1^TA^TAG_2 && \text{Using (7.3)}\\
&= G_1AG_1AG_2 && \text{Using (7.3)}\\
&= G_1AG_2 && \text{Using (7.2)}\\
&= G_1AG_2AG_2 && \text{Using (7.2)}\\
&= G_1AA^TG_2^TG_2 && \text{Using (7.4)}\\
&= A^TG_1^TA^TG_2^TG_2 && \text{Using (7.4)}\\
&= A^TG_2^TG_2 && \text{Using (7.1)}\\
&= G_2AG_2 && \text{Using (7.4)}\\
&= G_2 && \text{Using (7.2)}
\end{aligned} \tag{7.5}$$

□

By the above proof we conclude that A^{-1} is the unique Moore-Penrose inverse when A is a non-singluar square matrix.

7.3. Let A be an $m \times n$ matrix. Then, $A^+ = G_1AG_2$ where G_1 is a *minimum norm g-inverse* and G_2 is a *least squares g-inverse*.

Proof. Given,
G_1 follows (7.1) and (7.4).
G_2 follows (7.1) and (7.3).
Now,

$$\begin{aligned}
A(G_1AG_2)A &= (AG_1A)G_2A\\
&= AG_2A && \text{Using (7.1)}\\
&= A && \text{Using (7.1)}
\end{aligned}$$

$$\begin{aligned}
(G_1AG_2)A(G_1AG_2) &= G_1(AG_2A)(G_1AG_2)\\
&= G_1A(G_1AG_2) && \text{Using (7.1)}\\
&= G_1(AG_1A)G_2\\
&= G_1AG_2 && \text{Using (7.1)}
\end{aligned}$$

$$\begin{aligned}(A(G_1AG_2))^T &= ((AG_1A)G_2)^T \\ &= (AG_2)^T && \text{Using (7.1)} \\ &= (AG_2) && \text{Using (7.3)} \\ &= (AG_1AG_2) && \text{Using (7.1)}\end{aligned}$$

$$\begin{aligned}((G_1AG_2)A)^T &= (G_1(AG_2A))^T \\ &= (G_1A)^T && \text{Using (7.1)} \\ &= (G_1A) && \text{Using (7.4)} \\ &= ((G_1AG_2)A) && \text{Using (7.1)}\end{aligned}$$

Combining these results, G_1AG_2 is Moore-Penrose Inverse.

□

7.4. The *Moore-Penrose inverse* of the *Moore-Penrose inverse* of a matrix is the matrix itself, that is, $(A^+)^+ = A$.

Proof. To avoid confusion, let $B = A^+$ and $G = A$.
We know B is the *Moore-Penrose inverse* of G. $\therefore$ (7.1) to (7.4) follow as $GBG = G, BGB = B, (GB)^T = GB$ and $(BG)^T = BG$.
Combining these, we also get G as the *Moore-Penrose inverse* of B.
$\therefore (A^+)^+ = A$. □

7.5. Let A be an $m \times n$ matrix. Then, $(A^T)^+ = (A^+)^T$.

Proof. We verify that the transpose of A^+ is the *Moore-Penrose inverse* of the transpose of A.
To avoid confusion, let $B = A^T$ and $G = (A^+)^T$.
Now,

$$\begin{aligned}BGB &= A^T(A^+)^TA^T \\ &= A^T(AA^+)^T \\ &= (AA^+A)^T \\ &= A^T && \text{Using (7.1)} \\ &= B\end{aligned}$$

$$\begin{aligned} GBG &= (A^+)^T A^T (A^+)^T \\ &= (A^+)^T (A^+ A)^T \\ &= (A^+ A A^+)^T \\ &= (A^+)^T && \text{Using (7.2)} \\ &= G \end{aligned}$$

$$\begin{aligned} (BG)^T &= (A^T (A^+)^T)^T \\ &= A^+ A \\ &= A^T (A^+)^T && \text{Using (7.4)} \\ &= BG \end{aligned}$$

$$\begin{aligned} (GB)^T &= ((A^+)^T A^T)^T \\ &= AA^+ \\ &= (A^+)^T A^T && \text{Using (7.3)} \\ &= GB \end{aligned}$$

Combining these results, G is the *Moore-Penrose inverse* of B.
$\therefore (A^T)^+ = (A^+)^T$

□

7.6. Let A be an $m \times n$ matrix. Then, $A^+ = (A^T A)^+ A^T = A^T (AA^T)^+$

7.7. Let A be an $m \times n$ matrix such that it is symmetric and idempotent, that is, $A^T = A = A^2$. Then, $A^+ = A$.

Proof. Let $G = A$.

$$\begin{aligned} AGA &= A(AA) \\ &= AA \\ &= A \end{aligned}$$

$$\begin{aligned} GAG &= A(AA) \\ &= AA \\ &= A \end{aligned}$$

$$\begin{aligned}(AG)^T &= (AA)^T \\ &= (A)^T \\ &= A \\ (GA)^T &= (AA)^T \\ &= (A)^T \\ &= A\end{aligned}$$

Thus, (7.1) to (7.4) are verified.
$\therefore A^+ = A$.

□

7.8. Let A be an $m \times n$ matrix of rank r and its rank factorization is given by $A = BC$ where B is an $m \times r$ matrix of rank r and C is an $r \times n$ matrix of rank r. Then, $A^+ = C^+B^+ = C^T(B^TAC^T)^{-1}B^T$.

Proof. By 7.6 , $B^+ = (B^TB)^+B^T$ and $C^+ = C^T(CC^T)^+$.
B has full column rank $\implies$ B^TB is a non-singular square matrix.
C has full row rank $\implies$ CC^T is a non-singular square matrix.
$\therefore (B^TB)^+ = (B^TB)^{-1}$ and $(CC^T)^+ = (CC^T)^{-1}$.
$B^+ = (B^TB)^{-1}B^T$
$C^+ = C^T(CC^T)^{-1}$

$$\begin{aligned}C^+B^+ &= C^T(CC^T)^{-1}(B^TB)^{-1}B^T \\ &= C^T(B^TBCC^T)^{-1}B^T \\ &= C^T(B^TAC^T)^{-1}B^T\end{aligned}$$

Now, we verify that C^+B^+ is the *Moore-Penrose inverse* of A.

$$\begin{aligned}AC^+B^+A &= AC^T(CC^T)^{-1}(B^TB)^{-1}B^TA \\ &= BCC^T(CC^T)^{-1}(B^TB)^{-1}B^TBC \\ &= BIIC \\ &= BC \\ &= A\end{aligned}$$

$$
\begin{aligned}
C^+B^+AC^+B^+ &= C^T(CC^T)^{-1}(B^TB)^{-1}B^TBCC^T(CC^T)^{-1}(B^TB)^{-1}B^T \\
&= C^T(CC^T)^{-1}((B^TB)^{-1}(B^TB))((CC^T)(CC^T))^{-1}(B^TB)^{-1}B^T \\
&= C^T(CC^T)^{-1}II(B^TB)^{-1}B^T \\
&= C^T(CC^T)^{-1}(B^TB)^{-1}B^T \\
&= C^+B^+
\end{aligned}
$$

$$
\begin{aligned}
(AC^+B^+)^T &= (BCC^T(CC^T)^{-1}(B^TB)^{-1}B^T)^T \\
&= (BI(B^TB)^{-1}B^T)^T \\
&= (B(B^TB)^{-1}B^T)^T \\
&= (B^T)^T(B(B^TB)^{-1})^T \\
&= B((B^TB)^T)^{-1}B^T \\
&= B(B^TB)^{-1}B^T \\
&= BI(B^TB)^{-1}B^T \\
&= BCC^T(CC^T)^{-1}(B^TB)^{-1}B^T \\
&= AC^+B^+
\end{aligned}
$$

$$
\begin{aligned}
(C^+B^+A)^T &= (C^T(CC^T)^{-1}(B^TB)^{-1}B^TBC)^T \\
&= (C^T(CC^T)^{-1}IC)^T \\
&= (C^T(CC^T)^{-1}C)^T \\
&= C^T(C^T(CC^T)^{-1})^T \\
&= C^T((CC^T)^T)^{-1}C \\
&= C^T(CC^T)^{-1}C \\
&= C^T(CC^T)^{-1}IC \\
&= C^T(CC^T)^{-1}(B^TB)^{-1}B^TBC \\
&= C^+B^+A
\end{aligned}
$$

Thus, (7.1) to (7.4) are verified.
$\therefore A^+ = C^+B^+$.

$\square$

8. Computing Moore-Penrose Inverse

Listing 8.1: Matlab script for computing Moore-Penrose Inverse

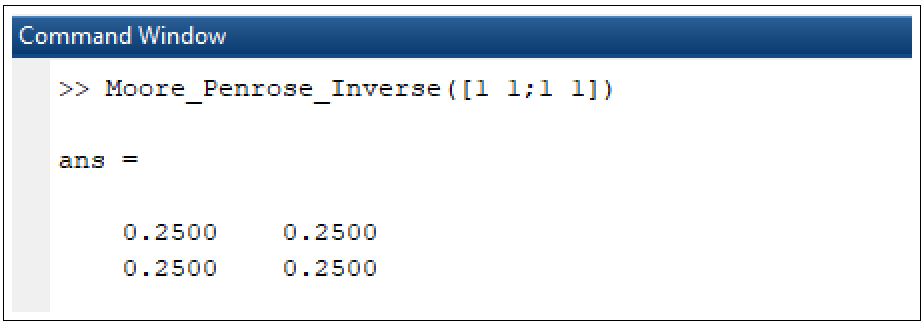
```
%%Find the Moore-Penrose Inverse of a given matrix.
function [MPI] = Moore_Penrose_Inverse(A)
D = size(A); %%size() is a function returns the two
   element row vector, i.e., D = [m,n]
m = D(1,1);
n = D(1,2);
X = A; %%Form a duplicate matrix
%%Row reduced form
for k = 1:(m-1)
    for i = (k+1):m
        if A(k,k)~=0
            x = (A(i,k)/A(k,k));
            for j = 1:n
                A(i,j) = A(i,j) - x*A(k,j);
            end
        end
    end
end
zero_rows = ~any(A,2); %%Gives the logical array for
    zero and non-zero rows
indices = find(zero_rows); %%Gives the positions of
   zero rows
rank = m - length(indices); %%Rank of the matrix =
   number of non-zero rows
```

```
%%Rank factorization %%A = BC
B = zeros(m,rank);
C = zeros(rank,n);
j = 1;
for i = 1:m
    if(i~=indices(:,1))
        C(j,:) = X(i,:); %%C contains the basis of row space
        j = j + 1;
    end
end
Ct = C.'; %%Transpose of Matrix C
for i = 1:m
    Xi = X(i,:).';
    Y = linsolve(Ct,Xi);
    B(i,:) = Y; %%B contains the constants of linear combinations
end
Bt = B.' ; %%Tranpose of Matrix B
P = (inv(Bt*B))*Bt; %%Moore-Penrose inverse of B
Q = Ct*(inv(C*Ct)); %%Moore-Penrose inverse of C
MPI = (Q*P);
```

We conside a simple example and verify the properties stated in 7. Let,

$$A = \begin{bmatrix} 1 & 1 \\ 1 & 1 \end{bmatrix}$$

Then, output of the above matlab script is



By 7.5, $(A^T)^+ = (A^+)^T$.
Let,

$$A = \begin{bmatrix} 1 & 2 & 3 \\ 4 & 5 & 6 \\ 7 & 8 & 9 \end{bmatrix}$$

$$A^T = \begin{bmatrix} 1 & 4 & 7 \\ 2 & 5 & 8 \\ 3 & 6 & 9 \end{bmatrix}$$

Then, output of the above matlab script is

```
>> Moore_Penrose_Inverse([1 2 3;4 5 6;7 8 9])

ans =

   -0.6389   -0.1667    0.3056
   -0.0556   -0.0000    0.0556
    0.5278    0.1667   -0.1944

>> Moore_Penrose_Inverse([1 4 7; 2 5 8; 3 6 9])

ans =

   -0.6389   -0.0556    0.5278
   -0.1667    0.0000    0.1667
    0.3056    0.0556   -0.1944
```

Hence, verified.

9. Applications

Research into statistical problems and data handling instigated several developments in many sub-fields of mathematics over time. Matrix theory is perhaps the best example depicting the interplay between statistics and mathematics.

The widely used statistical techniques are Analysis of variance and Regression and each can be treated by using matrix algebra. Generalized inverses have been defined in many forms and the solutions offered by this subject are hugely important and useful to the applied statisticians. The most frequently utlilized linear statistical model is mentioned here and we shall see how generalized inverses and its techniques are used to compute the problem.

There are n observations $(y_1, y_2, ..., y_n)$ for a process or experimental quantity.
The process or experiment has p elements $(x_1, x_2, ..., x_p)$, each of which has a fixed value for each of the n observations made.
In matrix notation, the model is expressed as

$$Y = X\beta + e$$

where,
Y is $n \times 1$ vector of observations with mean vector=$X\beta$ and variance-covariance matrix=$\sigma^2 I$.
X is $n \times p$ matrix of known constants.
β is $p \times 1$ matrix of unknown parameters.
σ^2 is the unknown variance of the individual observations.
e is the vector of error with mean vector=0 and variance-covariance

matrix=$\sigma^2 I$.
By the use of this general linear statistical model, the functional relationship between Y and X can be computed. Matrix theory is used in estimating the values of unknown parameters which would be useful in predicting values of Y or in explaining the variability of Y. The mathematical manipulation of a linear model draws conclusions about the process under study.

For example, considering the applications of least square generalized inverse(which will be studied in detail later on), the theory of least squares to the general linear model results in a set of linear equations called the normal or least sqaures equations.
Consider estimation of vector of unknown parameters, β. The method of least squares yields the set of equations,

$$X'X\hat{\beta} = X'Y$$

$\hat{\beta}$ = least squares estimator
Here, $\hat{\beta}$ can be calculated using generalized inverses depending upon the rank of X. The rank of X is usually not in favour of ordinary inverses when the linear model is applied to real-life experiments.

Furthermore, Moore-Penrose Inverses have many applications in solving real-time problems and forming algorithms. Some of its uses are given as follows-

(1.) Digital Image Restoration

(2.) Linear Regression

(3.) Multiple Regression

(4.) Bivariate Interpolation

(5.) Data Analysis

(6.) Principle Component Analysis Algorithm

10. Group Inverse

Let A be an $m \times n$ matrix. An $n \times m$ matrix G is said to be a *group inverse* of A if,

$$AGA = A$$
$$GAG = G$$
$$AG = GA$$

A *group inverse* is also denoted by $A^{\#}$.
Alike other pseudo-inverses, $A^{\#}$ is the unique inverse A^{-1} when A is a non-singular square matrix.

10.1. The *group inverse* of the *group inverse* of a matrix is the matrix itself, that is, $(A^{\#})^{\#} = A$.

Proof. To avoid confusion, let $B = A^{\#}$ and $G = A$.
We know B is the *group inverse* of G. $\therefore GBG = G, BGB = B$ and $GB = BG$.
Combining these, we also get G as the *group inverse* of B.
$\therefore (A^{\#})^{\#} = A$. $\square$

10.2. Let A be an $m \times n$ matrix. Then, $(A^T)^{\#} = (A^{\#})^T$.

Proof. We verify that the transpose of $A^{\#}$ is the *group inverse* of the transpose of A.
To avoid confusion, let $B = A^T$ and $G = (A^{\#})^T$.

Now,

$$\begin{aligned} BGB &= A^T(A^{\#})^T A^T \\ &= (AA^{\#}A)^T \\ &= (A)^T \\ &= B \end{aligned}$$

$$\begin{aligned} GBG &= (A^{\#})^T A^T (A^{\#})^T \\ &= (A^{\#}AA^{\#})^T \\ &= (A^{\#})^T \\ &= G \end{aligned}$$

$$\begin{aligned} BG &= A^T(A^{\#})^T \\ &= (A^{\#}A)^T \\ &= (AA^{\#})^T \\ &= (A^{\#})^T A^T \\ &= GB \end{aligned}$$

Combining these results, G is the *group inverse* of B.
$\therefore (A^T)^{\#} = (A^{\#})^T$ □

11. Drazin Inverse

11.1. Index of a matrix Index of a matrix is defined as the minimum k for which $\rho(A^{k+1}) = \rho(A^k)$.

11.2. Let A be an $m \times n$ matrix with $\text{index}(A) = k$. An $n \times m$ matrix G is said to be a *drazin inverse* of A if,

$$A^k G A = A^k \tag{11.1}$$

$$GAG = G \tag{11.2}$$

$$AG = GA \tag{11.3}$$

A *drazin inverse* is also denoted by A^D.

We conclude that *Drazin inverse* is a general form of *Group inverse*, group inverse being the special case when $\text{index}(A) = 1$.

11.3. The *drazin inverse* of the *drazin inverse* of a matrix is the matrix itself, that is, $(A^D)^D = A$ iff $\text{index}(A) = 1$.

Proof. If $\text{index}(A) = 1 \implies A^D = A^\#$.
$\therefore$ 10.1 $\implies (A^D)^D = A$.

Conversely, $(A^D)^D = A$.
$\therefore$ By (11.2), $AA^DA = A$.
We know, by (11.1), $A^kA^DA = A^k$.
Comparing both the equations we get, $k = 1$. $\square$

11.4. Let A be an $m \times n$ matrix. Then, $(A^T)^D = (A^D)^T$.

Proof. We verify that the transpose of A^D is the *drazin inverse* of the transpose of A.
To avoid confusion, let $B = A^T$ and $G = (A^D)^T$.
Now,

$$\begin{aligned} GBG &= (A^D)^T A^T (A^D)^T \\ &= (A^D A A^D)^T \\ &= (A^D)^T && \text{Using (11.2)} \\ &= G \end{aligned}$$

$$\begin{aligned} BG &= A^T (A^D)^T \\ &= (A^D A)^T \\ &= (A A^D)^T && \text{Using (11.3)} \\ &= (A^D)^T A^T \\ &= GB \end{aligned}$$

Combining these results, G is the *drazin inverse* of B.
$\therefore (A^T)^D = (A^D)^T$. $\square$

11.5. Let A be an $m \times n$ matrix and be defined in the form $A = XBX^{-1}$ for some matrix B and non-singular matrix X. Then, $A^D = XB^D X^{-1}$.

12. Appendix

Following are the implementations of algorithms presented in the book 'Linear Algebra by A. Ramachandra Rao and P. Bhimasankaram'.

12.1. Sweeping out a column

Let A be an $m \times n$ matrix with $a_{kl} \neq 0$. Following is the method to sweep out the l^{th} column with the $(k,l)-th$ element as pivot. That is, convert l^{th} column of A to e_k.

Listing 12.1: Matlab script for sweeping out the lth column

```
%% Sweep lth column with A(k,l) as pivot
function [S] = Sweep_column(A,k,l)
[m n] = size(A);
for i = 1:m
    for j = 1:n
    if i==k
        S(i,j) = A(i,j)/A(k,l);
    else
        S(i,j) = A(i,j) - (A(i,l)*A(k,j))/A(k,l);
    end
    end
end
```

Eg. Sweep 4^{th} column with $(3,4)-th$ element as pivot.

$$A = \begin{bmatrix} 2 & -3 & 4 & 3 & 6 \\ 0 & 1 & 3 & 0 & 2 \\ -1 & 0 & 1 & 2 & 4 \\ 5 & 7 & 0 & -2 & 0 \end{bmatrix}$$

Then, output of the above matlab script is

```
Command Window
>> Sweep_column([2 -3 4 3 6;0 1 3 0 2;-1 0 1 2 4;5 7 0 -2 0],3,4)

ans =

    3.5000   -3.0000    2.5000         0         0
         0    1.0000    3.0000         0    2.0000
   -0.5000         0    0.5000    1.0000    2.0000
    4.0000    7.0000    1.0000         0    4.0000
```

12.2. Sweeping out a row

Let A be an $m \times n$ matrix with $a_{kl} \neq 0$. Following is the method to sweep out the k^{th} row with the $(k,l)-th$ element as pivot.That is, convert k^{th} row of A to e_l.

Listing 12.2: Matlab script for sweeping out the kth row

```
%% Sweep kth row with A(k,l) as pivot
function [S] = Sweep_Row(A,k,l);
[m n] = size(A);
for j = 1:n
    for i = 1:m
        if j==l
            S(i,j) = A(i,j)/A(k,l);
        else
            S(i,j) = A(i,j) - (A(k,j)/A(k,l))*A(i,l);
        end
    end
end
```

Eg. Sweep 3^{rd} row with $(3,4)-th$ element as pivot.

$$A = \begin{bmatrix} 2 & -3 & 4 & 3 & 6 \\ 0 & 1 & 3 & 0 & 2 \\ -1 & 0 & 1 & 2 & 4 \\ 5 & 7 & 0 & -2 & 0 \end{bmatrix}$$

Then, output of the above matlab script is

```
Command Window
>> Sweep_Row([2 -3 4 3 6;0 1 3 0 2;-1 0 1 2 4;5 7 0 -2 0],3,4)

ans =

    3.5000   -3.0000    2.5000    1.5000         0
         0    1.0000    3.0000         0    2.0000
         0         0         0    1.0000         0
    4.0000    7.0000    1.0000   -1.0000    4.0000
```

12.3. Echelon form

Let A be an $m \times n$ matrix. Echelon form of a matrix is structured as

$$\begin{bmatrix} 0 & 1 & * & * & * & * \\ 0 & 0 & 0 & 1 & * & * \\ 0 & 0 & 0 & 0 & 1 & * \\ 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

Following is the method to reduce A into Echelon form.

Listing 12.3: Matlab script for reduction to echelon form

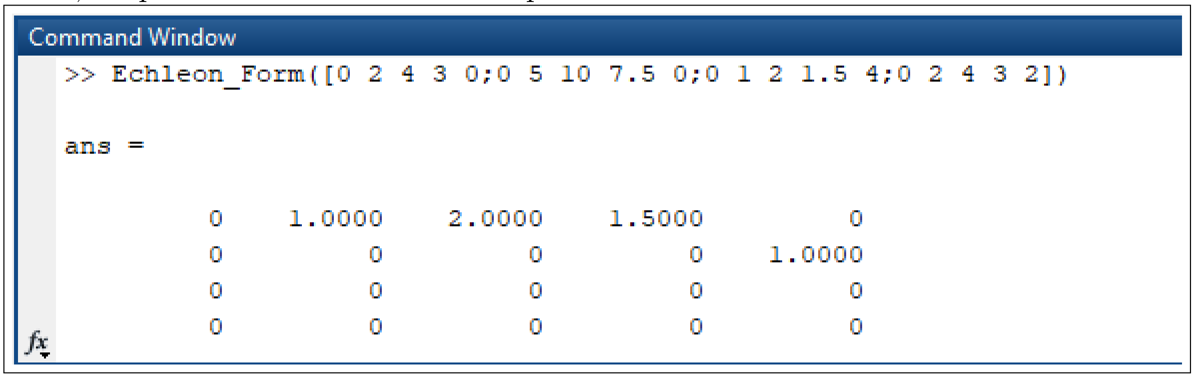
```
%%Row-reduced Echelon form
function [ECH] = Echelon_Form(A)
[m n] = size(A); %%size() is a function returns the two element row vector, i.e., D = [m,n]
i = 1;
j = 1;
l = 1;
flag = 0;
while 1
    if A(i,j) == 0
        if i~=m
            i = i+1;
            continue;
        else
            flag = 1;
        end
    else
        A([l,i],:) = A([i,l],:);
        A = Sweep_column(A,l,j);
        l = l+1;
        flag = 1;
    end
    if flag~=0
        if l==m+1 || j==n
            break;
        else
            j = j+1;
            i = 1;
        end
    end
```

```
end
ECH = A;
```

Eg. Reduce A to Echelon form.

$$A = \begin{bmatrix} 0 & 2 & 4 & 3.0 & 0 \\ 0 & 5 & 10 & 7.5 & 0 \\ 0 & 1 & 2 & 1.5 & 4 \\ 0 & 2 & 4 & 3.0 & 2 \end{bmatrix}$$

Then, output of the above matlab script is



12.4. Normal form

Let A be an $m \times n$ matrix. Normal form of a matrix is structured as

$$\begin{bmatrix} I_r & 0 \\ 0 & 0 \end{bmatrix}$$

Following is the method to reduce A into Normal form.

Listing 12.4: Matlab script for reduction to Normal form

```
%% Reduction to normal form
function [NF] = Normal_form(A)
[m n] = size(A);
i = 1;
j = 1;
k = 0;
while 1
    if A(i,j) == 0
        if i < m
            i = i + 1;
            continue;
        else
            if j < n
                j = j + 1;
                i = k + 1;
                continue;
            else
                break;
            end
        end
    else
        k = k+1;
        A([k,i],:) = A([i,k],:);
        A(:,[k,j]) = A(:,[j,k]);
        A = Sweep_column(A,k,k);
        A = Sweep_Row(A,k,k);
        if k ~= min(m,n)
            i = k+1;
            j = k+1;
            continue;
        else
            break;
```

```
            end
        end
end
NF = A;
```

Eg. Reduce A to Normal form.

$$A = \begin{bmatrix} 2 & 1 & 1 & 3 & 4 \\ 0 & 1 & -1 & 1 & 2 \\ 0 & 1 & -1 & 1 & 2 \\ 2 & 1 & 1 & 3 & 4 \end{bmatrix}$$

Then, output of the above matlab script is

```
Command Window

>> Normal_form([2 1 1 3 4;0 1 -1 1 2;0 1 -1 1 2;2 1 1 3 4])

ans =

     1     0     0     0     0
     0     1     0     0     0
     0     0     0     0     0
     0     0     0     0     0
```

Bibliography

[1] R.B. Bapat, *Linear Algebra and Linear Models*, Hindustan Book Agency (2012)
ISBN 9380250282.

[2] A. Ramachandra Rao and P. Bhimasankaram, *Linear Algebra*, Hindustan Book Agency (2000)
ISBN 8185931267.

[3] Adi Ben-Israel and Thomas N.E. Greville, *Generalized Inverses-Theory and Applications*, Springer-Verlag New York (2003)
ISBN 0-387-00293-6.

[4] Chapter-8, *Generalized Inverses*
`http://home.iitk.ac.in/ rksr/html/08Ginv.htm`

[5] C. Radhakrishna Rao and Sujit Kumar Mitra, *Generalized Inverse of a matrix and its applications*

[6] Charles A. Rohde, *Contibutions to the theory, computations and applications of Generalized imverses*
Mimeograph Series No. 392

Prepared in LaTeX 2_ε