# Introduction to Moore-Penrose Inverse

P. Sam Johnson

# Overview

For an $m \times n$ matrix $A$, consider the Moore-Penrose[1] (MP) equations

$$AXA = A \tag{MP1}$$

$$XAX = X \tag{MP2}$$

$$(AX)^* = AX \tag{MP3}$$

$$(XA)^* = XA. \tag{MP4}$$

We shall discuss the advantages of looking for "solutions ($X$)" that statisfy some of the MP-equations.

[1]The idea of a generalized inverse of a singular matrix goes back to **E.H. Moore** in a paper published in 1920. He investigated the idea of a "general reciprocal" of a matrix again in a paper in 1935. Independently, **R. Penrose** rediscovered Moore's idea in 1955.

# Why Moore's work was unknown in 1955?

Answer: Telegraphic style and idiosyncratic notation. Example:

(29.3) **Theorem.**

$$\mathfrak{U}^{C}\, \mathfrak{B}^{1\ \mathrm{II}}\, \mathfrak{B}^{2\ \mathrm{II}}\, \kappa^{12} \cdot) \cdot$$

$$\exists \mid \lambda^{21} \text{ type } \mathfrak{M}^{2}_{\kappa *}\ \overline{\mathfrak{M}}^{1}_{\kappa}\ \ni \cdot S^{2}\, \kappa^{12}\, \lambda^{21} = \delta^{11}_{\mathfrak{M}^{1}_{\kappa}} \cdot S^{1}\, \lambda^{21}\, \kappa^{12} = \delta^{22}_{\mathfrak{M}^{2}_{\kappa *}}$$

---

English translation:

(29.3) **Theorem.**

For every matrix $A$ there exists a unique matrix $X : R(A) \to R(A^*)$ such that

$$AX = P_{R(A)}\ ,\ XA = P_{R(A^*)}\ .$$

# Notation

We introduce some notations.

Let $A\{1\} = \{G : AGA = A\}$, $A\{2\} = \{H : HAH = H\}$, and so forth.

For example, $A\{1,2\} = A\{1\} \cap A\{2\}$. That is, a $\{1,2\}$-inverse of $A$ is a matrix that satisfies (MP1) and (MP2). Evidently we have the inclusions

$$A\{1,2,3,4\} \subseteq A\{1,2,3\} \subseteq A\{1,2\} \subseteq A\{1\}.$$

Of course, many other chains are also possible.

Even though the results can be proved over an arbitrary field, we deal with complex matrices. The set of complex $m \times n$ matrices will be denoted by $\mathbb{C}^{m \times n}$. We denote the range of $A$ by $R(A)$ and the null space of $A$ by $N(A)$. The conjugate transpose of $A$ is denoted by $A^*$. $\mathbb{O}$ and $\mathbb{I}$ represent the zero and the identity matrices respectively in appropriate orders.

# Outline of the talk

Given an $m \times n$ complex matrix $A$, we discuss properties, results and existence of the following inverses.

- $\{1\}$-inverses of $A$.
- $\{2\}$-inverses of $A$.
- $\{1, 2\}$-inverses of $A$.
- $\{1, 3\}$-inverses of $A$.
- $\{1, 4\}$-inverses of $A$.
- $\{1, 2, 3, 4\}$-inverse of $A$.
  We shall prove that $\{1, 2, 3, 4\}$-inverse of $A$ is unique ; called the **Moore-Penrose inverse of** $A$ and denoted by $A^\dagger$.

**Given an $m \times n$ complex matrix $A$, we now discuss the**

$\{1\}$-inverses of $A$.

$\{1\}$-inverses are the "equation solvers."

## Matrix equation : $Ax = b$

Let $A \in \mathbb{C}^{m \times n}$. Then a **natural question** is when we can solve

$$Ax = b \quad \text{for} \quad x \in \mathbb{C}^n, \quad \text{given} \quad b \in \mathbb{C}^m. \tag{1}$$

If $A$ is a square matrix $(m = n)$ and $A$ has an inverse, then (1) has a (unique) solution $x = A^{-1}b$.

**This gives a complete answer if $A$ is invertible.**
However, $A$ may be a square matrix that is not invertible, or $A$ may be $m \times n$ with $m \neq n$.

# When $A$ is not invertible, what happens?

**If $A$ is not invertible,** then

- (a) equation (1) **may have no solutions** (that is, $b$ may not be in $R(A)$), and
- (b) if there are solutions, then there may be **many different solutions.**

**Questions :**

1. Given $b$, is the system $Ax = b$ solvable?
2. If $Ax = b$ is solvable for a given $b$, how to find all possible solutions of $Ax = b$?

Answers to these questions can be found from the notion of a **$\{1\}$-inverse of $A$**. It is also called as a **generalized inverse $A$**.

# $\{1\}$-inverse

**Definition 1.**

*If $A$ is an $m \times n$ matrix, then* $G$ **is a** $\{1\}$**-inverse of** $A$ *if $G$ is an $n \times m$ matrix with*

$$AGA = A. \tag{2}$$

*If $A^{-1}$ exists in the usual sense, then $G = A^{-1}$. This justifies the term* **generalized inverse.**

**We will see later that every matrix $A$ has at least one $\{1\}$-inverse of $A$.**

However, unless $A$ is $n \times n$ and $A$ is invertible, there are many different $\{1\}$-inverses $G$, so that $G$ is not unique. We shall see that $\{1\}$-inverses are unique if we impose more conditions on $G$.

# $\{1\}$-inverse

The idea of a generalized inverse can be found in the book by **Baer** [1952][2].

Baer's idea was later developed by **Sheffield** [1958] in a paper [3].

An immediate consequence of the relation $AGA = A$ is that

$$AGAG = AG \text{ and } GAGA = GA.$$

**Thus both $AG_{m\times m}$ and $GA_{n\times n}$ are idempotent matrices.**

[2] Reinhold Baer, Linear Algebra and Projective Geometry, Academic Press, Inc., New York, (1952).

[3] R.D.Sheffield, A General Theory For Linear Systems, AMM, February, (1958), 109-111.

# Idempotent Matrix

A square matrix $P$ that satisfies $P^2 = P$ is called an **idempotent matrix**. If $P$ is an idempotent matrix, then $P = P^2$ implies $Py = P(Py)$ and $Pz = z$ for all $z = Py$ in the range of $P$. That is, if $P$ is $n \times n$, $P$ moves any $x \in \mathbb{C}^n$ into $R(P)$ and then keeps it at the same place.

Let $x \in \mathbb{C}^n$. Then $x = Px + (x - Px) = y + z$ satisfies $Py = P(Px) = Px = y$ and $Pz = P(x - Px) = 0$. Since $Px = P(y + z) = y$, we can say that $P$ projects $\mathbb{C}^n$ **onto** its range $V(= R(P))$ **along** the space $W = \{x : Px = 0\}(= N(P))$.

- Each idempotent $P$ on $\mathbb{C}^n$ decomposes $\mathbb{C}^n$ as the direct sum, $\mathbb{C}^n = R(P) + N(P)$, and vice-versa.
- Each projecton $P$ (idempotent and $P = P^*$) on $\mathbb{C}^n$ decomposes $\mathbb{C}^n$ as the orthogonal direct sum, $\mathbb{C}^n = R(P) \oplus N(P)$, and vice-versa.

The two idempotents $AG$ and $GA$ appearing in the next result are useful to show how $\{1\}$-inverses are used to solve matrix equations.

**Theorem 1.**

*Let $A$ be an $m \times n$ matrix and assume that $G$ is a $\{1\}$-inverse of $A$. Then, for any fixed $y \in \mathbb{C}^m$,*

(i) *the equation $Ax = b$ $(x \in \mathbb{C}^n)$ has a solution (the system of linear equations $Ax = b$ is solvable) iff $AGb = b$ (that is, iff $Gb$ is a solution of $Ax = b$, or $b \in R(AG) = R(A)$).*

(ii) *If a solution exists, they every solution is of the form*

$$x = Gb + (\mathbb{I} - GA)z$$

*where $z$ is an arbitrary element in $\mathbb{C}^n$.* LA-3(P-1)T-1

## Example 2.

*Let* $A = \begin{pmatrix} 1 & 2 \\ 3 & 6 \end{pmatrix}$. *Set* $G = \begin{pmatrix} 1 & 0 \\ 0 & 0 \end{pmatrix}$, *then* $AGA = A$.

*Also* $AG = \begin{bmatrix} 1 & 0 \\ 3 & 0 \end{bmatrix}$, $GA = \begin{bmatrix} 1 & 2 \\ 0 & 0 \end{bmatrix}$ *and* $I - GA = \begin{bmatrix} 0 & -2 \\ 0 & 1 \end{bmatrix}$.

$Ax = b$ *has a solution iff* $b \in R(A) = R(AG) = span\left\{ \begin{bmatrix} 1 \\ 3 \end{bmatrix} \right\}$. *By the Theorem 1, if* $b = c \begin{pmatrix} 1 \\ 3 \end{pmatrix}$, *then any solution of* $Ax = b$ *is exactly of the form:*

$$\begin{aligned} x =& Gc \begin{pmatrix} 1 \\ 3 \end{pmatrix} + (I - GA) \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \\ =& c \begin{pmatrix} 1 \\ 0 \end{pmatrix} + z_2 \begin{pmatrix} -2 \\ 1 \end{pmatrix}, \text{ for arbitrary } z_2. \end{aligned}$$

# Solving matrix equation using $\{1\}$-inverses

In fact, Theorem (1) can be proved for the **matrix equation**

$$AXB = C$$

where $A \in \mathbb{C}^{m\times n}$, $X \in \mathbb{C}^{n\times p}$, $B \in \mathbb{C}^{p\times q}$ and, necessarily, $C \in \mathbb{C}^{m\times q}$.

Here $A, B, C$ are given and we are to solve for $X$.

**Theorem 2.**

*Let $AXB = C$ be as above. Then this equation has a solution iff there exist an $A^g \in A\{1\}$ and a $B^g \in B\{1\}$ such that $AA^g CB^g B = C$ (the* **consistency condition***).*
*If solutions exist, they are all of the form*

$$X = A^g CB^g + W - A^g AWBB^g,$$

*where $W$ is arbitrary in $\mathbb{C}^{n\times p}$, for some $A^g \in A\{1\}$ and $B^g \in B\{1\}$.*
*LA-3(P-5)T-4*

**Corollary 3.**

*Consider the special case $AXA = A$. This equation always has solutions since the consistency condition*

$$AA^g AA^g A = A \quad \text{holds true for any } A^g \in A\{1\}.$$

*Moreover, $A\{1\} = \{X : X = A^g AA^g + W - A^g AWAA^g\}$, where $W$ is arbitrary and $A^g$ is some $\{1\}$-inverse of $A$.* LA-3(P-7)C-5

**Corollary 4.**

*Consider the matrix equation $AX = C$, $A \in \mathbb{C}^{m\times n}$, $X \in \mathbb{C}^{n\times p}$, and $C \in \mathbb{C}^{m\times p}$. This equation is solvable iff $AA^g C = C$ for some $A^g \in A\{1\}$ and then the general solution is*

$$X = A^g C + (\mathbb{I} - A^g A)W,$$

*where $W$ is arbitrary and $A^g$ is some $\{1\}$-inverse of $A$.*

# Solving matrix equation using $\{1\}$-inverses

## Corollary 5.

*Consider the matrix equation $XB = C$ where $X \in \mathbb{C}^{n\times p}$, $B \in \mathbb{C}^{p\times q}$, and $C \in \mathbb{C}^{n\times q}$. This equation is solvable iff $CB^g B = C$ for some $B^g \in B\{1\}$ and then the general solution is $X = CB^g + W(\mathbb{I} - BB^g)$, where $W$ is arbitrary.*

## Corollary 6.

*Consider the matrix equation $AX = \mathbb{O}$, where $A \in \mathbb{C}^{m\times n}$ and $X \in \mathbb{C}^{n\times p}$. Then this equation always has solutions since the consistency condition evidently holds for any $\{1\}$-inverse of $A$. The solutions are of the form $X = (\mathbb{I} - A^g A)W$, where $W$ is arbitrary and $A^g$ is some $\{1\}$-inverse of $A$.*

# Solving matrix equation using $\{1\}$-inverses

**Corollary 7.**

*Consider the matrix equation $XB = \mathbb{O}$. This equation always has solutions since the consistency condition evidently holds. The solutions are of the form $X = W(\mathbb{I} - BB^g)$, where $W$ is arbitrary and $B^g$ is some $\{1\}$-inverse of $B$.*

**Corollary 8.**

*Consider a system of linear equations $Ax = b$, where $A \in \mathbb{C}^{m\times n}$, $x \in \mathbb{C}^{n\times 1}$, and $b \in \mathbb{C}^{m\times 1}$. Then this system is solvable iff $AA^g b = b$ for any $A^g \in A\{1\}$ and the solutions are all of the form*

$$x = A^g b + (\mathbb{I} - A^g A)w,$$

*where $w \in \mathbb{C}^{n\times 1}$ is arbitrary and $A^g$ is some $\{1\}$-inverse of $A$.*

# Solving matrix equation using $\{1\}$-inverses

We recall that for a system of linear equations $Ax = b$, $A^{-1}$, if it exists, has the property that $A^{-1}b$ is a solution for every choice of $b$. It turns out that the $\{1\}$-inverses of $A$ generalize this property for arbitrary $A$.

**Theorem 3.**

*Consider the system of linear equations $Ax = b$, where $A \in \mathbb{C}^{m \times n}$. Then $Gb$ is a solution of this system for every $b \in R(A)$ iff $G \in A\{1\}$.* LA-3(P-8)T-6

As Campbell and Meyer put it, the "equation solving" generalized inverses of $A$ are exactly the $\{1\}$-inverses of $A$.

Next, we give **a way of generating $\{1\}$-inverses** of a given matrix. First, we need the following.

**Theorem 4.**

*If $S$ and $T$ are invertible matrices and $G$ is a $\{1\}$-inverse of $A$, then $T^{-1}GS^{-1}$ is a $\{1\}$-inverse of $SAT$ ($TGS$ is a $\{1\}$-inverse of $S^{-1}AT^{-1}$). Moreover, every $\{1\}$-inverse of $SAT$ is of this form.* LA-3(P-8)T-7

# Finding $\{1\}$-inverses of a given matrix

**Theorem 5 (ABCD Theorem).**

*Let $A$ be an $m \times n$ matrix with $\text{rank}(A) = r$. Then, after a suitable rearrangement of rows and columns, $A$ can be written in partitioned form as*

$$A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} \tag{3}$$

*where $a$ is an $r \times r$ invertible matrix. In that case $d = ca^{-1}b$, so that*

$$A = \begin{pmatrix} a & b \\ c & ca^{-1}b \end{pmatrix}. \tag{4}$$

*Note that $a, b, c, d$ are matrices, not numbers.* LA-3(P-2)R-2

# Observations

1. Some of the entries b,c,d in (3) may be empty, in which case they do not appear, for example if $m = n$ and $A$ is invertible.
2. (4) can also be written

$$A = \begin{pmatrix} \mathbb{I}_r \\ ca^{-1} \end{pmatrix} \begin{pmatrix} a & b \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} a^{-1} \begin{pmatrix} a & b \end{pmatrix} = \begin{pmatrix} a \\ c \end{pmatrix} \begin{pmatrix} \mathbb{I}_r & a^{-1} b \end{pmatrix}.$$

This can be viewed as a generalization of the representation $A = uv^T$ for an **outer product** of two vectors $u, v$.

# Finding $\{1\}$-inverses of a given matrix

**Theorem 6.**

*Let $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & ca^{-1}b \end{pmatrix}$ be an $m \times n$ matrix with $r = rank(A)$ where $a$ is an invertible $r \times r$ matrix, as in Theorem 5.*
*Let*

$$G = \begin{pmatrix} a^{-1} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix} \tag{5}$$

*where the $\mathbb{O}$s in (5) represent matrices of zeroes of dimension sufficient to make $G$ an $n \times m$ matrix. Then $G$ is a $\{1\}$-inverse of $A$.* LA-3(P-5)T-3

Two idempotents in this case are

$$AG = \begin{pmatrix} \mathbb{I}_r & \mathbb{O} \\ ca^{-1} & \mathbb{O} \end{pmatrix} \text{ and } GA = \begin{pmatrix} \mathbb{I}_r & a^{-1}b \\ \mathbb{O} & \mathbb{O} \end{pmatrix}.$$

The above theorem says that $Ax = b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ can be solved for $b_1 \in \mathbb{C}^r$, $b_2 \in \mathbb{C}^{m-r}$ iff

$$AGb = \begin{pmatrix} b_1 \\ ca^{-1}b_1 \end{pmatrix} = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} = b.$$

That is, $Ax = b = \begin{pmatrix} b_1 \\ b_2 \end{pmatrix}$ can be solved for $b_1 \in \mathbb{C}^r$, $b_2 \in \mathbb{C}^{m-r}$ iff $b_2 = ca^{-1}b_1$. In that case, the general solution of $Ax = b$ for $x \in \mathbb{C}^n$ is

$$\begin{aligned} x = \begin{pmatrix} x_1 \\ x_2 \end{pmatrix} &= Gb + (\mathbb{I}_m - GA)z \\ &= \begin{pmatrix} a^{-1} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \end{pmatrix} + \begin{pmatrix} \mathbb{O} & -a^{-1}b \\ \mathbb{O} & \mathbb{I}_{m-r} \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \\ &= \begin{pmatrix} a^{-1}y_1 \\ \mathbb{O} \end{pmatrix} + \begin{pmatrix} -a^{-1}bz_2 \\ z_2 \end{pmatrix} \end{aligned}$$

for arbitrary $z_2 \in \mathbb{C}^{m-r}$.

# If $rank(A) = r < \min\{m, n\}$, we have infinitely many $\{1\}$-inverses.

**Remark 9.**

*Let* $A = \begin{pmatrix} a & b \\ c & d \end{pmatrix} = \begin{pmatrix} a & b \\ c & ca^{-1}b \end{pmatrix}$ *be an* $m \times n$ *matrix with* $r = rank(A)$ *where* $A$ *is an invertible* $r \times r$ *matrix. Then* $A$ *has at least one* $\{1\}$*-inverse* $G$ *of the form* $G = \begin{pmatrix} a^{-1} & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{pmatrix}$*. Since often many different linearly independent set of* $r$ *rows can be permuted to the upper* $r$ *rows and many different linearly independent sets of* $r$ *columns can be permuted into the first* $r$ *column positions, a matrix* $A$ *with* $rank(A) = r < \min\{m, n\}$ *can have many different* $\{1\}$*-inverses of this form.*

# Rank Normal Form

**Theorem 7.**

*Any matrix $A \in \mathbb{C}^{m\times n}$ is equivalent to a unique matrix of the form*

$$\begin{cases} \mathbb{I}_r & \text{if } m = n = r, \\ \begin{bmatrix} \mathbb{I}_r & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix} & \text{if } m > r, n > r, \\ \begin{bmatrix} \mathbb{I}_r \ \vdots \ \mathbb{O} \end{bmatrix} & \text{if } m = r < n, \\ \begin{bmatrix} \mathbb{I}_r \\ \dots \\ \mathbb{O} \end{bmatrix} & \text{if } m > r = n \end{cases}$$

*called the* **rank normal form** *of $A$ and denoted by $RNF(A)$.*

# Rank Normal Form

**Theorem 8.**

*Let $A$ be an $m \times n$ matrix of rank $r$. Then there exist two invertible matrices $S$ (of size $m \times m$) and $T$ (of size $n \times n$) such that* $SAT = \begin{bmatrix} \mathbb{I}_r & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix}$.

**Example 9.**

Consider $A = \begin{bmatrix} 1 & 2 & 3 & 4 \\ 2 & 4 & 6 & 7 \\ 1 & 2 & 3 & 6 \end{bmatrix}$.

$$RNF(A) = \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix} = \begin{bmatrix} -7 & 4 & 0 \\ 2 & -1 & 0 \\ -5 & 2 & 1 \end{bmatrix} A \begin{bmatrix} 1 & 0 & -2 & -3 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}.$$

# Finding $\{1\}$-inverses of a given matrix

**Theorem 10 (Bose).**

*If $A \in \mathbb{C}_r^{m\times n}$, there exist invertible matrices $S$ and $T$ with $SAT = \begin{bmatrix} \mathbb{I}_r & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix}$ (Rank Normal Form).*

*A matrix $G$ is a $\{1\}$-inverse of $A$ iff $G = TNS$, where $N = \begin{bmatrix} \mathbb{I}_r & Y \\ X & W \end{bmatrix}$ where $X, Y$, and $W$ are arbitrary matrices of approximate size.* LA-3(P-9)T-8

We will approach $\{1\}$-inverses using the Hermite echelon form. First, we shall see what is Hermite echelon form and some properties of it.

There is another useful way to reduce a matrix, name in honor of the French mathematician **Charles Hermite**, that is very close to the RREF. However, it is only defined for **square matrices**. Statisticians have known about this for some time.

### Definition 11 (Hermite echelon form).

A matrix $H$ in $\mathbb{C}^{n \times n}$ in (upper) Hermite echelon form iff

1. $H$ is upper triangular.
2. The diagonal of $H$ consists only of zeros and ones.
3. If a row has a zero on the diagonal, then every element of that row is zero.
4. If a row has a 1 on the diagonal, then every other element in the column containing that 1 is zero.

## The Hermite Echelon Form

Algorithm for finding $HEF(A)$ (Hermite Echelon Form of $A$) is described as follows:

1. First we use elementary row operations to produce $RREF(A)$ (Row Reduced Echelon Form of $A$).
2. Then permute the rows of $RREF(A)$ until each first non-zero element of each nonzero row is a diagonal element.
3. The resulting matrix is in Hermite echelon form.
4. Use elementary row operations to produce

$$[A \mid \mathbb{I}] \to [HEF(A) \mid S].$$

5. Then $SA = HEF(A)$.

**Example 12.**

Consider $A = \begin{bmatrix} 3 & 6 & 9 \\ 1 & 2 & 5 \\ 2 & 4 & 10 \end{bmatrix}$.

$$RREF(A) = \begin{bmatrix} -5/6 & -3/2 & 0 \\ -1/6 & 1/2 & 0 \\ 0 & -1/2 & 1/4 \end{bmatrix} \begin{bmatrix} 3 & 6 & 9 \\ 1 & 2 & 5 \\ 2 & 4 & 10 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}.$$

To set Hermite echelon form, simply permute the second and third rows.

$$HEF(A) = \begin{bmatrix} -5/6 & -3/2 & 0 \\ 0 & -1/2 & 1/4 \\ -1/6 & 1/2 & 0 \end{bmatrix} \begin{bmatrix} 3 & 6 & 9 \\ 1 & 2 & 5 \\ 2 & 4 & 10 \end{bmatrix} = \begin{bmatrix} 1 & 2 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix}.$$

# The Hermite Echelon Form : Interesting Facts

1. A matrix in Hermite echelon form must be idempotent ($A^2 = A$).
2. Every matrix $A \in \mathbb{C}^{n \times n}$ can be brought into Hermite echelon form by using elemetary row operations.
3. For any $A \in \mathbb{C}^{n \times n}$, there exists a invertible matrix $S$ such that $SA$ is in Hermite echelon form.
4. The Hermite echelon form of a matrix is unique.
5. Let $A, B \in \mathbb{C}^{n \times n}$. Then $HEF(A) = HEF(B)$.
6. For any $A \in \mathbb{C}^{n \times n}$ and $S \in \mathbb{C}^{n \times n}$ invertible, $HEF(SA) = HEF(A)$ iff $R(A^*) = R(B^*)$. In other words, row equivalent matrices have the same Hermite echelon form.

Hermite Echelon form of square matrices helps to generate $\{1\}$-inverses, indeed invertible ones.

**Theorem 13 (Constructing $\{1\}$-inverse using Hermite echelon form).**

*Let $A \in \mathbb{C}^{n \times n}$ (square) and $S$ be invertible such that $SA = HEF(A)$. Then $S$ is a $\{1\}$-inverse of $A$.*

# $\{1\}$-inverse of a product of two matrices

In 1975, Robert E. Hartwig published a paper where he proposed a formula for the $\{1\}$-inverse of a product of two matrices. we present this next.

**Theorem 14 (Hartwig, 1975).**

*For conformal matrices $A$ and $B$,*

$$(AB)^g = B^g A^g - B^g(\mathbb{I} - A^g A)\left[(\mathbb{I} - BB^g)(\mathbb{I} - A^g A)\right]^g(\mathbb{I} - BB^g)A^g.$$

# Properties of $A\{1\}$

1. If $G_1, G_2 \in A\{1\}$ for some matrix $A$, then show that

$$\lambda G_1 + (1-\lambda) G_2 \in A\{1\}, \; 0 \leq \lambda \leq 1.$$

   In other words, prove that $A\{1\}$ is an affine set.

2. Suppose $G_1, G_2, \ldots, G_k$ are in $A\{1\}$ and $\lambda_1, \lambda_2, \ldots, \lambda_k$ are scalars that sum to 1. Prove that

$$\lambda_1 G_1 + \lambda_2 G_2 + \ldots + \lambda_k G_k \text{ is in } A\{1\}.$$

3. Is it true that any linear combination of $\{1\}$-inverses of a matrix $A$ is again a $\{1\}$-inverse of $A$?
4. Let $S$ and $T$ be invertible matrices. Then show that $T^{-1} A^g S^{-1}$ is a $\{1\}$-inverse of $SAT$ for any $A^g \in A\{1\}$.
5. Prove that $R(B) \subseteq R(A)$ iff $AA^g B = B$ for any $A^g \in A\{1\}$.

# Properties of $A\{1\}$

1. If $G \in A\{1\}$, then prove that $GA$ and $AG$ are both idempotent matrices that have the same rank as $A$. What direct sum decompositions do they generate?
2. If $G \in A\{1\}$, then prove that
   - (a) $rank(A) = rank(AG) = rank(GA) = trace(AG) \leq rank(G)$.
   - (b) Show that $rank(\mathbb{I} - AG) = m - rank(A)$ and $rank(\mathbb{I} - GA) = n - rank(A)$.

# Properties of $A\{1\}$

1. Let $A \in \mathbb{C}^{n\times n}$ and $H = HEF(A)$ (Hermite Echelon Form of $A$). Prove $A$ is idempotent if and only if $H$ is a $\{1\}$-inverse of $A$.
2. If $G$ is a $\{1\}$-inverse of $A$, prove that $GAG$ is also a $\{1\}$-inverse of $A$ and has the same rank as $A$.
3. Show that it is always possible to construct a $\{1\}$-inverse of $A \in \mathbb{C}_r^{m\times n}$ that has $rank = \min\{m, n\}$. In particular, prove that every square matrix has an invertible $\{1\}$-inverse.
4. What is $\mathbb{O}_{n\times n}\{1\}$?
   What is $\mathbb{I}_n\{1\}$?
   Let $E_{ij}$ be the $m \times n$-matrix with all entries zero except that the $(i, j)$ entry has a one. What is $E_{ij}\{1\}$?
5. Find a $\{1\}$-inverse for $\begin{bmatrix} 1 & 0 & 2 & 3 \\ 0 & 1 & 4 & 5 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{bmatrix}$.

# Properties of $A\{1\}$

1. Show that $A \in \mathbb{C}_r^{m\times n}$ can have a $\{1\}$-inverse of any rank between $r$ and $\min\{m, n\}$. [Hint: rank $\begin{bmatrix} \mathbb{I}_r & \mathbb{O} \\ \mathbb{O} & D \end{bmatrix} = r + rank(D)$.]
2. Prove that any square invertible matrix has a unique $\{1\}$-inverse.
3. If $G \in A\{1\}$, then show that $G^* \in A^*\{1\}$.
4. If $G \in A\{1\}$ and $\lambda \in \mathbb{C}$, then show that $\lambda^+ G \in (\lambda A)\{1\}$, where $\lambda^+ = \begin{cases} 0, & \text{if } \lambda = 0 \\ \lambda^{-1}, & \text{if } \lambda \neq 0 \end{cases}$.

# Properties of $A\{1\}$

1. If $G \in A\{1\}$, then prove that $rank(G) \geq rank(A)$.
2. If $G \in A\{1\}$, then prove that $rank(G) = rank(A)$ iff $G \in A\{2\}$.
3. Let $G \in A\{1\}$. Prove that $R(AG) = R(A)$, $N(AG) = N(A)$, and $R((GA)^*) = R(A^*)$.
4. Let $G \in A\{1\}$. Prove that
   - (a) $GA = \mathbb{I}$ iff $r = n$ iff $G$ is a left inverse of $A$ iff $A$ has a full column rank.
   - (b) $AG = \mathbb{I}$ iff $r = m$ iff $G$ is a right inverse of $A$ iff $A$ has a full row rank.
5. Suppose $G \in A\{1\}$ and $v \in N(A)$. Prove that $G_1 = [g_1|\dots|g_j + v|\dots|g_m] \in A\{1\}$.
6. Find matrices $G$ and $A$ such that $G \in A\{1\}$ but $A \notin G\{1\}$.

**Given an $m \times n$ complex matrix $A$, we now discuss the**

## $\{2\}$-inverses of $A$.

**($A$ is a $\{1\}$-inverse of "$\{2\}$-inverse of $A$." )**

# $\{2\}$-inverse

The problem of finding 2-inverses is a bit more challenging than that of describing $\{1\}$-inverses because it is a "quadratic" problem in the unknowns.

Let us look at the $2 \times 2$ case. Given $A = \begin{bmatrix} a & b \\ c & d \end{bmatrix}$, find $X = \begin{bmatrix} x & y \\ u & v \end{bmatrix}$ such that $XAX = X$. This amounts to solving the following equations for $x, y, u$ and $v$

$$\begin{aligned} x &= x^2a + ycx + xbu + ydu \\ y &= xay + y^2c + xbv + ydv \\ u &= uax + vcx + u^2b + vdu \\ v &= uay + vcy + ubv + v^2d \end{aligned}$$

which are quadratic in the unknowns. Recall that equations solving for $\{1\}$-inverses are linear in the unknowns.

# Generating $\{2\}$-inverses

**Theorem 15.**

*Let $A \in \mathbb{C}^{m\times n}$ and suppose $S$ and $T$ are invertible matrices of appropriate size.*
*Also suppose $X$ is a $\{2\}$-inverse of $A$. Then $T^{-1}XS^{-1}$ is a $\{2\}$-inverse of $SAT$. Moreover, every $\{2\}$-inverse of $SAT$ is of this form.*

We will approach $\{2\}$-inverses using the rank normal form.

# Constructing $\{2\}$-inverses using "rank normal form"

**Theorem 16 (Bailey, 2002).**

*Let* $A \in \mathbb{C}_r^{m\times n}$. *Suppose* $RNF(A) = SAT = \begin{bmatrix} \mathbb{I}_r & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix}$ *for appropriate invertible matrices* $S$ *and* $T$.
*A matrix* $X$ *is a* $\{2\}$*-inverse of* $A$ *if and only if* $X = TRS$, *where* $R = \begin{bmatrix} M & Y \\ W & WY \end{bmatrix}$ *where* $M^2 = M, MY = Y, WM = W$.

**Given an $m \times n$ complex matrix $A$, we now discuss the**

$\{1, 2\}$-inverses of $A$.

**(reflexive generalized inverses of $A$)**

# $\{1,2\}$-inverses

C.R. Rao in 1955 made use of a generalized inverse that satisfied (MP1) and (MP2). This type of inverse is sometimes called a **reflexive generalized inverse**.

We can describe the general form of $\{1, 2\}$-inverses as we did with $\{1\}$-inverses. It is interesting to see that extra ingredient that is needed. We take the constructive approach as usual.

**Theorem 17.**

*Let $A \in \mathbb{C}^{m \times n}$. There are matrices $S$ and $T$ with $SAT = \begin{bmatrix} \mathbb{I}_r & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix}$. A matrix $G$ is a $\{1, 2\}$-inverse of $A$ if and only if $G = TNS$, where $N = \begin{bmatrix} \mathbb{I}_r & Y \\ X & XY \end{bmatrix}$, where $X$ and $Y$ are arbitrary of appropriate size.*

*LA-3(P-12)T-10*

# Properties of $\{1,2\}$-inverses

1. Suppose $G$ is a $\{1\}$-inverse of $A$. Prove that $G$ is a $\{1,2\}$-inverse of $A$ iff $rank(G) = rank(A)$. That is, the rank $r$ 2-inverses of $A$ are exactly the $\{1,2\}$-inverses of $A$.
2. Prove that $G \in A\{1,2\}$ iff $G = G_1 A G_2$, where $G_1, G_2 \in A\{1\}$.
3. Prove that $G = E(HAE)^{-1}H$ belongs to $A\{2\}$, where $H$ and $E$ are selected judiciously so that $HAE$ is invertible.
4. Suppose $A$ and $B$ are $\{1,2\}$-inverses of each other. Prove that $AB$ is the projector onto $R(A)$ along $N(B)$ and $BA$ is the projector of $N(B)$ along $N(A)$.
5. Prove that $G \in A\{1,2\}$ iff there exist $S$ and $T$ invertible with $G = S \begin{bmatrix} \mathbb{I}_r & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix} T$ and $SAT = S \begin{bmatrix} \mathbb{I}_r & B \\ \mathbb{O} & \mathbb{O} \end{bmatrix}$.

**Given an $m \times n$ complex matrix $A$, we now discuss the**

## $\{1, 3\}$-inverses of $A$.

**(least-squares generalized inverses of $A$)**

Suppose $Ax = b$ is **inconsistent**. It seems reasonable to seek out a vector in $R(A)$ that is closest to $b$. In other words, $b \notin R(A)$, the vector $r(x) = Ax - b$, which we call the **residual vector**, is never zero. We shall try to minimize the length of this vector in the Euclidean norm.

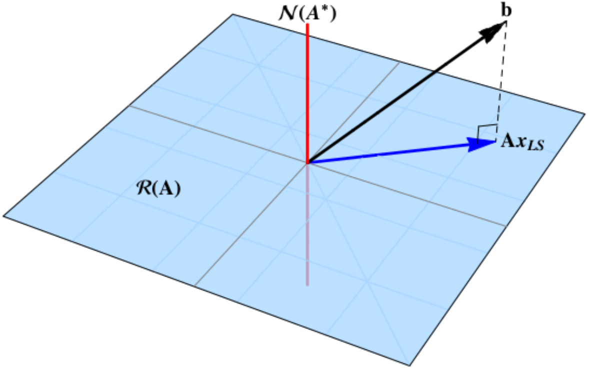
Statisticians do this all the time under the name "least squares."

**Definition 18.**

A vector $x_0$ is called a **least square solution** of the system of linear equations $Ax = b$ iff $\|Ax_0 - b\| \leq \|Ax - b\|$ for all vectors $x$.

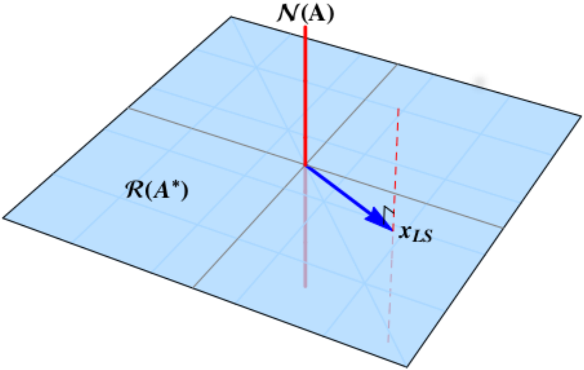
Remarkably, the connection here is with $\{1, 3\}$-inverses.

The least squares solution is shown below.



# Infinitely many least squares solutions

The least squares solution is the affine space represented by the dashed red line below. We shall prove later there exists a unique vector of smallest norm in $R(A^*)$.



# $\{1,3\}$-inverses

**Theorem 19.**

*Let $A \in \mathbb{C}^{m\times n}$ and $G \in A\{1,3\}$. Then*

1. $(\mathbb{I} - AG)^* = \mathbb{I} - AG = (\mathbb{I} - AG)^2$.
2. $(\mathbb{I} - AG)^* A = \mathbb{O}$.
3. $A^*(\mathbb{I} - AG) = \mathbb{O}$.

# $\{1,3\}$-inverses

**Theorem 20.**

*Suppose $G \in A\{1,3\}$. Then $x_0 = Gb$ is a least squares solution of the linear system $Ax = b$.* LA-3(P-14)T-11

We have seen that $Gb$ is a least squares solution of $Ax = b$.

If $x_1$ is a least squares solution, then $Ax_1$ and $AGb$ are of equal distance to $b$, which is shown below.

**Theorem 21.**

*Let $G \in A\{1,3\}$. Then $x_1$ is a least squares solution of the linear system $Ax = b$ iff $\|Ax_1 - b\| = \|AGb - b\|$.* LA-3(P-15)T-12

**Theorem 22.**

*Let $G \in A\{1,3\}$. Then $x_0$ is a least squares solution of $Ax = b$ iff $Ax_0 = AGb$.* LA-3(P-15)T-13

A linear system may have many least squares solutions for $Ax = b$. Any arbitrary least squares solutions $x_0$ is of the form $x_0 = Gb + u$, where $u \in N(A)$ because $\|A(x_0 - Gb)\| = 0$. However, we can describe them all without finding $N(A)$.

**Theorem 23.**

*Let $G \in A\{1,3\}$. Then all least squares solutions of $Ax = b$ are of the form $Gb + (\mathbb{I} - GA)z$ for any arbitrary $z$.* LA-3(P-16)T-14

It is nice that we can charaterize when a least squares solution is unique. This often happens in satistical examples.

**Theorem 24.**

*Let $A \in \mathbb{C}^{m\times n}$ and $G \in A\{1,3\}$. The system of linear equations $Ax = b$ has a unique least squares solution iff $\mathrm{rank}(A) = n$.* LA-3(P-17)T-15

# Unique least squares solution

**Theorem 25.**

*Consider the system $Ax = b$. The following are equivalent:*

1. $Ax = b$ *has a unique least squares solution.*
2. *The columns of* $A$ *are linearly independent.*
3. $rank(A) = n$.
4. $A^*A$ *is invertible.*
5. $A$ *has full column rank.*

1. Consider the system $\begin{pmatrix} 1 & 0 & 2 \\ 0 & 1 & -2 \\ 2 & 2 & 0 \\ 1 & 2 & -2 \end{pmatrix} \begin{pmatrix} x \\ y \\ z \end{pmatrix} = \begin{pmatrix} -1 \\ -1 \\ 2 \\ 1 \end{pmatrix}$. Verify that this system is inconsistent. Find all least squares solutions.
2. Prove that if $G \in AA^*\{1,2\}$, then $A^*G$ is a $\{1,3\}$-inverse of $A$.
3. Prove that $x_0$ is the minimum norm least squares solutions of $Ax = b$ iff
   - (a) $\|Ax_0 - b\| \leq \|Ax - b\|$, and
   - (b) $\|x_0\| \leq \|x\|$ for any $x \neq x_0$.
4. **Computing a least squares solution:** Prove that $x_0$ is the least squares solutions of $Ax = b$ iff $x_0$ is a solution to the always consistent (prove this) system $A^*Ax = A^*b$. These equations are often called the **normal equations**. Prove this latter is equivalent to $Ax - b \in N(A^*)$.
5. Suppose $A = FG$ is a full rank factorization. Then the normal equations are equivalent to $F^*Ax = F^*b$.

**Example 26 (Unique least squares solution).**

Consider the system $Ax = b$ where $A = \begin{pmatrix} 0 & 1 \\ 1 & 1 \\ 2 & 1 \end{pmatrix}$ and $b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$.

The only least squares solution is $\begin{pmatrix} -3 \\ 5 \end{pmatrix}$.

**Example 27 (Infinitely least squares solutions).**

Consider the system $Ax = b$ where $A = \begin{pmatrix} 1 & 0 & 1 \\ 1 & 1 & -1 \\ 1 & 2 & -3 \end{pmatrix}$ and $b = \begin{pmatrix} 6 \\ 0 \\ 0 \end{pmatrix}$.

There are infinitely many least squares soluions. Each least squares solution is of the form $\begin{pmatrix} -a+5 \\ 2a-3 \\ a \end{pmatrix}$, for some scalar $a$.

**Given an $m \times n$ complex matrix $A$, we now discuss the**

## $\{1, 4\}$-inverses of $A$.

**(minimum norm generalized inverses of $A$)**

We have seen that a consistent system of linear equations $Ax = b$ can have many solutions; indeed, there can be infinitely many solutions and they form an affine subspace.

Now we are in a position to ask, among all of these solutions, is there a shortest one? That is, there is a solution of minimum norm?

The first question is, which norm? For this section, we choose our familiar Euclidean norm, $\|x\|_2 = trace(x^*x)^{1/2}$.

**Definition 28.**

We say that $x_0$ is a minimum norm solution of $Ax = b$ iff $x_0$ is a solution and $\|x_0\| \leq \|x\|$ for all solutions $x$ of $Ax = b$.

**Theorem 29.**

*Let* $G \in A\{1,4\}$. *Then*

1. $(\mathbb{I} - GA)^* = (\mathbb{I} - GA) = (\mathbb{I} - GA)^2$.
2. $A(\mathbb{I} - GA)^* = \mathbb{O}$.
3. $(\mathbb{I} - GA)A^* = \mathbb{O}$.

# $\{1,4\}$-inverses

It turns out that the minimum norm issue is actually intimately connected with $\{1,4\}$-inverses.

**Theorem 30.**

*Suppose $Ax = b$ is consistent and $G \in A\{1,4\}$. Then $Gb$ is the unique solution of minimum norm.* LA-3(P-17)T-16

**Theorem 31 (Converse part).**

*Suppose $H \in \mathbb{C}^{n \times m}$ and, whenever $Ax = b$ is consistent, $AHb = b$ and $\|Hb\| < \|z\|$ for all solutions $z$ other than $Hb$; then $H \in A\{1,4\}$.*

Thus, $\{1,4\}$-inverses are characterized by giving the minimum norm solutions.

# Properties of $\{1,4\}$-inverses

1. Suppose $G \in A\{1,4\}$. Prove that $A\{1,4\} = \{H : HA = GA\}$.
2. Suppose $G \in A\{1,4\}$. Prove that $A\{1,4\} = \{G + W(\mathbb{I} - AG) : W \text{ is arbitrary }\}$.
3. Let $u$ and $v$ in $R(A^*)$ with $Au = Av$. Prove that $u = v$.

**Given an $m \times n$ complex matrix $A$, we now discuss the**

$$\{1,2,3,4\}\text{-inverse of } A.$$

**(Moore-Penrose inverse of $A$)**

We now develop a key concept, the **Moore-Penrose inverse** (MP inverse), also known as the **pseudoinverse**.

**Theorem 32 (the uniqueness theorem).**

*If $A \in \mathbb{C}^{m \times n}$ has a Moore-Penrose inverse at all, it must be unique. That is, there can be only one simultaneous solution to the four MP-equations.*
*LA-3(P-19)T-18*

In view of the uniqueness theorem for Moore-Penrose inverse, we use the notation $A^\dagger$ for the unique solution of the four MP-equations (when the solution exists, of course, which is yet to be established).

# Full Rank Factorization

Our approach to the Moore-Penrose inverse is to use the idea of **full rank factorization**; we build up from the factors of a full rank factroization.

Every non-null matrix can be written as a product of two full rank matrices. Matrices which are of full rank (either full row rank or full column rank) have several nice properties.

If $A$ has column rank $r$, then

- any $r$ linearly independent columns of $A$ form a basis for $R(A)$,
- every maximal linearly independent set of columns of $A$ contains exactly $r$ vectors,
- any $r$ columns of $A$ which generate $R(A)$ form a basis of $R(A)$.

# Full Rank Factorization

## Definition 33.

An $m \times n$ matrix $A$ is said to be of **full row rank** if its rows are linearly independent, that is, it its rank is $m$. Similarly $A$ is said to be of **full column rank** if its columns are linearly independent.

A **left inverse** of a matrix $A$ is any matrix $B$ such that $BA = \mathbb{I}$. A **right inverse** of $A$ is any matrix $C$ such that $AC = \mathbb{I}$.

A matrix $B$ is said to be an **inverse** of $A$ if it is both a left inverse and a right inverse of $A$.

# Full Rank Factorization

**Theorem 34.**

*Let $A \in \mathbb{C}^{m \times n}$. Then the following statements are equivalent.*

1. *$A$ has a right inverse.*
2. **Right cancellation law:** $XA = YA \Rightarrow X = Y$.
3. $XA = 0 \Rightarrow X = 0$.
4. *$A$ is of full row rank.*
5. **The linear transformation** $x \mapsto Ax$ **is onto.**

Question: If $A$ has a right inverse, how many right inverses does $A$ have ?

# Full Rank Factorization

**Theorem 35.**

*Let $A \in \mathbb{C}^{m \times n}$. Then the following statements are equivalent.*

1. *$A$ has a left inverse.*
2. **Left cancellation law:** $AX = AY \Rightarrow X = Y$.
3. $AX = 0 \Rightarrow X = 0$.
4. *$A$ is of full column rank.*
5. **The linear transformation** $f : x \mapsto Ax$ **is one-to-one.**

**Theorem 36.**

*If a matrix $A$ has a left inverse $B$ and a right inverse $C$, then the following are equivalent.*

1. *$A$ is square.*
2. *$B = C$.*
3. *$A$ has a unique left inverse, a unique right inverse and a unique inverse.*

# Full Rank Factorization

If a matrix $A$ has an inverse, then $A^{-1}$ is unique, $A$ is square and $AA^{-1} = A^{-1}A = I$.

**Theorem 37.**

*Let $A$ be a square matrix of order $n$. Then the following statements are equivalent:*

1. *$A$ has a right inverse.*
2. *rank of $A$ is $n$.*
3. *$A$ has a left inverse.*
4. *$A$ has an inverse.*

**Definition 38.**

Let $A$ be a $m \times n$ matrix with rank $r \geq 1$. Then $(P, Q)$ is said to be a **rank-factorization** of $A$ if $P$ is of order $m \times r$, $Q$ is of order $r \times n$ and $A = PQ$.

**Theorem 39.**

*Every non-null matrix has a rank-factorization.*

**Proof.** Let $A$ be a $m \times n$ matrix with rank $r$.

Let $B = [x_1 : x_2 : \cdots : x_r]$ be an $m \times r$ matrix whose columns form a basis of $R(A)$. Then for each $j = 1, 2, \ldots, n$, each column of $A$, $A_{*j}$ is a linear combination of the columns of $B$, so there exists an $r \times 1$ vector $y_j$ such that $A_{*j} = By_j$. $A_{i*}$ denotes the $i$-th row of $A$ and $A_{*j}$ denotes the $j$-th column of $A$.

# Full Rank Factorization

Now

$$\begin{aligned} A &= [A_{*1} : \cdots : A_{*n}] \\ &= [By_1 : \cdots : By_n] \\ &= B[y_1 : \cdots : y_n] \\ &= BC \end{aligned}$$

where $C = [y_1 : \cdots : y_n]$.

- **A null matrix cannot have a rank-factorization** since there cannot be a matrix with 0 rows.
- **Rank-factorization of a matrix is not unique.** The choice of the matrix $B$ is not unique because the columns of $B$ are coming from the column basis of $A$.

# When a factorization is a rank-factorization?

**Theorem 40.**

*Let $A = PQ$ where $P$ is a $m \times k$ matrix and $Q$ a $k \times n$ matrix. Then the rank of $A$ is at most $k$.*
*Moreover, the following are equivalent:*

- *the rank of $A$ is $k$.*
- *$(P, Q)$ is a rank-factorization of $A$.*
- *$P$ is of full column rank and $Q$ is of full row rank.*
- *the columns of $P$ form a basis of the column space of $A$.*
- *the row of $Q$ form a basis of the row space of $A$.*

**Corollary 41.**

*If $(P, Q)$ is a rank-factorization of $A$ then $R(P) = R(A)$, row space$(Q) =$ row space$(A)$ and $N(Q) = N(B)$.*

**Theorem 42.**

*If $A = A^2$, rank of $A$ equals trace of $A$.*

**Proof.** The result is trivial if the rank $r$ of $A$ is 0, so let $r \geq 1$.

Let $(P, Q)$ be a rank-factorization of $A$. Then $PQPQ = PQ = P\mathbb{I}_r Q$.

Since $P$ is of full column rank and $Q$ is of full row rank, left and cancellation laws are applied, we get $PA = \mathbb{I}_r$.

Hence rank of $A = r = tr(\mathbb{I}_r) = tr(QP) = tr(PQ) = tr(A)$.

# Full Rank Factorization

**Finding a rank-factorization of a matrix $A$ of rank $r$ is easy when $A$ is represented in the following nice form.**

**Theorem 43.**

*Let $A$ be an $m \times n$ matrix of rank $r \geq 1$. Then there exist permutation matrices $P$ and $Q$ such that*

$$A = P \begin{pmatrix} B & BC \\ DB & DBC \end{pmatrix} Q$$

*where $B$ is invertible matrix of order $r$ and, $C$ and $D$ are some matrices of orders $r \times (n-r)$ and $(m-r) \times r$ respectively.*

# Full Rank Factorization

When a matrix $A$ in the above form, can be factorized as $A = P_1 Q_1$ where

$$P_1 = P \begin{pmatrix} B \\ DB \end{pmatrix} \ \texttt{and} \ Q_1 = \begin{pmatrix} \mathbb{I}_r & : & C \end{pmatrix} Q.$$

Since $P_1$ is of order $m \times r$, it follows that $(P_1, Q_1)$ is a rank-factorization of $A$.

# Examples : Moore-Penrose Inverse

1. Clearly $\mathbb{I}_n^\dagger = \mathbb{I}_n$ for any $n$ and $\mathbb{O}_{m\times n}^\dagger = \mathbb{O}_{n\times m}$.
2. Suppose $A$ is square and invertible. Then $A^{-1} = A^\dagger$.
3. Suppose $P$ is a matrix such that $P = P^* = P^2$ (called a projection matrix, also known as a Hermitian idempotent). Then $P = P^\dagger$.
4. Let $D$ be a diagonal matrix, say $D = diag(d_1, d_2, \ldots, d_n)$. Then $D^\dagger = diag(d_1^\dagger, d_2^\dagger, \ldots, d_n^\dagger)$, where

$$\lambda^+ = \begin{cases} 0, & \text{if } \lambda = 0 \\ \lambda^{-1}, & \text{if } \lambda \neq 0 \end{cases}.$$

5. Let $A = \begin{pmatrix} a_1 \\ a_2 \\ \vdots \\ a_n \end{pmatrix}$. If $A \neq 0$, then $A^\dagger = \frac{1}{A^*A}A^*$.

We now see that for matrices of full row or column rank, the Moore-Penrose inverse picks out a specific left (right) inverse of the matrix as shown below.

**Theorem 44.**

1. *If* $F \in \mathbb{C}_r^{m\times r}$ *(*$F$ *has full column rank), then* $F^\dagger = (F^*F)^{-1}F^*$*.*
2. *If* $G \in \mathbb{C}_r^{r\times n}$ *(*$G$ *has full row rank), then* $G^\dagger = G^*(GG^*)^{-1}$*.*

Note that $F^\dagger F = \mathbb{I}_r$ and $GG^\dagger = \mathbb{I}_r$.

For an arbitrary matrix $A \in \mathbb{C}_r^{m \times n}$ with $r > 0$, we shall show how to construct the Moore-Penrose Inverse.

**Theorem 45 (Greville and A.S. Householder).**

*Let $A \in \mathbb{C}_r^{m \times n}$. Take any full rank factorization of $A = FG$. Then $A^\dagger = G^\dagger F^\dagger$. In otherwords, $A^\dagger = G^*(GG^*)^{-1}(F^*F)^{-1}F^*$.*

Moreover, $AA^\dagger = FF^\dagger$ and $A^\dagger A = G^\dagger G$, where $A = FG$ is any full rank factorization of $A$.

# Some More Properties of Moore-Penrose Inverse

We have already noted that a matrix $A$ in $\mathbb{C}_r^{m \times n}$ with $r > 0$ has infinitely many full rank factorizations. We even showed how to produce an infinite collection using invertible matrices. We show next that this is the only way to get full rank factorizations.

**Theorem 46.**

*Every matrix $A \in \mathbb{C}_r^{m \times n}$ with $r > 0$ has infinitely many full rank factorizations.*
*However, if $A = FG = F_1 G_1$ are two full rank factorizations of $A$, then there exists an invertible matrix $R$ in $\mathbb{C}^{r \times r}$ such that $F_1 = FR$ and $G_1 = R^{-1}G$.*
*Moreover, $(R^{-1}G)^\dagger = G^\dagger R$ and $(FR)^\dagger = R^{-1}F^\dagger$.*

Hence $A = F_1 G_1 = (FR)(R^{-1}G)$. So,
$A^\dagger = (R^{-1}G)^\dagger (FR)^\dagger = G^\dagger R R^{-1} F^\dagger = G^\dagger F^\dagger$.

# Some More Properties of Moore-Penrose Inverse

Let $A \in \mathbb{C}_r^{m\times n}$. Then the following are true.

1. $(AA^\dagger)^2 = AA^\dagger = (AA^\dagger)^*$.
2. $(\mathbb{I}_m - AA^\dagger)^2 = (\mathbb{I}_m - AA^\dagger) = (\mathbb{I}_m - AA^\dagger)^*$.
3. $(A^\dagger A)^2 = A^\dagger A = (A^\dagger A)^*$.
4. $(\mathbb{I}_n - A^\dagger A)^2 = (\mathbb{I}_n - A^\dagger A) = (\mathbb{I}_n - A^\dagger A)^*$.
5. $(\mathbb{I}_m - AA^\dagger)A = \mathbb{O}_{m\times n}$.
6. $(\mathbb{I}_n - A^\dagger A)A^\dagger = \mathbb{O}_{n\times m}$.
7. $A^{\dagger\dagger} = A$.
8. $(A^*)^\dagger = (A^\dagger)^*$.
9. $(A^*A)^\dagger = A^\dagger A^{*\dagger}$.
10. $A^* = A^*AA^\dagger = A^\dagger AA^*$.
11. $A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger$.
12. $(\lambda A)^\dagger = \lambda^\dagger A^\dagger$.

# Some More Properties of Moore-Penrose Inverse

1. Let $A = FG$ be a full rank factorization of $A$. Prove that $F^\dagger A = G, \quad FF^\dagger A = A, \ AG^\dagger = F$ and $AG^\dagger G = A$.
2. Suppose $A^L$ is a left inverse of $A$ — that is, $A^L A = \mathbb{I}$.
   Is $A^L = A^\dagger$ necessarily true?
   Suppose $A^*A = \mathbb{I}$. What can you say about $A^\dagger$?
3. Suppose $A^2 = A$ in $\mathbb{C}^{n\times n}$. Use a full rank factorization of $A$ to prove that rank $(A) = \mathit{trace}(A)$ (i.e., the rank of $A$ is just the trace of $A$ when $A$ is an idempotent matrix).
4. Prove that the row space of $A^\dagger$ is equal to the row space of $A^*$.
5. Prove that the column space of $A^\dagger$ is equal to the column space of $A^*$ and the column space of $A^\dagger A$.
6. Prove that $A, A^*, A^\dagger$, and $A^{\dagger *}$ all have the same rank.
7. Is $A\{2,3,4\}$ ever empty for some weird matrix $A$?

# Some More Properties of Moore-Penrose Inverse

1. Prove the following:

$$\begin{aligned} AA^{\dagger}A^{\dagger *} &= A^{\dagger *} \\ A^{\dagger *}A^{\dagger}A &= A^{\dagger *} \\ A^{*\dagger}A^{*}A &= A \\ AA^{*}A^{*\dagger} &= A \\ A^{*}A^{\dagger *}A^{\dagger} &= A^{\dagger} \\ A^{\dagger}A^{\dagger *}A^{*} &= A^{\dagger}. \end{aligned}$$

2. Prove

$$(AA^*)^{\dagger} = A^{\dagger *}A^{\dagger}, \quad (A^*A)^{\dagger} = A^{\dagger}A^{\dagger *} = A^{\dagger}A^{*\dagger}$$

and

$$(AA^*)^{\dagger}(AA^*) = AA^{\dagger}.$$

# Some More Properties of Moore-Penrose Inverse

1. Prove
$$A = AA^*(A^\dagger)^* = (A^\dagger)^*A^*A = AA^*(AA^*)^\dagger A$$
and
$$A^* = A^*AA^\dagger = A^\dagger AA^*.$$
2. Prove $A^\dagger = A^\dagger(A^\dagger)^*A^* = A^*(A^\dagger)^*A^\dagger$.
3. Prove $A^\dagger = (A^*A)^\dagger A^* = A^*(AA^*)^\dagger$ so that $AA^\dagger = A(A^*A)^\dagger A^*$.
4. Show that if $A = \sum A_t$ where $A_i^*A_j = \mathbb{O}$ whenever $i \neq j$. then $A^\dagger = \sum A_i^\dagger$.
5. Prove that all the following matrices have the same rank: $A, A^\dagger, AA^\dagger, A^\dagger A, AA^\dagger A$, and $A^\dagger AA^\dagger$. The rank is $trace(AA^\dagger)$.
6. Prove that $(-A)^\dagger = -A^\dagger$.
7. Suppose $A$ is $n \times m$ and $S$ is $m \times m$ invertible. Prove that $(AS)(AS)^\dagger = AA^\dagger$.

# Some More Properties of Moore-Penrose Inverse

1. Suppose $A^*A = AA^*$. Prove that $A^\dagger A = AA^\dagger$ and for any natural number $n$,
$$(A^n)^\dagger = (A^\dagger)^n.$$
What can you say if $A = A^*$?
2. Prove that $A^\dagger = A^*$ if and only if $A^*A$ is idempotent.
3. If $A = \begin{bmatrix} B & \mathbb{O} \\ \mathbb{O} & C \end{bmatrix}$, find a formula for $A^\dagger$.
4. Suppose $A$ is a matrix and $X$ is a matrix such that
$$AXA = A, XAX = X \;\text{ and }\; AX = XA.$$
Prove that if $X$ exists, it must be unique.
5. Let $A = FG$ be a full rank factorization of $A$ in $\mathbb{C}_r^{m\times n}$. Prove that $F^*AG^*$ is invertible and $A^\dagger = G^*(F^*AG^*)^{-1}F^*$.
(Hint: First prove that $F^*AG^*$ is in fact invertible.) Note $F^*AG^* = (F^*F)(GG^*)$ and those two matrices are $r \times r$ of rank $r$ hence invertible. Then $(F^*AG^*)^{-1} = (GG^*)^{-1}(F^*F)^{-1}$.

# Some More Properties of Moore-Penrose Inverse

1. Let $x = \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix}$ and $y = \begin{bmatrix} y_1 \\ \vdots \\ y_n \end{bmatrix}$. Show $(xy^*)^\dagger = (x^*x)^\dagger (y^*y)^\dagger yx^*$.
2. Find the MP inverse of a $2 \times 2$ matrix $\begin{bmatrix} a & b \\ c & d \end{bmatrix}$.
3. Find examples of matrices $A$ and $B$ with $(AB)^\dagger = B^\dagger A^\dagger$ and $A$ and $B$ with $(AB)^\dagger \neq B^\dagger A^\dagger$. Then prove Greville's [1996] result that $(AB)^\dagger = B^\dagger A^\dagger$ iff $A^\dagger A$ and $BB^\dagger$ commute.
4. Find $\begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{bmatrix}^\dagger$.
5. Remember the matrix units $E_{ij}$? What is $E_{ij}^\dagger$?

# Some More Properties of Moore-Penrose Inverse

1. Find necessary and sufficient conditions for $A^\dagger = A$.
2. Prove that the following statements are equivalent for $m \times n$ matrices $A$ and $B$:
   (i) $Col \begin{bmatrix} A \\ A^\dagger A \end{bmatrix} = Col \begin{bmatrix} B \\ B^\dagger B \end{bmatrix}$
   (ii) $Col \begin{bmatrix} A \\ A^* A \end{bmatrix} = Col \begin{bmatrix} B \\ B^* B \end{bmatrix}$
   (iii) $A = B$.

# Some More Properties of Moore-Penrose Inverse

1. In this exercise, we introduce the idea of a **circulant matrix**. An $n \times n$ matrix $A$ is called a circulant matrix if its first row is arbitrary but its subsequent rows are cyclical permutations of the previous row. So, if the first row is $(a_1 a_2 a_3 \cdots a_n)$, the second row $(a_n a_1 a_2 \cdots a_{n-1})$, and the last row is $(a_2 a_3 a_4 \cdots a_n a_1)$. There are entire books written on these kinds of matrices. Evidently, if you know the first row, you know the matrix. Write a typical $3 \times 3$ circulant matrix. Is the identity matrix a circulant matrix? Let $C$ be the circulant matrix whose first row is $(0100 \cdots 0)$. Prove that all powers of $C$ are also circulant matrices. Moreover, prove that if $A$ is any circulant matrix with first row $(a_1 a_2 a_3 \cdots a_n)$, then $A = a_1 I + a_2 C + a_3 C^2 + \cdots + a_n C^{n-1}$.
2. Continuing the problem above. prove that $A$ is a circulant matrix iff $AC = CA$.
3. Suppose $A$ is a circulant matrix. Prove that $A^\dagger$ is also circulant and $A^\dagger$ commutes with $A$.

# Some More Properties of Moore-Penrose Inverse

1. (Cline 1964) If $AB$ is defined, prove that $(AB)^\dagger = B_1^\dagger A_1^\dagger$ where $AB = A_1 B_1$, $B_1 = A^\dagger AB$ and $A_1 = AB_1 B_1^\dagger$.
2. If $rank(A) = 1$, prove that $A^\dagger = (tr(AA^*)^{-1})A^*$.
3. Prove that $AB = \mathbb{O}$ implies $B^\dagger A^\dagger = \mathbb{O}$.
4. Prove that $A^* B = \mathbb{O}$ iff $A^\dagger B = \mathbb{O}$.
5. Suppose $A^*AB = A^*C$. Prove that $AB = AA^\dagger C$.
6. Suppose $BB^*$ is invertible. Prove that $(AB)(AB)^\dagger = AA^\dagger$.
7. Suppose that $AB^* = \mathbb{O}$. Prove that $(A+B)^\dagger = A^\dagger + (\mathbb{I}_n - A^\dagger B)[C^\dagger + (\mathbb{I} - C^\dagger C)MB^*(A^\dagger)^* A^\dagger (\mathbb{I} - BC^\dagger)]$ where $C = (\mathbb{I}_m - AA^\dagger)B$ and $M = [\mathbb{I}_n + (\mathbb{I}_n - C^\dagger C)B^*(A^\dagger)A^\dagger B(\mathbb{I}_n - C^\dagger C)]^{-1}$.

1. Prove that $\begin{bmatrix} A \\ \cdots \\ B \end{bmatrix}^{\dagger} = [A^{\dagger} - TBA^{\dagger} | T]$ where $T = E^{\dagger} + (\mathbb{I}_n - E^{\dagger}B)A^{\dagger}(A^{\dagger})^{*}B^{*}K(\mathbb{I}_p - EE^{\dagger})$ with $E = B(\mathbb{I}_n - A^{\dagger}A)$ and $K = [\mathbb{I}_p + (I_p - EE^{\dagger})BA^{\dagger}(A^{\dagger})^{*}B^{*}(\mathbb{I} - EE^{\dagger})]^{-1}$.
2. Prove that $[A : B]^{\dagger} = \begin{bmatrix} A^{\dagger} - A^{\dagger}B(C^{\dagger} + D) \\ C^{\dagger} + D \end{bmatrix}$ where $C = (\mathbb{I}_m - AA^{\dagger})B$ and $D = (\mathbb{I}_p - C^{\dagger}C)[\mathbb{I}_p + (I_p - C^{\dagger}C)B^{*}(A^{\dagger})^{*}A^{\dagger}B(\mathbb{I}_p - C^{\dagger}C)]^{-1}B^{*}(A^{\dagger})^{*}A^{\dagger}(\mathbb{I}_m - BC^{\dagger})$.

# Some More Properties of Moore-Penrose Inverse

1. Prove Greville's [1966] results: $(AB)^\dagger = B^\dagger A^\dagger$ iff any one of the following holds true:
   - (a) $A^\dagger ABB^* A^* = BB^* A^*$ and $BB^\dagger A^* AB = A^* AB$.
   - (b) $A^\dagger ABB^*$ and $A^* ABB^\dagger$ are self adjoint.
   - (c) $A^\dagger ABB^* A^* ABB^\dagger = BB^* A^* A$.
   - (d) $A^\dagger AB = B(AB)^\dagger AB$ and $BB^\dagger A^* = A^* AB(AB)^\dagger$.
2. Suppose $A$ is $m \times n$ and $B$ is $n \times p$ and $rank(A) = rank(B) = n$. Prove that $(AB)^\dagger = B^\dagger A^\dagger$.

Recall that $\{1\}$-inverses are the "equation solvers." We have established the consistency condition:$Ax = b$ is consistent iff $AGb = b$ for some $G \in A\{1\}$.

In this case, all solutions can be described by $x = Gb + (\mathbb{I} - GA)z$, where $z$ is arbitray in $\mathbb{C}^m$.

The first thing we establish is that there is, in fact, **a minimum norm solution to any consistent system of linear equations and it is unique**.

**Theorem 47.**

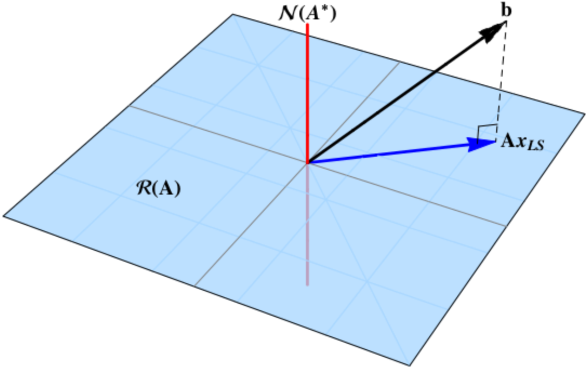
*Suppose $Ax = b$ is a consistent system of linear equations (i.e., $b \in R(A)$). Then there exists a unique solution of $Ax = b$ of minimum norm. In fact, it lies in $R(A^*)$.* LA-3(P-17)T-16+R-17

We proved in Theorem 47 that for every $b \in R(A)$, there exists a unique element $A^\dagger b \in R(A^*)$ which gives the unique norm solution for the system $Ax = b$.

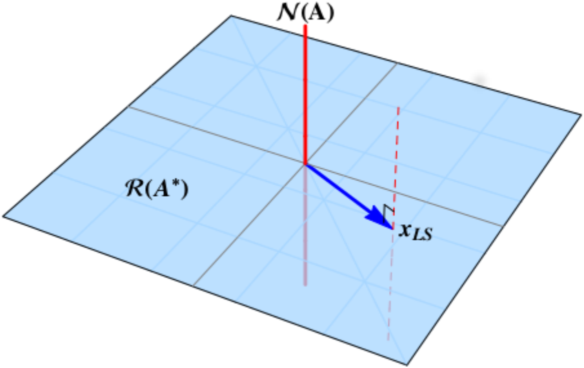
**Theorem 48 (Moore-1935, Penrose-1955).**

*For every matrix $A$, there exists a unique matrix $A^\dagger : R(A) \to R(A^*)$ such that $AA^\dagger = P_{R(A)}$ and $A^\dagger A = P_{R(A^*)}$.*

The least squares solution is shown below.



The least squares solution is the affine space represented by the dashed red line below.



# Solving Consistent Systems of Linear Equations

With the Moore-Penrose inverse in hand, we consider an arbitrary system of linear equations $Ax = b$.

**Theorem 49 (consistent system).**

*$Ax = b$ has a solution iff $AA^{\dagger}b = b$.*
*If a solution exists at all, every solution is of the form*

$$x = A^{\dagger}b + (\mathbb{I} - A^{\dagger}A)w,$$

*where $w$ is an arbitray matrix.*
*Indeed, a consistent system always has $A^{\dagger}b$ as a particular solution.*

# Solving Inconsistent Systems of Linear Equations

Consider the linear system $Ax = b$. There can be many least squares solutions (it certainly happens, when $rank(A) = n$, or $A$ has no column rank) to an inconsistent system $Ax = b$ (consistency condition fails, $AGb \neq b$) of linear equations.

We may ask, among all of these least squares solutions, is there one only element of minimum norm?

Yes. $A^\dagger b$ is the only one vector which has minimum norm, as shown in the following theorem.

**Theorem 50 (inconsistent system).**

*Among the least squares solutions of $Ax = b$, $A^\dagger b$ is the only one vector of minimum norm.*

$A^\dagger b$ is called the **best approximate solution** of $Ax = b$.

- When $Ax = b$ has a solution, $A^\dagger b$ is the **unique solution of minimum norm**.
- When $Ax = b$ does not have a solution, $A^\dagger b$ is the **unique least squares solution of minimum norm**.

**Theorem 51.**

*If $G$ has the property that $Gb$ is the minimum norm least squares solution for all $b$, then $G = A^\dagger$.*

**Example 52.**

Consider the system $\begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} x \\ y \end{pmatrix} = \begin{pmatrix} 1 \\ 0 \\ 3/4 \end{pmatrix} = b$.

A full rank factorization of $A$ is given by $A = FG = \begin{pmatrix} 1 & 1 \\ 1 & -1 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

So $A^\dagger = \begin{pmatrix} 1/2 & 1/2 & 0 \\ 1/3 & -1/3 & 1/3 \end{pmatrix}$. The consistency condition $AA^\dagger b = b$ fails, so the system is inconsistent.

The **best approximate solution** is $x_0 = A^\dagger b = \begin{pmatrix} 1/2 \\ 7/12 \end{pmatrix}$.

# Observations

$\{1,3\}$-inverses of $G$ have the property that no matter which one you choose, $AG$ is always the same, namely $AA^\dagger$. In fact, more can be said.

**Theorem 53.**

*$G \in A\{1,3\}$ iff $AG = AA^\dagger$. In particular, if $G \in A\{1,3\}$, then $Ax_0 = AGb = AA^\dagger b$, for any least squares solution $x_0$ of $Ax = b$, by Theorem 21.*

$\{1,4\}$-inverses of $G$ have the property that no matter which one you choose, $GA$ is always the same, namely $A^\dagger A$. In fact, more can be said.

**Theorem 54.**

*$G \in A\{1,4\}$ iff $AG = A^\dagger A$. In particular, if $G \in A\{1,4\}$, then $Gb = A^\dagger b$, for any $b \in R(A)$.*

# Exercises

This exercise refers back to the Hermite echelon form. Suppose we desire the solutions of $Ax = b$ where $A$ is square but not necessarily invertible. We have showed how to use the Moore-Penrose inverse to describe all solutions to $Ax = b$ if it exists. In this exercise, we consider a different approach. First form the augmented matrix $[A \mid b]$. There is an invertible matrix $S$ such that $SA = H = HEF(A)$ so form $[A \mid b] \to [SA \mid Sb] = [H \mid Sb]$.

(a) Prove that rank $(A) = rank(H) =$ the number of ones on the diagonal of $H$.
(b) Prove that $Ax = b$ is consistent iff $Sb$ has nonzero components only in the rows where $H$ has ones.
(c) If $Ax = b$ is consistent, prove $Sb$ is a particular solution to the system.
(d) Prove that if $H$ has $r$ ones down its diagonal, then $\mathbb{I} - SA$ has exactly $n - r$ nonzero columns and these nonzero columns span $N(A)$ and hence form a basis for $N(A)$.
(e) Prove that all solutions of $Ax = b$ are described by $x = Sb + (\mathbb{I} - SA)D$, where $D$ is a diagonal matrix containing $n - r$ free parameters.

# Connection of the $\{2\}$-inverse to other generalized inverses of a matrix

Let's fix an $m$-by-$n$ matrix $A$ of rank $r$ over $\mathbb{C}$. Now choose arbitrary matrices $E$ in $\mathbb{C}^{n \times k}$ and $H$ in $\mathbb{C}^{k \times m}$ and form the $k$-by-$k$ matrix $HAE$.

Suppose $A = FG$ is a full rank factorization of $A$. Note $HF$ is $k$-by-$r$ and $GE$ is $r$-by-$k$.

**Theorem 10.**

*Suppose $k = r = rank(A)$. Then $HAE$ is invertible iff $HF$ and $GE$ are invertible.*

**Theorem 11.**

*Let $k = r = rank(A)$ and choose $H$ and $E$ so that $HAE$ is invertible. Then $X = E(HAE)^{-1}H$ is a $\{1, 2\}$-inverse of $A$.*

# Connection of the $\{2\}$-inverse to other generalized inverses of a matrix

**Theorem 12.**

*With the hypotheses of the previous theorem, add that we choose $H = F^*$. Then $X$ is a $\{1, 2, 3\}$-inverse of $A$.*

**Theorem 13.**

*In the previous theorem, choose $E = G^*$ instead of $H = F^*$. Then $X$ is a $\{1, 2, 4\}$-inverse of $A$.*

**Theorem 14.**

*In the previous theorem, choose both $H = F^*$ and $E = G^*$. Then $X = A^\dagger$.*

# Connection of the $\{2\}$-inverse to other generalized inverses of a matrix

Above we wrote $\{2\}$-inverses as $X = E(HAE)^{-1}H$. But how did we know we could ever find any matrices $E$ and $H$ so that $HAE$ is invertible?

We have the following theorem.

**Theorem 15.**

*Let $A \in \mathbb{C}^{m\times n}$. Then $X$ is a $\{2\}$-inverse of $A$ if and only if there exists $E$ and $H$ where $E$ has full column rank, $H$ has full row rank, $Col(X) = Col(E)$, $Col(X^*) = Col(H^*)$, $HAE$ is invertible, and $X = E(HAE)^{-1}H$.*

**Exercise 16.**

*Suppose $A = FG$ is a full rank factorization of $A$. Prove that $F(GF)^{-1}G$ and $F(GF)^{\dagger}G$ are $\{2\}$-inverses of $A$.*

# Constructing Other Generalized Inverses

We take a constructive approach to build a variety of generalized inverses of a given matrix.

The approach we adopt goes back to the fundumental idea of reducing a matrix to row echelon form.

1. Given $A$ in $\mathbb{C}_r^{m\times n}$, form $RA = \begin{bmatrix} G \\ \dots \\ \mathbb{O} \end{bmatrix}$. Then $A^{g_1} = \left[G^\dagger + (\mathbb{I} - G^\dagger G)X \vdots V\right] R \in A\{1\}$, where $X$ and $V$ are arbitrary and $G^\dagger = G^*(GG^*)^{-1}$, we can see that $A^{g_1} \in A\{1\}$.
2. $A^{g_{14}} = \left[G^\dagger \vdots V\right] R \in A\{1,4\}$, where $X$ was chosen to be $\mathbb{O}$ and $V$ is still arbitrary. We can see that $A^{g_{14}} \in A\{1,4\}$.
3. $A^{g_{124}} = \left[G^\dagger \vdots G^\dagger W\right] R \in A\{1,2,4\}$, where $V$ is chosen as $G^\dagger W$, where $W$ is arbitrary. We can see that $A^{g_{124}} \in A\{1,2,4\}$.
4. $A^{g_{1234}} = G^\dagger F^\dagger$, where $F = AG^\dagger$ and $F^\dagger = (F^*F)^{-1}F^*$. We can see that $A^{g_{1234}}$ is the Moore-Penrose inverse of $A$.

# Now let's play the same game with $A^*$ instead of $A$.

Reasoning as before, we see the special choice $V = WF^\dagger$ yields

$A^{g_{123}} = S \begin{bmatrix} F^\dagger \\ \dots \\ WF^\dagger \end{bmatrix}$ as a $\{1, 2, 3\}$-inverse of $A$.

To get $A^\dagger$, we look at $G = F^\dagger A = F^\dagger \left[ F \vdots \mathbb{O} \right] S^{-1} = \left[ F^\dagger F \vdots \mathbb{O} \right] \begin{bmatrix} S_1 \\ \dots \\ S_2 \end{bmatrix} = S_1$,

which has full row rank so $G^\dagger = G^*(GG^*)^{-1}$.

One can verify that $A^\dagger$ can be reproduced as above.

# Constructing Other Generalized Inverses

1. Given $A$ in $\mathbb{C}_r^{m\times n}$, form $A^* = (S^*)^{-1}\begin{bmatrix} F^* \\ \cdots \\ \mathbb{O} \end{bmatrix}$ to get $A = [F \vdots \mathbb{O}]S^{-1}$.

   Then
   $A^{g_1} = S\begin{bmatrix} F^\dagger + X(\mathbb{I} - FF^\dagger) \\ \cdots\cdots\cdots\cdots\cdots \\ V \end{bmatrix} \in A\{1\}$, where $X$ and $V$ are arbitrary of appropriate size and $F^\dagger = (F^*F)^{-1}F^*$.

2. $A^{g_{13}} = S\begin{bmatrix} F^\dagger \\ \cdots \\ V \end{bmatrix} \in A\{1,3\}$, where $X$ was chosen to be $\mathbb{O}$ and $V$ is still arbitrary.

3. $A^{g_{123}} = S\begin{bmatrix} F^\dagger \\ \cdots \\ WF^\dagger \end{bmatrix} \in A\{1,2,3\}$, where $V$ is chosen as $WF^\dagger$, where $W$ is arbitrary.

4. $A^{g_{1234}} = A^\dagger = G^\dagger F^\dagger$, where $G = F^\dagger A$ and $G^\dagger = G^*(GG^*)^{-1}$.

# Constructing Other Generalized Inverses of a Specified Rank

We indicate next how to get generalized inverses of a specified rank.

We use the notation $A\{i,j,k,\ell\}_s$ for the set of all $\{i,j,k,\ell\}$-inverses of rank $s$. We begin with $\{2\}$-inverses.

**Theorem 17 (G. W. Stewart, R.E. Fundrelic).**

*Let $A \in \mathbb{C}_r^{m\times n}$ and $0 < s \leq r$. Then $A\{2\}_s = \{X | X = YZ$, where $Y \in \mathbb{C}^{n\times s}, Z \in \mathbb{C}^{s\times m}, ZAY = \mathbb{I}_s\}$.*

**Corollary 18.**

*Let $A \in \mathbb{C}_r^{m\times n}$. Then $A\{1,2\} = \{FG | F \in \mathbb{C}^{n\times r}, G \in \mathbb{C}^{r\times m}, GAF = \mathbb{I}_r\}$*

**Corollary 19.**

*If $GAF = \mathbb{I}_s$, then $G \in (AF)\{1,2,4\}$.*

**Theorem 20.**

*Let* $A \in \mathbb{C}_r^{m \times n}$ *and* $0 < s < r$*. Then* $A\{2,3\}_s = \{Y(AY)^\dagger \mid AY \in \mathbb{C}_s^{m \times s}\}$

**Theorem 21.**

*Let* $A \in \mathbb{C}_r^{m \times n}$ *and* $0 < s \leq r$*. Then* $A\{2,4\}_s = \{(YA)^\dagger Y \mid YA \in \mathbb{C}_s^{s \times n}\}$.

# Exercises

1. Prove that $A\{1,2,3,4\} \subseteq A\{1,2,3\} \subseteq A\{1,2\} \subseteq A\{1\}$, with equality holding throughout if and only if $A$ is invertible.
2. Suppose $G$ is a $\{1,2,3\}$-inverse of $A$. Prove $rank(G) = rank(A^\dagger) = rank(A)$.
3. Prove that the following statements are all equivalent:
   (i) $A^*B = \mathbb{O}$.
   (ii) $GB = \mathbb{O}$, where $G \in A\{1,2,3\}$.
   (iii) $HA = \mathbb{O}$, where $H \in B\{1,2,3\}$.
4. Prove that a matrix $G$ is in $A\{1,2,3\}$ if and only if $G = HA^*$, where $H$ is in $A^*A\{1\}$.
5. Prove that a matrix $G$ is in $A\{1,2,4\}$ if and only if $G = A^*H$, where $H$ is in $A^*A\{1\}$.
6. Let $B = A^*(AA^*)^{g_{12}}$. Prove that $B$ is a $\{1,2,4\}$-inverse of $A$.
7. Let $C = (A^*A)^{g_{12}}A^*$. Prove that $C$ is a $\{1,2,3\}$-inverse of $A$.
8. Let $B \in A\{1,2,4\}$ and $C \in A\{1,2,3\}$. Prove that $BAC = A^\dagger$. Is it good enough to assume $B \in A\{1,4\}$ and $C \in A\{1,3\}$?

# Exercises

1. Suppose $A \in \mathbb{C}_r^{m\times n}$ and $SAT = \begin{bmatrix} B & \mathbb{O} \\ \mathbb{O} & \mathbb{O} \end{bmatrix}$, where $B \in \mathbb{C}^{r\times r}$ is invertible. Then let $G = TNS$, where $N = \begin{bmatrix} Z & X \\ Y & W \end{bmatrix}$.

   Then prove

   (i) $G \in A\{1\}$ iff $Z = B^{-1}$.

   (ii) $G \in A\{1,2\}$ iff $Z = B^{-1}$ and $W = YBX$.

   (iii) $G \in A\{1,2,3\}$ iff $Z = B^{-1}$, $X = -B^{-1}S_1S_2$, and $W = -YS_1S_2$, where $S = \begin{bmatrix} S_1 \\ \cdots \\ S_2 \end{bmatrix}$.

   (iv) $G \in A\{1,2,4\}$ iff $Z = B^{-1}$, $Y = -T_2 + T_1B^{-1}$, and $W = -T_2 + T_1X$, where $T = [T_1 \vdots T_2]$.

   (v) $G = A^\dagger$ iff $Z = B^{-1}$, $X = -B^{-1}S_1S_2^\dagger$, $Y = -T_2^\dagger T_1B^{-1}$, and $W = T_2 + T_1B^{-1}S_1S_2^\dagger$.

   (vi) Let $G \in A^*A\{1\}$. Prove that $GA^* \in A\{1,2,3\}$. Let $H \in AA^*\{1\}$. Prove that $A^*H \in A\{1,2,4\}$.

2. (Urguhart) Let $G \in A\{1,4\}$ and $H \in A\{1,3\}$. Prove that $GAH = A^\dagger$.

# References


Robert Piziak and P. L. Odell, *Matrix Theory - From Generalized Inverses to Jordan Form*, Chapman & Hall, London.

Reinhold Baer, *Linear Algebra and Projective Geometry*, Academic Press, Inc., New York, (1952).

R.D.Sheffield, *A General Theory For Linear Systems*, AMM, February, (1958), 109-111.

A. Ramachandra Rao and P. Bhimasankaram, *Linear Algebra*, Hindustan Book Agency, 2000.