MTH207A: Matrix Theory and Linear Estimation (Module II)

Satya Prakash Singh

Indian Institute of Technology Kanpur, India

September 24, 2022

g-inverse

Definition 1

Let A be an $m \times n$ matrix. A matrix G of order $n \times m$ is said to be a generalized inverse (or a g–inverse) of A if $\mathrm{AGA} = \mathrm{A}$.

Remark 1: If A is square and nonsingular, then A^{-1} is the unique g–inverse of A. Otherwise, A has infinitely many g–inverses.

g-inverse

Theorem 2

Let A, G be matrices of order $m \times n$ and $n \times m$ respectively. Then the following conditions are equivalent:

(i) G *is a g-inverse of* A.

(ii) *For any* $\mathbf{y} \in \mathcal{C}(\mathrm{A})$, $\mathbf{x} = \mathrm{G}\mathbf{y}$ *is a solution of* $\mathrm{A}\mathbf{x} = \mathbf{y}$.

g-inverse

Proof: $(i) \Rightarrow (ii)$. Any $\mathbf{y} \in \mathcal{C}(\mathrm{A})$ is of the form $\mathbf{y} = \mathrm{A}\mathbf{z}$ for some $\mathbf{z}$. Then $\mathrm{A}(\mathrm{G}\mathbf{y}) = \mathrm{AGA}\mathbf{z} = \mathrm{A}\mathbf{z} = \mathbf{y}$.

$(ii) \Rightarrow (i)$. Since $\mathrm{AG}\mathbf{y} = \mathbf{y}$ for any $\mathbf{y} \in \mathcal{C}(\mathrm{A})$ we have $\mathrm{AGA}\mathbf{z} = \mathrm{A}\mathbf{z}$ for all $\mathbf{z}$. In particular, if we let $\mathbf{z}$ be the ith column of the identity matrix, then we see that the ith columns of AGA and A are identical. Therefore, $\mathrm{AGA} = \mathrm{A}$.

Factorization theorem

Theorem 3

Let A *be an* $m \times n$ *matrix of rank* r, $r \neq 0$. *Then there exist matrices* B *and* C *of order* $m \times r$ *and* $r \times n$, *respectively, such that* $Rank(\mathrm{B}) = Rank(\mathrm{C}) = r$ *and* $\mathrm{A} = \mathrm{BC}$. *This decomposition is called a rank factorization of* A.

Left and right inverse

Theorem 4

Let B *be an* $m \times r$ *matrix of rank* r*. Then there exists a matrix* X *(called a left inverse of* B*), such that* $\mathrm{XB} = \mathrm{I}$*.*

Theorem 5

Let C *be an* $r \times n$ *matrix of rank* r*. Then there exists a matrix* Y *(called a right inverse of* C*), such that* $\mathrm{CY} = \mathrm{I}$*.*

Proof of left inverse

Proof: If $m = r$, then B is nonsingular and admits an inverse. So suppose $r < m$. The columns of B are linearly independent. Thus we can find a set of $m - r$ columns that together with the columns of B form a basis for $\mathbb{R}^m$. In other words, we can find a matrix U of order $m \times (m - r)$ such that $[\mathrm{B}, \mathrm{U}]$ is nonsingular. Let the inverse of $[\mathrm{B}, \mathrm{U}]$ be partitioned as $\begin{bmatrix} \mathrm{X} \\ \mathrm{V} \end{bmatrix}$, where X is $r \times m$. Since

$$\begin{bmatrix} \mathrm{X} \\ \mathrm{V} \end{bmatrix} \begin{bmatrix} \mathrm{B} & \mathrm{U} \end{bmatrix} = \mathrm{I}, \tag{1}$$

we have $\mathrm{XB} = \mathrm{I}$.

Left inverse

Theorem 6

Let B be an $m \times r$ matrix of rank r. Then there exists a nonsingular matrix P such that

$$\mathrm{PB} = \begin{bmatrix} \mathrm{I} \\ 0 \end{bmatrix}.$$

Proof: Directly follows from (1).

Right inverse

Theorem 7

Let C be an $r \times n$ matrix of rank r. Then there exists a nonsingular matrix Q such that

$$\mathrm{CQ} = \begin{bmatrix} \mathrm{I} & 0 \end{bmatrix}.$$

Result

Theorem 8

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

$$\mathrm{PAQ} = \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Left inverse

Proof: By Factorization Theorem 3 there exist matrices B and C of order $m \times r$ and $r \times n$, respectively, such that $Rank(\mathrm{B}) = Rank(\mathrm{C}) = r$ and $\mathrm{A} = \mathrm{BC}$. Now by Theorems 6 and 7 we have

$$\mathrm{PB} = \begin{bmatrix} \mathrm{I} \\ 0 \end{bmatrix} \quad \text{and} \quad \mathrm{CQ} = \begin{bmatrix} \mathrm{I} & 0 \end{bmatrix}$$

where P and Q are some nonsingular matrices. Therefore,

$$\mathrm{PAQ} = \mathrm{PBCQ} = \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{bmatrix}.$$

Results

Result: Let A be a matrix with of rank r and $\mathrm{A} = \mathrm{BC}$ be its rank factorization. Then $G = \mathrm{C}_r^- \mathrm{B}_l^-$ is a g–inverse of A where B_l^- and C_r^- are left and right inverses of B and C respectively.

Results

Result: Consider the matrices U, V and W of appropriate dimensions. It can be easily verified that $\begin{bmatrix} \mathrm{I} & \mathrm{U} \\ \mathrm{V} & \mathrm{W} \end{bmatrix}$ is a *g*–inverse of $\begin{bmatrix} \mathrm{I} & 0 \\ 0 & 0 \end{bmatrix}$. By Theorem 8 any matrix A of rank *r* can be written as $\mathrm{A} = \mathrm{P} \begin{bmatrix} \mathrm{I}_r & 0 \\ 0 & 0 \end{bmatrix} \mathrm{Q}$ for some nonsingular matrices P and Q. Now it can verified that $\mathrm{Q}^{-1} \begin{bmatrix} \mathrm{I} & \mathrm{U} \\ \mathrm{V} & \mathrm{W} \end{bmatrix} \mathrm{P}^{-1}$ is a *g*–inverse of A.

Results

Result: Let A be of rank r. Choose any $r \times r$ nonsingular submatrix of A such that

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

where A_{11} is $r \times r$ and nonsingular. Then it can verified that

$$\begin{bmatrix} \mathrm{A}_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$$

is a g–inverse of A.

Results

Theorem 9

If G *is a g–inverse of* A*, then* $\mathcal{R}(\mathrm{A}) = \mathcal{R}(\mathrm{AG}) = \mathcal{R}(\mathrm{GA})$*.*

Proof: $\mathcal{R}(\mathrm{A}) = \mathcal{R}(\mathrm{AGA}) \leq \mathcal{R}(\mathrm{AG}) \leq \mathcal{R}(\mathrm{A})$. Similarly,
$\mathcal{R}(\mathrm{A}) = \mathcal{R}(\mathrm{AGA}) \leq \mathcal{R}(\mathrm{GA}) \leq \mathcal{R}(\mathrm{A})$.

Reflexive g–inverse

Definition 10

A g–inverse of A is called a *reflexive g–inverse* if it also satisfies $\mathrm{GAG} = \mathrm{G}$. Observe that if G is any g–inverse of A, then GAG is a reflexive g–inverse of A.

Reflexive g–inverse

Theorem 11

Let G be a g–inverse of A. Then $\mathcal{R}_n(\mathrm{A}) \leq \mathcal{R}_n(\mathrm{G})$. Furthermore, equality holds if and only if G is reflexive.

Proof: For any g–inverse G we have
$\mathcal{R}_n(\mathrm{A}) = \mathcal{R}_n(\mathrm{AGA}) \leq \mathcal{R}_n(\mathrm{G})$. If G is reflexive, then
$\mathcal{R}_n(\mathrm{G}) = \mathcal{R}_n(\mathrm{GAG}) \leq \mathcal{R}_n(\mathrm{A})$ and hence $\mathcal{R}_n(\mathrm{A}) = \mathcal{R}_n(\mathrm{G})$.

Reflexive g–inverse

Conversely, suppose $\mathcal{R}_n(\mathrm{A}) = \mathcal{R}_n(\mathrm{G})$. First observe that $\mathcal{C}(\mathrm{GA}) \subset \mathcal{C}(\mathrm{G})$. By Theorem 9, $\mathcal{R}_n(\mathrm{G}) = \mathcal{R}_n(\mathrm{GA})$, and hence $\mathcal{C}(\mathrm{G}) = \mathcal{C}(\mathrm{GA})$. Therefore, $\mathrm{G} = \mathrm{GAX}$ for some X. Now,

$$\mathrm{GAG} = \mathrm{GAGAX} = \mathrm{GAX} = \mathrm{G},$$

and G is reflexive.

g–inverse

Theorem 12

Let A *be a* $m \times n$ *matrix, let* G *be a* $g-$ *inverse of* A *and let* $\mathbf{y} \in \mathcal{C}(\mathrm{A})$. *Then the class of solutions of* $\mathrm{A}\mathbf{x} = \mathbf{y}$ *is given by* $\mathrm{G}\mathbf{y} + (\mathrm{I} - \mathrm{GA})\mathbf{z}$, *where* $\mathbf{z}$ *is arbitrary.*

Proof: For any $\mathbf{z}$,

$$\mathrm{A}\{\mathrm{G}\mathbf{y} + (\mathrm{I} - \mathrm{GA})\mathbf{z}\} = \mathrm{AG}\mathbf{y} = \mathbf{y},$$

since $\mathbf{y} \in \mathcal{C}(\mathrm{A})$, and hence $\mathrm{G}\mathbf{y} + (\mathrm{I} - \mathrm{GA})\mathbf{z}$ is a solution.

g–inverse

Conversely, if $\mathbf{u}$ is a solution i.e. $\mathrm{A}\mathbf{u} = \mathbf{y}$, then set $\mathbf{z} = \mathbf{u} - \mathrm{G}\mathbf{y}$ and calculate

$$
\begin{aligned}
\mathrm{G}\mathbf{y} + (\mathrm{I} - \mathrm{GA})\mathbf{z} &= \mathrm{G}\mathbf{y} + (\mathrm{I} - \mathrm{GA})(\mathbf{u} - \mathrm{G}\mathbf{y}) \\
&= \mathrm{G}\mathbf{y} + \mathbf{u} - \mathrm{GA}\mathbf{u} - \mathrm{G}\mathbf{y} + \mathrm{GAG}\mathbf{y} \\
&= \mathbf{u} - \mathrm{G}\mathbf{y} + \mathrm{GAGA}\mathbf{u} \\
&= \mathbf{u} - \mathrm{G}\mathbf{y} + \mathrm{GA}\mathbf{u} \\
&= \mathbf{u} - \mathrm{G}\mathbf{y} + \mathrm{G}\mathbf{y} = \mathbf{u}.
\end{aligned}
$$

Minimum norm g–inverse

Definition 13

A g–inverse G of A is said to be a minimum norm g–inverse of A if in addition to $\mathrm{AGA} = \mathrm{A}$, it satisfies $(\mathrm{GA})' = \mathrm{GA}$.

Minimum norm g–inverse

Theorem 14

Let A *be an* $m \times n$ *matrix. Then the following conditions are equivalent:*

(i) G *is a minimum norm* g*–inverse of* A.

(ii) *For any* $\mathbf{y} \in \mathcal{C}(\mathrm{A})$, $\mathbf{x} = \mathrm{G}\mathbf{y}$ *is a solution of* $\mathrm{A}\mathbf{x} = \mathbf{y}$ *with minimum norm.*

Minimum norm g–inverse

Proof: $(i) \Rightarrow (ii)$ We know from Theorem 12 that the class of solutions of $\mathrm{A}\mathbf{x} = \mathbf{y}$ is $\mathrm{G}\mathbf{y} + (\mathrm{I} - \mathrm{GA})\mathbf{z}$ for arbitrary $\mathbf{z}$. We have to show that

$$||\mathrm{G}\mathbf{y}|| \leq ||\mathrm{G}\mathbf{y} + (\mathrm{I} - \mathrm{GA})\mathbf{z}|| \tag{2}$$

for any $\mathbf{y} \in \mathcal{C}(\mathrm{A})$ and any $\mathbf{z}$.
We have

$$||\mathrm{G}\mathbf{y}+(\mathrm{I}-\mathrm{GA})\mathbf{z}||^2 = ||\mathrm{G}\mathbf{y}||^2+||(\mathrm{I}-\mathrm{GA})\mathbf{z}||^2+2\mathbf{y}'\mathrm{G}'(\mathrm{I}-\mathrm{GA})\mathbf{z}. \tag{3}$$

Since $\mathbf{y} \in \mathcal{C}(\mathrm{A})$, then $\mathbf{y} = \mathrm{A}\mathbf{u}$ for some $\mathbf{u}$. Hence

$$\mathbf{y}'\mathrm{G}'(\mathrm{I} - \mathrm{GA})\mathbf{z} = \mathbf{u}'\mathrm{A}'\mathrm{G}'(\mathrm{I} - \mathrm{GA})\mathbf{z} = \mathbf{u}'\mathrm{GA}(\mathrm{I} - \mathrm{GA})\mathbf{z} = 0,$$

since $(\mathrm{GA})' = \mathrm{GA}$ and $\mathrm{AGA} = \mathrm{A}$. Inserting this in (3) we get (2).

Minimum norm g–inverse

$(ii) \Rightarrow (i)$. Since for any $\mathbf{y} \in \mathcal{C}(\mathrm{A})$, $\mathbf{x} = \mathrm{G}\mathbf{y}$ is a solution of $\mathrm{A}\mathbf{x} = \mathbf{y}$, by Theorem 2, G is a g–inverse of A. Now from (2) and (3) we have

$$0 \leq |(\mathrm{I} - \mathrm{GA})\mathbf{z}||^2 + 2\mathbf{u}'\mathrm{A}'\mathrm{G}'(\mathrm{I} - \mathrm{GA})\mathbf{z} \tag{4}$$

for all $\mathbf{u}$ and $\mathbf{z}$.

Replace $\mathbf{u}$ by $\alpha\mathbf{u}$ in (4). If $\mathbf{u}'\mathrm{A}'\mathrm{G}'(\mathrm{I} - \mathrm{GA})\mathbf{z} < 0$, then choosing α large and positive we get a contradiction to (4).

Minimum norm g–inverse

Similarly, if $\mathbf{u}'\mathrm{A}'\mathrm{G}'(\mathrm{I}-\mathrm{GA})\mathbf{z} > 0$ then choosing α large and negative we get a contradiction. We therefore conclude that

$$\mathbf{u}'\mathrm{A}'\mathrm{G}'(\mathrm{I}-\mathrm{GA})\mathbf{z} = 0$$

for all $\mathbf{u}$, $\mathbf{z}$ and hence $\mathrm{A}'\mathrm{G}'(\mathrm{I}-\mathrm{GA}) = 0$. Thus $\mathrm{A}'\mathrm{G}'$ equals $(\mathrm{GA})'\mathrm{GA}$, which is symmetric.

Least squares g–inverse

Definition 15

A g–inverse G of A is said to be a least squares g–inverse of A if in addition to $\mathrm{AGA} = \mathrm{A}$, it satisfies $(\mathrm{AG})' = \mathrm{AG}$.

Theorem 16

Let A be an $m \times n$ matrix. Then the following conditions are equivalent:

(i) *G is a least squares g–inverse of A.*

(ii) *For any $\mathbf{x}$, $\mathbf{y}$, $||\mathrm{AG}\mathbf{y} - \mathbf{y}|| \leq ||\mathrm{A}\mathbf{x} - \mathbf{y}||$.*

Least squares g–inverse

Proof: $(i) \Rightarrow (ii)$ Let $\mathbf{x} - \mathrm{G}\mathbf{y} = \mathbf{w}$. We have to show that

$$||\mathrm{AG}\mathbf{y} - \mathbf{y}|| \leq ||\mathrm{AG}\mathbf{y} - \mathbf{y} + \mathrm{A}\mathbf{w}||. \tag{5}$$

We have

$$||\mathrm{AG}\mathbf{y} - \mathbf{y} + \mathrm{A}\mathbf{w}||^2 = ||\mathrm{AG}\mathbf{y} - \mathbf{y}||^2 + ||\mathrm{A}\mathbf{w}||^2 + 2\mathbf{w}'\mathrm{A}'(\mathrm{AG} - \mathrm{I})\mathbf{y}. \tag{6}$$

But

$$\mathbf{w}'\mathrm{A}'(\mathrm{AG} - \mathrm{I})\mathbf{y} = \mathbf{w}'(\mathrm{A}'\mathrm{G}'\mathrm{A}' - \mathrm{A}')\mathbf{y} = 0,$$

since $(\mathrm{AG})' = \mathrm{AG}$. Inserting this in (6) we get (5).

Least squares g–inverse

Proof: $(ii) \Rightarrow (i)$. For any vector $\mathbf{x}$, set $\mathbf{y} = \mathrm{A}\mathbf{x}$ in (ii). Then we see that

$$||\mathrm{AGA}\mathbf{x} - \mathrm{A}\mathbf{x}|| \leq ||\mathrm{A}\mathbf{x} - \mathrm{A}\mathbf{x}|| = 0$$

and hence $\mathrm{AGA}\mathbf{x} = \mathrm{A}\mathbf{x}$. Since $\mathbf{x}$ is arbitrary, $\mathrm{AGA} = \mathrm{A}$, and therefore G is a g–inverse of A. The remaining part of the proof parallels that of $(ii) \Rightarrow (i)$ of Theorem 14.

Least squares g–inverse

Theorem 17

An $n \times m$ matrix G is a least squares generalized inverse of an $m \times n$ matrix A if and only if $\mathrm{A}'\mathrm{A}\mathrm{G} = \mathrm{A}'$, or, equivalently, if and only if $\mathrm{G}'\mathrm{A}'\mathrm{A} = \mathrm{A}$.

Least squares g–inverse

Proof: Suppose that $\mathrm{A}'\mathrm{A}\mathrm{G} = \mathrm{A}'$ or, equivalently, that $\mathrm{G}'\mathrm{A}'\mathrm{A} = \mathrm{A}$. Then, $\mathrm{A}'\mathrm{A}\mathrm{G}\mathrm{A} = \mathrm{A}'\mathrm{A}$, implying that $\mathrm{A}\mathrm{G}\mathrm{A} = \mathrm{A}$ (verify!). Moreover, $\mathrm{A}\mathrm{G} = \mathrm{G}'\mathrm{A}'\mathrm{A}\mathrm{G} = (\mathrm{A}\mathrm{G})'\mathrm{A}\mathrm{G}$, so that $\mathrm{A}\mathrm{G}$ is symmetric. Thus, G is a least squares generalized inverse of A.

Conversely, if G is a least squares generalized inverse of A, then

$$\mathrm{A}'\mathrm{A}\mathrm{G} = \mathrm{A}'(\mathrm{A}\mathrm{G})' = (\mathrm{A}\mathrm{G}\mathrm{A})' = \mathrm{A}'$$

.

Least squares g–inverse

Proof: Verification that $\mathrm{A}'\mathrm{AGA} = \mathrm{A}'\mathrm{A}$, imply $\mathrm{AGA} = \mathrm{A}$.

Note that

$$(\mathrm{AGA} - \mathrm{A})'(\mathrm{AGA} - \mathrm{A}) = (\mathrm{GA} - \mathrm{I})'(\mathrm{A}'\mathrm{AGA} - \mathrm{A}'\mathrm{A}) = 0.$$

Thus, $\mathrm{AGA} - \mathrm{A} = 0$.

Moore–Penrose (MP) inverse

Definition 18

If G is a reflexive *g*–inverse of A that is both minimum norm and least squares then it is called a *Moore–Penrose inverse* of A. In other words, G is a Moore– Penrose inverse of A if it satisfies

$$\mathrm{AGA} = \mathrm{A},\ \mathrm{GAG} = \mathrm{G},\ (\mathrm{AG})' = \mathrm{AG},\ (\mathrm{GA})' = \mathrm{GA}. \tag{7}$$

Moore–Penrose inverse:Uniqueness

Let G_1 and G_2 are two MP inverses of A. Then both G_1 and G_2 will satisfy (7).

$$
\begin{aligned}
\mathrm{G}_1 &= \mathrm{G}_1\mathrm{A}\mathrm{G}_1 = \mathrm{G}_1\mathrm{G}_1'\mathrm{A}' = \mathrm{G}_1\mathrm{G}_1'\mathrm{A}'\mathrm{G}_2'\mathrm{A}' \\
&= \mathrm{G}_1\mathrm{G}_1'\mathrm{A}'\mathrm{A}\mathrm{G}_2 = \mathrm{G}_1\mathrm{A}\mathrm{G}_1\mathrm{A}\mathrm{G}_2 = \mathrm{G}_1\mathrm{A}\mathrm{G}_2 \\
&= \mathrm{G}_1\mathrm{A}\mathrm{G}_2\mathrm{A}\mathrm{G}_2 = \mathrm{G}_1\mathrm{A}\mathrm{A}'\mathrm{G}_2'\mathrm{G}_2 = \mathrm{A}'\mathrm{G}_1'\mathrm{A}'\mathrm{G}_2'\mathrm{G}_2 \\
&= \mathrm{A}'\mathrm{G}_2'\mathrm{G}_2 = \mathrm{G}_2\mathrm{A}\mathrm{G}_2 = \mathrm{G}_2.
\end{aligned}
$$

Moore–Penrose inverse:Existence

We will denote the MP inverse of A by A^+. We now show the existence. Let $\mathrm{A} = \mathrm{BC}$ be a rank factorization. Then it can be easily verified that

$$\mathrm{B}^+ = (\mathrm{B}'\mathrm{B})^{-1}\mathrm{B}', \quad \mathrm{C}^+ = \mathrm{C}'(\mathrm{CC}')^{-1},$$

and then

$$\mathrm{A}^+ = \mathrm{C}^+\mathrm{B}^+.$$