Mathematical modelling

Lecture 2, February 26, 2021

Faculty of Computer and Information Science
University of Ljubljana

2020/2021

Definition

A generalized inverse of a matrix $A \in \mathbb{R}^{n \times m}$ is a matrix $G \in \mathbb{R}^{m \times n}$ such that

$$AGA = A. \tag{1}$$

Remark

Note that the dimension of A and its generalized inverse are transposed to each other. This is the only way which enables the multiplication $A \cdot * \cdot A$.

Proposition

If A is invertible, it has a unique generalized inverse, which is equal to A^{-1}.

Proof.

Let G be a generalized inverse of A, i.e., (1) holds. Multiplying (1) with A^{-1} from the left and the right side we obtain:

$$\text{Left hand side (LHS):} \quad A^{-1}AGAA^{-1} = IGI = G,$$
$$\text{Right hand side (RHS):} \quad A^{-1}AA^{-1} = IA^{-1} = A^{-1},$$

where I is the identity matrix. The equality LHS=RHS implies that $G = A^{-1}$. □

Theorem

Every matrix $A \in \mathbb{R}^{n\times m}$ has a generalized inverse.

Proof.

Let r be the rank of A.

Case 1. $\operatorname{rank} A = \operatorname{rank} A_{11}$, where

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

and $A_{11} \in \mathbb{R}^{r\times r}, A_{12} \in \mathbb{R}^{r\times(m-r)}, A_{21} \in \mathbb{R}^{(n-r)\times r}$, $A_{22} \in \mathbb{R}^{(n-r)\times(m-r)}$.
We claim that

$$G = \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix},$$

where 0s denote zero matrices of appropriate sizes, is the generalized inverse of A. To prove this claim we need to check that

$$AGA = A.$$

$$AGA = \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} \begin{bmatrix} A_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix} = \begin{bmatrix} I & 0 \\ A_{21}A_{11}^{-1} & 0 \end{bmatrix} \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{22} \end{bmatrix}$$
$$= \begin{bmatrix} A_{11} & A_{12} \\ A_{21} & A_{21}A_{11}^{-1}A_{12} \end{bmatrix}.$$

For AGA to be equal to A we must have

$$A_{21}A_{11}^{-1}A_{12} = A_{22}. \tag{2}$$

It remains to prove (2). Since we are in Case 1, it follows that every column of $\begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix}$ is in the column space of $\begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix}$. Hence, there is a cofficient matrix $W \in \mathbb{R}^{r \times (m-r)}$ such that

$$\begin{bmatrix} A_{12} \\ A_{22} \end{bmatrix} = \begin{bmatrix} A_{11} \\ A_{21} \end{bmatrix} W = \begin{bmatrix} A_{11}W \\ A_{21}W \end{bmatrix}.$$

We obtain the equations $A_{11}W = A_{12}$ and $A_{21}W = A_{22}$. Since A_{11} is invertible, we get $W = A_{11}^{-1}A_{12}$ and hence $A_{21}A_{11}^{-1}A_{12} = A_{22}$, which is (2).

Case 2. *The upper left $r \times r$ submatrix of A is not invertible.*

One way to handle this case is to use permutation matrices P and Q, such that $PAQ = \begin{bmatrix} \widetilde{A}_{11} & \widetilde{A}_{12} \\ \widetilde{A}_{21} & \widetilde{A}_{22} \end{bmatrix}$, $\widetilde{A}_{11} \in \mathbb{R}^{r \times r}$ and $\operatorname{rank} \widetilde{A}_{11} = r$. By Case 1 we have that the generalized inverse $(PAQ)^g$ of PAQ equals to $\begin{bmatrix} \widetilde{A}_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix}$.

Thus,

$$(PAQ) \begin{bmatrix} \widetilde{A}_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} (PAQ) = PAQ. \tag{3}$$

Multiplying (3) from the left by P^{-1} and from the right by Q^{-1} we get

$$A \left(Q \begin{bmatrix} \widetilde{A}_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} P \right) A = A.$$

So, $Q \begin{bmatrix} \widetilde{A}_{11}^{-1} & 0 \\ 0 & 0 \end{bmatrix} P = \left(P \begin{bmatrix} \left(\widetilde{A}_{11}^{-1} \right)^T & 0 \\ 0 & 0 \end{bmatrix} Q \right)^T$ is a generalized inverse of A. □

Algorithm for computing a generalized inverse of A

Let r be the rank of A.

1. Find any nonsingular submatrix B in A of order $r \times r$,
2. in A substitute
   - elements of the submatrix B for corresponding elements of $(B^{-1})^T$,
   - all other elements with 0,
3. the transpose of the obtained matrix is a generalized inverse G.

Example

Compute at least one generalized inverse of

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

- Note that $\operatorname{rank} A = 2$. For B from the algorithm one of the possibilities is

$$B = \begin{bmatrix} 1 & 0 \\ 1 & 4 \end{bmatrix},$$

i.e., the submatrix in the right lower corner.

- Computing B^{-1} we get $B^{-1} = \begin{bmatrix} 1 & 0 \\ -\frac{1}{4} & \frac{1}{4} \end{bmatrix}$ and hence

$$\left(B^{-1}\right)^T = \begin{bmatrix} 1 & -\frac{1}{4} \\ 0 & \frac{1}{4} \end{bmatrix}.$$

- A generalized inverse of A is then

$$G = \begin{bmatrix} 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & -\frac{1}{4} \\ 0 & 0 & 0 & \frac{1}{4} \end{bmatrix}^T = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}.$$

Generalized inverses of a matrix A play a similar role as the usual inverse (when it exists) in solving a linear system $Ax = b$.

Theorem

Let $A \in \mathbb{R}^{n \times m}$ and $b \in \mathbb{R}^m$. If the system

$$Ax = b \tag{4}$$

is solvable (that is, $b \in \mathcal{C}(A)$) and G is a generalized inverse of A, then

$$x = Gb \tag{5}$$

is a solution of the system (4).

Moreover, all solutions of the system (4) are exaclty vectors of the form

$$x_z = Gb + (GA - I)z, \tag{6}$$

where z varies over all vectors from $\mathbb{R}^m$.

Proof.

We write A in the column form

$$A = \begin{bmatrix} a_1 & a_2 & \dots & a_m \end{bmatrix},$$

where a_i are column vectors of A. Since the system (4) is solvable, there exist real numbers $\alpha_1, \dots, \alpha_m \in \mathbb{R}$ such that

$$\sum_{i=1}^{m} \alpha_i a_i = b. \tag{7}$$

First we will prove that Gb also solves (4). Multiplying (7) with G we get

$$Gb = \sum_{i=1}^{m} \alpha_i G a_i. \tag{8}$$

Multiplying (9) with A the left side becomes $A(Gb)$, so we have to check that

$$\sum_{i=1}^{m} \alpha_i A G a_i = b. \tag{9}$$

Since G is a generalized inverse of A, we have that $AGA = A$ or restricting to columns of the left hand side we get

$$AGa_i = a_i \quad \text{for every } i = 1, \ldots, m.$$

Plugging this into the left side of (9) we get exactly (7), which holds and proves (9).

For the moreover part we have to prove two facts:

(i) Any x_z of the form (6) solves (4).

(ii) If $A\tilde{x} = b$, then $\tilde{x}$ is of the form x_z for some $z \in \mathbb{R}^m$.

(i) is easy to check:

$$\begin{aligned} Ax_z &= A\big(Gb + (GA - I)z\big) = AGb + A(GA - I)z \\ &= b + (AGA - A)z = b. \end{aligned}$$

To prove (ii) note that

$$A(\tilde{x} - Gb) = 0,$$

which implies that

$$\tilde{x} - Gb \in \ker A.$$

It remains to check that

$$\ker A = \left\{ (GA - I)z \colon z \in \mathbb{R}^m \right\}. \tag{10}$$

The inclusion ($\supseteq$) of (10) is straightforward:

$$A((GA - I)z) = (AGA - A)z = 0.$$

For the inclusion ($\subseteq$) of (10) we have to notice that any $v \in \ker A$ is equal to $(GA - I)z$ for $z = -v$:

$$(GA - I)(-v) = -GAv + v = 0 + v = v. \quad \square$$

Example

Find all solutions of the system

$$Ax = b,$$

where $A = \begin{bmatrix} 0 & 0 & 2 & 0 \\ 0 & 0 & 1 & 0 \\ 2 & 0 & 1 & 4 \end{bmatrix}$ and $b = \begin{bmatrix} 2 \\ 1 \\ 4 \end{bmatrix}$.

- Recall from the example a few slides above that $G = \begin{bmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & -\frac{1}{4} & \frac{1}{4} \end{bmatrix}$.
- Calculating Gb and $GA - I$ we get

$$Gb = \begin{bmatrix} 0 \\ 0 \\ 1 \\ \frac{3}{4} \end{bmatrix} \quad \text{and} \quad A = \begin{bmatrix} -1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ \frac{1}{2} & 0 & 0 & 0 \end{bmatrix}.$$

- Hence,

$$x_z = \begin{bmatrix} -z_1 & -z_2 & 1 & \frac{3}{4} + \frac{1}{2} z_1 \end{bmatrix}^T$$

where z_1, z_2 vary over $\mathbb{R}$.

1.2 The Moore-Penrose generalized inverse

Among all generalized inverses of a matrix A, one has especially nice properties.

Definition

The Moore-Penrose generalized inverse, or shortly the MP inverse of $A \in \mathbb{R}^{n\times m}$ is any matrix $A^+ \in \mathbb{R}^{m\times n}$ satifying the following four conditions:

1. A^+ is a generalized inverse of A: $AA^+A = A$.
2. A is a generalized inverse of A^+: $A^+AA^+ = A^+$.
3. The square matrix $AA^+ \in \mathbb{R}^{n\times n}$ is symmetric: $(AA^+)^T = AA^+$.
4. The square matrix $A^+A \in \mathbb{R}^{m\times m}$ is symmetric: $(A^+A)^T = A^+A$.

Remark

There are two natural questions arising after defining the MP inverse:

- *Does every matrix admit a MP inverse? Yes.*
- *Is the MP inverse unique? Yes.*

Theorem

The MP inverse A^+ of a matrix A is unique.

Proof.

Assume that there are two matrices M_1 and M_2 that satisfy the four conditions in the definition of MP inverse of A. Then,

$$
\begin{aligned}
AM_1 &= (AM_2A)M_1 && \text{by property (1)} \\
&= (AM_2)(AM_1) = (AM_2)^T(AM_1)^T && \text{by property (3)} \\
&= M_2^T(AM_1A)^T = M_2^TA^T && \text{by property (1)} \\
&= (AM_2)^T = AM_2 && \text{by property (3)}
\end{aligned}
$$

A similar argument involving properties (2) and (4) shows that

$$M_1A = M_2A,$$

and so

$$M_1 = M_1AM_1 = M_1AM_2 = M_2AM_2 = M_2.$$

□

Remark

Let us assume that A^+ exists (we will shortly prove this fact). Then the following properties are true:

- *If A is a square invertible matrix, then it $A^+ = A^{-1}$.*
- $(A^+)^+ = A$.
- $(A^T)^+ = (A^+)^T$.

In the rest of this chapter we will be interested in two obvious questions:

- How do we compute A^+?
- Why would we want to compute A^+?

To answer the first question, we will begin by three special cases.

Construction of the MP inverse of $A \in \mathbb{R}^{n\times m}$:

Case 1: $A^TA \in \mathbb{R}^{m\times m}$ *is an invertible matrix. (In particular,* $m \leq n$*.)*

In this case $A^+ = (A^TA)^{-1}A^T$.

To see this, we have to show that the matrix $(A^TA)^{-1}A^T$ satisfies properties (1) to (4):

1. $AMA = A(A^TA)^{-1}A^TA = A(A^TA)^{-1}(A^TA) = A.$
2. $MAM = (A^TA)^{-1}A^TA(A^TA)^{-1}A^T = (A^TA)^{-1}A^T = M.$
3. $$(AM)^T = \left(A(A^TA)^{-1}A^T\right)^T = A\left(\left(A^TA\right)^{-1}\right)^T A^T =$$ $$= A\left(\left(A^TA\right)^T\right)^{-1} A^T = A(A^TA)^{-1}A^T = AM.$$
4. Analoguous to the previous fact.

Case 2: AA^T *is an invertible matrix. (In particular,* $n \leq m$*.)*

In this case A^T satisfies the condition for Case 1, so $(A^T)^+ = (AA^T)^{-1}A$.

Since $(A^T)^+ = (A^+)^T$ it follows that

$$A^+ = \left((A^+)^T\right)^T = \left((AA^T)^{-1}A\right)^T = A^T\left((AA^T)^{-1}\right)^T$$
$$= A^T\left((AA^T)^{-T}\right)^{-1} = A^T(AA^T)^{-1}.$$

Hence, $\underline{A^+ = A^T(AA^T)^{-1}}$.

Case 3: $\Sigma \in \mathbb{R}^{n\times m}$ *is a diagonal matrix of the form*

$$\Sigma = \begin{bmatrix} \sigma_1 & & & & \\ & \sigma_2 & & & \\ & & \ddots & & \\ & & & \sigma_n & \end{bmatrix} \quad \text{or} \quad \widetilde{\Sigma} = \begin{bmatrix} \sigma_1 & & & \\ & \sigma_2 & & \\ & & \ddots & \\ & & & \sigma_m \\ & & & \\ & & & \end{bmatrix}.$$

The MP inverse is

$$\Sigma^+ = \begin{bmatrix} \sigma_1^+ & & & \\ & \sigma_2^+ & & \\ & & \ddots & \\ & & & \sigma_n^+ \\ & & & \\ & & & \end{bmatrix} \quad \text{or} \quad \widetilde{\Sigma}^+ = \begin{bmatrix} \sigma_1^+ & & & & \\ & \sigma_2^+ & & & \\ & & \ddots & & \\ & & & \sigma_m^+ & \end{bmatrix},$$

where $\sigma_i^+ = \begin{cases} \frac{1}{\sigma_i}, & \sigma_i \neq 0, \\ 0, & \sigma_i = 0. \end{cases}$