# THE LINEAR MODEL

## • Generalized Inverses

Definition: A generalized inverse of the matrix $\boldsymbol{A}: m \times n$ is a matrix $\boldsymbol{A}^*: n \times m$ such that $\boldsymbol{x} = \boldsymbol{A}^*\boldsymbol{y}$ is a solution of the system of equations $\boldsymbol{Ax} = \boldsymbol{y}$ for all $\boldsymbol{y}$ for which the equations are consistent.

(In algebra, a system of equations (either linear or nonlinear) is called consistent if there is at least one set of values for the unknowns that satisfies each equation in the system—that is, when substituted into each of the equations, they make each equation hold true as an identity.)

Theorem: $\boldsymbol{A}^*$ is a generalized inverse of $\boldsymbol{A} \Leftrightarrow \boldsymbol{AA}^*\boldsymbol{A} = \boldsymbol{A}$.

Proof: Suppose that $\boldsymbol{A}^*$ is a generalized inverse of $\boldsymbol{A}$.

The system of equations

$$\boldsymbol{Ax} = \boldsymbol{a}_i$$

where $\boldsymbol{A} = (\boldsymbol{a}_1, \boldsymbol{a}_2, \cdots, \boldsymbol{a}_n)$, is consistent, since $\text{rank}\,(\boldsymbol{A}) = \text{rank}\,(\boldsymbol{A}, \boldsymbol{a}_i)$.

> Given the linear system $Ax = B$ and the augmented matrix $(A|B)$.
>
> 1. If $rank(A) = rank(A|B) =$ the number of rows in $x$, then the system has a unique solution.
> 2. If $rank(A) = rank(A|B) <$ the number of rows in $x$, then the system has $\infty$-many solutions.
> 3. If $rank(A) < rank(A|B)$, then the system is inconsistent.

Therefore $\boldsymbol{x} = \boldsymbol{A}^*\boldsymbol{a}_i$ is a solution of the equations for all $i$. Therefore $\boldsymbol{AA}^*\boldsymbol{a}_i = \boldsymbol{a}_i$, or $\boldsymbol{AA}^*\boldsymbol{A} = \boldsymbol{A}$.

Conversely, suppose that $\boldsymbol{A}^*$ is any matrix such that $\boldsymbol{AA}^*\boldsymbol{A} = \boldsymbol{A}$. It follows that:

$$\boldsymbol{A}(\boldsymbol{A}^*\boldsymbol{y}) = \boldsymbol{AA}^*\boldsymbol{Ax} = \boldsymbol{Ax} = \boldsymbol{y}$$

and $\boldsymbol{A}^*\boldsymbol{y}$ is a solution for all $\boldsymbol{y}$ for which the equations are consistent.

## Examples

1.

$$\boldsymbol{A} = \begin{pmatrix} 2 & 0 \\ 1 & 1 \\ 3 & 1 \end{pmatrix}$$

$$\boldsymbol{Ax} = \boldsymbol{y} \Leftrightarrow \begin{aligned} 2x_1 + 0x_2 &= y_1 \\ 1x_1 + 1x_2 &= y_2 \\ 3x_1 + 1x_2 &= y_3 \end{aligned}$$

A solution is:

$$\boldsymbol{x} = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \end{pmatrix} \boldsymbol{y} \text{ and } \boldsymbol{A}^* = \begin{pmatrix} \frac{1}{2} & 0 & 0 \\ -\frac{1}{2} & 1 & 0 \end{pmatrix}.$$

The equations are consistent $\Leftrightarrow y_1 + y_2 = y_3$. By substitition of the latter in, for instance $x_1$, another solution can be obtained:

$$\boldsymbol{x} = \begin{pmatrix} 0 & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & 1 & 0 \end{pmatrix} \boldsymbol{y} \text{ with } \boldsymbol{A}^* = \begin{pmatrix} 0 & -\frac{1}{2} & \frac{1}{2} \\ -\frac{1}{2} & 1 & 0 \end{pmatrix}$$

In both cases it follows that:

$$\boldsymbol{AA}^*\boldsymbol{A} = \boldsymbol{A}.$$

R provides a function ginv(.) for calculating the generalized inverse of a matrix, available in the package MASS.

Example

(i) A non-square matrix not of full rank

```
> library(MASS)
> options(digits=2)
>  A<-matrix(c(3,2,1,4,2,0,5,
+ 2,-1,-1,0,1),4,3,byrow=T)
> A
     [,1] [,2] [,3]
[1,]    3    2    1
[2,]    4    2    0
[3,]    5    2   -1
[4,]   -1    0    1
```

```
>
> M<-ginv(A)
> M
     [,1] [,2]  [,3]  [,4]
[1,]  0    0.06  0.11 -0.06
[2,] 0.20 0.06 -0.06  0.11
[3,] 0.30 0.06 -0.22  0.28
>
```

2.

$$\boldsymbol{A} = \begin{pmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 0 & 0 \\ 1 & 3 & 1 & 1 \end{pmatrix}$$

$$\boldsymbol{Ax} = \boldsymbol{y} \Leftrightarrow \begin{matrix} 1x_1 + 1x_2 + 1x_3 + 1x_4 = y_1 \\ 0x_1 + 1x_2 + 0x_3 + 0x_4 = y_2 \\ 1x_1 + 3x_2 + 1x_3 + 1x_4 = y_3 \end{matrix}$$

The general solution is:

$$\boldsymbol{x} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{y} + \begin{pmatrix} -1 \\ 0 \\ 1 \\ 0 \end{pmatrix} c + \begin{pmatrix} -1 \\ 0 \\ 0 \\ 1 \end{pmatrix} d$$

for arbitrary $c$ and $d$. For a solution, set $c = d = 0$. Then:

$$\boldsymbol{x} = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix} \boldsymbol{y}$$

and $\boldsymbol{A}^* = \begin{pmatrix} 1 & -1 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{pmatrix}$.

The equations are consistent $\Leftrightarrow$ $y_1 + 2y_2 = y_3$. It follows that:

$$\boldsymbol{AA^*A} = \boldsymbol{A}.$$

3.

$$\boldsymbol{A} = \begin{pmatrix} 0 & 4 & 8 & -16 & -12 \\ 3 & 2 & -5 & -17 & -3 \\ 4 & 2 & -8 & -20 & -2 \\ 1 & -3 & -7 & 13 & 8 \end{pmatrix}$$

$$\boldsymbol{Ax} = \boldsymbol{y} \Leftrightarrow \begin{matrix} 0x_1 + 4x_2 + 8x_3 - 16x_4 - 12x_5 = y_1 \\ 3x_1 + 2x_2 - 5x_3 - 17x_4 - 3x_5 = y_2 \\ 4x_1 + 2x_2 - 8x_3 - 20x_4 - 2x_5 = y_3 \\ 1x_1 + -3x_2 - 7x_3 + 13x_4 + 8x_5 = y_4 \end{matrix}$$

$$
\begin{aligned}
0x_1 + 1x_2 + 2x_3 - 4x_4 - 3x_5 &= \frac{1}{4}y_1 \\
1x_1 + 0x_2 - 3x_3 - 3x_4 + 1x_5 &= \frac{1}{3}y_2 - \frac{1}{6}y_1 \\
0x_1 + 2x_2 + 4x_3 - 8x_4 - 6x_5 &= y_3 - \frac{4}{3}y_2 + \frac{2}{3}y_1 \\
0x_1 + -3x_2 - 4x_3 + 16x_4 + 7x_5 &= y_4 - \frac{1}{3}y_2 + \frac{1}{6}y_1
\end{aligned}
$$

$$
\begin{aligned}
0x_1 + 1x_2 + 2x_3 - 4x_4 - 3x_5 &= \frac{1}{4}y_1 \\
1x_1 + 0x_2 - 3x_3 - 3x_4 + 1x_5 &= \frac{1}{3}y_2 - \frac{1}{6}y_1 \\
0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 &= \frac{1}{6}y_1 - \frac{4}{3}y_2 + y_3 \\
0x_1 + 0x_2 + 1x_3 + 2x_4 - 1x_5 &= \frac{11}{24}y_1 - \frac{1}{6}y_2 + \frac{1}{2}y_4
\end{aligned}
$$

$$
\begin{aligned}
0x_1 + 1x_2 + 0x_3 - 8x_4 - 1x_5 &= -\frac{2}{3}y_1 + \frac{1}{3}y_2 - y_4 \\
1x_1 + 0x_2 + 0x_3 + 3x_4 - 2x_5 &= \frac{29}{24}y_1 - \frac{1}{6}y_2 + \frac{3}{2}y_4 \\
0x_1 + 0x_2 + 0x_3 + 0x_4 + 0x_5 &= \frac{1}{6}y_1 - \frac{4}{3}y_2 + y_3 \\
0x_1 + 0x_2 + 1x_3 + 2x_4 - 1x_5 &= \frac{11}{24}y_1 - \frac{1}{6}y_2 + \frac{1}{2}y_4
\end{aligned}
$$

The equations are consistent if and only if

$$\frac{1}{6}y_1 - \frac{4}{3}y_2 + y_3 = 0.$$

The general solution is:

$$x = \begin{pmatrix} \frac{29}{24} & -\frac{1}{6} & 0 & \frac{3}{2} \\ -\frac{2}{3} & \frac{1}{3} & 0 & -1 \\ \frac{11}{24} & -\frac{1}{6} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} y + \begin{pmatrix} -3 \\ 8 \\ -2 \\ 1 \\ 0 \end{pmatrix} c + \begin{pmatrix} 2 \\ 1 \\ 1 \\ 0 \\ 1 \end{pmatrix} d$$

for arbitrary $c$ and $d$. For a solution, set $c = d = 0$. Then:

$$x = \begin{pmatrix} \frac{29}{24} & -\frac{1}{6} & 0 & \frac{3}{2} \\ -\frac{2}{3} & \frac{1}{3} & 0 & -1 \\ \frac{11}{24} & -\frac{1}{6} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix} y$$

$$A^* = \begin{pmatrix} \frac{29}{24} & -\frac{1}{6} & 0 & \frac{3}{2} \\ -\frac{2}{3} & \frac{1}{3} & 0 & -1 \\ \frac{11}{24} & -\frac{1}{6} & 0 & \frac{1}{2} \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{pmatrix}.$$

Again it follows that:

$$AA^*A = A.$$

Theorem: Suppose that $P_1$ and $P_2$ are two non-singular matrices and let

$$B = P_1 A P_2.$$

The following **one-to-one** relationship exists between a generalized inverse of $A$ and a generalized inverse of $B$ : $A^* = P_2 B^* P_1$.

(The implication of this theorem is that, for a given generalized inverse of $B$, there exists a corresponding generalized inverse of $A$ and vice versa.)

Proof:

Suppose that $\boldsymbol{B}^*$ is a generalized inverse of $\boldsymbol{B}$. Then

$$\boldsymbol{B}\boldsymbol{B}^*\boldsymbol{B} = \boldsymbol{B}$$

$$\boldsymbol{P_1 A P_2 B^* P_1 A P_2} = \boldsymbol{P_1 A P_2}.$$

Since $\boldsymbol{P_1}$ and $\boldsymbol{P_2}$ are non-singular, it follows that

$$\boldsymbol{A P_2 B^* P_1 A} = \boldsymbol{A}$$

and $\boldsymbol{P_2 B^* P_1}$ is therefore a generalized inverse of $\boldsymbol{A}$.

Conversely, suppose that $\boldsymbol{A}^*$ is a generalized inverse of $\boldsymbol{A}$. Then

$$\boldsymbol{A A^* A} = \boldsymbol{A}$$

$$\boldsymbol{P_1^{-1} B P_2^{-1} A^* P_1^{-1} B P_2^{-1}} = \boldsymbol{P_1^{-1} B P_2^{-1}}$$

and

$$\boldsymbol{B P_2^{-1} A^* P_1^{-1} B} = \boldsymbol{B}.$$

Therefore $\boldsymbol{P_2^{-1} A^* P_1^{-1}}$ is a generalized inverse of $\boldsymbol{B}$.

Special case: $\boldsymbol{B}, \boldsymbol{P_1}$ and $\boldsymbol{P_2}$ are defined as follows: For any $\boldsymbol{A}$ there always exist two non-singular matrices $\boldsymbol{P_1}$ and $\boldsymbol{P_2}$ such that

$$\boldsymbol{P_1 A P_2} = \begin{pmatrix} \boldsymbol{I_r} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} = \boldsymbol{B}$$

where $r = \text{rank}\ (\boldsymbol{A})$. Let

$$\boldsymbol{D} = \begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{pmatrix}$$

where $\boldsymbol{X}: r \times r, \boldsymbol{U}: r \times (m-r), \boldsymbol{V}: (n-r) \times r$ and $\boldsymbol{W}: (n-r) \times (m-r)$. It then follows that:

$$\begin{aligned} \boldsymbol{BDB} &= \begin{pmatrix} \boldsymbol{I_r} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{pmatrix} \begin{pmatrix} \boldsymbol{I_r} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{X} & \boldsymbol{U} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \begin{pmatrix} \boldsymbol{I_r} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix} \\ &= \begin{pmatrix} \boldsymbol{X} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{pmatrix}. \end{aligned}$$

A necessary and sufficient condition for $\boldsymbol{D}$ to be a generalized inverse of $\boldsymbol{B}$ is that $\boldsymbol{X} = \boldsymbol{I_r}$.

The general form of a generalized inverse of $\boldsymbol{B}$ is therefore:

$$\boldsymbol{B}^* = \begin{pmatrix} \boldsymbol{I}_r & \boldsymbol{U} \\ \boldsymbol{V} & \boldsymbol{W} \end{pmatrix}$$

where the elements of $\boldsymbol{U}, \boldsymbol{V}$ and $\boldsymbol{W}$ are arbitrary.

## Corollary:

1. Any matrix has at least one generalized inverse.

2. In general, a matrix can have an infinite number of generalized inverses.

3. If the matrix $\boldsymbol{A}$ is non-singular, then:

$$\boldsymbol{A}\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{A}, \quad \boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{A}^{-1} = \boldsymbol{A}^{-1}\boldsymbol{A}\boldsymbol{A}^{-1} \quad \text{and} \quad \boldsymbol{A}^* = \boldsymbol{A}^{-1}.$$

In this case $\boldsymbol{A}^* = \boldsymbol{A}^{-1}$ is determined uniquely.

4. It can be shown that for any matrix $\boldsymbol{A}$ there is a unique generalized inverse matrix, $\boldsymbol{C}$, called the Moore-Penrose inverse, which satisfies the following four conditions:

$$\begin{aligned} \boldsymbol{ACA} &= \boldsymbol{A} & (\boldsymbol{i}) \\ \boldsymbol{CAC} &= \boldsymbol{C} & (\boldsymbol{ii}) \\ (\boldsymbol{CA})' &= \boldsymbol{CA} & (\boldsymbol{iii}) \\ (\boldsymbol{AC})' &= \boldsymbol{AC} & (\boldsymbol{iv}) \end{aligned}$$

5. If $\boldsymbol{A} : n \times n$ is symmetrical and $\boldsymbol{A}^*$ is a generalized inverse of $\boldsymbol{A}$, then:

$$\boldsymbol{A}\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{A}.$$

By transposing both sides, it follows that:

$$\boldsymbol{A}'\boldsymbol{A}^{*\prime}\boldsymbol{A}' = \boldsymbol{A}' \quad \text{or} \quad \boldsymbol{A}\boldsymbol{A}^{*\prime}\boldsymbol{A} = \boldsymbol{A}.$$

Therefore, although $\boldsymbol{A}^*$ is not necessarily symmetrical, $\boldsymbol{A}^{*\prime}$ is also a generalized inverse of $\boldsymbol{A}$.

6. It is always possible to choose a symmetrical generalized inverse for a symmetrical matrix $\boldsymbol{A}$, namely $\frac{1}{2}(\boldsymbol{A}^* + \boldsymbol{A}^{*\prime})$, since $\boldsymbol{A} \cdot \frac{1}{2}(\boldsymbol{A}^* + \boldsymbol{A}^{*\prime}) \cdot \boldsymbol{A} = \frac{1}{2}\boldsymbol{A} + \frac{1}{2}\boldsymbol{A} = \boldsymbol{A}$.

7. In the case of the matrix $\boldsymbol{A}: m \times n$ it follows that $\boldsymbol{A}^*\boldsymbol{A}$ is idempotent and that the columns of $\boldsymbol{I} - \boldsymbol{A}^*\boldsymbol{A}$ generate the vector space orthogonal to the rows of $\boldsymbol{A}$. The general solution of the system

$$\boldsymbol{Ax} = \boldsymbol{0}$$

is

$$\boldsymbol{x} = (\boldsymbol{I} - \boldsymbol{A}^{*}\boldsymbol{A})\boldsymbol{z}$$

with the elements of $\boldsymbol{z}$ arbitrary. The general solution of the system

$$\boldsymbol{A}\boldsymbol{x} = \boldsymbol{y}$$

is

$$\boldsymbol{x} = \boldsymbol{A}^{*}\boldsymbol{y} + (\boldsymbol{I} - \boldsymbol{A}^{*}\boldsymbol{A})\boldsymbol{z}$$

with the elements of $\boldsymbol{z}$ arbitrary.