Statistical Sciences Honours
Matrix Methods
Lecture 5 – Generalised Inverses and Linear Equations

Stefan S. Britz
stefan.britz@uct.ac.za

Department of Statistical Sciences
University of Cape Town

5.1 Solutions to different types of sets of linear equations

- In lecture 3 we considered the inverse of a matrix as a way to solve a system of linear equations.
- But what if the $m \times n$ matrix $\boldsymbol{A}$ is not invertible?
- This could either be because $\boldsymbol{A}$ is square but singular, or non-square (rectangular).
- Can we still say something about solutions to $\boldsymbol{Ax} = \boldsymbol{b}$, for fixed $\boldsymbol{b}$, where $\boldsymbol{x}$ and $\boldsymbol{b}$ are $n-$ and $m-$vectors respectively?
- Let's look at the 3 different scenarios where $m < n$, $n > m$, and $n = m$.

5.1 $m < n$ and $rank(\boldsymbol{A}) = m$

- Note that we assume $\boldsymbol{A}$ is of full row rank.
- We have m linear equations, so we will aim to solve for m elements of $\boldsymbol{x}$, by arbitrarily setting $n - m$ elements to 0.
- We will now partition $\boldsymbol{A}$ as follows:

$$\boldsymbol{A} = [{}_m\boldsymbol{B}_m \quad {}_m\boldsymbol{N}_{n-m}]$$

  where $\boldsymbol{B}$ forms the so-called basis for $\boldsymbol{A}$ and the remaining columns non-basic.

5.1 $m < n$ and $rank(\boldsymbol{A}) = m$

Now we can write:

$$\begin{aligned}
\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b} \\
\begin{bmatrix} \boldsymbol{B} & \boldsymbol{N} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{bmatrix} &= \boldsymbol{b} \\
\begin{bmatrix} \boldsymbol{B} & \boldsymbol{N} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_B \\ \boldsymbol{0} \end{bmatrix} &= \boldsymbol{b} \\
\boldsymbol{B}\boldsymbol{x}_B + \boldsymbol{0} &= \boldsymbol{b} \\
\boldsymbol{x}_B &= \boldsymbol{B}^{-1}\boldsymbol{b}
\end{aligned}$$

5.1 $m < n$ and $rank(\boldsymbol{A}) = m$

Now we can write:

$$
\begin{aligned}
\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{b} \\
\begin{bmatrix} \boldsymbol{B} & \boldsymbol{N} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_B \\ \boldsymbol{x}_N \end{bmatrix} &= \boldsymbol{b} \\
\begin{bmatrix} \boldsymbol{B} & \boldsymbol{N} \end{bmatrix} \begin{bmatrix} \boldsymbol{x}_B \\ \boldsymbol{0} \end{bmatrix} &= \boldsymbol{b} \\
\boldsymbol{B}\boldsymbol{x}_B + \boldsymbol{0} &= \boldsymbol{b} \\
\boldsymbol{x}_B &= \boldsymbol{B}^{-1}\boldsymbol{b}
\end{aligned}
$$

Adding the zeros back in, we can express this as

$$
\begin{aligned}
\boldsymbol{x} &= \boldsymbol{G}\boldsymbol{b} \\
&= \begin{bmatrix} \boldsymbol{B}^{-1} \\ \boldsymbol{0} \end{bmatrix} \begin{bmatrix} b_1 \\ \vdots \\ b_m \end{bmatrix}
\end{aligned}
$$

Exercises

Exercise 5.1

Show that $\boldsymbol{AGA} = \boldsymbol{A}$ and that $\boldsymbol{GAG} = \boldsymbol{G}$ where $rank(\boldsymbol{A}) = m$.

5.1 $m > n$ and $rank(\boldsymbol{A}) = n$

- Note that we assume $\boldsymbol{A}$ is of full column rank.
- Since we have more linear equations (m) than variables (n) to solve, there will be no general solution unless some redundancy exists.
- We could find a "closest" approximation to a solution in a least squares sense, by finding the vector $\boldsymbol{x}$ that minimises $(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})'(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x})$.

5.1 $m > n$ and $rank(\boldsymbol{A}) = n$

- Differentiating with respect to $\boldsymbol{x}$ (i.e. with respect to each element of $\boldsymbol{x}$ in turn, and expressing the results in vector form) gives the condition:

  $$2\boldsymbol{A}'(\boldsymbol{b} - \boldsymbol{A}\boldsymbol{x}) = \boldsymbol{0}$$

  Yielding

  $$\begin{aligned} \boldsymbol{A}'\boldsymbol{b} - \boldsymbol{A}'\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{0} \\ \boldsymbol{A}'\boldsymbol{A}\boldsymbol{x} &= \boldsymbol{A}'\boldsymbol{b} \\ \boldsymbol{x} &= (\boldsymbol{A}'\boldsymbol{A})^{-1}\boldsymbol{A}'\boldsymbol{b} \\ &= \boldsymbol{G}\boldsymbol{b} \end{aligned}$$

- Note that it can be shown that the $n \times n$ matrix $\boldsymbol{A}'\boldsymbol{A}$ has rank n and is therefore invertible.

Exercises

Exercise 5.2

Show that $\boldsymbol{AGA} = \boldsymbol{A}$ and that $\boldsymbol{GAG} = \boldsymbol{G}$ where $rank(\boldsymbol{A}) = n$.

5.1 $m = n$, i.e. $\boldsymbol{A}$ is square, but $rank(\boldsymbol{A}) < m$

- We know that a solution to $\boldsymbol{Ax} = \boldsymbol{b}$ exists for $\boldsymbol{b} \neq \boldsymbol{0}$ if $\boldsymbol{A}$ is of full rank.
- When $rank(\boldsymbol{A}) < m$, a solution will only exist if the linear relationships between the rows of $\boldsymbol{A}$ are mirrored by equivalent relationships between the elements of $\boldsymbol{b}$.
- This is illustrated in the following example.

5.1 $m = n$, i.e. $\boldsymbol{A}$ is square, but $rank(\boldsymbol{A}) < m$

Let

$$\boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 1 & 2 \\ 3 & 5 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 14 \\ 6 \\ 22 \end{bmatrix}$$

- Since $\boldsymbol{a}_{(3)} = 2\boldsymbol{a}_{(1)} - \boldsymbol{a}_{(2)}$, we know that $rank(\boldsymbol{A}) < 3$.
- However, on the RHS we also have $22 = 2(14) - 6$.
- Now we effectively have 2 equations and 3 unknowns, so solutions can certainly be found.
- In general, if $rank([\boldsymbol{A} \quad \boldsymbol{b}]) = rank(\boldsymbol{A})$, then the aforementioned mirrored relationship is present.
- This ensures the existence of a solution to the equation $\boldsymbol{A}\boldsymbol{x} = \boldsymbol{b}$, which is then said to be consistent.

5.2 Generalised inverse definitions

Any $m \times n$ matrix $\boldsymbol{G}$ related to $\boldsymbol{A}$ through $\boldsymbol{AGA} = \boldsymbol{A}$ and/or $\boldsymbol{GAG} = \boldsymbol{G}$ plays an important role in finding solutions to $\boldsymbol{Ax} = \boldsymbol{b}$.

Note the following definitions:

5.2 Generalised inverse definitions

Any $m \times n$ matrix $\boldsymbol{G}$ related to $\boldsymbol{A}$ through $\boldsymbol{AGA} = \boldsymbol{A}$ and/or $\boldsymbol{GAG} = \boldsymbol{G}$ plays an important role in finding solutions to $\boldsymbol{Ax} = \boldsymbol{b}$.

Note the following definitions:

- Generalised inverse
  If $\boldsymbol{AGA} = \boldsymbol{A}$ then $\boldsymbol{G}$ is a generalised inverse of $\boldsymbol{A}$, often written as $\boldsymbol{A}^{-}$.

5.2 Generalised inverse definitions

Any $m \times n$ matrix $\boldsymbol{G}$ related to $\boldsymbol{A}$ through $\boldsymbol{AGA} = \boldsymbol{A}$ and/or $\boldsymbol{GAG} = \boldsymbol{G}$ plays an important role in finding solutions to $\boldsymbol{Ax} = \boldsymbol{b}$.

Note the following definitions:

- Generalised inverse
  If $\boldsymbol{AGA} = \boldsymbol{A}$ then $\boldsymbol{G}$ is a generalised inverse of $\boldsymbol{A}$, often written as $\boldsymbol{A}^{-}$.

- Reflexive generalised inverse
  If $\boldsymbol{AGA} = \boldsymbol{A}$ and $\boldsymbol{GAG} = \boldsymbol{G}$, then $\boldsymbol{G}$ is a generalised inverse of $\boldsymbol{A}$ and $\boldsymbol{A}$ is a generalised inverse of $\boldsymbol{G}$. We then say that $\boldsymbol{G}$ is a reflexive generalised inverse of $\boldsymbol{A}$.

5.2 Generalised inverse definitions

Any $m \times n$ matrix $\boldsymbol{G}$ related to $\boldsymbol{A}$ through $\boldsymbol{AGA} = \boldsymbol{A}$ and/or $\boldsymbol{GAG} = \boldsymbol{G}$ plays an important role in finding solutions to $\boldsymbol{Ax} = \boldsymbol{b}$.

Note the following definitions:

- Generalised inverse
  If $\boldsymbol{AGA} = \boldsymbol{A}$ then $\boldsymbol{G}$ is a generalised inverse of $\boldsymbol{A}$, often written as $\boldsymbol{A}^-$.
- Reflexive generalised inverse
  If $\boldsymbol{AGA} = \boldsymbol{A}$ and $\boldsymbol{GAG} = \boldsymbol{G}$, then $\boldsymbol{G}$ is a generalised inverse of $\boldsymbol{A}$ and $\boldsymbol{A}$ is a generalised inverse of $\boldsymbol{G}$. We then say that $\boldsymbol{G}$ is a reflexive generalised inverse of $\boldsymbol{A}$.
- Pseudo inverse (Moore-Penrose inverse)
  If $\boldsymbol{G}$ is a reflexive generalized inverse of $\boldsymbol{A}$ and if $\boldsymbol{AG}$ and $\boldsymbol{GA}$ are symmetric, then $\boldsymbol{G}$ is the pseudo inverse of $\boldsymbol{A}$.

5.3 Construction

One method of constructing a generalized inverse starts by reducing $\boldsymbol{A}$ to the form:

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{D}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

where $r = rank(\boldsymbol{A})$ such that $\boldsymbol{D}_r$ is of full rank.

5.3 Construction

One method of constructing a generalized inverse starts by reducing $\boldsymbol{A}$ to the form:

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{D}_r & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$$

where $r = rank(\boldsymbol{A})$ such that $\boldsymbol{D}_r$ is of full rank.

- Note that this is done in the same way as when we reduced a matrix to canonical form, expect that $\boldsymbol{D}_r$ isn't necessarily $\boldsymbol{I}_r$.
- Therefore, $\boldsymbol{R} = \boldsymbol{PAQ}$, where $\boldsymbol{P}$ and $\boldsymbol{Q}$ are the products of the relevant elementary operators.

5.3 Construction

One method of constructing a generalized inverse starts by reducing $\boldsymbol{A}$ to the form:

$$\boldsymbol{R} = \begin{bmatrix} \boldsymbol{D}_r & \mathbf{0} \\ \mathbf{0} & \mathbf{0} \end{bmatrix}$$

where $r = rank(\boldsymbol{A})$ such that $\boldsymbol{D}_r$ is of full rank.

- Note that this is done in the same way as when we reduced a matrix to canonical form, expect that $\boldsymbol{D}_r$ isn't necessarily $\boldsymbol{I}_r$.
- Therefore, $\boldsymbol{R} = \boldsymbol{PAQ}$, where $\boldsymbol{P}$ and $\boldsymbol{Q}$ are the products of the relevant elementary operators.
- Then

  $$\boldsymbol{G} = \boldsymbol{Q} \begin{bmatrix} \boldsymbol{D}_r^{-1} & \boldsymbol{X} \\ \boldsymbol{Y} & \boldsymbol{Z} \end{bmatrix} \boldsymbol{P}$$

  is a generalised inverse of $\boldsymbol{A}$ for any arbitrary matrices (of the required sizes) $\boldsymbol{X}$, $\boldsymbol{Y}$ and $\boldsymbol{Z}$.

Exercise 5.3

a) Demonstrate that $\boldsymbol{G}$ as defined above is a generalised inverse of $\boldsymbol{A}$, i.e. $\boldsymbol{AGA} = \boldsymbol{A}$, by invoking the relationship $\boldsymbol{A} = \boldsymbol{P}^{-1}\boldsymbol{R}\boldsymbol{Q}^{-1}$.

b) Also show that $\boldsymbol{G}$ is a reflexive generalised inverse if and only if $\boldsymbol{Z} = \boldsymbol{Y}\boldsymbol{D}_r\boldsymbol{X}$.

5.3 Construction

Consider the following special case:

Let

$$\boldsymbol{A} = \begin{bmatrix} \boldsymbol{B} & \boldsymbol{S} \\ \boldsymbol{T} & \boldsymbol{U} \end{bmatrix}$$

where $\boldsymbol{B}: r \times r$ is the non-singular sub-matrix of $\boldsymbol{A}$.

If $\boldsymbol{U} = \boldsymbol{T}\boldsymbol{B}^{-1}\boldsymbol{S}$, then

$$\boldsymbol{G} = \begin{bmatrix} \boldsymbol{B}^{-1} & \boldsymbol{0} \\ \boldsymbol{0} & \boldsymbol{0} \end{bmatrix}$$

is a generalised inverse of $\boldsymbol{A}$.

5.4 Solutions to $\boldsymbol{Ax} = \boldsymbol{b}$

- First consider non-trivial solutions to $\boldsymbol{Ax} = \boldsymbol{0}$.
- If $\boldsymbol{A}$ is non-singular, then the trivial solution is unique.
- A general solution can be expressed as

$$\boldsymbol{x} = (\boldsymbol{GA} - \boldsymbol{I})\boldsymbol{z}$$

for any arbitrary vector $\boldsymbol{z}$, where $\boldsymbol{G}$ is a generalised inverse of $\boldsymbol{A}$.

5.4 Solutions to $\boldsymbol{Ax} = \boldsymbol{b}$

- First consider non-trivial solutions to $\boldsymbol{Ax} = \boldsymbol{0}$.
- If $\boldsymbol{A}$ is non-singular, then the trivial solution is unique.
- A general solution can be expressed as

$$\boldsymbol{x} = (\boldsymbol{GA} - \boldsymbol{I})\boldsymbol{z}$$

  for any arbitrary vector $\boldsymbol{z}$, where $\boldsymbol{G}$ is a generalised inverse of $\boldsymbol{A}$.
- Showing that this is a solution:

$$\boldsymbol{Ax} = \boldsymbol{A}(\boldsymbol{GA} - \boldsymbol{I})\boldsymbol{z} = (\boldsymbol{AGA} - \boldsymbol{A})\boldsymbol{z} = \boldsymbol{0}$$

- If $\boldsymbol{A}$ is of less than full column rank, then this solution will in general be non-trivial.

5.4 Solutions to $\boldsymbol{Ax} = \boldsymbol{b}$

- Now let's return to $\boldsymbol{Ax} = \boldsymbol{b}$, which we suppose is consistent.
- For any arbitrary matrix $\boldsymbol{G}$ of the appropriate size, we have

$$\boldsymbol{AGAx} = \boldsymbol{AGb}$$

5.4 Solutions to $\boldsymbol{Ax} = \boldsymbol{b}$

- Now let's return to $\boldsymbol{Ax} = \boldsymbol{b}$, which we suppose is consistent.
- For any arbitrary matrix $\boldsymbol{G}$ of the appropriate size, we have

$$\boldsymbol{AGAx} = \boldsymbol{AGb}$$

- If $\boldsymbol{G}$ is a generalised inverse of $\boldsymbol{A}$, then this becomes

$$\begin{aligned} \boldsymbol{Ax} &= \boldsymbol{AGb} \\ \boldsymbol{b} &= \boldsymbol{A}(\boldsymbol{Gb}) \end{aligned}$$

- Thus if the equations are consistent, then a solution is given by $\boldsymbol{x} = \boldsymbol{Gb}$.

5.4 Solutions to $\boldsymbol{Ax} = \boldsymbol{b}$

- Now, since $\boldsymbol{x} = (\boldsymbol{GA} - \boldsymbol{I})\boldsymbol{z}$ is a solution to $\boldsymbol{Ax} = \boldsymbol{0}$ for any arbitrary vector $\boldsymbol{z}$, and $\boldsymbol{G}$ is a generalised inverse of $\boldsymbol{A}$, $\boldsymbol{Gb} + (\boldsymbol{GA} - \boldsymbol{I})\boldsymbol{z}$ is also a solution.
- This characterises all solutions which can exist.
- The solution is unique if and only if $\boldsymbol{GA} = \boldsymbol{I}$.

Class exercise

Consider again

$$\boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 1 & 2 \\ 3 & 5 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 14 \\ 6 \\ 22 \end{bmatrix}$$

Class exercise

Consider again

$$\boldsymbol{A}\boldsymbol{x} = \begin{bmatrix} 2 & 3 & 1 \\ 1 & 1 & 2 \\ 3 & 5 & 0 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} 14 \\ 6 \\ 22 \end{bmatrix}$$