Further Mathematical Methods (Linear Algebra) 2002

Lecture 18: Weak Generalised Inverses

We now extend our notion of 'inverse' even further and examine some of the other properties that such matrices may have. We start by defining a new class of 'inverses' which, incidentally, includes those that we introduced in the last lecture:

Definition 18.1 *Let* A *be an arbitrary* $m \times n$ *matrix. A weak generalised inverse (WGI) of* A*, denoted by* A^g*, is any* $n \times m$ *matrix such that*

$$\mathsf{A}\mathsf{A}^g\mathsf{A} = \mathsf{A}.$$

Note: In general, a given matrix A will have *many* WGIs. For instance, if B is an $n \times m$ matrix such that $\mathsf{AB} = \mathsf{0}$ and A^g is a WGI of A, then as

$$\mathsf{A}(\mathsf{A}^g + \mathsf{B})\mathsf{A} = \mathsf{A}\mathsf{A}^g\mathsf{A} + \mathsf{ABA} = \mathsf{A} + \mathsf{0} = \mathsf{A},$$

the matrix $\mathsf{A}^g + \mathsf{B}$ will be a WGI of A too.

For example: All left and right inverses (if they exist) are weak generalised inverses. This is because for any matrix A, if it has a left inverse L, then we have:

$$\mathsf{ALA} = \mathsf{A}(\mathsf{LA}) = \mathsf{AI} = \mathsf{A},$$

whereas, if it has a right inverse R, then

$$\mathsf{ARA} = (\mathsf{AR})\mathsf{A} = \mathsf{IA} = \mathsf{A},$$

as desired. ♣

In fact, WGIs have a rather surprising property, namely that:

Theorem 18.2 *If* A^g *is a WGI of the matrix* $m \times n$ *matrix* A*, then*

- $\mathsf{A}\mathsf{A}^g$ *projects* $\mathbb{R}^m$ *onto* $R(\mathsf{A})$*.*
- $\mathsf{A}^g\mathsf{A}$ *projects* $\mathbb{R}^n$ *parallel to* $N(\mathsf{A})$*.*

Proof: Firstly, to show that $\mathsf{A}\mathsf{A}^g$ projects $\mathbb{R}^m$ onto $R(\mathsf{A})$, we note that:

- $\mathsf{A}\mathsf{A}^g$ is idempotent since:

  $$(\mathsf{A}\mathsf{A}^g)^2 = \mathsf{A}\mathsf{A}^g\mathsf{A}\mathsf{A}^g = (\mathsf{A}\mathsf{A}^g\mathsf{A})\mathsf{A}^g = \mathsf{A}\mathsf{A}^g,$$

  and so $\mathsf{A}\mathsf{A}^g$ represents a projection.

- $R(\mathsf{A}\mathsf{A}^g) = R(\mathsf{A})$ since:
  - For any $\mathbf{y} \in R(\mathsf{A})$, there exists an $\mathbf{x} \in \mathbb{R}^n$ such that $\mathbf{y} = \mathsf{A}\mathbf{x}$. However, as $\mathsf{A}\mathsf{A}^g\mathsf{A} = \mathsf{A}$, this is the same as $\mathbf{y} = \mathsf{A}\mathsf{A}^g\mathsf{A}\mathbf{x} = \mathsf{A}\mathsf{A}^g(\mathsf{A}\mathbf{x})$, and so $\mathbf{y} \in R(\mathsf{A}\mathsf{A}^g)$ too. Thus, $R(\mathsf{A}) \subseteq R(\mathsf{A}\mathsf{A}^g)$.
  - For any $\mathbf{y} \in R(\mathsf{A}\mathsf{A}^g)$, there exists an $\mathbf{x} \in \mathbb{R}^m$ such that $\mathbf{y} = (\mathsf{A}\mathsf{A}^g)\mathbf{x}$, or indeed, $\mathbf{y} = \mathsf{A}(\mathsf{A}^g\mathbf{x})$ and so, $\mathbf{y} \in R(\mathsf{A})$ too. Thus, $R(\mathsf{A}\mathsf{A}^g) \subseteq R(\mathsf{A})$.

  Consequently, $\mathsf{A}\mathsf{A}^g$ represents a projection onto $R(\mathsf{A})$.

So, as A^g is an $n \times m$ matrix, $\mathsf{A}\mathsf{A}^g$ projects $\mathbb{R}^m$ onto $R(\mathsf{A})$, as required.

Secondly, to show that $\mathsf{A}^g\mathsf{A}$ projects $\mathbb{R}^n$ parallel to $N(\mathsf{A})$, we note that:

- $\mathsf{A}^g\mathsf{A}$ is idempotent since:

$$(\mathsf{A}^g\mathsf{A})^2 = \mathsf{A}^g\mathsf{A}\mathsf{A}^g\mathsf{A} = \mathsf{A}^g(\mathsf{A}\mathsf{A}^g\mathsf{A}) = \mathsf{A}^g\mathsf{A},$$

  and so $\mathsf{A}^g\mathsf{A}$ represents a projection.

- $N(\mathsf{A}^g\mathsf{A}) = N(\mathsf{A})$ since:

  - For any $\mathbf{y} \in N(\mathsf{A})$, we have $\mathsf{A}\mathbf{y} = \mathbf{0}$, or indeed, $\mathsf{A}^g\mathsf{A}\mathbf{y} = \mathsf{A}^g\mathbf{0} = \mathbf{0}$, and so $\mathbf{y} \in N(\mathsf{A}^g\mathsf{A})$ too. Thus, $N(\mathsf{A}) \subseteq N(\mathsf{A}^g\mathsf{A})$.
  - For any $\mathbf{y} \in N(\mathsf{A}^g\mathsf{A})$, we have $\mathsf{A}^g\mathsf{A}\mathbf{y} = \mathbf{0}$, or indeed, $\mathsf{A}\mathsf{A}^g\mathsf{A}\mathbf{y} = \mathbf{0}$. However, as $\mathsf{A}\mathsf{A}^g\mathsf{A} = \mathsf{A}$, this is the same as $\mathsf{A}\mathbf{y} = \mathbf{0}$, and so $\mathbf{y} \in N(\mathsf{A})$ too. Thus, $N(\mathsf{A}^g\mathsf{A}) \subseteq N(\mathsf{A})$.

  Consequently, $\mathsf{A}\mathsf{A}^g$ represents a projection parallel to $N(\mathsf{A})$.

So, as A is an $m \times n$ matrix, $\mathsf{A}^g\mathsf{A}$ projects $\mathbb{R}^n$ parallel to $N(\mathsf{A})$, as required. ♠

Like the left and right inverses which we saw in the last lecture, WGIs are closely linked to the solubility of matrix equations. To see this, we note the following theorem:

Theorem 18.3 *For an $m \times n$ matrix A, the matrix equation $\mathsf{A}\mathbf{x} = \mathbf{b}$ is consistent iff $\mathbf{b} = \mathsf{A}\mathsf{A}^g\mathbf{b}$. Further, when it is consistent, its solutions are given by the formula:*

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

where $\mathbf{w}$ is any vector in $\mathbb{R}^n$.

Proof: We start by proving that the matrix equation $\mathsf{A}\mathbf{x} = \mathbf{b}$ is consistent iff $\mathbf{b} = \mathsf{A}\mathsf{A}^g\mathbf{b}$. As this is an 'iff' statement, we have to prove it 'both ways':

- **LTR:** If $\mathsf{A}\mathbf{x} = \mathbf{b}$ is consistent, then it must be the case that $\mathbf{b} \in R(\mathsf{A})$. So, as $\mathsf{A}\mathsf{A}^g$ projects onto $R(\mathsf{A})$, we have $\mathsf{A}\mathsf{A}^g\mathbf{b} = \mathbf{b}$.

- **RTL:** If $\mathsf{A}\mathsf{A}^g\mathbf{b} = \mathbf{b}$, then $\mathbf{x} = \mathsf{A}^g\mathbf{b}$ is clearly a solution of the matrix equation $\mathsf{A}\mathbf{x} = \mathbf{b}$. That is, this matrix equation has a solution and so it is consistent.

as required.

Further, we have to prove that when this matrix equation is consistent, its solutions are given by the formula:

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

where $\mathbf{w}$ is any vector in $\mathbb{R}^n$. To do this, we note that if the matrix equation $\mathsf{A}\mathbf{x} = \mathbf{b}$ is consistent, then from the previous part of this theorem $\mathbf{b} = \mathsf{A}\mathsf{A}^g\mathbf{b}$ and so $\mathbf{x} = \mathsf{A}^g\mathbf{b}$ is a particular solution to this matrix equation. Consequently, $\mathsf{A}\mathbf{x} = \mathbf{b}$ will have a general solution of the form $\mathbf{x} = \mathbf{y} + \mathbf{z}$ where $\mathbf{z} \in N(\mathsf{A})$. But, $\mathsf{I} - \mathsf{A}^g\mathsf{A}$ is a projection onto $N(\mathsf{A})$ and so $\mathbf{z}$ will be a vector of the form $(\mathsf{I} - \mathsf{A}^g\mathsf{A})\mathbf{w}$ where $\mathbf{w}$ is any vector in $\mathbb{R}^n$. Consequently, a general solution of the matrix equation $\mathsf{A}\mathbf{x} = \mathbf{b}$ [when it is consistent] will be

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

where $\mathbf{w}$ is any vector in $\mathbb{R}^n$ (as required). ♠

So, using this theorem, we can see that if the matrix equation $\mathsf{A}\mathbf{x} = \mathbf{b}$ has a solution, we can use a WGI to find it. Indeed, we can easily verify that this formula does give us solutions by noting that

$$\mathsf{A}\left[\mathsf{A}^g\mathbf{b} + (\mathsf{I} - \mathsf{A}^g\mathsf{A})\mathbf{w}\right] = \mathsf{A}\mathsf{A}^g\mathbf{b} + (\mathsf{A}\mathsf{I} - \mathsf{A}\mathsf{A}^g\mathsf{A})\mathbf{w} = \mathbf{b} + (\mathsf{A} - \mathsf{A})\mathbf{w} = \mathbf{b},$$

where we have used the theorem above, i.e. if there are solutions, then $\mathsf{A}\mathsf{A}^g\mathbf{b} = \mathbf{b}$.

Note: The existence of such a formula does not imply that WGIs provide the best way of solving matrix equations or the corresponding systems of linear equations. Indeed, this will usually not be the case.

Note: You will see how to calculate WGIs, or at least left and right inverses which are examples of them, in Questions 1,2 and 3 of Problem Sheet 10.