MATH1030 Homogeneous system associated to a system of linear equations

0. Reminder on notations: $\mathbb{R}^n$ stands for the set of all (column) vectors with n entries.

1. **Definition.**
   Let A be an $(m \times n)$-matrix with real entries.
   (a) *The system of linear equations $\mathcal{LS}(A, \mathbf{0})$ is called the homogeneous system with coefficient matrix A.*
   (b) *For each $\mathbf{b} \in \mathbb{R}^m$, the system $\mathcal{LS}(A, \mathbf{0})$ is referred to as the homogeneous system associated to the system of linear equations $\mathcal{LS}(A, \mathbf{b})$.*

2. **Theorem (1).**
   Let A be an $(m \times n)$-matrix. Let $\mathbf{b} \in \mathbb{R}^m$.
   (a) *For any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$, if '$\mathbf{x} = \mathbf{u}$', '$\mathbf{x} = \mathbf{v}$' are solutions of $\mathcal{LS}(A, \mathbf{b})$, then '$\mathbf{x} = \mathbf{v} - \mathbf{u}$' is a solution of $\mathcal{LS}(A, \mathbf{0})$.*
   (b) *For any $\mathbf{u}, \mathbf{h} \in \mathbb{R}^n$, if '$\mathbf{x} = \mathbf{u}$' is a solution of $\mathcal{LS}(A, \mathbf{b})$ and '$\mathbf{x} = \mathbf{h}$' is a solution of $\mathcal{LS}(A, \mathbf{0})$, then '$\mathbf{x} = \mathbf{u} + \mathbf{h}$' is a solution of $\mathcal{LS}(A, b)$.*

3. **Proof of Theorem (1).**
   Let A be an $(m \times n)$-matrix. Let $\mathbf{b} \in \mathbb{R}^m$.
   (a) Pick any $\mathbf{u}, \mathbf{v} \in \mathbb{R}^n$.
   Suppose '$\mathbf{x} = \mathbf{u}$', '$\mathbf{x} = \mathbf{v}$' are solutions of $\mathcal{LS}(A, \mathbf{b})$.
   Then $A\mathbf{u} = \mathbf{b}$ and $A\mathbf{v} = \mathbf{b}$.
   Therefore $A(\mathbf{v} - \mathbf{u}) = A\mathbf{v} - A\mathbf{u} = \mathbf{b} - \mathbf{b} = \mathbf{0}$.
   Hence '$\mathbf{x} = \mathbf{v} - \mathbf{u}$' is a solution of $\mathcal{LS}(A, \mathbf{0})$.
   (b) Pick any $\mathbf{u}, \mathbf{h} \in \mathbb{R}^n$.
   Suppose '$\mathbf{x} = \mathbf{u}$' is a solution of $\mathcal{LS}(A, \mathbf{b})$ and '$\mathbf{x} = \mathbf{h}$' is a solution of $\mathcal{LS}(A, \mathbf{0})$.
   Then $A\mathbf{u} = \mathbf{b}$ and $A\mathbf{h} = \mathbf{0}$.
   Therefore $A(\mathbf{u} + \mathbf{h}) = A\mathbf{u} + A\mathbf{h} = \mathbf{b} + \mathbf{0} = \mathbf{b}$.
   Hence '$\mathbf{x} = \mathbf{u} + \mathbf{h}$' is a solution of $\mathcal{LS}(A, b)$.

4. **Theorem (2). (Set-theoretic re-formulation of Theorem (1).)**
   Let A be an $(m \times n)$-matrix. For each $\mathbf{c} \in \mathbb{R}^m$, denote the solution set of $\mathcal{LS}(A, c)$ by Π_c. (Note that $\Pi_{\mathbf{0}} = \mathcal{N}(A)$.)
   For each $\mathbf{b} \in \mathbb{R}^m$, for each $\mathbf{u} \in \Pi_{\mathbf{b}}$, the equality
   $$\Pi_{\mathbf{b}} = \left\{ \mathbf{t} \in \mathbb{R}^n : \begin{array}{l} \text{There exists some } \mathbf{h} \in \mathcal{N}(A) \\ \text{such that } \mathbf{t} = \mathbf{u} + \mathbf{h} \end{array} \right\}$$
   holds.
   Remark. What the set equality in the conclusion of Theorem (2) says is:
   (†) For any $\mathbf{v} \in \mathbb{R}^n$, if
   $$\mathbf{v} \text{ belongs to the solution set of } \mathcal{LS}(A, \mathbf{b})$$
   then
   $$\text{there exists some } \mathbf{h} \in \mathcal{N}(A) \text{ such that } \mathbf{v} = \mathbf{u} + \mathbf{h}.$$
   (‡) For any $\mathbf{v} \in \mathbb{R}^n$, if
   $$\text{there exists some } \mathbf{h} \in \mathcal{N}(A) \text{ such that } \mathbf{v} = \mathbf{u} + \mathbf{h}$$
   then
   $$\mathbf{v} \text{ belongs to the solution set of } \mathcal{LS}(A, \mathbf{b}).$$
   Further remark. Some people like to 'abuse notation' to present the conclusion in this result as:
   For each $\mathbf{b} \in \mathbb{R}^m$, for each $\mathbf{u} \in \Pi_{\mathbf{b}}$, the 'equality' $\Pi_{\mathbf{b}} = \mathbf{u} + \Pi_{\mathbf{0}}$ holds.
   The 'equality' $\Pi_{\mathbf{b}} = \mathbf{u} + \Pi_{\mathbf{0}}$, when put in plain words, is to be interpreted as:

the 'totality of all solutions of $\mathcal{LS}(A,\,\mathbf{b})$' can be regarded as

$$\left(\begin{array}{l}\text{one 'particular solution'}\\ \text{of } \mathcal{LS}(A,\,\mathbf{b})\end{array}\right)+\left(\begin{array}{l}\text{totality of all solutions}\\ \text{of } \mathcal{LS}(A,\,\mathbf{0})\end{array}\right)$$

5. **Illustration (1) of the idea in Theorem (1) and Theorem (2), from school maths.**

Let $A = \begin{bmatrix} -1 & 2 \end{bmatrix}$.

For each real number b, $\mathcal{LS}(A,\,b)$ is simply the one linear equation two unknowns $-x_1 + 2x_2 = b$.

For each b, the solution set of $\mathcal{LS}(A,\,b)$ is the 'entire line' in the coordinate plane.

The null space of A is the line ℓ_0 passing through the origin with slope $\dfrac{1}{2}$.

Now suppose $b \neq 0$ (for the sake of visual illustration).

(a) The solution set of $\mathcal{LS}(A,\,b)$ is the line ℓ_b parallel to ℓ_0 and distinct from ℓ_0, passing through the point $(-b, 0)$. '$(x_1, x_2) = (-b, 0)$' is a solution of $\mathcal{LS}(A,\,b)$.

(b) Now identify column vectors of size 2 with the points on the coordinate plane in the natural way. We see that:

- $\mathbf{v}$ belongs to ℓ_b if and only if $\mathbf{v} = \begin{bmatrix} -b \\ 0 \end{bmatrix} + \mathbf{h}$ for some $\mathbf{h} \in \mathcal{N}(A)$.

We can visualize this relation geometrically as follows:—

- We can obtain ℓ_b from ℓ_0 by applying to every point of ℓ_0 a 'translation' by $\begin{bmatrix} -b \\ 0 \end{bmatrix}$.

(c) In general, given any point (u_1, u_2) on ℓ_b, it happens that:

- $\mathbf{v}$ belongs to ℓ_b if and only if $\mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \end{bmatrix} + \mathbf{g}$ for some $\mathbf{g} \in \mathcal{N}(A)$.

We can visualize this relation geometrically as follows:—

- We can obtain ℓ_b from ℓ_0 by applying to every point of ℓ_0 a 'translation' by $\begin{bmatrix} u_1 \\ u_2 \end{bmatrix}$.

(d) For instance, note that ℓ_b passes through $(0, b/2)$. Then '$(x_1, x_2) = (0, b/2)$' is also a solution of $\mathcal{LS}(A,\,b)$. It so happens that:

- $\mathbf{v}$ belongs to ℓ_b if and only if $\mathbf{v} = \begin{bmatrix} 0 \\ b/2 \end{bmatrix} + \mathbf{f}$ for some $\mathbf{f} \in \mathcal{N}(A)$.

6. **Illustration (2) of the idea in Theorem (1) and Theorem (2).**

Let $A = \begin{bmatrix} 1/3 & 1/2 & -1 \end{bmatrix}$.

For each real number b, $\mathcal{LS}(A,\,b)$ is simply the one linear equation with unknowns $\dfrac{1}{3}x_1 + \dfrac{1}{2}x_2 - x_3 = b$.

For each b, the solution set of $\mathcal{LS}(A,\,b)$ is the 'entire line' in the coordinate plane.

The null space of A is the plane Π_0 passing through the origin with 'normal direction' parallel to the vector $\begin{bmatrix} 1/3 \\ 1/2 \\ -1 \end{bmatrix}$.

It is explicitly given by $\mathcal{N}(A) = \left\{ \mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = s \begin{bmatrix} 3 \\ 0 \\ 1 \end{bmatrix} + t \begin{bmatrix} 0 \\ 2 \\ 1 \end{bmatrix} \text{ for some } s, t\mathbb{R} \right\}$.

Now suppose $b \neq 0$ (for the sake of visual illustration).

(a) The solution set of $\mathcal{LS}(A,\,b)$ is the line ℓ_b parallel to Π_0 and distinct from Π_0, passing through the point $(3b, 0, 0)$.
'$(x_1, x_2, x_3) = (3b, 0, 0)$' is a solution of $\mathcal{LS}(A,\,b)$.

(b) Now identify column vectors of size 3 with the points on the coordinate space in the natural way. We see that:

- $\mathbf{v}$ belongs to Π_b if and only if $\mathbf{v} = \begin{bmatrix} 3b \\ 0 \\ 0 \end{bmatrix} + \mathbf{h}$ for some $\mathbf{h} \in \mathcal{N}(A)$.

We can visualize this relation geometrically as follows:—

- We can obtain Π_b from Π_0 by applying to every point of Π_0 a 'translation' by $\begin{bmatrix} 3b \\ 0 \\ 0 \end{bmatrix}$.

(c) In general, given any point (u_1, u_2, u_3) on Π_b, it happens that:

- $\mathbf{v}$ belongs to Π_b if and only if $\mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \mathbf{g}$ for some $\mathbf{g} \in \mathcal{N}(A)$.

We can visualize this relation geometrically as follows:—

- We can obtain Π_b from Π_0 by applying to every point of Π_0 a 'translation' by $\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$.

(d) For instance, note that Π_b passes through $(0, 0, -b)$. Then '$(x_1, x_2, x_3) = (0, 0, -b)$' is also a solution of $\mathcal{LS}(A,\, b)$.

It so happens that:

- $\mathbf{v}$ belongs to Π_b if and only if $\mathbf{v} = \begin{bmatrix} 0 \\ 0 \\ -b \end{bmatrix} + \mathbf{f}$ for some $\mathbf{f} \in \mathcal{N}(A)$.

7. **Illustration (3) of the idea in Theorem (1) and Theorem (2).**

Let $A = \begin{bmatrix} 1 & 0 & -2 \\ 0 & 1 & -3 \end{bmatrix}$.

For each vector $\mathbf{b}$ in $\mathbb{R}^2$, $\mathcal{LS}(A,\, \mathbf{b})$ is the system of two linear equation with three unknowns $\left\{ \begin{array}{ccccccc} x_1 & & & - & 2x_3 & = & b_1 \\ & & x_2 & - & 3x_3 & = & b_2 \end{array} \right.$,

in which $\mathbf{b} = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$.

For each b_1, b_2, the solution set of $\mathcal{LS}(A,\, b)$ is a 'line' in the coordinate space.

The null space of A is the plane $\Lambda_{\mathbf{0}}$ passing through the origin in the 'direction' of the vector $\begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix}$

It is explicitly given by $\mathcal{N}(A) = \left\{ \mathbf{x} \in \mathbb{R}^3 : \mathbf{x} = t \begin{bmatrix} 2 \\ 3 \\ 1 \end{bmatrix} \text{ for some } t\mathbb{R} \right\}$.

Now suppose $b \neq 0$ (for the sake of visual illustration).

(a) The solution set of $\mathcal{LS}(A,\, \mathbf{b})$ is the line $\Lambda_{\mathbf{b}}$ parallel to $\Lambda_{\mathbf{0}}$ and distinct from $\Lambda_{\mathbf{0}}$, passing through the point $(b_1, b_2, 0)$.

'$(x_1, x_2, x_3) = (b_1, b_2, 0)$' is a solution of $\mathcal{LS}(A,\, \mathbf{b})$.

(b) Now identify column vectors of size 3 with the points on the coordinate space in the natural way. We see that:

- $\mathbf{v}$ belongs to $\Lambda_{\mathbf{b}}$ if and only if $\mathbf{v} = \begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix} + \mathbf{h}$ for some $\mathbf{h} \in \mathcal{N}(A)$.

- We can obtain $\Lambda_{\mathbf{b}}$ from $\Lambda_{\mathbf{0}}$ by applying to every point of $\Lambda_{\mathbf{0}}$ a 'translation' by $\begin{bmatrix} b_1 \\ b_2 \\ 0 \end{bmatrix}$.

(c) In general, given any point (u_1, u_2, u_3) on $\Lambda_{\mathbf{b}}$, it happens that:

- $\mathbf{v}$ belongs to $\Lambda_{\mathbf{b}}$ if and only if $\mathbf{v} = \begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix} + \mathbf{g}$ for some $\mathbf{g} \in \mathcal{N}(A)$.

We can visualize this relation geometrically as follows:—

- We can obtain $\Lambda_{\mathbf{b}}$ from $\Lambda_{\mathbf{0}}$ by applying to every point of $\Lambda_{\mathbf{0}}$ a 'translation' by $\begin{bmatrix} u_1 \\ u_2 \\ u_3 \end{bmatrix}$.

8. **Proof of Theorem (2).**

Let $\mathbf{b} \in \mathbb{R}^m$. Let $\mathbf{u} \in \Pi_{\mathbf{b}}$.

By assumption, $A\mathbf{u} = \mathbf{b}$.

Write $\Sigma_{\mathbf{b},\mathbf{u}} = \{\mathbf{t} \in \mathbb{R}^n : \mathbf{t} = \mathbf{u} + \mathbf{h} \text{ for some } \mathbf{h} \in \mathcal{N}(A)\}$.

[We want to prove $\Pi_{\mathbf{b}} = \Sigma_{\mathbf{b},\mathbf{u}}$.]

- [We want to prove that '*every vector in* Π_b *belongs to* $\Sigma_{\mathbf{b},\mathbf{u}}$.'

  What do we mean to prove, really? '*For any vector* $\mathbf{w}$, *if* $\mathbf{w} \in \Pi_b$, *then* $\mathbf{w} \in \Sigma_{\mathbf{b},\mathbf{u}}$.']

  Pick any vector $\mathbf{w}$. Suppose $\mathbf{w} \in \Pi_{\mathbf{b}}$. [Ask: Is it true that $\mathbf{w} \in \Sigma_{\mathbf{b},\mathbf{u}}$?]

By definition, $A\mathbf{w} = \mathbf{b}$.

(Recall that $A\mathbf{u} = \mathbf{b}$ also.)

Take $\mathbf{h} = \mathbf{w} - \mathbf{u}$. Then, by definition, $A\mathbf{h} = \mathbf{0}$. Therefore $\mathbf{h} \in \mathcal{N}(A)$. Also by definition, $\mathbf{w} = \mathbf{u} + \mathbf{h}$.

Hence $\mathbf{w} \in \Sigma_{\mathbf{b},\mathbf{u}}$.

- [We want to prove that '*every vector in* $\Sigma_{\mathbf{b},\mathbf{u}}$ *belongs to* Π_b.'

  What do we mean to prove, really? '*For any vector* $\mathbf{w}$, *if* $\mathbf{w} \in \Sigma_{\mathbf{b},\mathbf{u}}$, *then* $\mathbf{w} \in \Pi_b$.']

  Pick any vector $\mathbf{w}$. Suppose $\mathbf{w} \in \Sigma_{\mathbf{b},\mathbf{u}}$. [Ask: Is it true that $\mathbf{w} \in \Pi_{\mathbf{b}}$?]

  By definition, $\mathbf{w} = \mathbf{u} + \mathbf{h}$ for some $\mathbf{h} \in \mathcal{N}(A)$.

  By definition, $A\mathbf{h} = \mathbf{0}$.

  (Recall that $A\mathbf{u} = \mathbf{b}$.)

  Then $A\mathbf{w} = A(\mathbf{u} + \mathbf{h}) = A\mathbf{u} + A\mathbf{h} = \mathbf{b} + \mathbf{0} = \mathbf{b}$.

  Therefore $\mathbf{w} \in \Pi_{\mathbf{b}}$.

It follows that the equality $\Pi_{\mathbf{b}} = \Sigma_{\mathbf{b},\mathbf{u}}$ holds.

9. **Illustration (4) of the idea in Theorem (1) and Theorem (2).**

Let $A = \begin{bmatrix} 1 & 2 & 0 & 1 \\ 1 & 1 & 1 & -1 \\ 3 & 1 & 5 & -7 \end{bmatrix}$, $\mathbf{b} = \begin{bmatrix} 7 \\ 3 \\ 1 \end{bmatrix}$.

Consider the system $\mathcal{LS}(A, \mathbf{b})$ and its associated homogeneous system $\mathcal{LS}(A, \mathbf{0})$.

(a) The respective augmented matrix representations of $\mathcal{LS}(A, \mathbf{b})$ and $\mathcal{LS}(A, \mathbf{0})$ are $C_{\mathbf{b}} = [\ A \mid \mathbf{b}\ ]$ and $C_{\mathbf{0}} = [\ A \mid \mathbf{0}\ ]$.

The reduced row-echelon forms $C'_{\mathbf{b}}, C'_{\mathbf{0}}$ which are row-equivalent to $C_{\mathbf{b}}$, $C_{\mathbf{0}}$ respectively are given by

$$C'_{\mathbf{b}} = \left[\begin{array}{cccc|c} 1 & 0 & 2 & -3 & -1 \\ 0 & 1 & -1 & 2 & 4 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right], \qquad C'_{\mathbf{0}} = \left[\begin{array}{cccc|c} 1 & 0 & 2 & -3 & 0 \\ 0 & 1 & -1 & 2 & 0 \\ 0 & 0 & 0 & 0 & 0 \end{array}\right]$$

Write $\mathbf{h}_1 = \begin{bmatrix} -2 \\ 1 \\ 1 \\ 0 \end{bmatrix}$, $\mathbf{h}_2 = \begin{bmatrix} 3 \\ -2 \\ 0 \\ 1 \end{bmatrix}$, $\mathbf{u} = \begin{bmatrix} -1 \\ 4 \\ 0 \\ 0 \end{bmatrix}$.

(b) The solution set of $\mathcal{LS}(A, \mathbf{b})$ is given by $\Pi_{\mathbf{b}} = \{\mathbf{u} + c_1\mathbf{h}_1 + c_2\mathbf{h}_2 \mid c_1, c_2 \in \mathbb{R}\}$.

Note that '$\mathbf{x} = \mathbf{u}$' is a (particular) solution of $\mathcal{LS}(A, \mathbf{b})$.

(c) The solution set of $\mathcal{LS}(A, \mathbf{0})$ (or in other words, the null space of A) is given by $\mathcal{N}(A) = \{c_1\mathbf{h}_1 + c_2\mathbf{h}_2 \mid c_1, c_2 \in \mathbb{R}\}$.

(d) The difference of any vector in $\Pi_{\mathbf{b}}$ from any vector in $\Pi_{\mathbf{b}}$ is the vector $\alpha_1\mathbf{h}_1 + \alpha_2\mathbf{h}_2$ for some numbers α_1, α_2. It belongs to $\mathcal{N}(A)$.

(e) The sum of any vector in $\Pi_{\mathbf{b}}$ and any vector in $\mathcal{N}(A)$ is the vector $\beta_1\mathbf{h}_1 + \beta_2\mathbf{h}_2$ for some numbers β_1, β_2. It belongs to $\Pi_{\mathbf{b}}$.

(f) The relation between $\Pi_{\mathbf{b}}$ and $\mathcal{N}(A)$ can be visualized geometrically as follows:—

- We can obtain $\Pi_{\mathbf{b}}$ from $\mathcal{N}(A)$ by applying to every point of $\mathcal{N}(A)$ a 'translation' by $\mathbf{u}$.

10. **Theorem (3).**

Suppose A is an $(m \times n)$-matrix. Then:

(1) *$\mathcal{LS}(A, \mathbf{0})$ is consistent, with '$\mathbf{x} = \mathbf{0}$' as a solution for the system. (Or equivalently, $\mathcal{N}(A)$ contains some vector in $\mathbb{R}^n$, namely $\mathbf{0}$.)*

(2) *The statements below are logically equivalent:*

  (2a) *$\mathbf{0}$ is the only vector in $\mathcal{N}(A)$.*

  (2b) *$\mathcal{LS}(A, \mathbf{0})$ has a unique solution.*

  (2c) *For each vector $\mathbf{b} \in \mathbb{R}^m$, if $\mathcal{LS}(A, \mathbf{b})$ is consistent then $\mathcal{LS}(A, \mathbf{b})$ has a unique solution.*

(3) *The statements below are logically equivalent:*

  (3a) *$\mathcal{N}(A)$ contains some vector in $\mathbb{R}^n$ other than $\mathbf{0}$.*

  (3b) *$\mathcal{LS}(A, \mathbf{0})$ has at least two solutions.*

  (3c) *For each vector $\mathbf{b} \in \mathbb{R}^m$, if $\mathcal{LS}(A, \mathbf{b})$ is consistent then $\mathcal{LS}(A, \mathbf{b})$ has at least two solutions.*

Proof. Exercise. (Apply Theorem (1).)