MATH1030 Span of vectors and column space of a matrix

1. **Definition. (Column space of a matrix.)**

   Let H be a $(p \times q)$-matrix.

   The column space of the matrix H is defined to be the set

   $$\left\{\mathbf{y} \in \mathbb{R}^p : \begin{array}{l} \textit{There exist some } \mathbf{u} \in \mathbb{R}^q \\ \textit{such that } \mathbf{y} = H\mathbf{u}. \end{array}\right\}$$

   We denote this set by $\mathcal{C}(H)$.

   Remark. We are applying the method of specification, with 'selection criterion'

   (∗) '*there exist some $\mathbf{u} \in \mathbb{R}^q$ such that $\mathbf{y} = H\mathbf{u}$.*'

   to form a certain set of vectors in $\mathbb{R}^p$, called the column space of the matrix H.

   When put into plain words, the selection criterion (∗) reads:

   '*y is a vector in $\mathbb{R}^p$ which can be expressed as the product of H in the left and some vector in $\mathbb{R}^q$ in the right.*'

   According to this 'selection criterion'

   - Those vectors in $\mathbb{R}^p$ resultant from multiplying H from the left to some vector in $\mathbb{R}^q$ are collected.
   - Those vectors in $\mathbb{R}^p$ not resultant from multiplying H from the left to some vector in $\mathbb{R}^q$ are 'discarded'.

   For this reason, $\mathcal{C}(H)$ is simply the collection of all vectors in $\mathbb{R}^p$ which can be 'expressed in the form' $H\mathbf{u}$, and only such vectors.

   So very often the set $\mathcal{C}(H)$ is given the short-hand $\{H\mathbf{u} \mid \mathbf{u} \in \mathbb{R}^q\}$.

   Further remark. How to use the various versions of the definitions?

   Always remember, whenever $\mathbf{v} \in \mathbb{R}^p$, the statements below mean the same thing:

   (a) $\mathbf{v} \in \mathcal{C}(H)$.

   (b) *There exists some $\mathbf{u} \in \mathbb{R}^q$ such that $\mathbf{v} = H\mathbf{u}$.*

2. **Theorem (1). (Column space of a matrix as a 'subspace'.)**

   Suppose H is a $(p \times q)$-matrix. Then the statements below hold.

   (a) $\mathbf{0}_p \in \mathcal{C}(H)$.

   (b) *For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$, if $\mathbf{x} \in \mathcal{C}(H)$ and $\mathbf{y} \in \mathcal{C}(H)$ then $\mathbf{x} + \mathbf{y} \in \mathcal{C}(H)$.*

   (c) *For any $\mathbf{x} \in \mathbb{R}^p$, for any $\alpha \in \mathbb{R}$, if $\mathbf{x} \in \mathcal{C}(H)$ then $\alpha\mathbf{x} \in \mathcal{C}(H)$.*

   (d) *For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$, for any $\alpha, \beta \in \mathbb{R}$, if $\mathbf{x} \in \mathcal{C}(H)$ and $\mathbf{y} \in \mathcal{C}(H)$ then $\alpha\mathbf{x} + \beta\mathbf{y} \in \mathcal{C}(H)$.*

3. **Proof of Theorem (1).**

   Suppose H is a $(p \times q)$-matrix.

   (a) Note that $\mathbf{0}_p = H\mathbf{0}_q$, and $\mathbf{0}_q \in \mathbb{R}^q$.
   Then $\mathbf{0}_p \in \mathcal{C}(H)$.

   (b) Pick any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^p$. Suppose $\mathbf{x}, \mathbf{y} \in \mathcal{C}(H)$.
   [Ask: What to verify? Answer: 'There exist some $\mathbf{w} \in \mathbb{R}^q$ such that $\mathbf{x} + \mathbf{y} = H\mathbf{w}$.'
   Further ask: How comes such a vector $\mathbf{w}$? Answer: Make use of the information provided by '$\mathbf{x} \in \mathcal{C}(H)$' and '$\mathbf{y} \in \mathcal{C}(H)$'.]
   By definition of $\mathcal{C}(H)$, there exist some $\mathbf{u}, \mathbf{v} \in \mathbb{R}^q$ such that $\mathbf{x} = H\mathbf{u}$ and $\mathbf{y} = H\mathbf{v}$.
   Now $\mathbf{x} + \mathbf{y} = H\mathbf{u} + H\mathbf{v} = H(\mathbf{u} + \mathbf{v})$. Since $\mathbf{u}, \mathbf{v} \in \mathbb{R}^q$, it happens that $\mathbf{u} + \mathbf{v} \in \mathbb{R}^q$.
   Then by the definition of $\mathcal{C}(H)$, $\mathbf{x} + \mathbf{y} \in \mathcal{C}(H)$.

   (c) Pick any $\mathbf{x} \in \mathbb{R}^p$. Pick any $\alpha \in \mathbb{R}$. Suppose $\mathbf{x} \in \mathcal{C}(H)$.
   [Ask: What to verify?
   Answer. 'There exist some $\mathbf{w} \in \mathbb{R}^q$ such that $\alpha\mathbf{x} = H\mathbf{w}$ ']
   By definition of $\mathcal{C}(H)$, there exist some $\mathbf{u} \in \mathbb{R}^q$ such that $\mathbf{x} = H\mathbf{u}$.
   Now $\alpha\mathbf{x} = \alpha H\mathbf{u} = H(\alpha\mathbf{u})$. Since $\mathbf{u}, \mathbf{v} \in \mathbb{R}^q$, it happens that $\alpha\mathbf{u} \in \mathbb{R}^q$.
   Then by the definition of $\mathcal{C}(H)$, $\alpha\mathbf{x} \in \mathcal{C}(H)$.

(d) Exercise.

4. An alternative way of visualizing the notion of *column space* is through the notions of *linear combination* and *span* (which will be introduced shortly).

   Recall the definition for the notion of *linear combination*:

   Let $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$ *be vectors in* $\mathbb{R}^m$.
   Let $\mathbf{w}$ *be a vector in* $\mathbb{R}^m$.
   We say $\mathbf{v}$ *is a linear combination of* $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$ *if the statement* (†) *holds:*

   (†) There exist some real numbers $\alpha_1, \alpha_2, \cdots, \alpha_n$ such that $\mathbf{w} = \alpha_1\mathbf{z}_1 + \alpha_2\mathbf{z}_2 + \cdots + \alpha_n\mathbf{z}_n$.

   The expression $\alpha_1\mathbf{z}_1 + \alpha_2\mathbf{z}_2 + \cdots + \alpha_n\mathbf{z}_n$ on its own is called the linear combination of the vectors $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$ and the scalars $\alpha_1, \alpha_2, \cdots, \alpha_n$.

5. **Definition. (Span of a set of vectors in $\mathbb{R}^m$.)**

   Let $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$ *be ('finitely many') vectors in* $\mathbb{R}^m$.

   The span of (the set of vectors) $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$ *is defined to be the set*

   $$\{\mathbf{y} \in \mathbb{R}^m : \ \mathbf{y} \textit{ is a linear combination of } \mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n \ \}$$

   We denote this set by $\mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n\})$ *(or* $\langle\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n\}\rangle$*).*

   Remark. $\mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n\})$ is constructed with the help of the method of specification, with 'selection criterion'

   (⋆) '$\mathbf{y}$ *is a linear combination of* $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$,'

   when we collect those and only those vectors in $\mathbb{R}^m$ which are linear combinations of $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$.

   For this reason, $\mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n\})$ is simply the collection of all vectors in $\mathbb{R}^m$ which can be 'expressed' as linear combinations of $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$, and only such vectors.

   Further remark. How to use the various versions of the definitions?

   Always remember, whenever $\mathbf{v} \in \mathbb{R}^m$, the statements below mean the same thing:

   (♯) $\mathbf{y}$ *belongs to* $\mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n\})$.

   (♮) $\mathbf{y}$ *is a linear combination of* $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$.

   (♭) *There exist some real numbers* $\alpha_1, \alpha_2, \cdots, \alpha_n$ *such that* $\mathbf{y} = \alpha_1\mathbf{z}_1 + \alpha_2\mathbf{z}_2 + \cdots + \alpha_n\mathbf{z}_n$.

   Further remark on terminologies and symbols.

   (a) In some textbooks, it is emphasized that the notion of *span* is defined on sets of vectors; hence the brackets '{', '}' are used in the notation.

   (b) For convenience, we may read '$\mathbf{y} \in \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n\})$' as '$\mathbf{y}$ *is spanned by* $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$'.

   When a set of vectors, say, V, is equal to the set $\mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n\})$, we may read this set equality as '*the set* V *is spanned by* $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$'.

6. With the help of Lemma (A) (from the handout *linear combinations*), we are going to set up a 'dictionary' between the notion of *span* and the notion of *column space.*

   Recall Lemma (A):

   Let A *be an* $(m \times n)$*-matrix, and* $\mathbf{t}$ *be a vector in* $\mathbb{R}^n$.

   Suppose that for each $j = 1, 2, \cdots, n$, *the* j*-th column of* A *is* $\mathbf{a}_j$ *and the* j*-th entry of* $\mathbf{t}$ *is* t_j. *(So* $A = [\ \mathbf{a}_1 \mid \mathbf{a}_2 \mid \cdots \mid \mathbf{a}_n\ ]$ *and* $\mathbf{t} = \begin{bmatrix} t_1 \\ t_2 \\ \vdots \\ t_n \end{bmatrix}$.*)*

   Then $A\mathbf{t} = t_1\mathbf{a}_1 + t_2\mathbf{a}_2 + \cdots + t_n\mathbf{a}_n$.

7. **Theorem (D). ('Dictionary' between the notion of span and the notion of column space.)**

   Let $\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_q$ *be vectors in* $\mathbb{R}^p$, *and* H *be a* $(p \times q)$*-matrix.*

   Suppose that the j*-th column of* H *is* $\mathbf{h}_j$ *for each* j. *(So* $H = [\ \mathbf{h}_1 \mid \mathbf{h}_2 \mid \cdots \mid \mathbf{h}_q\ ]$.*)*

   Then $\mathcal{C}(H) = \mathsf{Span}\ (\{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_q\})$.

Remark. The significance of Theorem (D) is that every statement about spans of collections of finitely many vectors can be translated into a statement about column spaces of matrices, and vice versa.

Further remark. The equality '$\mathcal{C}(H) = \mathsf{Span}\ (\{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_q\})$' is a set equality. What such an equality means is that the statements (†), (‡) below hold simultaneously:

(†) For any $\mathbf{y} \in \mathbb{R}^p$, if $\mathbf{y} \in \mathcal{C}(H)$ then $\mathbf{y} \in \mathsf{Span}\ (\{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_q\})$.

(‡) For any $\mathbf{y} \in \mathbb{R}^p$, if $\mathbf{y} \in \mathsf{Span}\ (\{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_q\})$ then $\mathbf{y} \in \mathcal{C}(H)$.

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

Let $\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_q$ be vectors in $\mathbb{R}^p$, and H be a $(p \times q)$-matrix.

Suppose that the j-th column of H is $\mathbf{h}_j$ for each j. Then $H = [\ \mathbf{h}_1 \mid \mathbf{h}_2 \mid \cdots \mid \mathbf{h}_q\ ]$.

[We verify the statements (†), (‡):

(†) For any $\mathbf{y} \in \mathbb{R}^p$, if $\mathbf{y} \in \mathcal{C}(H)$ then $\mathbf{y} \in \mathsf{Span}\ (\{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_q\})$.

(‡) For any $\mathbf{y} \in \mathbb{R}^p$, if $\mathbf{y} \in \mathsf{Span}\ (\{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_q\})$ then $\mathbf{y} \in \mathcal{C}(H)$.

The arguments are given in two separate paragraphs, one for (†) and the other (‡).]

- Pick any $\mathbf{y} \in \mathbb{R}^p$. Suppose $\mathbf{y} \in \mathcal{C}(H)$.
  Then there exists some $\mathbf{u} \in \mathbb{R}^q$ such that $\mathbf{y} = H\mathbf{u}$.
  For each i, denote the i-th entry of $\mathbf{u}$ by u_i.
  Then, by Lemma (A), $\mathbf{y} = u_1\mathbf{h}_1 + u_2\mathbf{h}_2 + \cdots + u_q\mathbf{h}_q$.
  Therefore $\mathbf{y} \in \mathsf{Span}\ (\{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_q\})$
- Pick any $\mathbf{y} \in \mathbb{R}^p$. Suppose $\mathbf{y} \in \mathsf{Span}\ (\{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_q\})$.
  Then there exists some $u_1, u_2, \cdots, u_q \in \mathbb{R}$ such that $\mathbf{y} = u_1\mathbf{h}_1 + u_2\mathbf{h}_2 + \cdots + u_q\mathbf{h}_q$.
  Define the vector $\mathbf{u}$ in $\mathbb{R}^q$ by $\mathbf{u} = \begin{bmatrix} u_1 \\ u_2 \\ \vdots \\ u_q \end{bmatrix}$.
  Then by Lemma (A), we have $\mathbf{y} = H\mathbf{u}$.
  Therefore $\mathbf{y} \in \mathcal{C}(H)$.

It follows that $\mathcal{C}(H) = \mathsf{Span}\ (\{\mathbf{h}_1, \mathbf{h}_2, \cdots, \mathbf{h}_q\})$ holds.

9. **Illustrations of the content of Theorem (D).**

(a) $\mathcal{C}\left(\begin{bmatrix} 1 & 2 & 3 & 4 \\ 1 & 3 & 5 & 7 \\ 1 & 4 & 7 & 10 \end{bmatrix}\right) = \mathsf{Span}\ \left(\left\{\begin{bmatrix} 1 \\ 1 \\ 1 \end{bmatrix}, \begin{bmatrix} 2 \\ 3 \\ 4 \end{bmatrix}, \begin{bmatrix} 3 \\ 5 \\ 7 \end{bmatrix}, \begin{bmatrix} 4 \\ 7 \\ 10 \end{bmatrix}\right\}\right)$

(b) $\mathcal{C}\left(\begin{bmatrix} 1 & 0 & 9 \\ 0 & 2 & 8 \\ 1 & 4 & 7 \\ 0 & 6 & 6 \\ 1 & 8 & 5 \end{bmatrix}\right) = \mathsf{Span}\ \left(\left\{\begin{bmatrix} 1 \\ 0 \\ 1 \\ 0 \\ 1 \end{bmatrix}, \begin{bmatrix} 0 \\ 2 \\ 4 \\ 6 \\ 8 \end{bmatrix}, \begin{bmatrix} 9 \\ 8 \\ 7 \\ 6 \\ 5 \end{bmatrix}\right\}\right)$

10. **Theorem (2). (Span of vectors as a 'subspace'.)**

Suppose $\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n$ are vectors in $\mathbb{R}^m$. Write $V = \mathsf{Span}\ (\{\mathbf{z}_1, \mathbf{z}_2, \cdots, \mathbf{z}_n\})$.

The statements below hold:

(a) $\mathbf{0} \in V$.

(b) *For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, if $\mathbf{x} \in V$ and $\mathbf{y} \in V$ then $\mathbf{x} + \mathbf{y} \in V$.*

(c) *For any $\mathbf{x} \in \mathbb{R}^m$, for any $\alpha \in \mathbb{R}$, if $\mathbf{x} \in V$ then $\alpha\mathbf{x} \in V$.*

(d) *For any $\mathbf{x}, \mathbf{y} \in \mathbb{R}^m$, for any $\alpha, \beta \in \mathbb{R}$, if $\mathbf{x} \in V$ and $\mathbf{y} \in V$ then $\alpha\mathbf{x} + \beta\mathbf{y} \in V$.*

Proof of Theorem (2). This is a consequence of Theorem (1) and Theorem (D).