MATH1030 How bases for the same subspace of $\mathbb{R}^n$ relate to each other.

1. We know that each non-zero subspace of $\mathbb{R}^n$ has many bases. It is a natural question to ask how the bases for the same non-zero subspace of $\mathbb{R}^n$ are related to each other.

   As far as the subspace concerned is $\mathbb{R}^n$ itself, the answer to this question is simple.

   Theorem (1).

   Let $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_n, \mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ be vectors in $\mathbb{R}^n$. Suppose that $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_n$ constitute a basis for $\mathbb{R}^n$, and that $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ constitute a basis for $\mathbb{R}^n$.

   Define the $(n \times n)$-square matrices T, U by $T = \left[\begin{array}{c|c|c|c} \mathbf{t}_1 & \mathbf{t}_2 & \cdots & \mathbf{t}_n \end{array}\right]$, $U = \left[\begin{array}{c|c|c|c} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_n \end{array}\right]$.

   Then the statements below hold:

   (a) *T, U are invertible.*

   (b) *Denote by g_{ij} the (i,j)-th entry of the $(n \times n)$-square matrix $U^{-1}T$ by g_{ij}.*
   For any $k = 1, 2, \cdots, n$, the equality $\mathbf{t}_k = g_{1k}\mathbf{u}_1 + g_{2k}\mathbf{u}_2 + \cdots + g_{nk}\mathbf{u}_n$ holds.

   The proof of Theorem (1) is straightforward and is left as an exercise. The matrix $G = U^{-1}T$, which, when being multiplied from the left to the individual vectors in the basis $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$, 'transforms' this basis to the basis $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_n$, is usually referred to as the matrix of change-of-basis from $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_n$ to $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_n$.

2. But what can be said in the situation of general non-zero subspaces of $\mathbb{R}^n$?

   To answer this question, we shall have to make use of some general results concerned with relations amongst the null spaces and the column spaces of individual matrices and their products. They are Theorem (2), Theorem (3) below. To formulate these results, we need to recall the definition for the notions of *intersections of subspaces of* $\mathbb{R}^n$ and *sums of subspaces of* $\mathbb{R}^n$:

   Suppose V, W are subspaces of $\mathbb{R}^n$. Then:

   - *the intersection of V, W, denoted by $V \cap W$, is defined to be the collection of vectors in $\mathbb{R}^n$ given by*
   $$V \cap W = \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x} \in V \text{ and } \mathbf{x} \in W\},$$
   and
   - *the sum of V, W, denoted by $V + W$, is defined to be the collection of vectors in $\mathbb{R}^n$ given by*
   $$V + W = \left\{ \mathbf{x} \in \mathbb{R}^n : \begin{array}{l} \text{There exist some } \mathbf{s} \in V \text{ and } \mathbf{t} \in W \\ \text{such that } \mathbf{x} = \mathbf{s} + \mathbf{t}. \end{array} \right\}.$$

3. Now recall the result below from *Rank-nullity Formula*:

   ($\star$) *Suppose A is a $(p \times q)$-matrix, and B is an $(q \times s)$-matrix. Then $\mathcal{N}(B)$ is a subspace of $\mathcal{N}(AB)$.*

   We incorporate the result ($\star$) into something more comprehensive about null spaces of matrices:

   Theorem (2).

   Suppose A is an $(p \times q)$-matrix, and B is an $(q \times s)$-matrix. Then $\mathcal{N}(B)$ is a subspace of $\mathcal{N}(AB)$. Moreover, the statements below hold:

   (a) *Suppose $\mathcal{N}(AB)$ is the zero subspace of $\mathbb{R}^s$. Then $\mathcal{N}(B)$ is the zero subspace of $\mathbb{R}^s$.*

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

   i. *$\mathcal{N}(A) \cap \mathcal{C}(B)$ is the zero subspace of $\mathbb{R}^q$.*

   ii. *$\mathcal{N}(AB)$ is a subspace of $\mathcal{N}(B)$.*

   iii. *$\mathcal{N}(AB) = \mathcal{N}(B)$.*

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

   Suppose A is an $(p \times q)$-matrix, and B is an $(q \times s)$-matrix.

   We have already known that $\mathcal{N}(B)$ is a subspace of $\mathcal{N}(AB)$.

   (a) Suppose $\mathcal{N}(AB)$ is the zero subspace of $\mathbb{R}^s$.

   [We verify that the only vector in $\mathcal{N}(B)$ is the zero vector in $\mathbb{R}^s$.]

Pick any $\mathbf{x} \in \mathbb{R}^s$. Suppose $\mathbf{x} \in \mathcal{N}(B)$. Then $\mathbf{x} \in \mathcal{N}(AB)$. Now since $\mathcal{N}(AB) = \{\mathbf{0}_s\}$, we have $\mathbf{x} = \mathbf{0}_s$.

It follows that $\mathcal{N}(B)$ is the zero subspace of $\mathbb{R}^s$.

(b)
- Suppose $\mathcal{N}(A) \cap \mathcal{C}(B)$ is the zero subspace of $\mathbb{R}^q$.

  [We verify that $\mathcal{N}(AB)$ is a subspace of $\mathcal{N}(B)$. This is done by verify that for any $\mathbf{x} \in \mathbb{R}^s$, if $\mathbf{x} \in \mathcal{N}(AB)$ then $\mathbf{x} \in \mathcal{N}(B)$.]

  Pick any $\mathbf{x} \in \mathbb{R}^s$. Suppose $\mathbf{x} \in \mathcal{N}(AB)$. Then we have $A(B\mathbf{x}) = (AB)\mathbf{x} = \mathbf{0}_m$. Therefore $B\mathbf{x} \in \mathcal{N}(A)$.
  Now note that by definition, $B\mathbf{x} \in \mathcal{C}(B)$. Then $B\mathbf{x} \in \mathcal{N}(A) \cap \mathcal{C}(B)$.
  Since $\mathcal{N}(A) \cap \mathcal{C}(B)$ is the zero subspace of $\mathbb{R}^q$, we have $B\mathbf{x} = \mathbf{0}_q$.
  Then $\mathbf{x} \in \mathcal{N}(B)$.
  It follows that $\mathcal{N}(AB)$ is a subspace of $\mathcal{N}(B)$.
- Suppose $\mathcal{N}(AB)$ is a subspace of $\mathcal{N}(B)$. Then every vector in $\mathcal{N}(AB)$ belongs to $\mathcal{N}(B)$.
  Recall that $\mathcal{N}(B)$ is (definitely) a subspace of $\mathcal{N}(AB)$. Then every vector in $\mathcal{N}(B)$ belongs to $\mathcal{N}(AB)$.
  It follows that $\mathcal{N}(AB) = \mathcal{N}(B)$.
- Suppose $\mathcal{N}(AB) = \mathcal{N}(B)$.

  [We verify that $\mathcal{N}(A) \cap \mathcal{C}(B)$ is the zero subspace of $\mathbb{R}^q$. This is done by verify that for any $\mathbf{y} \in \mathbb{R}^q$, if $\mathbf{y} \in \mathcal{N}(A) \cap \mathcal{C}(B)$ then $\mathbf{y} = \mathbf{0}_q$.]

  Pick any $\mathbf{y} \in \mathcal{N}(A) \cap \mathcal{C}(B)$. We have $\mathbf{y} \in \mathcal{N}(A)$ and $\mathbf{y} \in \mathcal{C}(B)$.
  Since $\mathbf{y} \in \mathcal{N}(A)$, we have $A\mathbf{y} = \mathbf{0}_m$.
  Since $\mathbf{y} \in \mathcal{C}(B)$, there exists some $\mathbf{x} \in \mathbb{R}^s$ such that $\mathbf{y} = B\mathbf{x}$.
  Now we have $(AB)\mathbf{x} = A(B\mathbf{x}) = A\mathbf{y} = \mathbf{0}_m$. Then $\mathbf{x} \in \mathcal{N}(AB)$.
  Since $\mathcal{N}(AB) = \mathcal{N}(B)$, we have $\mathbf{x} \in \mathcal{N}(B)$. Then $\mathbf{y} = B\mathbf{x} = \mathbf{0}_q$.
  It follows that $\mathcal{N}(A) \cap \mathcal{C}(B)$ is the zero subspace of $\mathbb{R}^q$.

5. **Theorem (3).**

   Suppose A is an $(p \times q)$-matrix, and B is an $(q \times s)$-matrix. Then $\mathcal{C}(AB)$ is a subspace of $\mathcal{C}(A)$. Moreover, the statements below hold:

   (a) *Suppose $\mathcal{C}(AB) = \mathbb{R}^p$. Then $\mathcal{C}(A) = \mathbb{R}^p$.*

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

      i. $\mathcal{N}(A) + \mathcal{C}(B) = \mathbb{R}^q$.

      ii. $\mathcal{C}(A)$ *is a subspace of* $\mathcal{C}(AB)$.

      iii. $\mathcal{C}(AB) = \mathcal{C}(A)$.

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

   Suppose A is an $(p \times q)$-matrix, and B is an $(q \times s)$-matrix.

   Note that $\mathcal{C}(AB)$, $\mathcal{C}(A)$ are both subspaces of $\mathbb{R}^p$.

   [We verify that for any $\mathbf{z} \in \mathbb{R}^p$, if $\mathbf{z} \in \mathcal{C}(AB)$ then $\mathbf{z} \in \mathcal{C}(A)$.]

   Pick any $\mathbf{z} \in \mathcal{C}(AB)$. There exists some $\mathbf{x} \in \mathbb{R}^s$ such that $\mathbf{z} = (AB)\mathbf{x}$.

   For the same $\mathbf{x}$, we have $\mathbf{z} = A(B\mathbf{x})$ and $B\mathbf{x} \in \mathbb{R}^q$. Then $\mathbf{z} \in \mathcal{C}(A)$.

   It follows that $\mathcal{C}(AB)$ is a subspace of $\mathcal{C}(A)$.

   (a) Suppose $\mathcal{C}(AB) = \mathbb{R}^p$.

      [We verify that every vector in $\mathbb{R}^p$ belongs to $\mathcal{C}(A)$.]

      Pick any $\mathbf{z} \in \mathbb{R}^p$. By assumption, we have $\mathbf{z} \in \mathcal{C}(AB)$. Since $\mathcal{C}(AB)$ is a subspace of $\mathcal{C}(A)$, we have $\mathbf{z} \in \mathcal{C}(A)$.

      It follows that $\mathcal{C}(A) = \mathbb{R}^p$.

   (b)
   - Suppose $\mathcal{N}(A) + \mathcal{C}(B) = \mathbb{R}^q$.

     [We verify that $\mathcal{C}(A)$ is a subspace of $\mathcal{C}(AB)$. This is done by verifying that for any $\mathbf{z} \in \mathbb{R}^p$, if $\mathbf{z} \in \mathcal{C}(A)$ then $\mathbf{z} \in \mathcal{C}(AB)$.]

Pick any $\mathbf{z} \in \mathbb{R}^p$. Suppose $\mathbf{z} \in \mathcal{C}(A)$. Then there exists some $\mathbf{y} \in \mathbb{R}^q$ such that $\mathbf{z} = A\mathbf{y}$.
Since $\mathcal{N}(A) + \mathcal{C}(B) = \mathbb{R}^q$, there exist some $\mathbf{y}' \in \mathcal{N}(A)$, $\mathbf{y}'' \in \mathcal{C}(B)$ such that $\mathbf{y} = \mathbf{y}' + \mathbf{y}''$. Since $\mathbf{y}'' \in \mathcal{C}(B)$, there exists some $\mathbf{x} \in \mathbb{R}^s$ such that $\mathbf{y}'' = B\mathbf{x}$.
Now we have $\mathbf{z} = A\mathbf{y} = A(\mathbf{y}' + \mathbf{y}'') = A\mathbf{y}' + A\mathbf{y}'' = \mathbf{0}_p + A(B\mathbf{x}) = (AB)\mathbf{x}$. Then $\mathbf{z} \in \mathcal{C}(AB)$.
It follows that $\mathcal{C}(A)$ is a subspace of $\mathcal{C}(AB)$.

- Suppose $\mathcal{C}(A)$ is a subspace of $\mathcal{C}(AB)$. Then every vector of $\mathcal{C}(A)$ belongs to $\mathcal{C}(AB)$.
  Recall that $\mathcal{C}(AB)$ is a subspace of $\mathcal{C}(A)$. Then every vector of $\mathcal{C}(AB)$ belongs to $\mathcal{C}(A)$.
  It follows that $\mathcal{C}(AB) = \mathcal{C}(A)$.
- Suppose $\mathcal{C}(AB) = \mathcal{C}(A)$.

  [We verify that $\mathcal{N}(A) + \mathcal{C}(B) = \mathbb{R}^q$. This is done by verifying that for any $\mathbf{y} \in \mathbb{R}^q$, there exist some $\mathbf{y}' \in \mathcal{N}(A)$, $\mathbf{y}'' \in \mathcal{C}(B)$ such that $\mathbf{y} = \mathbf{y}' + \mathbf{y}''$.]

  Pick any $\mathbf{y} \in \mathbb{R}^q$. We have $A\mathbf{y} \in \mathcal{C}(A)$.
  Since $\mathcal{C}(AB) = \mathcal{C}(A)$, we have $\mathbf{y} \in \mathcal{C}(AB)$. Then there exists some $\mathbf{x} \in \mathbb{R}^s$ such that $A\mathbf{y} = (AB)\mathbf{x}$.
  Take $\mathbf{y}' = \mathbf{y} - B\mathbf{x}$ and $\mathbf{y}'' = B\mathbf{x}$.
  By definition, $\mathbf{y}'' \in \mathcal{C}(B)$.
  Also by definition, $\mathbf{y}' \in \mathbb{R}^q$ and $A\mathbf{y}' = A(\mathbf{y} - B\mathbf{x}) = A\mathbf{y} - AB\mathbf{x} = \mathbf{0}_q$. Then $\mathbf{y}' \in \mathcal{N}(A)$.
  By construction, $\mathbf{y} = \mathbf{y}' + B\mathbf{x} = \mathbf{y}' + \mathbf{y}''$.
  It follows that $\mathcal{N}(A) + \mathcal{C}(B) = \mathbb{R}^q$.

7. An immediate consequence of Theorem (2) and Theorem (3) is Theorem (4) below. Towards the end of Theorem (4), we have the complete answer for the question how two bases of the same subspace of $\mathbb{R}^n$ are related to each other.

   Theorem (4).

   Let $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_s, \mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ be vectors in $\mathbb{R}^p$, T be the $(p \times s)$-matrix defined by $T = \left[\begin{array}{c|c|c|c} \mathbf{t}_1 & \mathbf{t}_2 & \cdots & \mathbf{t}_s \end{array}\right]$, and U be the $(p \times q)$-matrix defined by $U = \left[\begin{array}{c|c|c|c} \mathbf{u}_1 & \mathbf{u}_2 & \cdots & \mathbf{u}_q \end{array}\right]$.

   Suppose that for each $k = 1, 2, \cdots, s$, there exist some $g_{1k}, g_{2k}, \cdots, g_{qk}$ such that $\mathbf{t}_k = g_{1k}\mathbf{u}_1 + g_{2k}\mathbf{u}_2 + \cdots + g_{qk}\mathbf{u}_q$.

   For each k, define $\mathbf{g}_k = \begin{bmatrix} g_{1k} \\ g_{2k} \\ \vdots \\ g_{qk} \end{bmatrix}$.

   Define the $(q \times s)$-matrix G by $G = \left[\begin{array}{c|c|c|c} \mathbf{g}_1 & \mathbf{g}_2 & \cdots & \mathbf{g}_s \end{array}\right]$.

   Then $T = UG$.

   Moreover, the statements below hold:

   (a) *Suppose $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_s$ are linearly independent.*
   Then $\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_s$ are linearly independent.

   (b) *Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ are linearly independent.*
   Then $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_s$ are linearly independent if and only if $\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_s$ are linearly independent.

   (c) $\mathsf{Span}\ (\{\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_s\})$ *is a subspace of* $\mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$.

   (d) *Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ are linearly independent.*
   Then $\mathsf{Span}\ (\{\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_s\}) = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$ *if and only if* $\mathsf{Span}\ (\{\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_s\}) = \mathbb{R}^q$.

   Write $W = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q\})$. Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ constitute a basis for W. Now further suppose $q = s$. (So $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ are linearly independent, and G is a $(q \times q)$-square matrix.)

   Then the statements below are logically equivalent:

   (a) *$\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_q$ constitute a basis for W.*

   (b) *$\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_q$ are linearly independent.*

   (c) $\mathsf{Span}\ (\{\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_q\}) = \mathbb{R}^q$.

   (d) *$\mathbf{g}_1, \mathbf{g}_2, \cdots, \mathbf{g}_q$ constitute a basis for $\mathbb{R}^q$.*

   (e) *G is invertible.*

Now suppose G is invertible indeed, with matrix inverse G^{-1}, whose (i, j)-th entry is denoted by h_{ij} for each i, j. Then for each $k = 1, 2, \cdots, k$, the equality $\mathbf{u}_k = h_{1k}\mathbf{t}_1 + h_{2k}\mathbf{t}_2 + \cdots + h_{qk}\mathbf{t}_q$ holds.

Remark. Where $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_q$ constitute a basis for W and $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ also constitute a basis for W, the invertible $(q \times q)$-square matrix G which, when being multiplied from the left to the individual vectors in the basis $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$, 'transforms' this basis to the basis $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_q$, is usually referred to as the matrix of change-of-basis from $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$ to $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_q$.