0. Here we going to give another proof for the Replacement Theorem (which is Theorem (F) in the handout *More on minimal spanning set.*)

   It will be a direct argument for the Replacement Theorem: we do not have to rely on what we have learnt about reduced row-echelon form and non-singular matrices. All we need will be the definitions for the notions of *linear combinations*, *span*, *linear independence*, and *bases*.

1. **Lemma (1). (Baby version of Replacement Theorem.)**

   Let V be a subspace of $\mathbb{R}^n$. Suppose $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_k$ constitute a basis for V.

   Let $\mathbf{u}$ be a non-zero vector in $\mathbb{R}^n$. Suppose $\mathbf{u}$ is a linear combination of $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_k$.

   Then, after relabelling the indices of $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_k$ if necessary, $\mathbf{u}, \mathbf{t}_2, \cdots, \mathbf{t}_k$ constitute a basis for V.

2. **Proof of Lemma (1).**

   Let V be a subspace of $\mathbb{R}^n$. Suppose $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_k$ constitute a basis for V.

   Let $\mathbf{u}$ be a non-zero vector in $\mathbb{R}^n$. Suppose each of $\mathbf{u}$ is a linear combination of $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_k$.

   By assumption, there exist some $\alpha_1, \alpha_2, \cdots, \alpha_k$ such that $\mathbf{u} = \alpha_1\mathbf{t}_1 + \alpha_2\mathbf{t}_2 + \cdots + \alpha_k\mathbf{t}_k$.

   By assumption $\mathbf{u} \neq \mathbf{0}$. Then at least one of $\alpha_1, \alpha_2, \cdots, \alpha_k$ is non-zero.

   Without loss of generality, suppose $\alpha_1 \neq 0$. (Otherwise, choose the first i for which α_i is non-zero. Then relabel $\alpha_1, \mathbf{t}_1$ as $\alpha_i, \mathbf{t}_i$ respectively, and $\alpha_i, \mathbf{t}_i$ as $\alpha_1, \mathbf{t}_1$ respectively.)

   We verify that $\mathbf{u}, \mathbf{t}_2, \mathbf{t}_3, \cdots, \mathbf{t}_k$ constitute a basis for V:

   - Pick any $\beta, \gamma_2, \gamma_3, \cdots, \gamma_k \in \mathbb{R}$.
     Suppose $\beta\mathbf{u} + \gamma_2\mathbf{t}_2 + \gamma_3\mathbf{t}_3 + \cdots + \gamma_k\mathbf{t}_k = \mathbf{0}$.
     Then

     $$\begin{aligned} \mathbf{0} &= \beta(\alpha_1\mathbf{t}_1 + \alpha_2\mathbf{t}_2 + \cdots + \alpha_k\mathbf{t}_k) + \gamma_2\mathbf{t}_2 + \gamma_3\mathbf{t}_3 + \cdots + \gamma_k\mathbf{t}_k \\ &= \beta\alpha_1\mathbf{t}_1 + (\beta\alpha_2 + \gamma_2)\mathbf{t}_2 + (\beta\alpha_3 + \gamma_3)\mathbf{t}_3 + \cdots + (\beta\alpha_k + \gamma_k)\mathbf{t}_k \end{aligned}$$

     By assumption $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_k$ are linearly independent. Then

     $$\beta\alpha_1 = 0 = \beta\alpha_2 + \gamma_2 = \beta\alpha_3 + \gamma_3 = \cdots = \beta\alpha_k + \gamma_k.$$

     Recall that $\alpha_1 \neq 0$. Then $\beta = 0$. Therefore $\gamma_2 = \gamma_3 = \cdots = \gamma_k = 0$.
     It follows that $\mathbf{u}, \mathbf{t}_2, \mathbf{t}_3, \cdots, \mathbf{t}_k$ are linearly independent.
   - By definition, $V = \mathsf{Span}\ (\{\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3 \cdots, \mathbf{t}_k\})$.
     By assumption $\mathbf{u}$ is a linear combination of $\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3 \cdots, \cdots, \mathbf{t}_k$. Then $V = \mathsf{Span}\ (\{\mathbf{t}_1, \mathbf{t}_2, \mathbf{t}_3, \cdots, \mathbf{t}_k, \mathbf{u}\})$.
     Note that $\mathbf{t}_1 = \dfrac{1}{\alpha_1}\mathbf{u} - \dfrac{\alpha_2}{\alpha_1}\mathbf{t}_2 - \dfrac{\alpha_3}{\alpha_1}\mathbf{t}_3 - \cdots - \dfrac{\alpha_k}{\alpha_1}\mathbf{t}_k$.
     Then $\mathbf{t}_1$ is a linear combination of $\mathbf{u}, \mathbf{t}_2, \mathbf{t}_3, \cdots, \mathbf{t}_k$.
     Therefore $V = \mathsf{Span}\ (\{\mathbf{u}, \mathbf{t}_2, \mathbf{t}_3, \cdots, \mathbf{t}_k\})$.

   It follows that $\mathbf{u}, \mathbf{t}_2, \mathbf{t}_3, \cdots, \mathbf{t}_k$ constitute a basis for V.

3. **Theorem (2). (Replacement Theorem.)**

   Let V be a subspace of $\mathbb{R}^n$. Suppose $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_p, \mathbf{t}_{p+1}, \cdots, \mathbf{t}_{p+s}$ constitute a basis for V.

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ be vectors in V. Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ are linearly independent.

   Then, after relabelling the indices of $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_p, \mathbf{t}_{p+1}, \cdots, \mathbf{t}_{p+s}$ if necessary, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p, \mathbf{t}_{p+1}, \cdots, \mathbf{t}_{p+s}$ constitute a basis for V.

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

   Let V be a subspace of $\mathbb{R}^n$. Suppose $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_p, \mathbf{t}_{p+1}, \cdots, \mathbf{t}_{p+s}$ constitute a basis for V.

   Let $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ be vectors in V. Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ are linearly independent.

   - $\mathbf{u}_1$ is a linear combination of $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_{p+s}$. Moreover, $\mathbf{u}_1 \neq \mathbf{0}$. (Why?)
     We apply Lemma (1): after relabelling the indices of $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_{p+s}$ if necessary, $\mathbf{u}_1, \mathbf{t}_2, \mathbf{t}_3, \cdots, \mathbf{t}_p, \mathbf{t}_{p+1} \cdots, \mathbf{t}_{p+s}$ constitute a basis for V.

- Suppose $1 \leq j < p$, and suppose that after relabelling the indices of $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_{p+s}$ if necessary, $\mathbf{u}_1, \cdots, \mathbf{u}_j, \mathbf{t}_{j+1}, \cdots, \mathbf{t}_{p+s}$ constitute a basis for V.

  $\mathbf{u}_{j+1}$ is a linear combination of $\mathbf{u}_1, \cdots, \mathbf{u}_j, \mathbf{t}_{j+1}, \mathbf{t}_{j+2}, \cdots, \mathbf{t}_{p+s}$.

  Then there exist some $\kappa_1, \cdots, \kappa_j, \lambda, \mu_{j+2}, \cdots, \cdots, \mu_{p+s} \in \mathbb{R}$ such that $\mathbf{u}_{j+1} = \kappa_1\mathbf{u}_1 + \cdots + \kappa_j\mathbf{u}_j + \lambda\mathbf{t}_{j+1} + \mu_{j+2}\mathbf{t}_{j+2} + \cdots + \mu_{p+s}\mathbf{t}_{p+s}$.

  $\lambda, \mu_{j+2}, \cdots, \cdots, \mu_{p+s}$ are not all zero. (Otherwise, $\mathbf{u}_{j+1} = \kappa_1\mathbf{u}_1 + \kappa_2\mathbf{u}_2 + \cdots + \kappa_j\mathbf{u}_j$. Then $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ would be linearly dependent.)

  Without loss of generality, suppose $\lambda \neq 0$.

  We verify that $\mathbf{u}_1, \cdots, \mathbf{u}_j, \mathbf{u}_{j+1}, \mathbf{t}_{j+2} \cdots, \mathbf{t}_{p+s}$ constitute a basis for V:

  * Pick any $\alpha_1, \cdots, \alpha_j, \beta, \gamma_{j+2}, \cdots, \gamma_{p+s} \in \mathbb{R}$.
    Suppose $\alpha_1\mathbf{u}_1 + \cdots + \alpha_j\mathbf{u}_j + \beta\mathbf{u}_{j+1} + \gamma_{j+2}\mathbf{t}_{j+2} + \cdots + \gamma_{p+s}\mathbf{t}_{p+s} = \mathbf{0}$.
    Then

    $$\begin{aligned}\mathbf{0} &= \alpha_1\mathbf{u}_1 + \cdots + \alpha_j\mathbf{u}_j \\ &\quad +\beta(\kappa_1\mathbf{u}_1 + \cdots + \kappa_j\mathbf{u}_j + \lambda\mathbf{t}_{j+1} + \mu_{j+2}\mathbf{t}_{j+2} + \cdots + \mu_{p+s}\mathbf{t}_{p+s}) \\ &\quad +\gamma_{j+2}\mathbf{t}_{j+2} + \cdots + \gamma_{p+s}\mathbf{t}_{p+s} \\ &= (\beta\kappa_1 + \alpha_1)\mathbf{t}_1 + \cdots + (\beta\kappa_j + \alpha_j)\mathbf{t}_j + \beta\lambda\mathbf{t}_{j+1} + (\beta\mu_{j+2} + \gamma_{j+2})\mathbf{t}_{j+2} + \cdots + (\beta\mu_{p+s} + \gamma_{p+s})\mathbf{t}_{p+s}\end{aligned}$$

    Note that $\mathbf{u}_1, \cdots, \mathbf{u}_j, \mathbf{t}_{j+1}, \cdots, \mathbf{t}_{p+s}$ are linearly independent. Then

    $$\beta\lambda = 0 = \beta\kappa_1 + \alpha_1 = \cdots = \beta\kappa_j + \alpha_j = \beta\mu_{j+2} + \gamma_{j+2} = \cdots = \beta\mu_{p+s} + \gamma_{p+s}.$$

    Recall that $\lambda \neq 0$. Then $\beta = 0$. Therefore $\alpha_1 = \cdots = \alpha_j = \gamma_{j+2} = \cdots = \gamma_{p+s} = 0$.
    It follows that $\mathbf{u}_1, \cdots, \mathbf{u}_j, \mathbf{u}_{j+1}, \mathbf{t}_{j+2}, \cdots, \mathbf{t}_{p+s}$ are linearly independent.
  * Note that $V = \mathsf{Span}\ (\{\mathbf{u}_1, \cdots, \mathbf{u}_j, \mathbf{t}_{j+1}, \cdots, \mathbf{t}_{p+s}\})$, and $\mathbf{u}_{j+1}$ is a linear combination of $\mathbf{u}_1, \cdots, \mathbf{u}_j, \mathbf{t}_{j+1} \cdots, \mathbf{t}_{p+s}$.
    Then $V = \mathsf{Span}\ (\{\mathbf{u}_1, \cdots, \mathbf{u}_j, \mathbf{t}_{j+1}, \cdots, \mathbf{t}_{p+s}, \mathbf{u}_{j+1}\})$.
    Note that $\mathbf{t}_{j+1} = \dfrac{1}{\lambda}\mathbf{u}_{j+1} - \dfrac{\kappa_1}{\lambda}\mathbf{u}_1 - \cdots - \dfrac{\kappa_j}{\lambda}\mathbf{u}_j - \dfrac{\mu_{j+2}}{\lambda}\mathbf{t}_{j+2} - \cdots - \dfrac{\mu_{p+s}}{\lambda}\mathbf{t}_{p+s}$.
    Then $\mathbf{t}_{j+1}$ is a linear combination of $\mathbf{u}_1, \cdots, \mathbf{u}_j, \mathbf{u}_{j+1}, \mathbf{t}_{j+2}, \cdots, \mathbf{t}_{p+s}$.
    Then $V = \mathsf{Span}\ (\{\mathbf{u}_1, \cdots, \mathbf{u}_j, \mathbf{u}_{j+1}, \mathbf{t}_{j+2}, \cdots, \mathbf{t}_{p+s}\})$.

  It follows that $\mathbf{u}_1, \cdots, \mathbf{u}_j, \mathbf{u}_{j+1}, \mathbf{t}_{j+2} \cdots, \mathbf{t}_{p+s}$ constitute a basis for V.

Hence inductively, we deduce that, after relabelling the indices of $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_p, \mathbf{t}_{p+1}, \cdots, t_{p+s}$ if necessary, $\mathbf{u}_1, \mathbf{u}_2, \mathbf{u}_3, \mathbf{u}_4, \mathbf{u}_5, \cdots, \mathbf{u}_p, \mathbf{t}_{p+1}, \cdots, \mathbf{t}_{p+s}$ constitute a basis for V.

5. **Comments on the arguments for the Replacement Theorem.**

   This is for students (such as MATH/BMED students) who need to take MATH2040.

   Refer also to Theorem (E) in the handout *Minimal spanning set*, and Theorem (F) in the handout *More on minimal spanning set*.

   (a) The argument above for the Replacement Theorem is 'general' in the sense that we do not rely on the specific nature of the vectors in $\mathbb{R}^n$ as 'column vectors with n entries'.
   For this reason, this argument can be adapted (with minimal change) to give a proof for the 'Replacement Theorem' to *abstract linear algebra* (in which the objects of study are no longer simply vectors, matrices, and systems of linear equations.)

   (b) The Replacement Theorem in this course can be seen as a consequence of Theorem (E) and the considerations on 'sums of subspaces of $\mathbb{R}^n$' leading towards Theorem (F) in the handout *More on minimal spanning set*. The argument there relies heavily on the specific nature of the vectors in $\mathbb{R}^n$ as 'column vectors with n entries'. (We need to form matrices with these vectors as various columns to do various manipulations.)

   For this reason, that argument cannot be immediately and directly adapted to *abstract linear algebra*. However, it provides an easy method of calculations in the context where vectors in $\mathbb{R}^n$ is involved.

   The comments above also apply to what we are going to do next: we give a direct argument for Theorem (B) below, as an application of the Replacement Theorem. In contrast to what has been done in the handout *More on minimal spanning set*, our argument here can be adapted immediately to *abstract linear algebra*.

6. **Theorem (B).**

   Any two bases for a subspace of $\mathbb{R}^n$ have the same number of vectors.

7. **Proof of Theorem (B).**

Let V be a subspace of $\mathbb{R}^n$. When V is the zero subspace of $\mathbb{R}^n$, the empty set is its one and only one basis, and there is nothing to prove here.

From now on we suppose V is not the zero subspace of $\mathbb{R}^n$.

Suppose $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p$ is a basis for V.

Also suppose $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_p'$ is a basis for V.

Further suppose it were true that $p \neq p'$. Without loss of generality, assume $p < p'$. Then there would be some positive integer q, namely $q = p' - p$, so that $p' = p + q$.

By assumption $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_p, \mathbf{y}_{p+1}, \cdots, \mathbf{y}_{p'}$ constitute a basis for V.

By assumption $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p$ constitute a basis for V. Then they are are vectors in V and they are linearly independent.

Therefore, by the Replacement Theorem, after relabelling the indices of $\mathbf{y}_1, \mathbf{y}_2, \cdots, \mathbf{y}_p, \mathbf{y}_{p+1}, \cdots, \mathbf{y}_{p+q}$ if necessary, $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p, \mathbf{y}_{p+1}, \cdots, \mathbf{y}_{p+q}$ would constitute a basis for V.

Therefore, in particular, $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p, \mathbf{y}_{p+1}, \cdots, \mathbf{y}_{p+q}$ would be linearly independent.

Again recall that $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p$ constitute a basis for V. Since $\mathbf{y}_{p+1}$ is a vector in V, $\mathbf{y}_{p+1}$ would be a linear combination of $\mathbf{x}_1, \mathbf{x}_2, \cdots, \mathbf{x}_p$. Contradiction arises.

Therefore it would be impossible for $p < p'$ to hold.

Similarly, it would be impossible for $p > p'$ to hold.

Hence $p = p'$ in the first place.