1. **Definition. (Subspaces of subspace of $\mathbb{R}^n$.)**

   Let V, W be subspaces of $\mathbb{R}^n$.

   We say V is a subspace of W if and only if the statement (†) holds:

   (†) *For any $\mathbf{x} \in \mathbb{R}^n$, if $\mathbf{x} \in V$ then $\mathbf{x} \in W$.*

   Remark. In plain words, the statement (†) reads: '*every vector of V belongs to W*'.

2. **Illustrations.**

   (a) Let $A = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 1 & 3 & 3 & 5 \\ 2 & 6 & 5 & 6 \end{bmatrix}$, $B = \begin{bmatrix} 1 & 2 & 2 & 4 \\ 1 & 3 & 3 & 5 \end{bmatrix}$, $C = \begin{bmatrix} 1 & 3 & 3 & 5 \\ 2 & 6 & 5 & 6 \end{bmatrix}$, $D = \begin{bmatrix} 1 & 3 & 3 & 5 \end{bmatrix}$.

   With a direct application of the definition, we can show that:
   - $\mathcal{N}(A)$ is a subspace of $\mathcal{N}(B)$,
   - $\mathcal{N}(A)$ is a subspace of $\mathcal{N}(C)$,
   - $\mathcal{N}(A)$ is a subspace of $\mathcal{N}(D)$,
   - $\mathcal{N}(B)$ is a subspace of $\mathcal{N}(D)$, and
   - $\mathcal{N}(C)$ is a subspace of $\mathcal{N}(D)$.

   After some harder work, we can also show that:
   - $\mathcal{N}(B)$ is not a subspace of $\mathcal{N}(C)$, and
   - $\mathcal{N}(C)$ is not a subspace of $\mathcal{N}(B)$.

   (b) Let $\mathbf{x}_1 = \begin{bmatrix} 2 \\ -1 \\ 3 \\ 1 \\ 2 \end{bmatrix}$, $\mathbf{x}_2 = \begin{bmatrix} 1 \\ 2 \\ -1 \\ 5 \\ 2 \end{bmatrix}$, $\mathbf{x}_3 = \begin{bmatrix} 2 \\ 1 \\ -3 \\ 6 \\ 1 \end{bmatrix}$, $\mathbf{x}_4 = \begin{bmatrix} -6 \\ 7 \\ -1 \\ 1 \\ 1 \end{bmatrix}$, and

   $$T = \mathsf{Span}\,(\{\mathbf{x}_1, \mathbf{x}_2\}), U = \mathsf{Span}\,(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3\}), V = \mathsf{Span}\,(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_4\}), W = \mathsf{Span}\,(\{\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \mathbf{x}_4\}).$$

   We can show that:
   - T is a subspace of U,
   - T is a subspace of V,
   - T is a subspace of W,
   - U is a subspace of W, and
   - V is a subspace of W.

   After some harder work, we can also show that:
   - U is not a subspace of V, and
   - V is not a subspace of U.

3. Again recall the Replacement Theorem (Theorem (F)) in the handout *More on minimal spanning set*:

   Let W be a subspace of $\mathbb{R}^n$.
   Let $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_p, \mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_q$ be vectors in W. Suppose none of these vectors is the zero vector.
   Suppose $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_p$ are linearly independent.
   Further suppose $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_q$ constitute a basis for W.
   Then, $q \geq p$, and there is a basis for W which is constituted by $\mathbf{s}_1, \mathbf{s}_2, \cdots, \mathbf{s}_p$ together with some $q - p$ vectors from amongst $\mathbf{t}_1, \mathbf{t}_2, \cdots, \mathbf{t}_q$.

   This result, combined with the notion of *subspaces of a subspace of* $\mathbb{R}^n$, leads to Theorem (I), which is a useful tool for comparing various subspaces of $\mathbb{R}^n$.

4. **Theorem (I).**

   Let V, W be subspaces of $\mathbb{R}^n$.

   Suppose V is a subspace of W.

   Then $\dim(V) \leq \dim(W)$. Moreover, equality holds if and only if $V = W$.

   Remark. From the handout *Dimension*, we have learnt that

   - every subspace of $\mathbb{R}^n$ is of dimension at most n, and
   - $\mathbb{R}^n$ is the one and only one n-dimensional subspace of $\mathbb{R}^n$.

What we have learnt earlier can be seen as a manifestation of Theorem (I) in the special case in which $W = \mathbb{R}^n$.

5. **Proof of Theorem (I).**

   Let V, W be subspaces of $\mathbb{R}^n$.

   Suppose V is a subspace of W.

   Write $\dim(V) = k$. Pick some basis for V, say, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$. Then by assumption, each of them belongs to W.

   (a) By definition, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are linearly independent.
   So they are k linearly independent vectors in W.
   Hence $\dim(V) = k \leq \dim(W)$ by Theorem (H).

   (b) Suppose $V = W$. Then $\dim(V) = \dim(W)$.

   (c) Suppose $\dim(V) = \dim(W)$. Then $\dim(W) = k$.
   Pick some basis for W, say, $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_k$.
   Recall that $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ are k linearly independent vectors in W.
   Then by the Replacement Theorem, $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$, together with possibly some vectors from amongst $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_k$, constitute a basis for W.
   Since $\dim(W) = k$, the k vectors $\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k$ already constitute a basis for W.
   It follows that $W = \mathsf{Span}\ (\{\mathbf{v}_1, \mathbf{v}_2, \cdots, \mathbf{v}_k\}) = V$.

6. **Corollary (1) to Theorem (I).**

   Let V, W be subspaces of $\mathbb{R}^n$. Suppose $\dim(W) = p$.

   Further suppose V is a subspace of W.

   Also suppose that there are p vectors in V which are linearly independent.

   Then $V = W$.

7. **Proof of Corollary (1) to Theorem (I).**

   Let V, W be subspaces of $\mathbb{R}^n$. Suppose $\dim(W) = p$.

   Further suppose V is a subspace of W.

   Also suppose that there are p vectors in V which are linearly independent.

   - Since V is a subspace of W, we have $\dim(V) \leq \dim(W)$.
   - Since there are p vectors in V which are linearly independent, we have $\dim(V) \geq p = \dim(W)$.

   Then $\dim(V) = \dim(W)$. Now, by Theorem (I), since V is a subspace of W, we have $V = W$.

8. **Corollary (2) to Theorem (I).**

   Let V, W be subspaces of $\mathbb{R}^n$. Suppose $\dim(V) = p$.

   Further suppose V is a subspace of W.

   Also suppose that there are p vectors of W so that every vector of W is a linear combination of these p vectors.

   Then $V = W$.

9. **Proof of Corollary (2) to Theorem (I).**

   Let V, W be subspaces of $\mathbb{R}^n$. Suppose $\dim(V) = p$.

   Further suppose V is a subspace of W.

   Also suppose that there are p vectors of W, say, $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_p$, so that every vector of W is a linear combination of these p vectors.

   - Since $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_p$ are all vectors in W, every linear combination of $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_p$ is a vector in W.
     Then $W = \mathsf{Span}\ \{\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_p\}$.
     Therefore there is a basis for W from amongst the p vectors $\mathbf{w}_1, \mathbf{w}_2, \cdots, \mathbf{w}_p$. Hence $\dim(W) \leq p = \dim(V)$.
   - Since V is a subspace of W, we have $\dim(V) \leq \dim(W)$.

   Then $\dim(V) = \dim(W)$. Now, by Theorem (I), since V is a subset of W, we have $V = W$.

10. **Theorem (J). (Re-formulation of the notion of basis in terms of dimension.)**

    Let W be a subspace of $\mathbb{R}^n$. Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ be vectors in W.

    Then the statements below are logically equivalent:

    ($\sharp$) *$\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ constitute a basis for W.*

    ($\natural$) *$\dim(W) = p$, and $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ are linearly independent.*

($\flat$) $\dim(W) = p$, *and every vector of* W *is a linear combination of* $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$.

11. **Proof of Theorem (J).**

    Let W be a subspace of $\mathbb{R}^n$. Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ be vectors in W.

    [The statement ($\sharp$) implies each of the statements ($\natural$), ($\flat$) immediately.
    What matters is whether whether each of the statements ($\natural$), ($\flat$) separately implies ($\sharp$).]

    - [We ask whether ($\natural$) implies ($\sharp$).]
      Suppose $\dim(W) = p$ and $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ are linearly independent.
      Define $V = \mathsf{Span}\ (\{\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p\})$. [We want to show that $V = W$.]
      By definition, $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ constitute a basis for the p-dimensional subspace V of $\mathbb{R}^n$.
      Now we have $\dim(V) = p = \dim(W)$.
      Since $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ are vectors in W, every linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ belongs to W.
      Then V is a subspace of W. Now, by Corollary (2) to Theorem (I), we have $V = W$.
      Hence $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ constitute a basis for W.
    - [We ask whether ($\flat$) implies ($\sharp$).]
      Suppose $\dim(W) = p$, and every vector in W is a linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$.
      Since $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ are vectors in W, every linear combination of $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ is a vector in W.
      Then there is a basis for W, with, say, q vectors, from amongst the vectors $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$. Without loss of generality, assume they are $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_q$. These q vectors are linearly independent.
      Since these q vectors constitute a basis for W, we have $\dim(W) = q$.
      Then $p = \dim(W) = q$. Therefore $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ are linearly independent.
      Hence $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ constitute a basis for W.

12. **Theorem (J'). (Re-formulation of Theorem (J) in terms of systems of equations.)**

    Let W *be a subspace of* $\mathbb{R}^n$.
    Suppose $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ *be vectors in* W, *and* U *is the* $(n \times p)$*-matrix given by* $U = [\ \mathbf{u}_1 \mid \mathbf{u}_2 \mid \cdots \mid \mathbf{u}_p\ ]$.
    Then the statements below are logically equivalent:

    (a) $\mathbf{u}_1, \mathbf{u}_2, \cdots, \mathbf{u}_p$ *is a basis for* W.
    (b) $\dim(W) = p$, *and the homogeneous system* $\mathcal{LS}(U, \mathbf{0})$ *has no non-trivial solution.*
    (c) $\dim(W) = p$, *and for any* $\mathbf{b} \in V$, *the system* $\mathcal{LS}(U, \mathbf{b})$ *is consistent.*